***"Contribution à l'amélioration de la prévision hydrométéorologique opérationnelle.***

***Pour l'usage des probabilités dans la communication entre acteurs"***

**Thèse défendue par Benoît Houdant**

**Ecole Nationale du Génie Rural des Eaux et des Forêts**

Date de la soutenance : lundi 3 mai 2004

Lieu : EDF-DTG, 21 avenue de l'Europe, Grenoble

Composition du jury :

| | |
|---|---|
| Rapporteurs : | **Roman KRZYSZTOFOWICZ**<br>University of Virginia (USA) |
| | **Denis HILTON**<br>Université de Toulouse II - le Mirail |
| Directeur de thèse : | **Eric PARENT**<br>Ecole Nationale du Génie Rural des Eaux et des Forêts, Laboratoire de Gestion du Risque en Sciences de l'Environnement |
| Co-encadrant : | **Rémy GARCON**<br>Electricité de France, Division Technique Générale |
| Examinateurs : | **Frédéric ATGER**<br>Météo-France |
| | **Vincent FORTIN**<br>Institut de recherche d'Hydro-Québec (Canada) |
| | **Jean-Luc WYBO**<br>Ecole Nationale Supérieure des Mines de Paris, Pôle Cindyniques |

## Remerciements

Mes remerciements s'adressent tout d'abord aux prévisionnistes d'EDF-DTG, ceux de Grenoble et ceux de Toulouse, avec qui j'ai eu beaucoup de plaisir à travailler, et qui ont été au cœur du sujet de ma recherche.

Je remercie Roman Krzysztofowicz, professeur de statistiques à l'Université de Virginie (USA) et éditeur de la revue "Journal of Hydrology", et Denis Hilton, professeur de Psychologie Sociale à l'Université de Toulouse II, d'avoir accepté la tâche et la responsabilité d'être rapporteurs de ma thèse. Je leur en suis très reconnaissant.

Je remercie mon directeur de thèse Eric Parent d'avoir pensé à moi, alors que j'étais en première année de l'ENGREF, pour traiter ce sujet de recherche passionnant. Merci à lui de m'avoir judicieusement orienté durant ces années de recherche, notamment en m'encourageant à aller voir ce qui se faisait outre Atlantique. Je remercie également ce spécialiste des statistiques bayésiennes de sa bienveillance à mon égard lorsque mon questionnement m'a conduit à chercher méthodes et concepts dans d'autres disciplines : systémique, économie ou psychologie.

Je remercie Rémy Garçon, expert à EDF en hydrologie statistique et en prévision hydrométéorologique, pour l'attention, la rigueur et la curiosité intellectuelle et avec laquelle il m'a co-encadré pendant ces trois ans de recherche.

Je remercie Frédéric Atger, spécialiste de la prévision d'ensemble à Météo-France, ancien prévisionniste, et passionné par la question de la formulation des incertitudes en prévision opérationnelle. Merci à lui de m'avoir encouragé sur le chemin des probabilités subjectives, et de m'avoir entraîné sur le terrain de la psychologie cognitive et sociale.

Je remercie Jean-Luc Wybo, directeur du Pôle Cindyniques à l'Ecole des Mines de Paris d'avoir ouvert mon questionnement à d'autres disciplines en sciences des risques et des dangers, notamment pour prendre en compte la dimension organisationnelle de la gestion des risques.

Je remercie Vincent Fortin, hydrologue et statisticien à l'Institut de Recherche d'Hydro-Québec, de m'avoir accueilli pour trois semaines de mission à Hydro-Québec. Ce travail en collaboration a été riche d'enseignements sur la formulation des incertitudes par les prévisionnistes, qu'ils soient français ou québécois.

Je remercie Robert Julian, chef de la branche CADE à EDF-DTG, pour son enthousiasme et sa curiosité à l'égard de ma recherche. J'ai pu mener mes travaux dans des conditions matérielles optimales ; je lui en suis infiniment reconnaissant.

Merci à toute la branche CADE d'EDF-DTG pour l'ambiance fantastique dans laquelle j'ai eu beaucoup de bonheur à travailler durant ces trois années.

Merci à Jean-Michel Tinland, directeur d'EDF-DTG de m'avoir accueilli au sein de son unité.

Merci à François Lévêque et à Marc Ringeisen, tous deux économistes à EDF (le premier au Centre d'Ingénierie Hydraulique ; le second au Centre Opérationnel d'Optimisation) pour leur collaboration aux travaux concernant l'estimation de la valeur économique de la prévision.

Je remercie mes nouveaux employeurs à EDF, Jean Comby, directeur de l'Unité de Production Sud-Ouest, et Marc Bourdel, directeur du Groupement d'Exploitation d'Hydraulique Adour et Gaves, de m'avoir accordé le temps nécessaire à la finalisation du manuscrit et à la préparation de la soutenance.

Je remercie Anne-Lise pour son soutien énergique, en particulier durant les trois dernières semaines de rédaction… et sa relecture critique et attentive du manuscrit.

Je remercie notre fils Siméon d'avoir eu la patience d'attendre la fin de la rédaction du manuscrit pour faire son premier sourire à la vie.

# *"Contribution à l'amélioration de la prévision hydrométéorologique opérationnelle ;*

# *Pour l'usage des probabilités dans la communication entre acteurs"*

*Remerciements* *3*

*Glossaire* *9*

*1 Introduction générale* *11*

*2 Un électricien aux prises avec les caprices de la météo* *13*

**2.1 Contexte** **14**
2.1.1 L'hydroélectricité, cette énergie qui nous vient du ciel 14
2.1.2 La vulnérabilité du système production-transport-consommation à l'aléa hydrométéorologique. 14
2.1.3 Les acteurs d'EDF et la gestion de l'eau. 15

**2.2 La prévision opérationnelle, ou l'art de composer avec l'incertain** **17**
2.2.1 Outils et incertitudes de l'expert prévisionniste 17
2.2.2 L'élaboration de la prévision opérationnelle 18

**2.3 Pourquoi proposer un système de prévision probabiliste ?** **20**
2.3.1 Oser communiquer les incertitudes de prévision. 20
2.3.2 La nécessité de quantifier le risque de se tromper 20
2.3.3 Plaidoyer pour l'usage des probabilités dans la communication entre acteurs 22

**2.4 Pourquoi la prévision déterministe est-elle si tenace ?** **24**
2.4.1 Introduction 24
2.4.2 Des raisons pratiques 24
2.4.3 Des raisons de psychologie sociale 24
2.4.4 Des raisons liées à la formation du prévisionniste 24
2.4.5 Des raisons liées au partage d'expérience entre experts 25

**2.5 Conclusions** **26**

**2.6 Bibliographie** **26**

*3 La valeur économique de la prévision opérationnelle* *27*

**3.1 Introduction** **28**

**3.2 Un exemple de valorisation des prévisions à court terme : les prévisions à J+4 sur la chaîne de l'Ain** **30**
3.2.1 Contexte 30
3.2.2 Que vaut la prévision ? Le retour d'expérience du prévisionniste. 33
3.2.3 Que vaut la prévision ? Le point de vue de l'usager 37
3.2.4 Résultats 38
3.2.5 Discussion et perspectives 40

**3.3 Un exemple de valorisation des prévisions à long terme : les prévisions saisonnières sur le lac de Serre-Ponçon** **40**
3.3.1 Contexte 40

3.3.2 L'aléa hydrologique au cours des années 2001 et 2002 41
3.3.3 Méthode d'estimation de la valeur des prévisions saisonnières sur Serre-Ponçon 44
3.3.4 La valeur de la prévision pour « une année à neige » 47
3.3.5 La valeur de la prévision pour « une année sans neige » 50
3.3.6 La gestion réelle du lac de Serre-Ponçon en 2001 et 2002 55
3.3.7 Discussion et perspectives 56

**3.4 Conclusions sur la valorisation économique des prévisions opérationnelles 57**

**3.5 Bibliographie 58**

***4 Le pari de la prévision probabiliste 61***

**4.1 Comment éliciter des probabilités subjectives ? 62**
4.1.1 Introduction : révéler le savoir tacite du prévisionniste 62
4.1.2 Experts et machines en compétition sur le terrain des incertitudes 63
4.1.3 Comment amener le prévisionniste aux probabilités subjectives ? 65
4.1.4 Le cas particulier des prévisions de précipitation : des pdf mixtes 65

**4.2 L'introduction des probabilités subjectives en salle de prévision 66**
4.2.1 Objectifs 66
4.2.2 Le stage de sensibilisation à l'usage des probabilités en prévision opérationnelle 67
Expérience menée au CHA 68
Expérience menée au CHPMC 68
4.2.3 La prévision probabiliste au fil de l'eau 69

**4.3 Conclusion 71**

**4.4 Discussion : ce que nous croyons, ce que nous connaissons, et ce que nous croyons connaître… 71**

**4.5 Bibliographie 72**

***5 L'usage des probabilités est sujet à controverses 75***

**5.1 Le point de vue des systémiciens et sociologues des organisations 76**

**5.2 Le point de vue des économistes 77**

**5.3 Le point de vue avancé dans cette thèse 80**
5.3.1 Des probabilités pour analyser la perception des incertitudes lors de l'élaboration de la prévision 80
5.3.2 Des probabilités pour mesurer la valeur de la prévision pour l'usager en charge de l'optimisation de la production 81
5.3.3 Les probabilités, un nouvel outil pour une meilleure gestion du risque lié à la Sûreté Hydraulique 81
5.3.4 Conclusions 82

**5.4 Bibliographie 83**

***6 Les probabilités : une langue qui révèle et améliore la performance du système de prévision face aux enjeux de Sûreté 85***

**6.1 Sait-on mesurer la performance du système de prévision en matière de Sûreté Hydraulique? 86**
6.1.1 La vulnérabilité des installations hydroélectriques au risque de crue 86
6.1.2 Mesure de la Sûreté : l'approche systémique 86

6.1.3 Application de l'approche systémique à la gestion d'un ouvrage en crue 88
6.1.4 Mesure de la Sûreté par le « potentiel de danger » 89
6.1.5 Mesure de la Sûreté par la « manœuvrabilité » 89
6.1.6 Mesure de la Sûreté à l'aide de la notion de résilience 92
6.1.7 Conclusion 94

**6.2 Amélioration du système de prévision face aux événements extrêmes 94**
6.2.1 En quoi et comment la prévision des crues contribue-t-elle à la Sûreté Hydraulique ? 94
6.2.2 Les réponses des décideurs, et les conséquences pour notre étude 95
6.2.3 La performance de la prévision des épisodes de pluie intense 96
6.2.4 L'aptitude des experts à anticiper les épisodes de pluie intense 98
6.2.5 Comparaison à la performance des analogues 100
6.2.6 Conclusion 101

**6.3 Discussion 102**

**6.4 Bibliographie 103**

***7 Une langue qui révèle la sous-estimation des incertitudes de l'expert 105***

**7.1 Introduction 106**

**7.2 Un exemple de surconfiance dans un centre de prévision opérationnelle 109**
7.2.1 Contexte 109
7.2.2 Résultats 110
7.2.3 Interprétation et discussion 112

**7.3 Sensibiliser et former les prévisionnistes pour réduire la surconfiance 114**
7.3.1 Contexte 114
7.3.2 Résultats 115
7.3.3 Interprétation et discussion 115

**7.4 Assurer la pérennité de la fiabilité de l'estimation des incertitudes par les prévisionnistes 116**
7.4.1 Contexte 116
7.4.2 Résultats 117
7.4.3 Interprétation 119

**7.5 Discussion : Quand le prévisionniste se met à la place du décideur… 121**

**7.6 Résumé, conclusions et perspectives 123**

**7.7 Bibliographie 125**

***8 La relation entre surconfiance et biais : résultats expérimentaux 129***

**8.1 Introduction 130**

**8.2 Surconfiance et biais dans les réponses à un questionnaire de culture générale : pièges dans les questions ou biais cognitif ? 131**
8.2.1 Contexte 131
8.2.2 On observe de la surconfiance dans les réponses au questionnaire 132
8.2.3 A-t-on choisi des questions-pièges ? 134
8.2.4 Pièges dressés par l'expérimentateur ou surconfiance des sujets ? 137
8.2.5 Conclusions sur le test de culture générale 138

**8.3 Détection d'un biais dans le jugement des prévisionnistes. Doit-on y voir la source de la surconfiance ? 139**

8.3.1 Résultats à EDF 139
8.3.2 Résultats à Hydro-Québec 142

**8.4 Interprétation 143**

**8.5 Résumé, et conséquences pratiques en psychologie expérimentale. 145**

**8.6 Bibliographie 147**

***9 Une langue adaptée au contexte opérationnel ? 149***

**9.1 Le prévisionniste confronté à ce nouveau langage : premiers balbutiements. 150**
9.1.1 Introduction 150
9.1.2 Contexte 150
Protocole suivi au CHA 150
Protocole suivi au CHPMC 151
L'inconsistance produite par le mode opératoire 152
9.1.3 Résultats 152
9.1.4 Interprétation 153
Un protocole expérimental peu représentatif des conditions de travail habituelles des prévisionnistes 153
Des défauts de représentativité plus importants au CHA qu'au CHPMC 154
9.1.5 Conclusion 156
9.1.6 Discussion : faut-il placer le prévisionniste hors du monde ? 157
9.1.7 De la nécessité d'intégrer les enjeux de l'usager dans le processus d'élaboration de la prévision 159

**9.2 Interlude : peut-on séparer parole et pensée ? 160**

**9.3 Le prévisionniste face au pari : "Entre risque et incertitude, mon cœur balance". 162**
9.3.1 Introduction 162
9.3.2 Protocole 163
9.3.3 Résultats obtenus au CHPMC 163
9.3.4 Résultats obtenus au CHA 165
9.3.5 Interprétation et discussion 166

**9.4 Bibliographie 168**

***10 Conclusion générale 171***

**10.1 Bilan : finalement, que vaut le système de prévision ? 171**

**10.2 Voies d'amélioration du système de prévision 174**
10.2.1 Piloter les améliorations du système de prévision par leur performance pour l'usager 174
10.2.2 Améliorer l'efficacité du vecteur de communication de la prévision 174
10.2.3 Choisir des priors informatifs et utiles au prévisionniste 175

***Table des illustrations 179***

***Annexes 181***

## Glossaire

ANALOG :modèle de prévision probabiliste de précipitations d'EDF.

ARPEGE : modèle de prévision déterministe de Météo-France

CEPMMT : Centre européen de prévision météorologique à moyen terme. Par extension, modèle de prévision de ce centre (souvent abrégé par CEP).

CHA : Centre hydrométéorologique Alpes. Cellule de prévision opérationnelle d'EDF basée à Grenoble.

CHPMC : Centre hydrométéorologique Pyrénées et Massif Central. Cellule de prévision opérationnelle d'EDF basée à Toulouse.

CIH : Centre d'ingénierie hydraulique. Unité d'EDF réalisant les travaux d'ingénierie pour le parc hydraulique du groupe EDF, en France et à l'international.

CNPE : Centre nucléaire de production d'électricité, que l'on appelle plus couramment "centrales nucléaires".

COOP : Centre opérationnel d'optimisation de la production. Unité d'EDF en charge de l'optimisation des moyens de production à toutes échéances (de l'infra journalier au pluriannuel).

DRIRE : Direction régionale de l'industrie, de la recherche et de l'environnement. Organisme de tutelle des sites industriels, et notamment des installations hydroélectriques.

DEGS : Direction d'EDF-GDF Services. Entité d'EDF et de GDF, en charge de la distribution de l'électricité et du gaz aux particuliers et aux industriels.

DPFT : Différence première des fonctions de transfert. Un des modèles hydrologiques pluie-débit utilisés à EDF.

DPN : Direction de la production nucléaire. Entité d'EDF qui coordonne à l'échelle nationale les actions concernant la gestion du parc de production nucléaire.

DTG : Division Technique Générale. Unité d'EDF (450 personnes réparties sur 7 sites) réalisant prévisions, contrôles, essais, surveillance, diagnostic et conseils pour les exploitants du parc de production électrique national, et pour les filiales du groupe à l'international.

EDF : Electricité de France. Principal producteur français d'électricité, et distributeur national. Au 1er mars 2004, c'est un établissement public.

ENGREF, Ecole Nationale du Génie Rural des Eaux et des Forêts. Ecole d'ingénieur et 3e cycle universitaire (école d'application de l'Ecole Polytechnique, et de l'Institut National d'Agronomie de Paris-Grignon).

ESSH : Evénement significatif concertant la Sûreté Hydraulique. Il s'agit d'incidents ou d'accidents d'exploitation qui sont analysés et collectés au niveau national pour servir de support à des retours d'expérience.

GDF : Gaz de France. Premier acteur du gaz naturel en France : exploration-production, négoce, transport, stockage, distribution. Au 1er mars 2004, c'est un établissement public.

HQ : Hydro-Québec. Producteur national d'électricité au Québec (Canada).

MF : Météo-France. Service de prévision météorologique national.

MH : Mission hydraulique. Entité d'EDF qui coordonne à l'échelle nationale les actions concernant gestion du parc hydraulique.

MORDOR : Modèle à réservoirs de détermination optimale du ruissellement.Un des modèles hydrologiques utilisés à EDF.

pdf : probability density function. Fonction de densité de probabilité d'une variable aléatoire continue. (note n°103, p154).

PHV : poste hydraulique de vallée. C'est l'entité d'EDF qui gère (depuis un poste de téléconduite) la production d'une chaîne de centrales hydroélectriques sur une ou plusieurs vallées.

Q : variable qui désigne classiquement le débit, en hydrologie.

QMJ : débit moyen journalier.

Quantile de probabilité :

> Soit X une variable aléatoire sur [0, +∞[. Le quantile P sur cette variable aléatoire X, noté $X_P$ est le réel positif tel que : $Prob(X < X_P)=P\%$
>
> En modélisant un débit futur par une variable aléatoire Q sur [0, +∞[, on parle alors par exemple de « quantile 90 » (noté $Q_{90}$) sur cette prévision. $Prob(Q < Q_{90})=90\%$
>
> Le quantile 90 est la valeur qui n'aura, selon le prévisionniste, qu'une chance sur dix d'être dépassée.

RTE : Gestionnaire du réseau de transport électrique français. Entité indépendante d'EDF (au 1er mars 2004) en charge du transport d'électricité (lignes haute tension et très haute tension).

SH : Sûreté Hydraulique. "La Sûreté Hydraulique à EDF est constituée de l'ensemble des dispositions prises lors de la conception et de la gestion des aménagements hydroélectriques pour assurer la protection des personnes, des biens et de l'environnement contre les risques liés à l'eau et dus à la présence ou au fonctionnement des ouvrages" (*in* La politique de Sûreté Hydraulique, EDF, document interne).

TEM : taux d'événements manqués. C'est le nombre d'événements (crues, pluies intenses, etc.) dont l'occurrence n'avait pas été prévue sur le nombre total d'occurrences de ces événements durant une période donnée.

TFA : taux de fausses alertes. C'est le nombre d'événements prévus qui ne se sont pas réalisés sur le nombre total de fois où l'événement était prévu.

Trans-Energie : Gestionnaire du réseau de transport électrique québécois. Entité indépendante d'Hydro-Québec en charge du transport d'électricité (lignes haute tension et très haute tension)

# 1 Introduction générale

Pour gérer son parc de production d'hydroélectricité, EDF a développé son propre système de prévision hydrométéorologique. Deux centres de prévision opérationnelle, l'un à Grenoble, l'autre à Toulouse, se partagent la couverture du territoire national.

L'objet de notre recherche était de répondre à la question suivante:

**Quelle est la performance du système de prévision hydrométéorologique d'EDF, et comment peut-on l'améliorer?**

Il est de coutume de préciser, au départ d'un travail de thèse, la discipline choisie. L'exercice est ici délicat, étant donné la multiplicité des domaines explorés pour répondre à la question de recherche initiale. Nous citerons tous ces domaines, en précisant le cas échéant l'auteur ou l'ouvrage qui a plus particulièrement orienté notre cheminement.

Pour appréhender le fonctionnement du système de prévision d'EDF, nous avons tout d'abord travaillé en compagnie de ses prévisionnistes en **hydrométéorologie opérationnelle**. Il était en effet indispensable d'acquérir la compréhension du processus d'élaboration de la prévision, de pressentir ses diverses sources d'incertitudes, et surtout de cerner son utilité pour ses usagers (Chapitre 2).

L'**économie** (Walliser) nous a permis de mesurer la valeur de l'information pour l'usager, lorsque celui-ci avait un objectif d'optimisation de la production (Chapitre 3).

Il nous est cependant vite apparu qu'une bonne partie de l'information délivrée à l'usager ne passait pas par le canal officiel, tracé, et quantitatif des bulletins écrits, notamment quand les besoins de ce dernier concernaient la Sûreté Hydraulique. Des discussions téléphoniques entre prévisionniste et usager permettaient alors de préciser le risque hydrométéorologique (crue, tempête, neige collante, etc.).

Pour mesurer la valorisation de telles informations, il a fallu tout d'abord les recueillir et les conserver sur un support qui permette de réaliser a posteriori des retours d'expérience. La lecture (ou la fréquentation) de chercheurs en **météorologie** (Murphy, Atger) et en **hydrologie** (Krzysztofowicz), tous convaincus de l'importance de savoir correctement quantifier l'incertitude en prévision opérationnelle, nous a alors entraînés sur le chemin de la **prévision probabiliste** (Chapitre 4).

L'usage des probabilités faisant cependant débat entre chercheurs en **sciences humaines et sociales**, nous avons cherché à mettre en lumière la posture épistémologique de chaque protagoniste du débat (Chapitre 5). Nous avons ainsi pu retirer de ces diverses approches des éléments utiles pour répondre à notre question de recherche.

En nous plaçant dans le champ de la **systémique** (Piaget), nous pouvons ainsi mieux cerner en quoi la prévision est utile à l'usager en charge de la Sûreté Hydraulique. La prévision probabiliste peut constituer, pour ce type d'usager, une nouvelle information plus riche et mieux adaptée à la gestion des risques et des crises (Chapitre 6).

Les probabilités sont également un outil qui permet de mieux investiguer la perception puis la communication des incertitudes par les prévisionnistes. Une telle recherche relève de la **psychologie cognitive** (Kahneman, Slovic, Tversky), et met en lumière certains défauts cognitifs des prévisionnistes :sous-estimation des incertitudes, et biais dans les prévisions

émises. La **psychologie sociale** (Hilton) pose un autre regard sur le comportement des prévisionnistes en situation d'incertitude. Elle nous invite à étudier les signaux que les prévisionnistes perçoivent de leur environnement social et qui les poussent à agir de telle ou telle manière (Chapitres 7 et 8).

Pour que les informations quantitatives sur l'incertitude fournies par les prévisionnistes puissent effectivement se traduire en probabilité, il nous faudra néanmoins vérifier que les prévisionnistes respectent bien les axiomes comportementaux qui fondent la **statistique décisionnelle** (Savage) (Chapitre 9).

Si une telle profusion de disciplines a été abordée, nous n'avons pourtant pas le sentiment de nous être éparpillés. Avec toujours une seule question en tête, celle de la performance du système de prévision, nous avons recherché le maximum d'approches et d'outils qui pouvaient nous être profitable pour y répondre.

Nous convenons que notre démarche tenait du pari, mais la rencontre de chercheurs qui, sur une même question (le comportement d'un individu en univers incertain) ont des approches variées (normatives ou descriptives), avec des velléités divergentes (apprendre des individus ou enseigner aux individus), et des méthodologies différentes (analytique ou systémique) constituait une telle richesse que nous n'avons pas pu nous résigner à rester cloisonnés à une discipline.

Au terme de cette recherche, nous proposerons d'introduire l'outil des probabilités subjectives dans la production des prévisions opérationnelles fournies par les prévisionnistes d'EDF aux gestionnaires du parc de production. Ce mémoire de thèse a pour objet d'en démontrer et d'en mesurer la valeur ajoutée par rapport aux prévisions déterministes classiques. Il sélectionne des méthodes que nous avons voulues simples et intuitives pour une meilleure appropriation de ce nouveau langage par les prévisionnistes. Il en trace aussi les limites et cherche à identifier les écueils à éviter lors du déploiement de ces méthodes.

Atger, F., 2000, La prévision du temps à moyenne échéance en France, La Météorologie, 8$^{e}$ série, n°30.

Hilton, D., 1995. The social context of reasoning: conversational inference and rational judgment. Psychological Bulletin, vol. 118, 2, pp 248-271.

Kahneman, D., Slovic, P., Tversky, A., 1982. Judgment under uncertainty: Heuristics and biases. Cambridge University Press.

Krzysztofowicz, R., 2001. The case for probabilistic forecasting in hydrology, Journal of Hydrology, Vol. 249, 2-9.

Murphy, A. H., 1991. Probabilities, odds, and forecast of rare events, Weather and Forecasting, vol. 6, no 2, 302-307.

Piaget, J., 1967. Biologie et connaissance. Gallimard, Paris.

Savage, L. J., 1954, The Foundations of statistics. Wiley.

Walliser, 2000. L'Economie Cognitive, Odile Jacob.

# 2 Un électricien aux prises avec les caprices de la météo

*EDF, 1er producteur européen d'électricité, est particulièrement concernée par les événements météorologiques : la température influe sur la consommation, la pluie sur la production, les tempêtes sur le transport et la distribution d'électricité.*

*Pour anticiper au mieux ces évènements, EDF s'est dotée de deux cellules de prévision hydrométéorologique sur le territoire national. Mais si la prévision hydrométéorologique opérationnelle est l'art de composer avec des incertitudes de toutes natures, elle reste pourtant délivrée sous forme essentiellement déterministe.*

*Si nous souhaitons donner la place qu'elle mérite à la prévision probabiliste, il nous faut d'abord comprendre pourquoi prévisionnistes et usagers continuent à utiliser un mode d'expression déterministe.*

## 2.1 Contexte

### 2.1.1 L'hydroélectricité, cette énergie qui nous vient du ciel

Le parc hydroélectrique d'EDF produit de l'énergie renouvelable à la mesure de ce que le ciel lui offre. Avec ses 450 centrales hydroélectriques, le cycle de l'eau, des lacs de montagne jusqu'à la mer, est mis à profit pour valoriser les écoulements en énergie. L'électricien français contrôle en effet 75% des eaux de surface du territoire national.

Contrairement à l'électricité, l'eau peut se stocker. C'est l'un des atouts majeurs de l'électricité d'origine hydraulique au sein du mix énergétique français : moduler la production et répondre rapidement aux pointes de consommation électrique[1]. Ainsi, l'usine hydroélectrique du Pouget, dans l'Aveyron, est capable de fournir au réseau de transport électrique 300 MW en moins de 3 minutes ; sa proche voisine, la centrale à charbon d'Albi, met quant à elle 7h pour fournir au réseau 250 MW (avec une consommation horaire de 100 tonnes de charbon). Les usines hydroélectriques contribuent ainsi à la sûreté du système électrique, en démarrant ou en modulant à chaque instant leur production en fonction de l'équilibre production-consommation. Notons que l'exploitant d'une tranche nucléaire module peu la production de celle-ci, afin de limiter les contraintes mécaniques sur les cuves du réacteur. Une telle gestion contribue à accroître la durée de vie des centrales françaises.

### 2.1.2 La vulnérabilité du système production-transport-consommation à l'aléa hydrométéorologique.

La production, le transport et la consommation d'électricité sont chacun soumis aux divers aléas météorologiques.

Production

Les centrales nucléaires hydroréfrigérées ne peuvent produire que si le débit et la température du fleuve (qui constitue leur source froide) restent dans un intervalle établi par l'autorité de tutelle. En particulier, les centrales nucléaires sont contraintes d'arrêter leur production dans les cas suivants :

- étiage sévère : augmentation de la concentration des polluants physico-chimiques émis dans le milieu naturel ;
- crue : risque de perte de la source froide à cause d'embâcles ;
- température trop élevée : risque de mortalité piscicole due à la surchauffe de l'eau du fleuve par la centrale ;
- gel du fleuve : perte de la source froide.

Une vague de froid au cœur de l'hiver peut faire geler des cours d'eau comme la Loire, affectant alors la production des centrales nucléaires hydroréfrigérées, à une période où la demande est très forte.

Contrairement à ce type d'aléa identifié dès la conception des centrales nucléaires françaises, la concomitance de la canicule et de la sécheresse de l'été 2003 a constitué un aléa qui n'avait jamais été pointé comme critique pour l'équilibre offre-demande. La faiblesse des débits et la

[1] L'éolien, autre source d'énergie renouvelable, a pour inconvénient majeur d'être une source d'énergie non programmable : le vent, pas plus que l'électricité ne se peut se stocker. D'où l'idée actuelle de valoriser les usines hydroélectriques réversibles (pompe-turbine) pour stocker et réguler l'énergie éolienne produite. On rappelle que le rendement global du cycle pompe-turbine vaut, sur le parc français actuel, 80%.

température élevée de l'eau ne permettaient plus d'assurer le refroidissement des centrales dans le respect des règles environnementales.

Les usines hydroélectriques sont quant à elles particulièrement vulnérables aux crues, qui entraînent souvent l'arrêt de la production, et peuvent parfois mettre à mal la Sûreté Hydraulique (6.1.1). En effet, comme tout industriel, et quelles que soient ses conditions d'exploitation (en crue ou non), EDF doit « exercer son activité dans des conditions de nature à prévenir les risques auxquels son activité est susceptible d'exposer des tiers ». C'est ce qu'a rappelé le 12 juin 1998 la Cour d'Appel de Grenoble, lors du procès consécutif à la catastrophe du Drac[2].

Que ce soit pour la sécurité des tiers ou pour le respect de l'environnement, EDF est tenue de respecter la règle suivante : Ne pas aggraver les conséquences d'une crue du fait de l'existence ou du fonctionnement de ses ouvrages.

Transport et distribution

La tempête du 27 décembre 1999 (durant laquelle en quelques heures 3,5 millions de foyers français furent privés d'électricité) ou des épisodes de neige collante (comme celui du 4 décembre 2003 dans l'Ariège, privant plus de 10 000 foyers d'électricité) montrent à quel point le réseau de transport électrique est lui aussi vulnérable à l'aléa météorologique.

Consommation

Si la météorologie conditionne les capacités de production, elle a également un impact sur la consommation d'électricité. En plein cœur de l'hiver, une baisse de 1°C sur le territoire national entraîne une augmentation de la consommation pouvant aller jusqu'à 1500 MW (principalement du fait du chauffage électrique). En plein été, une augmentation de 1°C provoque une surconsommation de 400 MW[3].

### 2.1.3 Les acteurs d'EDF et la gestion de l'eau.

Nous avons distingué deux types d'enjeux liés à l'eau :

- valoriser au mieux la ressource en eau en fonction de la consommation (2.1.1) ;
- exploiter et maintenir l'outil de production, ou le système de transport d'électricité en toute sécurité et dans le respect de l'environnement(2.1.2).

Pour gérer ces enjeux, nous pouvons identifier deux types d'acteurs :

◊ ***Max*** *travaille à l'optimisation de la production électrique. On l'appelle ainsi car il cherche à maximiser la valorisation en énergie de la ressource en eau. Ses objectifs sont d'une part de ne pas déverser au barrage des $m^3$ qui auraient pu être turbinés à l'usine en contrebas s'il avait anticipé la montée des débits ; d'autre part, il cherche à optimiser le stockage de l'eau pour pouvoir la turbiner lorsque le coût marginal de production d'électricité est le plus élevé (ou que le prix d'échange du kWh sur la bourse de l'énergie est le plus élevé).*

[2] Le 5 décembre 1995, 6 enfants et leur accompagnatrice sont morts lors d'un lâcher d'eau d'un barrage hydroélectrique sur la rivière Drac au sud de Grenoble.

[3] Jusqu'à l'été 2003, cet "effet climatisation" était dû principalement à l'augmentation de la consommation des réfrigérateurs et congélateurs domestiques. Depuis la canicule qui a frappé la France en août 2003, entraînant une surmortalité de 15 000 personnes, on peut penser que le nombre de climatiseurs domestiques a augmenté. Il est toutefois difficile de d'évaluer la surconsommation induite par ces nouveaux équipements.

◊ ***Jéo*** *travaille à la maintenance du parc de production hydroélectrique. Si on le nomme ainsi, c'est parce qu'il constitue un maillon essentiel de la Sûreté Hydraulique, et qu'il est très sensible aux dangers de toutes sortes qui peuvent la mettre en péril (jeopardy signifie en anglais danger, péril). Jéo s'intéresse à des événements hydrologique ou météorologique peu probables, mais très dommageables à l'entreprise s'ils se produisent.*

Dans leurs missions respectives, Max et Jéo utilisent les prévisions élaborées par les prévisionnistes des deux centres hydrométéorologiques d'EDF[4], basés à Grenoble et à Toulouse : le CHA (Centre Hydrométéorologique Alpes), et le CHPMC (Centre Hydrométéorologique Pyrénées Massif Central).

La double compétence des prévisionnistes en hydrologie et en météorologie leur permet d'accroître leurs capacités d'anticipation du risque de crue : ils peuvent prévoir, dans une certaine mesure (6.2.4), les précipitations à venir, et ne se contentent pas de suivre les processus hydrologiques une fois les pluies observées au sol. Cette double compétence confère également au système de prévision souplesse et réactivité pour répondre à de nouveaux besoins des usagers. Ainsi, dès le début du mois de juin 2003, les CH ont anticipé sur la première vague de chaleur qui s'est abattue sur la France en proposant de leur propre initiative des prévisions de température de l'eau des fleuves au droit des centrales nucléaires hydroréfrigérées. Quelques semaines plus tard, cette information allait devenir vitale pour l'exploitant des centrales nucléaires, dans ses discussions avec les autorités de tutelle (Ministère de l'Industrie, Ministère de l'Ecologie et du Développement Durable).

Notons que les CH bénéficient de l'avantage indéniable d'avoir des usagers identifiés pour chacune de leurs prévisions. Cela permet donc une contractualisation de l'offre, pour fournir des prévisions « sur mesure » aux différents usagers, en terme de variable à prévoir, d'échéance, de délais d'anticipation, de niveau d'incertitude, de fréquence de remise à jour de l'information, etc.[5] Une telle connaissance de l'usager permet également de réaliser des retours d'expérience sur épisode (crue, tempête, étiage, etc.), pour analyser la performance du système de prévision et la valorisation qui en est faite par l'usager, et trouver ensemble des voies d'amélioration.

Cette collaboration est également fructueuse en temps réel. Durant une crue par exemple, l'usager peut joindre le prévisionniste d'astreinte par téléphone, et obtenir des compléments d'information pour gérer son installation en fonction de contraintes ponctuelles (défaillance d'un automate, embâcles dans la retenue, personnel empêché, etc.).

[4] Ces deux centres sont dotés des mêmes outils, et réalisent les mêmes prestations, chaque centre s'occupant de la moitié du territoire national.

[5] Malheureusement le dialogue sur l'adéquation de l'offre au regard des objectifs de l'usager et de l'état de l'art en météorologie n'est pas encore assez abouti pour aborder de tels sujets avec chacun des usagers (6.2).

## 2.2 La prévision opérationnelle, ou l'art de composer avec l'incertain

### 2.2.1 Outils et incertitudes de l'expert prévisionniste

Pour répondre aux enjeux de Max et de Jéo, les prévisionnistes disposent de différents outils de mesure et de modélisation, tant en météorologie qu'en hydrologie (Tourasse, 1991) :

- un réseau de mesure au sol (pluviomètres, thermomètres, perches à neige, stations hydrométriques), dont les données enregistrées sont relevées périodiquement ou télétransmises en temps réel ;
- des images Radar et satellitaires sur la situation atmosphérique ;
- des modèles météorologiques : ceux de Météo-France (ARPEGE et ALADIN, Atger, 2000), celui du CEPMMT (Centre européen de prévision à moyen terme, Atger, 2000), et le modèle ANALOG (Obled et al. 2002).
- des modèles hydrologiques sur environ 50 bassins

Aux incertitudes des mesures au sol s'ajoutent les difficultés d'interpolation entre les mesures ponctuelles, l'hétérogénéité spatiale des précipitations étant particulièrement marquée sur les bassins de montagne. Le Radar, qui complète utilement les images satellitaires pour donner une vision spatiale de l'épisode en cours, souffre actuellement de plusieurs limites. Il est d'une part très onéreux et ne peut pas de ce fait être installé dans chaque vallée de montagne. D'autre part, il connaît encore aujourd'hui de gros problèmes de calibration qui rendent extrêmement délicate l'évaluation automatique de la lame d'eau spatiale par un traitement de l'image Radar (Garçon, 2002).

A ces incertitudes concernant les observations atmosphériques viennent bien entendu s'ajouter les incertitudes de modélisation météorologique et hydrologique.

Ajoutons également les aléas technologiques : défaillance des télétransmissions, indisponibilité du réseau de mesure au sol, des sorties de modèles météorologiques, des observations Radar ou satellitaires, etc. Cette accumulation de défaillances, que Krzysztofowicz (1999) nomme "operational uncertainty" avant de l'écarter de son champ d'analyse, est parfois la principale source d'incertitude[6]. En gestion de crue, ces défaillances conditionnent bien souvent la décision, prise en univers particulièrement incertain. C'est l'une des raisons pour lesquelles la production opérationnelle de prévisions est confiée à un prévisionniste expérimenté : celui-ci peut plus facilement rétablir une connexion, s'enquérir par d'autres canaux d'information sur les conditions locales, se faire une idée de l'évolution même en situation dégradée, etc. Le prévisionniste contribue à la robustesse du système de gestion de crue.

Toutes ces incertitudes rendent la prévision extrêmement délicate. Atger (2001) en donne un exemple pour la détection des pluies intenses à J+1 par des modèles déterministes et ensemblistes. Le prévisionniste croisera quant à lui les différentes sorties d'un modèle (température pseudo-adiabatique du thermomètre mouillé, vitesse verticale ou tourbillon absolu des masses d'air, etc.), et même différents modèles atmosphériques, pour fournir son avis d'expert sur le risque d'occurrence de pluie intense à 24h. Tout son art consistera à

[6] F. Atger (comm. pers.) estime que durant les crues du Gard des 8 et 9 septembre 2002 (22 victimes), les défaillances de mesure furent limitées, de même que les défaillances du réseau de transmission. Pourtant, les prévisionnistes n'ont pas pu avoir accès en temps réel à la totalité des données d'observation. Le logiciel de récupération des données était mal adapté, et les procédures étaient méconnues ou mal documentées. Il y a eu également des problèmes d'accès à des données pourtant disponibles, du fait de faiblesses des logiciels de visualisation.

fournir une information valorisable compte tenu de l'incertitude résiduelle et incompressible inhérente à l'état de l'art actuel en matière de prévision météorologique.

### 2.2.2 L'élaboration de la prévision opérationnelle

Comme le rappelle Krzysztofowicz (2001), prévoir n'est pas prédire. L'art de la prévision opérationnelle n'est donc pas seulement de tenter d'annoncer dans un délai suffisant les événements futurs. Il s'agit également d'intégrer les incertitudes de toutes natures et de livrer un avis d'expert sur l'incertitude liée à la prévision, sur la vraisemblance de telle ou telle évolution météorologique ou hydrologique, sur la probabilité d'occurrence de tel ou tel phénomène dit « à risque » : crue, orage, neige collante, vents violents, etc. D'où l'importance des explications qualitatives qui accompagnent la communication des prévisions déterministes.

Pour ce faire, le prévisionniste exerce son expertise et son esprit critique à deux niveaux : en sortie des modèles météorologiques, puis, en aval, en sortie des modèles hydrologiques (pour le cas des bassins modélisés, Figure 1).

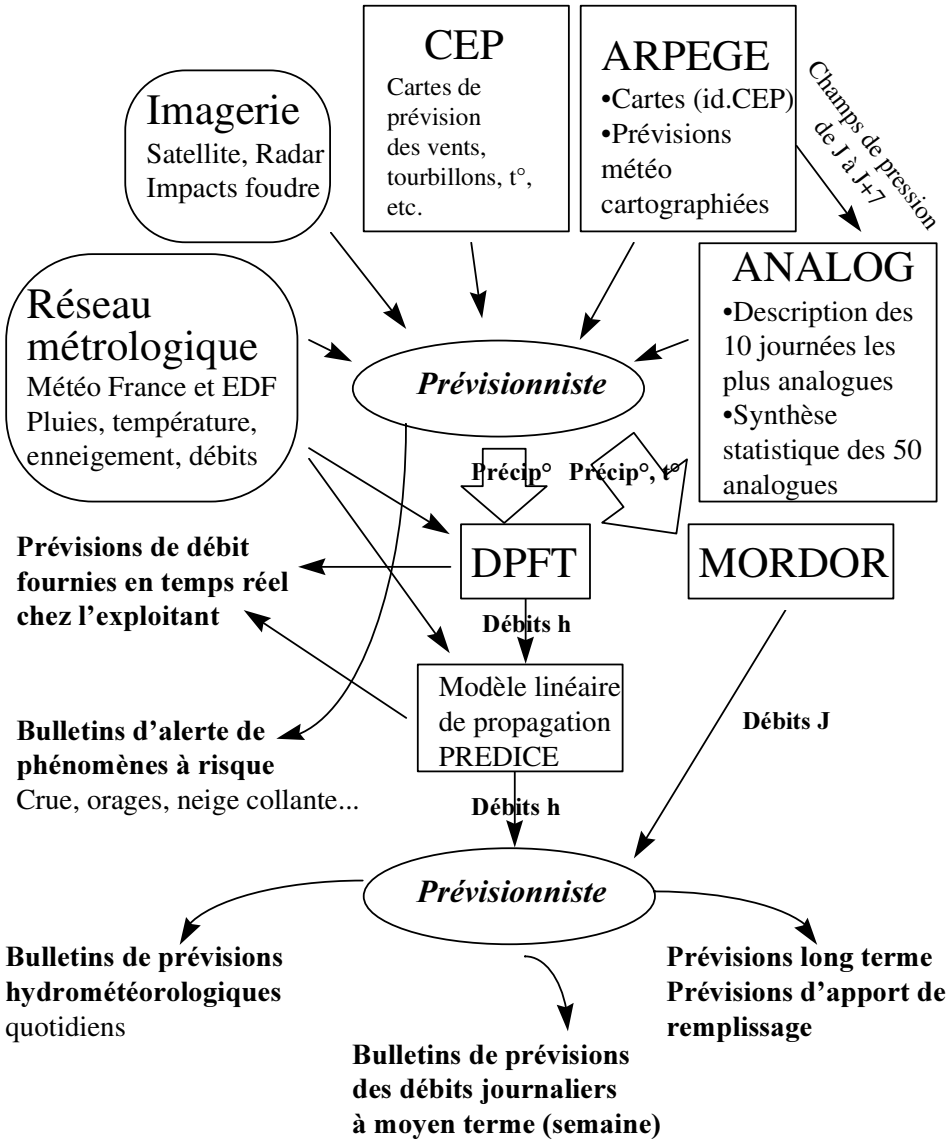



**Figure 1 : Elaboration des prévisions opérationnelles à EDF**

Il existe une seule boucle en temps réel sans intervention humaine (sauf réinitialisation ou invalidation ponctuelle des données) : c'est celle qui à partir des données observées en temps réel sur le réseau envoie directement sur serveur chez l'exploitant les vecteurs de prévisions de débits (après traitement des données par un ou plusieurs modèles hydrologiques).

Toutes les autres prévisions découlent d'un minutieux travail d'élaboration de la prévision. Décrivons-en à présent les grandes étapes.

Première étape

Le prévisionniste commence par prendre connaissance de l'évolution récente de la situation atmosphérique, grâce au réseau de mesure au sol, et à l'imagerie satellitaire et Radar. Il compare la situation en cours aux sorties du modèle de la veille, et à sa propre prévision de la veille ; il réalise ainsi un rapide retour d'expérience. Ce retour d'expérience constitue, jour après jour, un apprentissage du comportement des modèles atmosphériques (CEPMMT, ARPEGE et ALADIN). Le prévisionniste apprend aussi à connaître ses propres réactions face à tel ou tel épisode météorologique.

Deuxième étape

Le prévisionniste analyse ensuite les différentes cartes de sortie des modèles de prévision, pour faire le point sur la situation météorologique en cours et sur ses évolutions probables, en considérant divers paramètres physiques (pression au niveau de la mer, épaisseur de la couche 1000-700 hPa, température pseudo-adiabatique du thermomètre mouillé, instabilité absolue, tourbillons absolus, etc.).

Troisième étape

Le prévisionniste peut alors se concentrer sur les variables qui l'intéressent : principalement le champ de précipitation (et éventuellement, en hiver, l'altitude de l'isotherme 0°C). Il consulte ainsi les sorties du modèle physique ARPEGE, ainsi que celles du modèle statistique ANALOG, qui lui fournit les hauteurs de précipitations provoquées, dans le passé, par des situations météorologiques assez proches de celle en cours[7].

Le prévisionniste prévoit également d'autres variables du temps sensible intéressant l'usager : la température, la nébulosité, la vitesse du vent en rafale, les orages, et, en hiver, la neige collante.

On peut dès lors faire la remarque suivante : l'hydrométéorologiste élabore sa prévision grâce à un savant équilibre entre la physique du phénomène, et l'analogie qu'il fait avec d'autres situations plus ou moins semblables rencontrées dans le passé.

NB : On peut présupposer que l'analyse physique conduit le prévisionniste à adopter une vision déterministe de l'avenir. L'approche par analogues l'inviterait en revanche à considérer l'avenir comme un ensemble de possibles plus ou moins probables, et de ce fait à adopter une vision probabiliste de l'avenir.

Notons toutefois que dans son analyse physique, le prévisionniste réalise également des analogies en référence à des cas d'école théoriques (genèse d'un orage pseudo stationnaire, système en V, retour d'Est, etc.).

[7] L'indice de proximité est construit sur les niveaux des surfaces 1000 et 700 hPa chaque jour en 37 points de l'Europe de l'ouest et de la bordure atlantique (Duband, 1981, Obled et al., 2002).

Quatrième étape

Le prévisionniste traduit en débit ses prévisions de pluie et de température, à l'aide des modèles hydrologiques ou sur la base de ses propres heuristiques sur les bassins non modélisés.

## 2.3 Pourquoi proposer un système de prévision probabiliste ?

### 2.3.1 Oser communiquer les incertitudes de prévision.

Lorsque l'expert est sollicité pour une prévision météorologique, il évoque rarement les incertitudes que lui-même pressent sur son propre jugement. Quand bien même ces incertitudes sont évoquées, celles-ci ne servent bien souvent qu'à nuancer qualitativement son avis, ou à le conditionner à d'autres sources d'informations ou à de futures expertises. A titre d'exemple, on trouve classiquement dans un bulletin de prévision en météorologie ou en hydrologie des phrases du type : « Risque à confirmer dans notre prochain bulletin (demain 9h) ».

Bien rares sont les fois où l'expert va jusqu'à formaliser ses propres incertitudes dans l'état de connaissance qui est le sien au moment où il doit rendre son jugement[8]. Pourtant, à mieux quantifier ses propres incertitudes, le prévisionniste transmettrait, en plus d'une donnée, **un savoir**. Il fiabiliserait ainsi la communication avec l'usager.

Transmettre un savoir sur l'état de l'art de la science hydrométéorologique

« L'hydrométéorologie est-elle une science exacte ? ». Si la réponse est non, il convient de s'interroger sur l'ordre de grandeur des incertitudes associées aux chiffres annoncés.

Krzysztofowicz (2001) donne un exemple éloquent à propos de la crue de l'été 1997 sur la Red River, à Grand Forks dans le Dakota : les prévisionnistes avaient annoncé un pic de crue à 15 m (sachant que le niveau d'alerte de crue était de 8,5 m). En fait, le niveau de la rivière a atteint la cote de 16,5 m. La digue, construite conformément aux strictes prévisions, a donc été submergée, et la ville de Grand Forks dévastée par l'inondation.

Après coup, on imagine facilement l'étonnement d'un hydrologue à la lecture du témoignage d'un conseiller municipal de Grand Forks : *"The National Weather Service continued to predict that the river's crest at Grand Forks would be 49 ft (15m). […] If someone had told us that these estimates were not an exact science, […] we may have been better prepared."*

### 2.3.2 La nécessité de quantifier le risque de se tromper

Les deux cellules de prévision opérationnelle (CHA et CHPMC) délivrent des prévisions hydrométéorologiques que l'on peut qualifier de déterministes. L'incertitude n'est quantifiée que pour les prévisions hydrologiques à long terme (plusieurs mois), par l'indication des

[8] Il est vrai que bien souvent, on ne lui en donne pas les moyens, en lui demandant de rendre son jugement sous un format imposé : un chiffre, une réponse binaire (« Y a-t-il risque oui ou non ? »). Par exemple, les prévisionnistes de Météo-France ne sont pas encouragés à indiquer le niveau d'incertitude de leurs prévisions dans les bulletins destinés au grand public (à l'exception de l'indice de confiance valable pour les moyennes échéances). Il leur est demandé au contraire d'utiliser un style reflétant le déterminisme de la prévision et l'absence de doute quant à la réalisation des faits évoqués (par exemple, utilisation du présent de l'indicatif dans les bulletins enregistrés). (Atger, comm. pers.)

quantiles 10, 50 et 90. Les autres prévisions hydrologiques ou météorologiques ne donnent que des indications qualitatives sur le niveau d'incertitude qu'elles comportent.

Dans les bulletins émis quotidiennement, on trouve par exemple « fort risque de franchissement du seuil des 25 m3/s » pour une prévision hydrologique, ou « 10-30mm » pour une lame d'eau spatiale. Qu'entend-on par fort risque : plus d'une chance sur deux, ou neuf chances sur dix ? Que met-on derrière cette fourchette sur la lame d'eau spatiale : un intervalle de confiance à 80%, à 50% ? Les différents prévisionnistes de l'équipe ont-ils bien les mêmes concepts lorsqu'ils donnent cette fourchette ? L'utilisateur a-t-il la même référence en tête ?

Tant le prévisionniste que l'usager sont pourtant des habitués de la gestion des incertitudes liées aux prévisions hydrométéorologiques, parce qu'il a toujours fallu « faire avec », et prendre des décisions malgré ce flou persistant et irréductible. Même pour un vecteur de prévision déterministe, comme un débit moyen journalier de J à J+5, l'utilisateur de la prévision a bien dû faire l'apprentissage, avec le temps, des incertitudes liées à ce type de prévision. Il a intuitivement des notions de l'écart « classique » entre le chiffre annoncé et le chiffre mesuré suivant l'échéance de la prévision. L'utilisateur n'agit pas comme si la prévision allait être parfaite, mais se fait une idée - incontrôlable par le prévisionniste - de la fiabilité de l'information qu'il utilise dans sa prise de décision (par exemple en prenant des « marges de sécurité »).

Cependant, on comprend vite que dans ces conditions, le prévisionniste n'a aucune prise sur la perception de l'incertitude de la prévision par l'usager. Si aujourd'hui, le prévisionniste est bien plus (ou bien moins) sûr de sa prévision quotidienne que la semaine dernière, comment le dire, et faire en sorte que l'usager puisse en tenir compte ?

A ne rien communiquer sur l'incertitude, le seul recul dont dispose l'usager d'une prévision déterministe est la crédibilité de cette prévision à ses yeux. Celle-ci est le fruit des rapprochements quotidiens entre les prévisions de la veille et ses observations du matin, mais aussi de la concordance entre le bulletin hydrométéorologique et ses autres sources d'information (notamment son observation du temps sensible).

Le prévisionniste est donc condamné à constater l'évolution (parfois irréversible) de l'appréciation de sa prévision par l'exploitant hydraulique. Il aurait pourtant pu ménager la crédibilité du système de prévision, ainsi que sa légitimité d'aide à la décision, s'il avait mentionné le degré d'incertitude pressenti pour chaque prévision, et pour chaque échéance[9].

D'où l'intérêt de mettre en place une méthode pour aider le prévisionniste à formaliser, puis à communiquer à chaque type d'usager des informations sur ce niveau d'incertitude. En d'autres termes, lui faire part non seulement de ses connaissances, mais également de ses connaissances sur ses connaissances ; mettre en lumière les limites de ce qu'il avance, en acceptant de regarder les frontières de son savoir, tout en donnant un message clair et concis, qui apporte à l'usager un gain de lucidité. Celui-ci pourra alors prendre une décision rapide et efficace, en (mé)connaissance de cause.

[9] Weber et Hilton (1990) ont montré que chaque individu associe aux mots "risque", "probable", "possible", etc., des probabilités qui dépendent du contexte dans lequel il se trouve. Ainsi, à défaut de préciser quantitativement un niveau de risque (i.e. une probabilité d'occurrence), le prévisionniste perd le contrôle de l'estimation de ses incertitudes de prévision. Cette estimation est laissée à l'interprétation de l'usager et à l'arbitraire des circonstances qui l'entourent.

### 2.3.3 Plaidoyer pour l'usage des probabilités dans la communication entre acteurs

Tant que l'usager ne disposera pas d'une prévision parfaite, il devra procéder à des arbitrages entre ses diverses exigences envers le système de prévision hydrométéorologique. Nous allons voir dans cette section que ces arbitrages ne sont possibles que si des incertitudes de prévisions quantifiées lui sont communiquées.

Le premier arbitrage, universel en matière de prévision, est celui du couple (**délai d'anticipation** ; **niveau d'incertitude**). L'usager doit se prononcer sur un niveau d'incertitude acceptable en fonction du délai d'anticipation souhaité, ou réciproquement, un délai d'anticipation acceptable en fonction du niveau de certitude souhaité. En somme, définir ses priorités et connaître ses besoins et ses enjeux pour demander une prévision adaptée. Reste que si les délais d'anticipation souhaités et réels sont faciles à mesurer (un chronomètre suffit !), la mesure du niveau d'incertitude d'une prévision demande une réflexion bien plus complexe pour l'usager et le prévisionniste :

- si la prévision est déterministe et quantitative (vecteur de prévision de lames d'eau spatiales journalières ou de débits journaliers de J à J+5), une mesure a posteriori de l'incertitude de prévision pourrait être la moyenne ou la borne supérieure des écarts entre les prévus et les observés ;
- si la prévision est qualitative, et porte sur des événements hydrométéorologiques dangereux (crue, orage, vent violent, neige collante, etc.), comment juger de la pertinence de communications telles que « risque de précipitations intenses dans la soirée » ?

Si certains usagers veulent connaître la tendance centrale de la prévision, pour optimiser la gestion de la ressource en eau par exemple, tous veulent connaître les risques que la situation météorologique fait planer sur eux, même si les niveaux de risques sont faibles[10]. Les responsables de la maintenance des installations hydroélectriques et de la sûreté des ouvrages d'EDF sont particulièrement attentifs à ce deuxième type de prévision. Ils réagissent à des **franchissements de seuil** (de débit, de pluie, de vitesse de vent par exemple) qui entraînent diverses manœuvres d'exploitation ou de protection des personnes et des biens : évacuation d'un chantier dans le lit d'une rivière, abaissement du plan d'eau de la retenue jusqu'à la cote minimale d'exploitation, communication à la Sécurité Civile, etc.

Ces usagers ont des exigences incontournables quant aux délais d'anticipation : ils connaissent bien le temps qu'il leur faut pour engager les manœuvres de sécurité, préparer les équipes d'astreinte, mettre en place une cellule de crise, etc. Ces délais d'anticipations fixés, les usagers sont donc confrontés à l'incertitude irréductible que connaissent les modèles et les hydrométéorologistes pour des prévisions à ces échéances. Cette incertitude se traduira cette fois-ci, non par un écart entre le prévu et l'observé, mais par les deux défauts caractéristiques de ce type de prévision :

- une certaine proportion de franchissements de seuils qui n'avaient pas été anticipés dans les délais impartis (des **événements manqués**),
- une proportion de **fausses alertes**.

[10] Les usagers des prévisions hydrométéorologiques ne sont pas tous demandeurs d'information sur les risques en prévision, mais pourtant chacun d'entre eux est susceptible d'en vouloir aux prévisionnistes après l'occurrence d'un épisode météorologique sévère. Après diverses catastrophes, c'est souvent le premier chef d'accusation, notamment contre Météo-France : « ne pas avoir averti de l'existence d'un risque alors qu'on en avait connaissance »

En l'état actuel de la science, il faut donc mettre l'usager en face d'un deuxième arbitrage : est-il prêt à tolérer de nombreuses fausses alertes pour obtenir un taux très bas d'événements manqués (exemple de la Sûreté Hydraulique), ou au contraire refuse-t-il les fausses alertes, qui lui coûteraient plus cher qu'un événement non anticipé ?

Atger (2001) montre comment l'usage de la prévision probabiliste permet d'optimiser les prévisions qui sont transmises aux usagers, chacun d'eux étant défini par sa sensibilité aux événements manqués comparativement à sa sensibilité aux fausses alertes. Nous montrerons également en 6.2.4 que la prévision probabiliste permet de régler le couple (taux de fausses alertes, taux d'événements manqués)[11] le moins défavorable à l'usager.

En résumé :

- Si l'usager s'intéresse à l'optimisation de la production (comme Max), les délais d'anticipation nécessaires pour valoriser l'information imposent le niveau d'incertitude de la prévision. Ce niveau d'incertitude se traduira a posteriori en terme d'écarts entre prévus et observés. Une indication sur ces écarts (moyenne, borne supérieure, etc.) lui permettra de connaître les limites de la prévision aux diverses échéances qu'il a fixées, et donc de tirer le meilleur profit du système de prévision.
  La donnée quotidienne, a priori, du niveau d'incertitude liée à la prévision lui permettra de prendre selon les jours des marges de sécurité variables dans l'optimisation de la production, et de ne pas faillir si les observations ne correspondent pas rigoureusement aux prévisions. Il pourra ainsi, si le niveau d'incertitude est jugé trop grand, mobiliser certains moyens de productions alternatifs (turbines à combustion), qu'il n'utilisera que si les apports en eaux sont plus faibles que ceux escomptés.
- Si l'usager s'intéresse à l'annonce d'un événement dommageable (comme Jéo), les délais d'anticipation nécessaires pour prendre des mesures de protection imposeront l'ensemble des couples (TFA, TEM)[12] entre lesquels il peut choisir. L'usager sera orienté dans son choix par sa vulnérabilité aux événements manqués, et par le coût d'une fausse alerte. La donnée quotidienne, a priori, du niveau de risque attendu (c'est à dire la probabilité de franchissement d'un seuil dans les jours à venir) lui permettra de prendre chaque jour des décisions adéquates pour assurer la Sûreté Hydraulique sur ses aménagements.

Seules les prévisions probabilistes permettent de quantifier a priori et en temps réel le niveau de performance du système de prévision. Ce n'est que par la communication des incertitudes de prévision que l'usager peut être placé face à ses responsabilités, dans son utilisation quotidienne de la prévision hydrométéorologique, outil performant, mais encore perfectible.

Dans de telles conditions, on peut alors se demander pourquoi, dans la grande majorité des centres de prévision opérationnels, la prévision est encore communiquée sous forme déterministe.

---

[11] Le taux de fausses alertes (TFA) est la proportion du nombre d'alertes données à tort sur le nombre total d'alertes communiquées.
Le taux d'événements manqués (TEM) est la proportion du nombre d'événements observés qui n'avaient pas été anticipés sur le nombre total d'événements observés.

[12] Pour chaque délai d'anticipation, à niveau de performance des modèles et/ou des prévisionnistes fixé, il existe une infinité de couples (TFA, TEM) -variant entre (0,1) et (1,0)- que peut produire le système de prévision. Ces deux paramètres sont liés par une relation biunivoque, ce qui revient, pour l'usager, à choisir un TFA, et à se voir imposer un TEM, ou réciproquement.

## 2.4 Pourquoi la prévision déterministe est-elle si tenace ?

### 2.4.1 Introduction

Si la prévision hydrologique à long-terme est probabiliste, tous les autres produits de prévision délivrés sont de nature déterministe. Dans la littérature, on s'en étonne, voire on s'en désole (Krzysztofowicz, 2001). La prévision probabiliste semble représenter la quintessence de l'information communicable, l'information la plus riche et la mieux utilisable par divers types d'usagers (Atger, 2001). Pourtant la majorité des systèmes de prévision délivre encore des prévisions déterministes, même si celles-ci sont parfois accompagnées d'indications qualitatives concernant le niveau d'incertitude, ou d'un indice de performance qui n'est pas à proprement parler une prévision probabiliste.

Il est pourtant entendu que la météorologie est l'un des domaines où les incertitudes sont importantes et les ratés lourds de conséquences (section 2.3) ! Le biais du déterminisme, qui consiste à ne pas même évoquer les incertitudes de prévisions, a de multiples sources[13]. Les hypothèses ci-dessous pourraient expliquer la ténacité de la prévision déterministe dès lors qu'il s'agit de prévision opérationnelle.

### 2.4.2 Des raisons pratiques

L'information déterministe est plus facile à comprendre pour l'usager.

Que ce soit sur des considérations déterministes ou probabilistes, l'usager finira toujours par prendre une décision déterministe : turbiner ou non, se protéger ou non, etc. On lui complique donc sérieusement la tâche en lui délivrant plutôt, qu'un simple chiffre, une information probabiliste telle qu'une pdf (fonction de densité de probabilité), où la totalité des états de la nature sont assignés d'une vraisemblance.

### 2.4.3 Des raisons de psychologie sociale

En délivrant une prévision déterministe, qui ne fait aucune mention des incertitudes associées, le prévisionniste peut conforter son statut d'expert. Fischhoff (1982, p 439) a montré certaines motivations qui conduisent les experts à négliger d'évoquer leurs incertitudes : « Bien entendu, si on fait appel à de tels experts (et si on les paye) à la mesure de la confiance qu'ils inspirent, ils pourraient être tentés de masquer l'étendue de leur ignorance ».

### 2.4.4 Des raisons liées à la formation du prévisionniste

L'enseignement académique, donné à un futur expert (le météorologiste en particulier), bien souvent axé sur le triptyque sémiotique-diagnostic-pronostic, peut favoriser sa tendance au déterminisme.

- **Sémiotique** : on lui apprend à reconnaître les signes qui vont lui permettre de réaliser un diagnostic. Ces signes sont, pour le météorologiste, des observations atmosphériques telles que la nébulosité ou les précipitations, ou des mesures atmosphériques, comme la pression ou la température.

[13] Nous verrons au Chapitre 7 que la tendance au déterminisme entraîne une sous-estimation des incertitudes de prévision lorsqu'on demande explicitement au prévisionniste d'évaluer celles-ci.

- **Diagnostic** : l'expert apprend à définir une situation météorologique en référence à un cas d'école : front chaud, système en V, etc.. Si la sémiotique est descriptive, le diagnostic est le fruit de la réflexion de l'expert, qui va "repérer des phénomènes météorologiques, objets conceptuels appartenant à une théorie et reconnaissables par un ensemble de manifestations atmosphériques" (Atger, 2000).
- **Pronostic**: l'expert apprend à prévoir l'évolution de ces phénomènes à l'aide de modèles allant de simples techniques d'extrapolation à des modèles très sophistiqués. Le météorologiste apprend à interpréter et utiliser les sorties des modèles atmosphériques pour élaborer sa propre prévision météo. La dernière étape de cette phase de pronostic consiste à expliciter l'impact de la prévision pour l'usager : Quel sera le temps sensible ? Quelle est la lame d'eau spatiale attendue ? etc.

Ainsi, la sémiotique joue un rôle clé dans la formation qui est donnée au futur expert : c'est elle qui va lui permettre d'accéder à tous les travaux de modélisation de ses précurseurs et de ses pairs. Leurs modèles, concepts et théories lui seront en effet indispensables pour élaborer son diagnostic de la situation. L'insistance sur la reconnaissance des signes révélateurs de chaque situation météorologique est telle que ces signes deviennent des attracteurs pour le météorologiste : à défaut de bien comprendre la situation atmosphérique, il se rattache à des signes qu'il croit voir, simplement parce qu'il les connaît.

C'est ainsi que peut apparaître le paradoxe suivant : plus grande est l'échéance de la prévision, moins le prévisionniste est capable de prendre en compte ses incertitudes, et plus il se concentre sur un scénario qui lui semble « le bon ». Lorsque la situation est complexe ou l'échéance lointaine, traiter conjointement deux tâches (prévision des évolutions atmosphériques et évaluation des incertitudes sur sa prévision) lui demande un travail de computation trop difficile.

Cela explique que plus les questions sont difficiles (c'est à dire plus l'échéance de la prévision augmente), plus l'illusion du déterminisme est importante. On montrera au Chapitre 7 (Tableau 18) que le prévisionniste peut avoir tendance à fournir des intervalles de confiance plus étroits à J+4 qu'à J.

### 2.4.5 Des raisons liées au partage d'expérience entre experts

L'apprentissage individuel et collectif participe également à cette tendance au déterminisme de l'expert.

La mémoire individuelle ne peut s'exercer que sur des objets pensés, compris et interprétés. De simples observations, avant d'être intelligibles (c'est à dire devenues noumènes, modélisées) ne permettent que très difficilement au travail de mémoire de s'exercer. Une phase de sémiotique est donc nécessaire avant tout apprentissage individuel. Ce travail de sémiotique, par l'insistance exagérée qu'il porte aux différents signes distinctifs qui permettent une classification, accentue l'illusion du déterminisme de l'expert, c'est à dire l'illusion de pouvoir prédire le temps avec certitude.

Le partage d'expérience, moteur de l'apprentissage collectif, ne peut fonctionner que sur la base d'un langage commun. Or qu'est-ce qu'un langage sinon un système de signes ? La sémiotique, qui entraîne avec elle une part indubitable de caricature, est donc le préalable indispensable à tout partage d'expérience au sein d'une équipe d'experts.

## 2.5 Conclusions

Répondre aux enjeux évoqués en 2.1.1 et 2.1.2 nécessite d'adopter un mode de communication de la prévision fondé sur l'usage des probabilités (Chapitre 6), même si les théoriciens des différentes sciences sociales ne sont pas unanimes sur la question (Chapitre 5). Il nous faudra donc mettre en place un protocole adapté au contexte opérationnel, afin d'aider les prévisionnistes à produire des prévisions probabilistes expertisées (Chapitre 4).

Les probabilités permettront en outre au prévisionniste de mieux se connaître, afin de mieux gérer son éventuelle sous-estimation chronique des incertitudes (Chapitre 7), le biais de ses prévisions (Chapitre 8), ou enfin son comportement en univers risqué ou incertain (Chapitre 9). C'est en essayant de mieux se comprendre qu'il pourra mieux se faire comprendre, notamment lorsqu'il parle de risque.

En ce qui concerne l'utilisation de la prévision pour l'optimisation de la production hydroélectrique, voyons dès à présent ce que vaut le système de prévision actuel (Chapitre 3).

## 2.6 Bibliographie

Atger, F., 2001. Verification of intense precipitation forecasts from single models and ensemble prediction systems. Non Linear Processes in Geophysics, 8, 401-417.

Atger, F., 2000, La prévision du temps à moyenne échéance en France, La Météorologie, 8$^{e}$ série, n°30.

Duband, D., 1981. Prévision spatiale des hauteurs de précipitations journalières, La Houille Blanche, n° 7/8.

Garçon, R., 2002. L'imagerie satellitaire et Radar pour la prévision hydrologique : l'expérience d'EDF, La Houille Blanche, n°2-2002.

Garçon, R., 1996. Prévision opérationnelle des apports de la Durance à Serre-Ponçon à l'aide du modèle MORDOR, bilan de l'année 1994-1995. La Houille Blanche, vol.5.

Krzysztofowicz R., 1999. Bayesian theory of probabilistic forecasting via deterministic hydrologic model, Water Resources Research, Vol. 35, n°9, 2739-2750.

Krzysztofowicz, R., 2001. The case for probabilistic forecasting in hydrology, Journal of Hydrology, Vol. 249, 2-9.

Obled, C., Bontron, G., Garçon, R., 2002. Quantitative precipitation forecasts: a statistical adaptation of model outputs through an analogues sorting approach, Atmospheric Research, vol. 63.

Paquet, E., Garçon, R., 2002. Caprices du climat et de l'hydrologie en Haute Durance : nos prévisions d'apports plurimensuels sont-elles encore fiables ?, La Houille Blanche, vol.8-2002.

Tourasse, P., 1991. Prévision et Télésurveillance Hydrométéorologique à EDF, La Houille Blanche, n°2-1991.

Weber, E., U., Hilton, D., J., 1990. Contextual effect in the interpretations of probability words: perceived base rate and severity of events, Journal of experimental psychology, Human perception and performance, vol. 16, 4, 781-789.

# 3 La valeur économique de la prévision opérationnelle

*Que valent nos prévisions ? Pour répondre à cette question légitime, les prévisionnistes ont tôt fait de se lancer dans des retours d'expérience et du contrôle de la qualité des informations émises, grâce à des critères statistiques de distance entre prévus et observés. Si cela permet de pointer certaines défaillances du système de prévision, et d'en tracer les améliorations, cela ne permet pas de répondre aux questions suivantes :*

- *Comment l'usager se débrouille-t-il sans prévision ?*
- *Quel est le potentiel du système de prévision ? Autrement dit, à supposer que l'information délivrée soit parfaite, comment l'usager saurait-il la valoriser ? Combien lui rapporterait-elle ?*
- *Combien rapportent effectivement nos prévisions à l'usager ? Quel serait son manque à gagner s'il ne disposait pas de nos informations ?*

*Autant de questions qui nous amèneront à faire le grand saut : sortir de la salle de prévision, et rejouer, avec l'usager, la gestion de la ressource en eau qu'il aurait eue suivant le type d'informations fournies sur les débits à venir : prévision dite naïve (modèle de persistance, ou distribution historique des débits), prévision parfaite (avenir connu), ou quelque part entre ces deux extrêmes****, la prévision opérationnelle.***

## 3.1 Introduction

Il ne fait aucun doute que la prévision météorologique est utile à de nombreux usagers qui exercent des activités aussi diverses que la navigation aéronautique ou la production agricole. Dès le début des années 1960, c'est à dire avant l'arrivée des modèles numériques de l'atmosphère[14], de nombreux auteurs ont cherché à mesurer cette utilité pour l'usager (Johnson et Holt, 1997, pour une revue complète de la littérature sur le sujet). Dans la plupart des études, les usagers modélisés étaient des céréaliers ou des maraîchers utilisant diverses prévisions météorologiques (prévision de gel, de pluie intense, ou de température min. et max.).

L'estimation de la valeur économique de la prévision hydrologique a débuté peu après (entre autres Davis et al., 1979, et Krzysztofowicz, 2001). Dans tous les cas, il s'agissait de répondre à la question de prime abord très simple : « Combien vaut la prévision pour l'usager ? ».

Les premières études visant à répondre à cette question ont adopté un **point de vue normatif**, qui reste aujourd'hui encore majoritaire. Dans le cadre de la théorie de la décision en univers incertain (Savage, 1954), l'usager est supposé rationnel (au sens des 4 axiomes énoncés au Chapitre 5). Les enjeux de l'usager sont souvent modélisés de manière très schématique :

- soit on affecte à l'usager une fonction de coût proportionnelle à l'écart quadratique entre prévus et observés (Krzysztofowicz, 2001) ;
- soit (plus généralement) on suppose l'usager sensible à un seuil unique pour lequel il engage des actions de protection. Les coûts associés sont alors le coût de la protection, et le coût des dommages occasionnés par le franchissement du seuil en l'absence de protection (littérature abondante, de Lave, 1963 à Atger, 2001 ; Johnson et Holt, 1997, pour une revue complète).

Ces modélisations, très normatives sur la rationalité du décideur et réductrices sur ses enjeux, permettent néanmoins de construire des cas d'école. Ceux-ci aboutissent à des ordres de grandeur sur la valeur potentielle de la prévision, en supposant l'usager parfait (Stewart, 1997). Ces modélisations ont également une vocation prescriptive, pour montrer aux usagers comment utiliser au mieux une information qui leur est fournie (en utilisant par exemple le théorème de Bayes, Krzysztofowicz, 2001).

Face à ces modèles normatifs, plusieurs **modèles descriptifs** ont été proposés (Stewart, 1997 et références incluses). Si ceux-ci ont rarement permis de produire des estimations quantitatives de la valeur de l'information, les résultats obtenus ont suggéré que les usagers ne valorisent pas au mieux l'information de prévision qui leur est fournie, et ce, pour plusieurs raisons (Stewart, 1997) :

- manque de souplesse dans le processus décisionnel : la force des habitudes dans le cas de décisions individuelles ou l'inertie de la hiérarchie dans le cas de décisions d'entreprises peuvent être un obstacle à la prise en compte de l'information ;
- manque de prise de conscience de l'existence de l'information et des moyens d'y accéder ;

[14] A cette époque, on utilisait pour la prévision à moyen terme (i.e. au delà de J+2) « des méthodes statistiques reposant sur l'analogie entre différentes situations météorologiques. Sans grand succès. » (Atger, 2000).

- erreurs de compréhension ou d'interprétation de la prévision ;
- utilisation d'informations locales plus accessibles (évolutions du temps sensible).

Ces études sont délicates, dans la mesure où il est difficile de discerner, à partir des décisions prises par l'usager "dans la vraie vie", sur quelles informations celui-ci s'est effectivement appuyé (McQuigg, 1971).

D'autre part, il est peu probable de rencontrer un usager qui aurait complètement explicité son processus d'intégration de la prévision dans sa prise de décision, et saurait chiffrer la valeur de cette information. Ainsi, le principal obstacle à la compréhension de la règle de décision de l'usager est le manque de représentativité[15] (au sens de Brunswik, 1956) des expérimentations visant à comprendre les mécanismes du processus décisionnel.

Une solution alternative à ces expérimentations serait de demander directement à l'usager de la prévision d'estimer la valeur que celui-ci accorde à cette information, par une **méthode d'évaluation contingente**. Cette méthode est utilisée principalement en économie de l'environnement pour quantifier la valeur de biens non marchands (la qualité de l'air, la préservation d'une espèce en voie de disparition, etc.). Elle consisterait dans notre cas à interroger l'usager sur son **consentement à payer** ("Willingness to pay") pour obtenir l'information de prévision. Ce trade-off, supposé traduire la valeur de la prévision pour l'usager, pourrait cependant souffrir des mêmes limites que celles rencontrées en économie de l'environnement. Bateman *et al*. (1995) ont en effet montré que ce type d'évaluation traduisait parfois l'idée d'un don symbolique de la part de la personne interrogée(« je fais un geste »). Stewart (1997) rapporte que Brown et Murphy (non publié) ont interrogé des entreprises de gaz naturel sur leur consentement à payer pour obtenir des prévisions climatiques. Les réponses obtenues ne leur ont pas permis d'établir un estimateur valide de la valeur de telles prévisions, soit que les estimations étaient éminemment variables selon les personnes ou les compagnies interrogées, soit que les personnes refusaient tout simplement de répondre. Cette faiblesse de l'évaluation contingente semble être la raison pour laquelle cette méthode reste marginale, et confinée à l'économie de l'environnement.

Quel que soit le mode d'expérimentation adopté, les échanges entre services météorologiques et usagers semblent difficiles, et de ce fait rares. Il existe en effet une grande diversité d'usagers utilisant la même prévision, avec des enjeux spécifiques à chacun, et des règles de décision informulées, fondées sur un sentiment insaisissable (l'intuition), et parfois versatiles[16].

Les défauts de communication semblent provenir de la césure entre services météorologiques d'une part et les divers usagers d'autre part.

Rares sont les entités comme EDF qui, de la prévision météorologique jusqu'à la gestion en temps réel d'un aménagement, intègrent toute la chaîne de décision. C'est cela qui nous a permis de sortir de la cellule de prévision opérationnelle pour aller évaluer avec les outils de

[15] La représentativité est ici employée dans le sens d'adéquation avec l'environnement naturel de l'usager (Brunswik, 1956), afin que celui-ci dispose des mêmes critères de décision, règles et heuristiques que lors de ses prises de décision habituelles.

[16] Par exemple Mroz et Raven (1993) rapportent de leur étude sur la compréhension des prévisions météorologiques par le grand public le constat suivant : "les prévisions météorologiques –quelle qu'en soit leur nature- sont généralement peu comprises, souvent mal interprétées, voire parfois sciemment manipulées par le grand public".

l'usager la valeur d'une prévision hydrologique en terme d'optimisation de la production hydroélectrique.

Nous étudions dans les deux sections suivantes (3.2 et 3.3) la valeur de prévisions opérationnelles, c'est à dire effectivement fournies à l'usager. Il s'agit d'une part d'un exemple de prévisions déterministes à moyen terme (du jour J au jour J+4), et d'autre part de prévisions probabilistes à long terme (données sous forme de 50 chroniques équiprobables).

## 3.2 Un exemple de valorisation des prévisions à court terme : les prévisions à J+4 sur la chaîne de l'Ain

### 3.2.1 Contexte

L'aménagement hydroélectrique de l'Ain, avec ses 6 usines équipées d'une puissance totale de 450 MW, est un complexe hydroélectrique d'intérêt national. La chaîne de production hydroélectrique de la vallée de l'Ain est particulièrement délicate à gérer pour les raisons suivantes :

- il s'agit d'une « vallée en Y » sur laquelle les aménagements hydroélectriques sont répartis sur le cours d'eau principal (Vouglans, Saut-Mortier, Coiselet, Bolozon et Allement), et sur l'un de ses affluents (Moux, sur l'Oignin), Figure 3 ;
- l'usine de Vouglans, « suréquipée », peut turbiner jusqu'à 360$m^3$/s (lors des pointes de consommations), alors que l'usine de Saut Mortier est limitée à 220$m^3$/s, le réservoir tampon ayant de plus une capacité extrêmement limitée. La chaîne de production hydroélectrique est saturée en quelques heures lorsque l'usine de Vouglans tourne à plein régime. Notons que le producteur a interdiction de déverser[17] sur les aménagements en aval de Vouglans, sauf bien entendu en période de crue ;
- il existe entre Vouglans, le lac de tête, et la retenue de Cize-Bolozon, un bassin versant intermédiaire (BVI) important (Figure 2 ) ;
- EDF doit en outre intégrer dans sa gestion deux contraintes réglementaires : l'entreprise doit d'une part maintenir la cote du réservoir de tête (celui de Vouglans), au-dessus d'une cote qui permette la pratique d'activités récréatives du 1er juillet au 31 août. D'autre part, EDF doit assurer en permanence en aval du dernier aménagement de ce complexe (celui d'Allement) un débit minimal de 12 $m^3$/s du 1er juin au 30 novembre (débit réservé), et de 28$m^3$/s du 1er décembre au 31 mai (convention pêche).

Dans ces conditions, on comprend que le gestionnaire attende des prévisionnistes une prévision performante à moyen terme. Les bulletins de prévision, émis simultanément pour les débits entrant à Vouglans et pour les apports du bassin versant intermédiaire (que nous nommerons par la suite "BVI Bolozon"), permettent d'anticiper au mieux les éventuels problèmes de saturation en aval de Vouglans (en particulier après la confluence avec la Bienne). Emis deux fois par semaine, ils présentent des vecteurs de prévision des débits moyens journaliers de J à J+4.

[17] Il y a déversement lorsque les capacités de turbinage sont insuffisantes, et que la retenue est remplie : les débits déversés transitent alors par les évacuateurs de crue qui équipent le barrage.

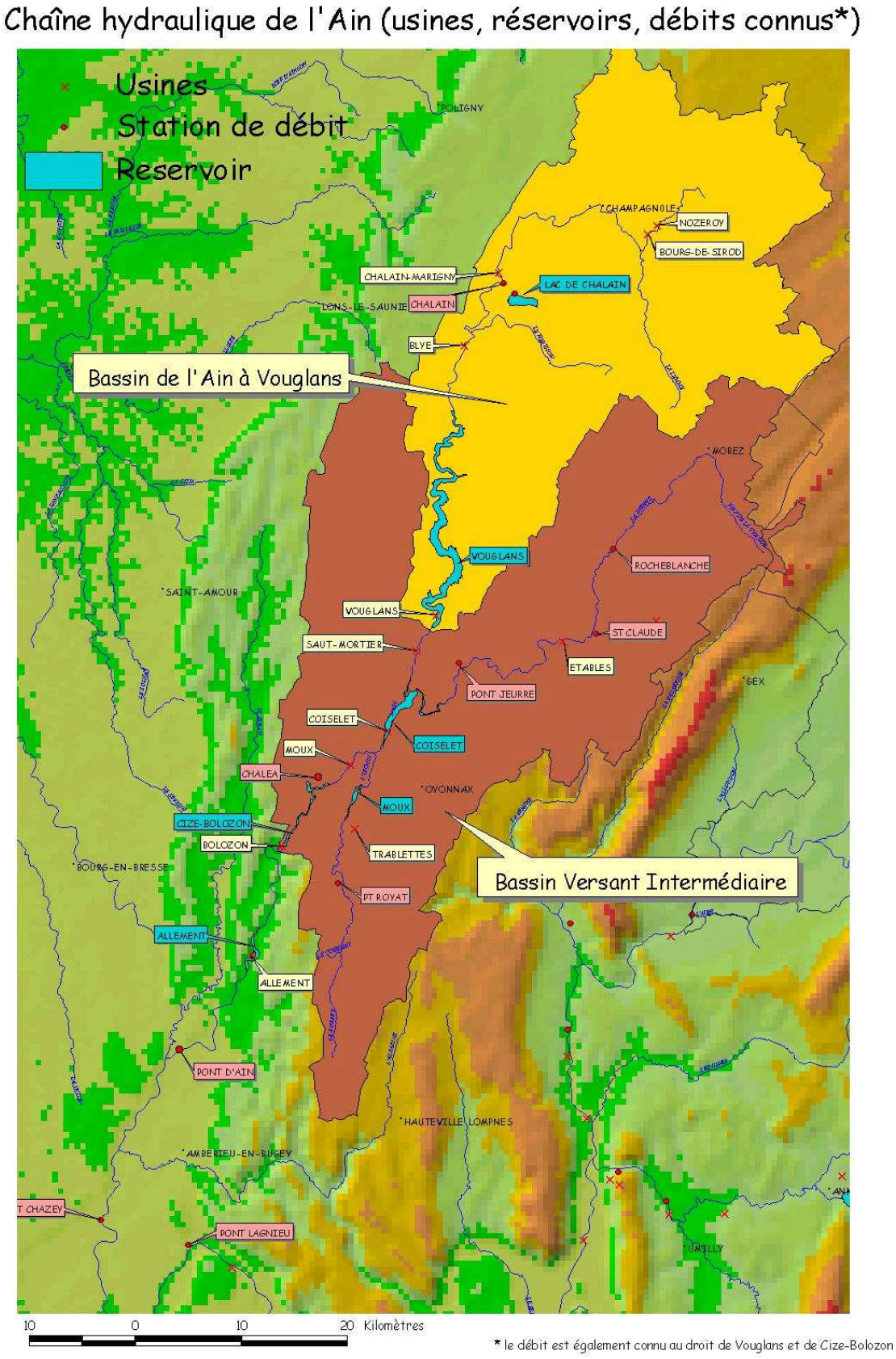



**Figure 2 : Localisation du bassin versant de l'Ain à Vouglans et du BVI à Bolozon.**

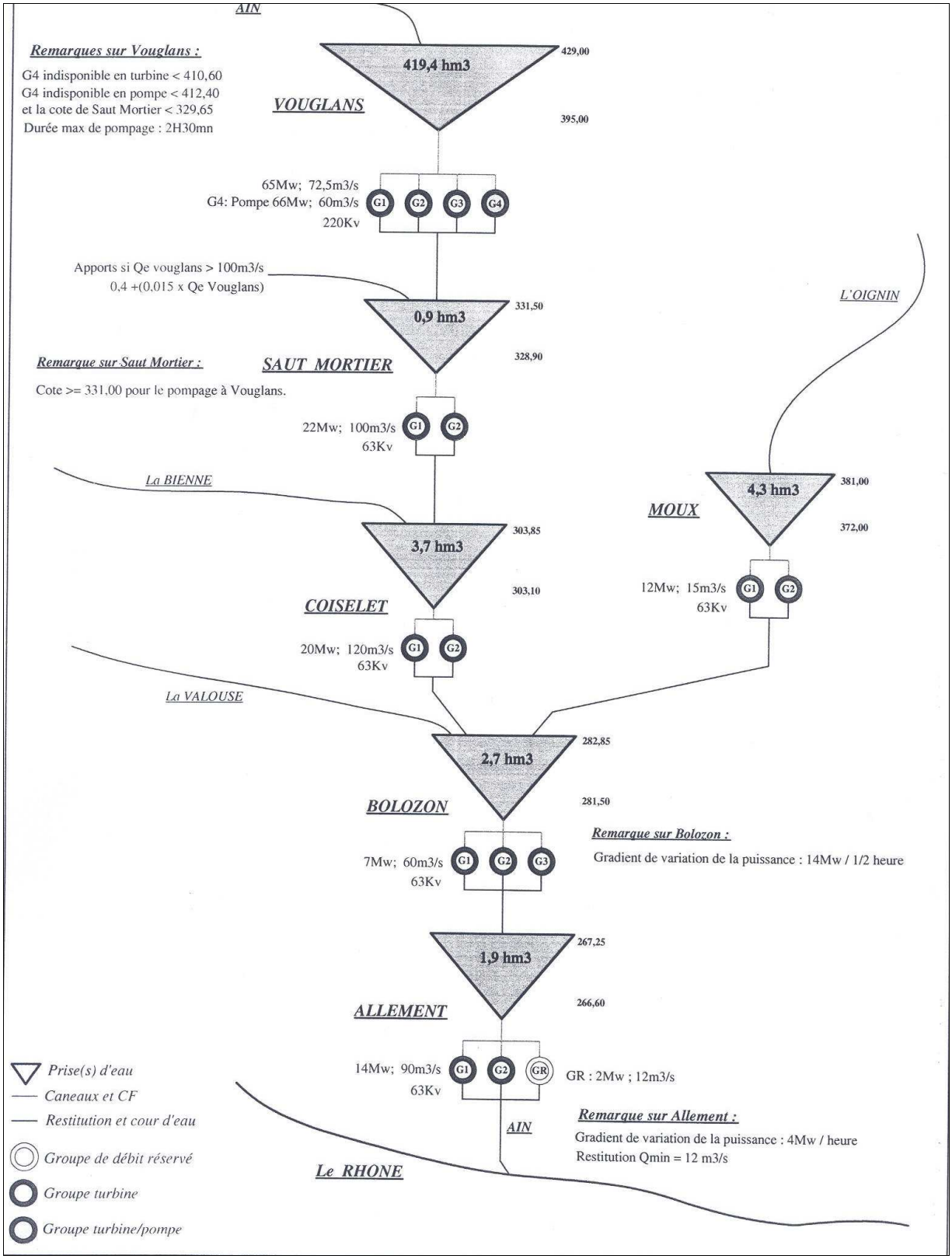



**Figure 3 : Schéma hydraulique de la vallée de l'Ain.**

### 3.2.2 Que vaut la prévision ? Le retour d'expérience du prévisionniste.

Intéressons-nous par exemple aux prévisions émises sur l'Ain durant les années 2000 et 2001. Le prévisionniste désirant faire un retour d'expérience sur la valeur de sa prévision (Bonnaire, 2002), commence par observer les différents événements qui ont marqué l'hydrologie des bassins étudiés (Figure 4 et Figure 5).

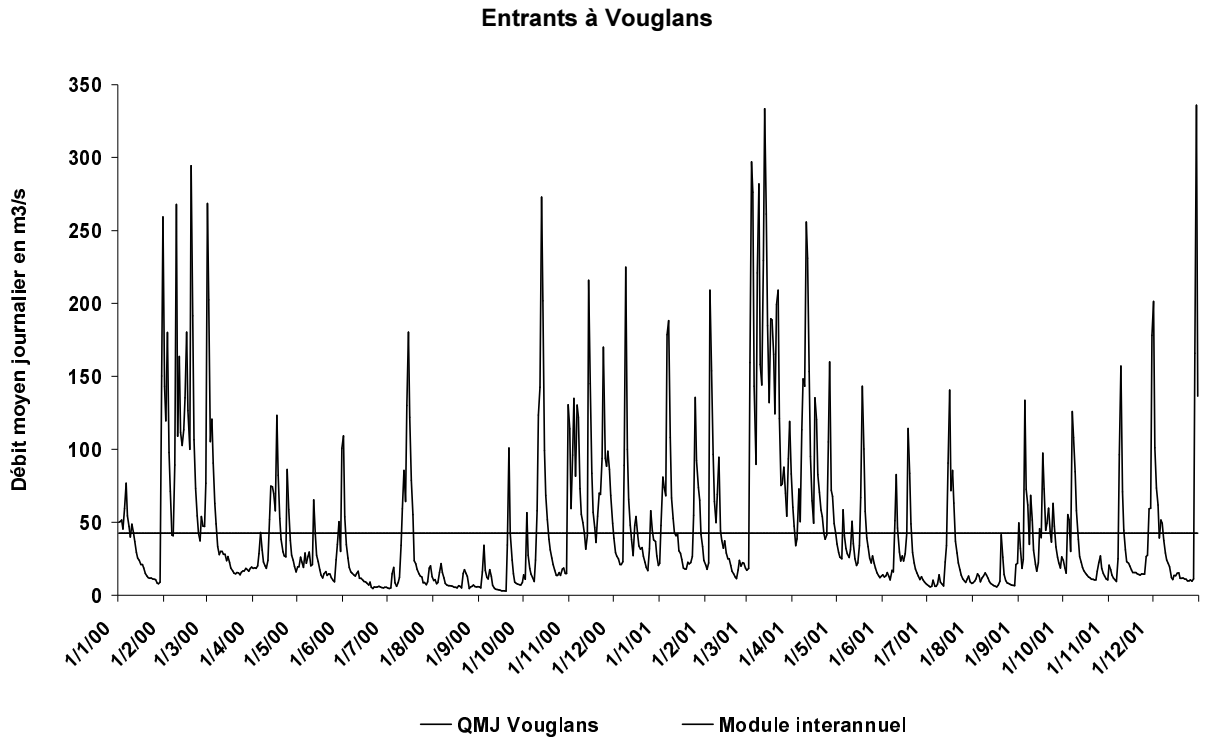



**Figure 4 : Chronique des débits moyens journaliers sur l'Ain à Vouglans, durant les années 2000 et 2001**

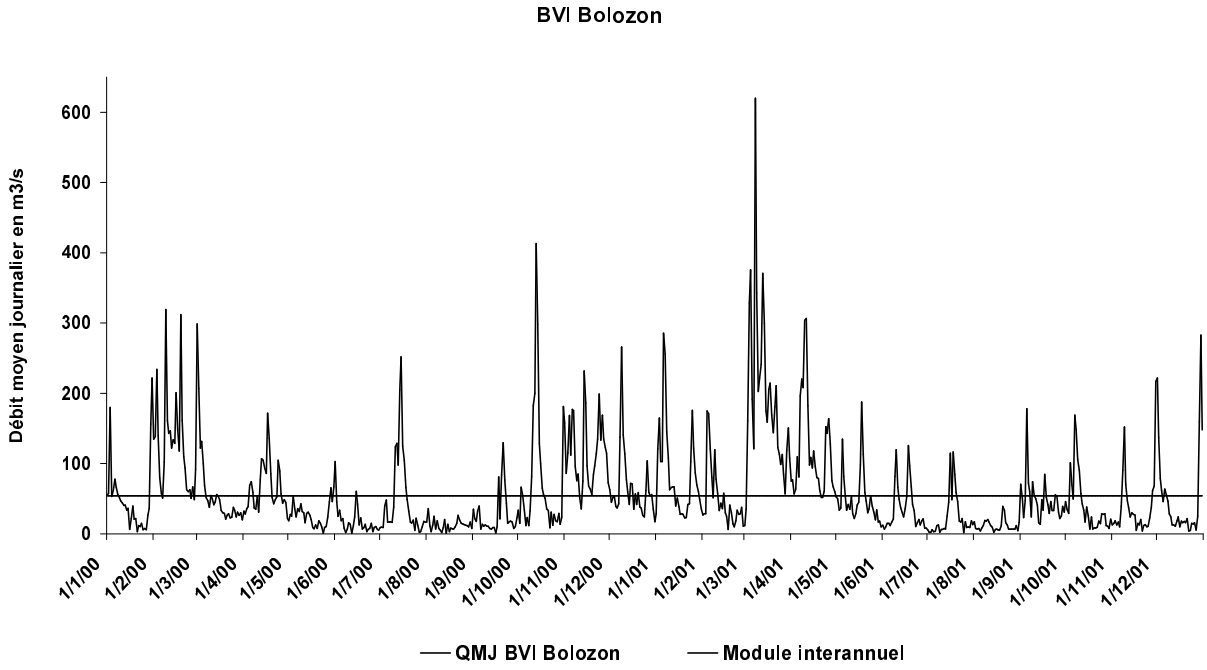



**Figure 5 : Chronique des débits moyens journaliers du bassin versant intermédiaire de l'Ain à Bolozon, durant les années 2000 et 2001**

*NB :Les deux chroniques données Figure 4 et Figure 5 illustrent la forte réactivité du bassin de l'Ain ; la part du débit de base est très faible, et le débit retombe très vite après la pluie. Sur un tel bassin, le prévisionniste ne peut pas tirer parti de la persistance du phénomène, et doit composer au quotidien avec la météorologie.*

Les débits de pointe horaire les plus forts ont eu lieu :

- pour les apports du BVI Bolozon le 8 mars 2001, avec 620 m3/s ;
- pour les débits entrant à Vouglans le 13 Mars 2001, avec 330 m3/s.

En remontant aux observations météorologiques, le prévisionniste observe que ces deux années se caractérisent par :

- des précipitations annuelles supérieures à la normale, avec certains mois particulièrement excédentaires : mars 2001 avec 320% de la normale, février 2000, et juillet 2000 avec 200% de la normale ;
- des températures plus hautes que la normale durant l'année 2001 (+2.5 °C en moyenne).

Il résulte de ces paramètres une hydraulicité supérieure à la normale, en particulier pour l'année 2001 (Tableau 1).

| | **2000** | **2001** |
|---|---|---|
| Entrants Vouglans | 42,9 $m^3/s$ (1,01) | 50,8 $m^3/s$ (1,18) |
| BVI Bolozon | 57,0 $m^3/s$ (1,06) | 62,3 $m^3/s$ (1,16) |

**Tableau 1 : Moyenne des débits moyens journaliers (entre parenthèses le rapport à la normale).**

Dans sa démarche de retour d'expérience, le prévisionniste cible alors quelques périodes durant lesquelles les variations de débit ont été notables, et analyse la bonne adéquation des vecteurs de prévision aux observations durant ces épisodes. Pour cela, le prévisionniste visionne des graphes de type "poils de chat", superposant les vecteurs de prévision émis à la courbe des débits observés (Figure 6 et Figure 7).

Il peut ainsi déceler des crues mal anticipées, ou au contraire, certaines fausses alertes (comme les prévisions du 14/12/00 à J+2 sur Vouglans et le BVI Bolozon).

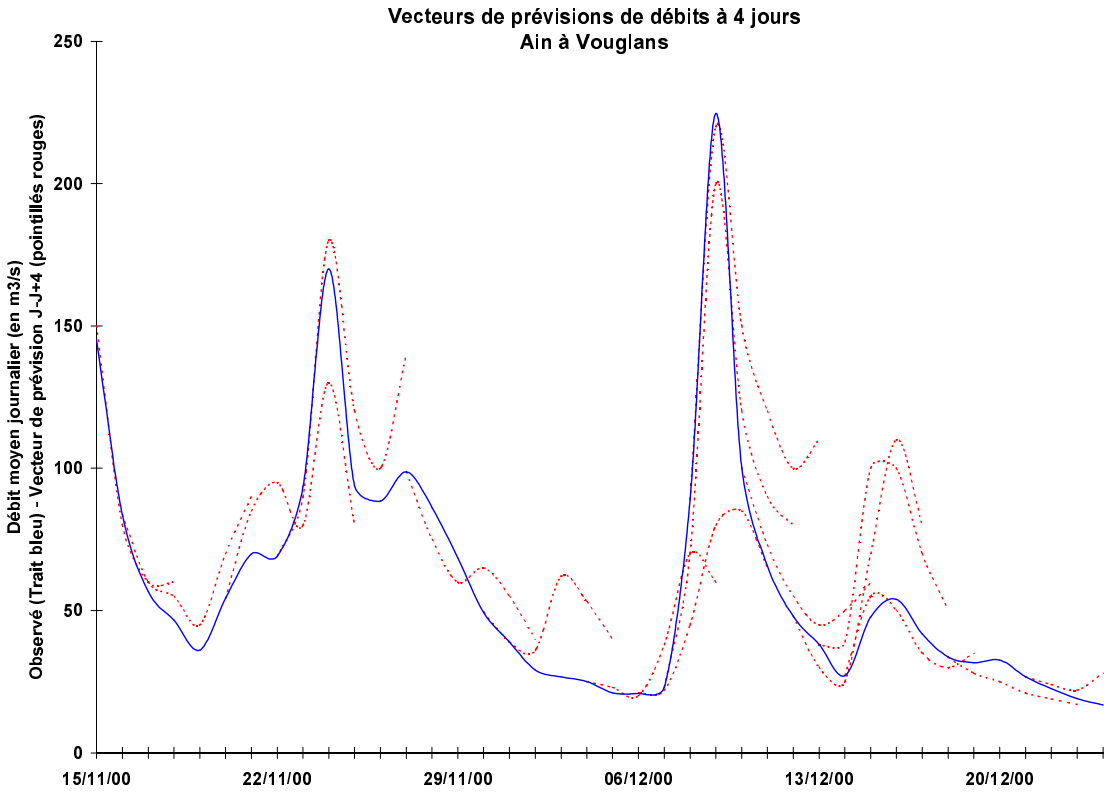



**Figure 6 : Exemple de vecteurs de prévision du jour J au jour J+4.**
**Débit moyen journalier de l'Ain à Vouglans.**

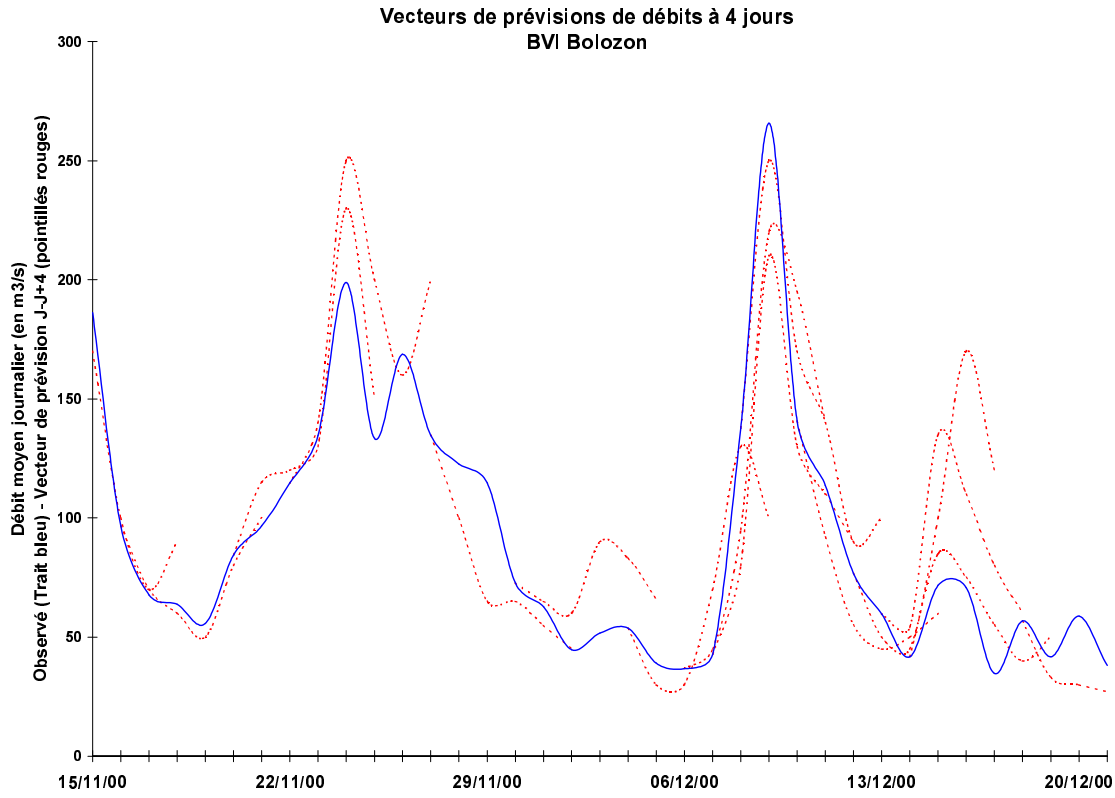



**Figure 7 : Exemple de vecteurs de prévision du jour J au jour J+4.**
**Débit moyen journalier du bassin versant intermédiaire (BVI) de l'Ain à Bolozon.**

Pour avoir une vision plus globale de la qualité moyenne des prévisions émises, le prévisionniste adopte différents types de mesure de l'écart entre prévisions et observations, comme le coefficient de corrélation ou l'erreur quadratique moyenne. Le Tableau 2 donne par exemple la performance de la prévision sur l'Ain à Vouglans selon ces deux critères, pour les échéances J à J+4.

| Echéance des prévisions | J | J+1 | J+2 | J+3 | J+4 |
|---|---|---|---|---|---|
| Erreur quadratique moyenne | 20 m3/s | 26 m3/s | 36 m3/s | 41 m3/s | 48 m3/s |
| Coefficient de corrélation R | 0.96 (233) | 0.92 (232) | 0.83 (232) | 0.77 (232) | 0.67 (231) |

**Tableau 2 : Erreur quadratique moyenne, et corrélation entre débits observés et débits prévus (entre parenthèses le nombre de données utilisées dans la régression). Prévisions sur l'Ain à Vouglans durant les années 2000 et 2001.**

Conscient des attentes de l'usager, le prévisionniste ne s'arrête pas à ce niveau d'analyse. Ce qui intéresse l'usager, c'est d'avoir une information sur la tendance hydrologique. En cas d'absence de prévision, l'usager suppose en effet que le débit ne varie pas. Cette information a priori est dite prévision naïve. C'est ici une prévision de persistance. Le prévisionniste doit, sur les deux critères cités ci-dessus, apporter une valeur ajoutée par rapport à cette information a priori.

Le prévisionniste termine alors son retour d'expérience en construisant le score :

$$S = 1 - \frac{EQM(\mathrm{Pr}\,evision; Observation)}{EQM(Persis\,\mathrm{tan}\,ce; Observation)}$$

Cet indicateur est négatif si la qualité des prévisions est - en moyenne - inférieure à celle des prévisions issues d'un modèle de persistance, égal à 0 si elle est équivalente, et égal à 1 si toutes les prévisions sont parfaites.

| Echéance des prévisions | J | J+1 | J+2 | J+3 | J+4 |
|---|---|---|---|---|---|
| Score | 0.80 | 0.82 | 0.70 | 0.59 | 0.45 |

**Tableau 3 : Score de performance des prévisions émises. Prévisions sur l'Ain à Vouglans durant les années 2000 et 2001.**

Du Tableau 3, le prévisionniste peut tirer plusieurs conclusions :

- pour les échéances rapprochées (J et J+1), sa prévision semble très proche du maximum réalisable ;
- entre J et J+1, la performance relative de sa prévision par rapport au modèle de persistance augmente (du fait d'une chute de la performance du modèle de persistance à J+1) ;

- à J+4, le score de performance de sa prévision est largement positif. Cela laisse supposer qu'aux échéances ultérieures, la prévision présente encore un avantage par rapport à la prévision de persistance.

De tels retours d'expérience mettent en lumière des éléments sur la performance du système de prévision, et surtout, en tracent les évolutions au fil des ans. On est cependant loin de prendre en compte les enjeux opérationnels de l'usager : comment l'aide-t-on à optimiser sa gestion de la ressource en eau ? Quelle gestion aurait-il eue sans cette information de prévision ? Quelle gestion aurait-il eue s'il avait disposé d'une prévision parfaite ? Et surtout, quels sont les gains économiques associés à chacune de ces gestions ?

Autant de questions qui restent sans réponse tant qu'on n'est pas en mesure de modéliser la règle de décision de l'usager.

### 3.2.3 Que vaut la prévision ? Le point de vue de l'usager

Nous avons travaillé en collaboration avec l'usager pour rejouer ces deux années 2000 et 2001, avec la vision de l'avenir qu'il a en temps réel.

L'usager reçoit des prévisions à J+4. A échéance plus lointaine (à partir de J+5), sa vision se fonde généralement sur l'historique des débits : la distribution des possibles pour un jour donné dans le futur est la distribution des débits observés ce jour là pour toutes les années de l'historique.

A échéance plus proche (J à J+4), on modélise les différentes visions que peut avoir l'usager de l'avenir :

- la vision "historique", en supposant que le débit du lendemain est représenté par la distribution historique des débits à la même date ;
- la persistance sur le dernier débit connu (celui de la veille reconduit jusqu'à J+4) ;
- l'utilisation des prévisions expertisées fournies de manière opérationnelle ;
- la parfaite connaissance de l'avenir, jusqu'à J+4.

Ainsi, de J+4 à J+5, l'usager raboute brutalement deux types de connaissance distincts : l'un fondé sur l'état hydrique du bassin et la météo en cours, l'autre fondé sur la connaissances des débits écoulés durant les autres années.

Notons d'ores et déjà que cette vision (que l'usager adopte de manière opérationnelle) est une vision très myope de l'avenir. Au delà de J+4, l'usager voit l'avenir de façon très floue, sur la seule base de ses connaissances historiques, sans tenir compte des circonstances courantes sur le bassin (crue, étiage, etc.).

Pour étudier la valeur des différents types de prévision pour l'usager nous avons repris son modèle opérationnel, SKIPPY (Ringeisen, 2003), qu'il utilise chaque jour pour gérer la ressource en eau sur la vallée de l'Ain. Il élabore ainsi le programme de production de ses usines en fonction des prévisions hydrologiques et du contexte économique (prix de vente de l'électricité sur le marché de l'énergie, coût du moyen de production alternatif disponible, etc.).

Le modèle intègre toutes les caractéristiques techniques des usines et les capacités des barrages, ainsi que les contraintes sur la cote touristique de Vouglans, et le débit minimal à Allement. Nous avons réalisé les simulations pour chaque année 2000 et 2001

### 3.2.4 Résultats

Les Figure 8 et Figure 9 montrent les différentes trajectoires de la cote du lac de tête (Vouglans) proposées par SKIPPY, pour les années 2000 et 2001, suivant le type de prévision utilisée : gestion basée sur l'historique, la persistance, la prévision DTG, et l'avenir certain.

On observe que les différences entre les différentes gestions sont très minimes, ce qui s'explique par la "myopie" de l'usager. De J+5 jusqu'à la fin de l'année en cours, tous les types de gestion l'amènent à utiliser les mêmes chroniques.

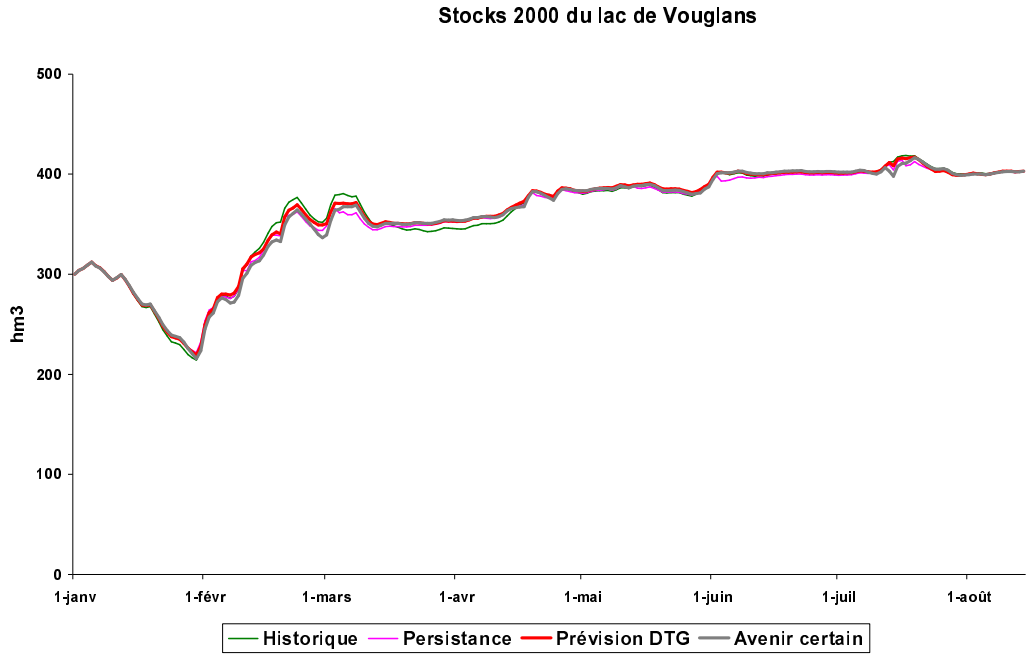



**Figure 8 : Trajectoires des stocks du lac de Vouglans du 1er janvier au 15 août 2000.**

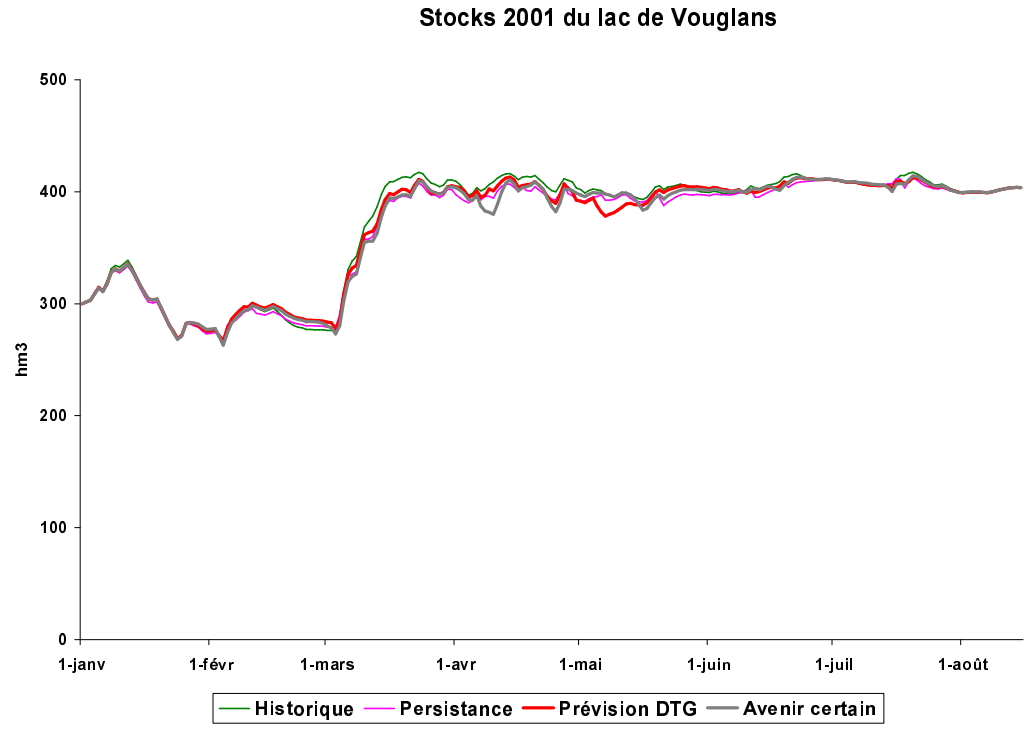



**Figure 9 : Trajectoires des stocks du lac de Vouglans du 1er janvier au 15 août 2001.**

Le Tableau 4 fait pourtant apparaître des différences substantielles entre les gains réalisés avec les différents types de gestion. **Les gains se font donc dans l'anticipation à court terme des problèmes de saturation de la chaîne à l'aval de l'usine de Vouglans**, et des déversements sur les barrages de Bolozon et d'Allement.

| Gestion du 1er janvier au 15 août | Connaissances historiques | Prévisions de persistance | Prévisions expertisées | Avenir certain |
|---|---|---|---|---|
| **Année 2000** | 100 | 100,2 | 100,4 | 101,5 |
| **Année 2001** | 100 | 101,1 | 101,6 | 102,8 |

**Tableau 4 : Gains (normalisés) obtenus avec la gestion réalisée à l'aide des connaissances historiques (base 100), de la persistance sur 5 jours, de la prévision expertisée de J à J+4, et en avenir certain sur 5 jours glissants.**

Le Tableau 4 nous amène aux conclusions suivantes :

- dans l'ignorance, on a toujours intérêt à utiliser une prévision de persistance pendant quelques jours plutôt que d'utiliser directement une prévision fondée sur l'historique. Ce serait dommage de se priver de ce type de prévision naïve (et gratuite !) ;
- pour chacune des années 2000 et 2001, la prévision opérationnelle permet chaque année de réaliser des gains substantiellement plus élevés qu'en utilisant le modèle de persistance ;
- le potentiel de valorisation de la prévision (les gains réalisables en avenir certain) est très variable d'une année sur l'autre (nous n'avons trouvé aucune interprétation d'ordre hydrologique ou économique pour expliquer la différence d'ordre de grandeur entre les gains des années 2000 et 2001).

On extrait de ce Tableau 4 le score de performance qui indique la valeur relative du modèle de persistance et de la prévision opérationnelle, entre prévision naïve "historique" et prévision parfaite (Tableau 5).

| Gestion du 1er janvier au 15 août | Gain relatif obtenu avec la simple persistance | Gain relatif obtenu avec la prévision expertisée |
|---|---|---|
| **Année 2000** | 11% | 31% |
| **Année 2001** | 39% | 56% |

**Tableau 5 : Valeur relative de la prévision opérationnelle et de la persistance : Entre connaissance historique (0%) et prévision parfaite (100%). Années 2000 et 2001.**

Le score du Tableau 5 correspond à la note qu'aurait effectivement attribuée l'usager au système de prévision, et donne la marge de progrès vers la prévision parfaite à 5 jours (de J à J+4). La valeur économique de l'information est donnée quant à elle dans le Tableau 4 : c'est la somme monétaire que le prévisionniste a effectivement pu faire gagner à l'usager en lui communiquant régulièrement des prévisions.

### 3.2.5 Discussion et perspectives

Ainsi, même si ces prévisions ont une faible portée (seulement J+4), et aussi imparfaites et imprécises qu'elles soient, elles permettent à l'usager de réaliser des gains substantiels.

Ajoutons que l'on n'a pas tenu compte des prévisions qualitatives de crue, qui sont diffusées au fil de l'eau dans des bulletins spéciaux ou par téléphone. De telles informations sont également valorisables dans la mesure où elles permettent d'anticiper sur l'indisponibilité éventuelle des groupes de production.

La variabilité de ces résultats suivant l'année considérée nous amène à nous demander si les années 2001 et 2002 n'ont pas été atypiques. L'extension de cette analyse à d'autres années est nécessaire pour connaître la robustesse de nos résultats.

## 3.3 Un exemple de valorisation des prévisions à long terme : les prévisions saisonnières sur le lac de Serre-Ponçon

### 3.3.1 Contexte

Nous étudions ici la valeur de la prévision d'apport de remplissage des réservoirs saisonniers. C'est pour ce type de prévision à long terme que la prévision probabiliste s'est imposée à EDF dès la fin des années 50 (E. Paquet, comm. pers.). A plusieurs mois, il semblait très peu crédible d'afficher un chiffre unique, du fait de l'ignorance des pluies futures. Cependant, les prévisionnistes, disposant d'une estimation du stock nival du bassin à la fin de l'hiver semblaient en mesure d'apporter une information plus utile que les seules données climatiques des dernières années.

C'est ainsi qu'ils ont tout d'abord produit des prévisions d'apport, par une méthode de régression linéaire multiple avec résidu gaussien entre la hauteur de neige mesurée en divers points et les apports printaniers.

Les prévisions saisonnières ont été affinées à la fin des années 1980 par l'arrivée en salle de prévision d'un modèle hydrologique à réservoirs : le modèle MORDOR (Garçon, 1996). Les prévisionnistes ont alors pu injecter dans le modèle opérationnel les chroniques climatiques de chaque année dont ils disposaient (historique depuis 1948), pour obtenir une prévision d'ensemble à l'échelle saisonnière[18].

Chaque année, ces prévisions ont été communiquées de manière opérationnelle à l'usager à compter du 1er avril, et actualisées mensuellement. Récemment, pour répondre aux nouvelles demandes de l'usager, ces prévisions sont émises à compter du 1er février, voire du 1er janvier, avec une actualisation bimensuelle.

Nous étudions dans cette section 3.3 la valeur des prévisions telles qu'elles ont été communiquées à l'usager, en intégrant ses règles de décision. Pour nos simulations, nous avons donc utilisé PARSIFAL (Bourgueil, 1989, Moclinda et al., 1998), son logiciel de planification et d'optimisation de la production hydroélectrique.

[18] Chaque membre de l'ensemble est une chronique de débits moyens journaliers depuis la date T de l'émission de la prévision jusqu'à une date variable (1er juillet, 1er août voire 31 décembre de l'année en cours).

En fonction des connaissances disponibles au jour J (données historiques, et prévisions de DTG), PARSIFAL optimise la production d'un ouvrage hydroélectrique en fonction du contexte économique (prix du kWh échangé sur le marché[19]), et calcule le gain qui en résulte.

Nous cherchons avant tout à obtenir un ordre de grandeur sur la valeur de la prévision pour l'usager. Pour cela :

- nous avons choisi une échéance importante aux dires de l'usager : la prévision émise au 1er avril de l'année en cours. Cette prévision probabiliste est constituée de 50 chroniques équiprobables du 1er avril au 31 décembre (les 50 membres de l'ensemble de la prévision probabiliste) ;
- nous avons choisi deux années contrastées : 2001, avec de très forts apports de printemps (+44% par rapport à la normale pour les apports du 1er avril au 30 juin, apports qui correspondent au quantile 92% des apports historiques de 1948 à 1999), et 2002, avec de très faibles apports (-38%, quantile 6% des apports historiques) (Tableau 6).

Nous comparons les conséquences monétaires des différentes gestions que l'usager aurait faites sur la base de chacune de ces informations :

- prévision de DTG expertisée et communiquée opérationnellement ;
- prévision automatique non expertisée : sorties brutes du modèle hydrologique ;
- aucune prévision : utilisation des séries historiques connues ;
- prévision parfaite.

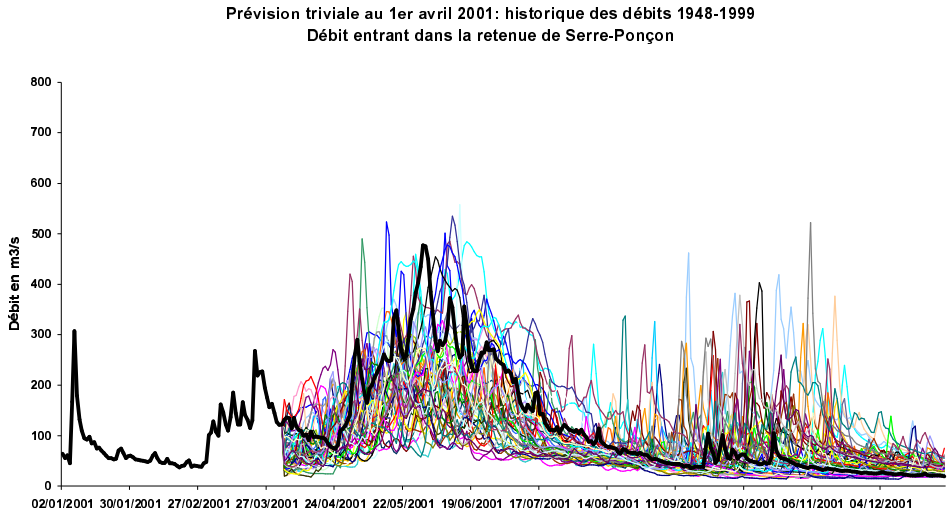
Nous pouvons ensuite estimer la valeur de la prévision expertisée et de la prévision brute par rapport à celle de la prévision naïve (connaissances historiques de l'usager qui ne dispose d'aucune prévision) et celle de la prévision parfaite ( connaissance des débits futurs).

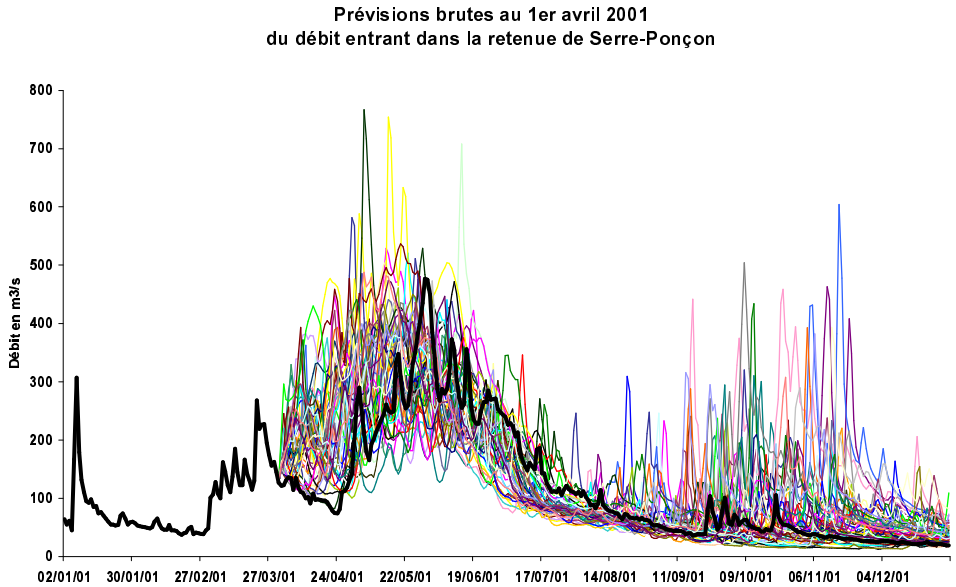
### 3.3.2 L'aléa hydrologique au cours des années 2001 et 2002

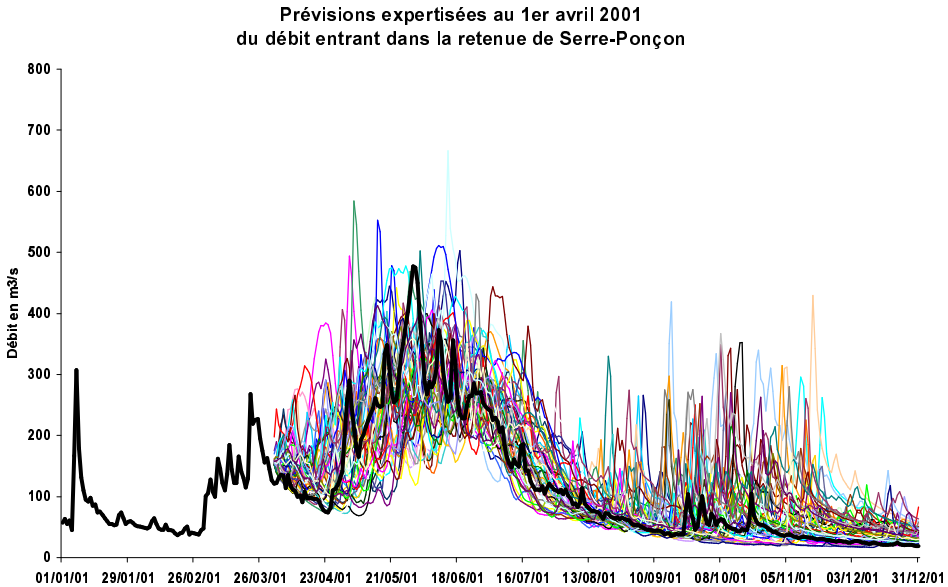
| Apports 4-6 en hm3 | Apports observés | Espérance des apports Prévision expertisée | Espérance des apports Prévision Brute | Quantile 10% apports historiques | Moyenne des apports historiques | Quantile 90% apports historiques |
|---|---|---|---|---|---|---|
| Année 2001 | 1790 | 1754 | 2043 | 789 | 1245 | 1711 |
| Année 2002 | 771 | 733 | 688 | | | |

**Tableau 6 : Apports cumulés du 1er avril au 30 juin dans la retenue de Serre-Ponçon. Observés, espérance des prévus et moyenne des historiques.**

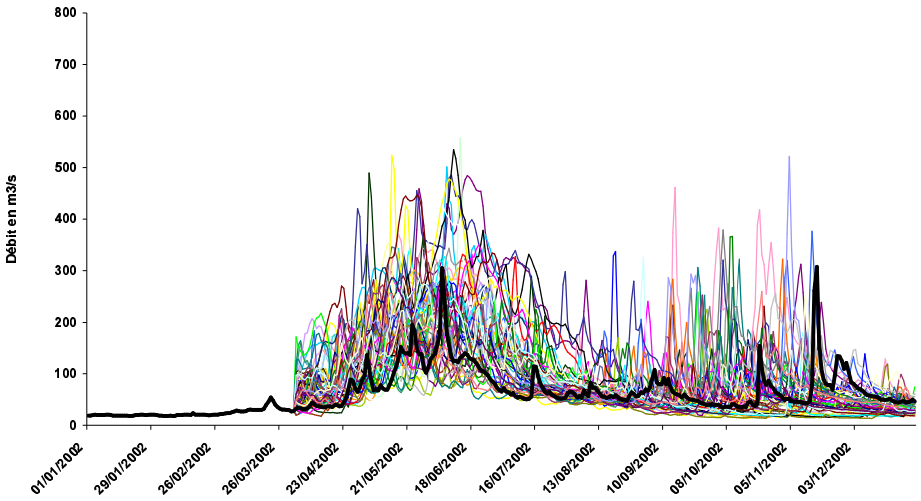
[19] Dans le cas qui nous intéresse, le marché n'étant pas encore ouvert complètement à la concurrence (l'ouverture totale est prévue pour le 1er juillet 2007), le contexte économique est donné à chaque instant par le coût marginal de production. Ces coûts marginaux traduisent à chaque instant l'aléa sur la demande de consommation et sur la disponibilité des groupes de production. Pour répondre à la demande, on commence par exploiter les moyens de production disponibles ayant les coûts proportionnels les plus faibles, puis on démarre les moyens de production de plus en plus dispendieux (ce qu'on appelle « empilement des moyens de production »). Le coût marginal de production d'électricité est le coût proportionnel du plus cher moyen de production mobilisé pour satisfaire la demande.


Prévision triviale au 1er avril 2001: historique des débits 1948-1999
Débit entrant dans la retenue de Serre-Ponçon
Débit en m3/s
800
700
600
500
400
300
200
100
0
02/01/2001 30/01/2001 27/02/2001 27/03/2001 24/04/2001 22/05/2001 19/06/2001 17/07/2001 14/08/2001 11/09/2001 09/10/2001 06/11/2001 04/12/2001

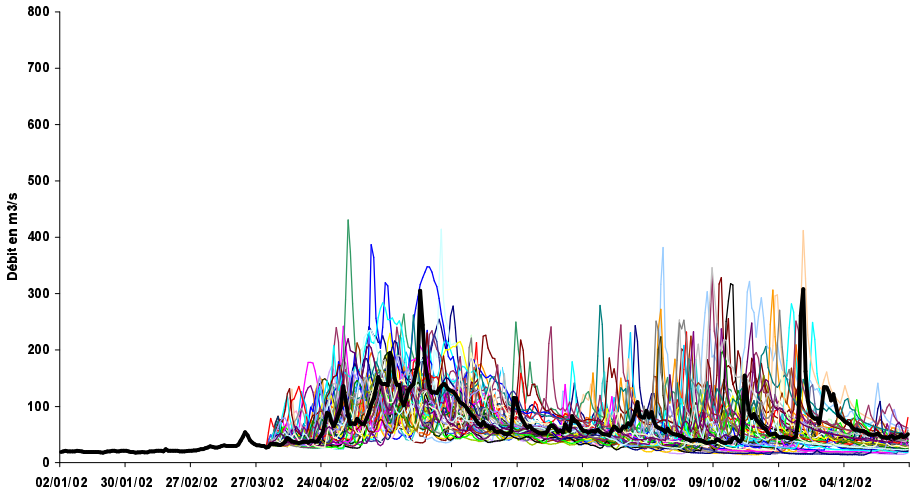

Prévisions brutes au 1er avril 2001
du débit entrant dans la retenue de Serre-Ponçon
Débit en m3/s
800
700
600
500
400
300
200
100
0
02/01/01 30/01/01 27/02/01 27/03/01 24/04/01 22/05/01 19/06/01 17/07/01 14/08/01 11/09/01 09/10/01 06/11/01 04/12/01


Prévisions expertisées au 1er avril 2001
du débit entrant dans la retenue de Serre-Ponçon
Débit en m3/s
800
700
600
500
400
300
200
100
0
01/01/01 29/01/01 26/02/01 26/03/01 23/04/01 21/05/01 18/06/01 16/07/01 13/08/01 10/09/01 08/10/01 05/11/01 03/12/01 31/12/01

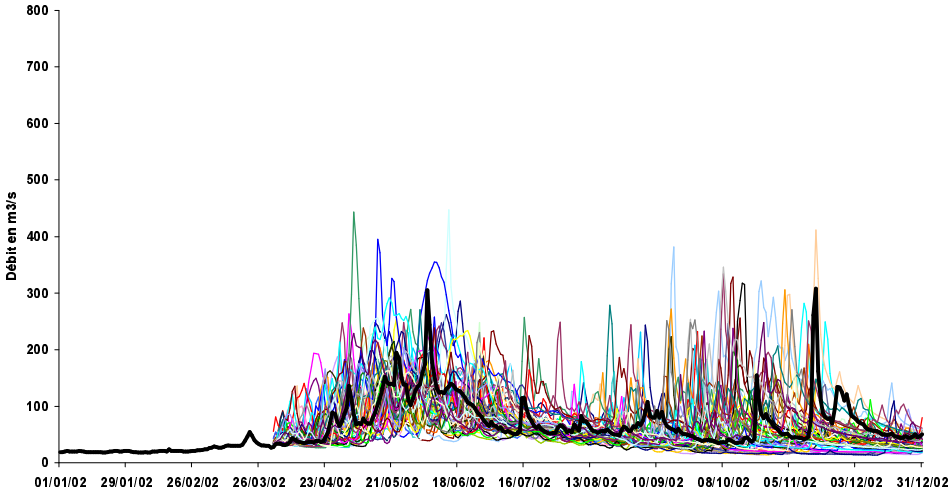
**Prévision triviale au 1er avril 2002: historique des débits 1948-1999**
**Débit entrant dans la retenue de Serre-Ponçon**




**Prévisions brutes au 1er avril 2002**
**du débit entrant dans la retenue de Serre-Ponçon**




**Prévisions expertisées au 1er avril 2002**
**du débit entrant dans la retenue de Serre-Ponçon**



### 3.3.3 Méthode d'estimation de la valeur des prévisions saisonnières sur Serre-Ponçon

Nous estimons ici la valeur des prévisions d'apport sur la retenue de Serre-Ponçon (département 05). Cet exemple est emblématique à plusieurs points de vue :

- cette retenue constitue le plus gros réservoir géré par EDF : 1030 $hm^3$ de capacité utile ;
- la vallée à l'aval de la retenue est très équipée: 2000 MW de puissance installée sur la chaîne Durance (l'équivalent de 2 réacteurs nucléaires) ;
- le bassin versant est conséquent : 3580 km², avec un fort influencement nival : altitude comprise entre 780 m et 4000 m, avec une altitude médiane de 2030 m (Paquet et Garçon, 2001) ;
- les enjeux touristiques sont très forts (activités nautiques), d'où l'exigence (réelle, mais non contractuelle) de la part des pouvoirs publics qu'au 1er juillet la retenue atteigne une cote admissible pour les activités touristiques, dite « cote touristique ». En pratique, EDF s'est fixé pour objectif d'atteindre la cote 778 (qui correspond au volume utile de 965 $hm^3$) au 1er juillet avec une probabilité de 9/10. Nous analyserons la valeur des prévisions avec et sans cette contrainte de cote touristique.
- EDF doit également assurer un débit minimal durant l'été, pour subvenir aux besoins de l'irrigation en basse Durance.

De tels enjeux ont incité les prévisionnistes à porter une attention toute particulière à la prévision saisonnière. Le modèle hydrologique MORDOR produit des prévisions sur lesquelles ils exercent leur expertise hydrologique (compte tenu de leur connaissance du comportement du modèle durant les dernières années, de l'historique du modèle durant l'année en cours, des différents états de remplissage des stocks hydriques modélisés, etc).

En pratique, pour des raisons de temps de calcul, nous avons modélisé dans le logiciel PARSIFAL le seul réservoir de Serre-Ponçon, sans tenir compte de la chaîne de production hydroélectrique en aval, sur la Durance. Les caractéristiques de la retenue et de l'usine tels que nous les avons modélisées sont les suivantes :

- une capacité utile de 1030$hm^3$, la cote variant entre 722 m et 780 m (altitude NGF) ;
- un débit maximum turbinable de 380$m^3/s$ ;
- un coefficient énergétique constant K égal à 990 kW/($m^3/s$). On ne tient pas compte de l'augmentation du coefficient énergétique avec la hauteur de chute (cote de la retenue).

PARSIFAL nous permet de modéliser la gestion opérationnelle qu'aurait programmée l'usager du 1er avril au 31 décembre d'une année donnée compte tenu des informations dont il dispose au 1er avril (connaissances historiques, prévisions brutes ou expertisées, ou connaissance parfaite du futur) Le simulateur détermine à chaque pas de temps quel est le volume optimal à lâcher en comparant le gain immédiat procuré par le turbinage de ces $m^3$ à l'espérance de gain qu'il pourra en obtenir dans le futur.

Nous comparons les gains auxquels auraient mené les gestions fondées sur les diverses connaissances du futur au 1er avril, en intégrant ou non la contrainte touristique dans la règle de décision[20]...

La prise en compte de la cote touristique

Nous utilisons la règle de décision effective qui permet à l'usager d'intégrer cette contrainte.

Tout d'abord, l'usager construit des courbes de remplissage par rétroduction, sur la base de ses connaissances du futur. Supposons par exemple qu'il ne dispose d'aucune prévision, et qu'il utilise les 50 années d'historique. Pour être à la cote touristique au 1er juillet de l'année en cours à 9 chances sur 10, il fait le raisonnement approximatif suivant :

- 1 semaine avant, i.e. au 24 juin, il doit être à la cote pour laquelle 9 fois sur 10 les apports historiques lui permettent de remplir la retenue en 1 semaine
- 2 semaines avant, il doit être à la cote pour laquelle 9 fois sur 10 les apports historiques lui permettent de remplir la retenue en 2 semaines
- etc.

Ainsi, point par point, il construit ce que l'on nomme une « courbe de remplissage »[21] (Figure 10, Figure 11).

Il modifie ensuite les valeurs de l'eau issues de la phase d'optimisation, en imposant une valeur de l'eau forte pour les $m^3$ situés au-dessous de cette courbe de remplissage.

Durant la phase de simulation, le modèle ne fera aucun turbinage énergétique si le stock en début de semaine et les apports prévus sur la semaine ne lui permettent pas de se trouver, en fin de semaine, au-dessus de cette courbe-guide.

[20] En réalité, l'usager actualise sa programmation à chaque réception de bulletin bimensuel. Nous nous contenterons ici d'optimiser la gestion de l'ouvrage à partir d'une seule prévision (celle du 1er avril), pour estimer la valeur intrinsèque de cette information.

[21] Pour établir ces courbes de remplissage, on utilise la série des observés du 1er janvier au 31 mars, raboutée par les différentes séries (avenir certain, historique, prévisions brutes et expertisées) du 1er avril au 30 juin (i.e. à partir de la semaine 13). Cela explique que de la semaine 1 à 13 les courbes de remplissage historiques diffèrent entre les années 2001 et 2002, les apports observés du 1er janvier au 31 mars ayant été très différents pour ces deux années (2001 très humide et 2002 très sèche).

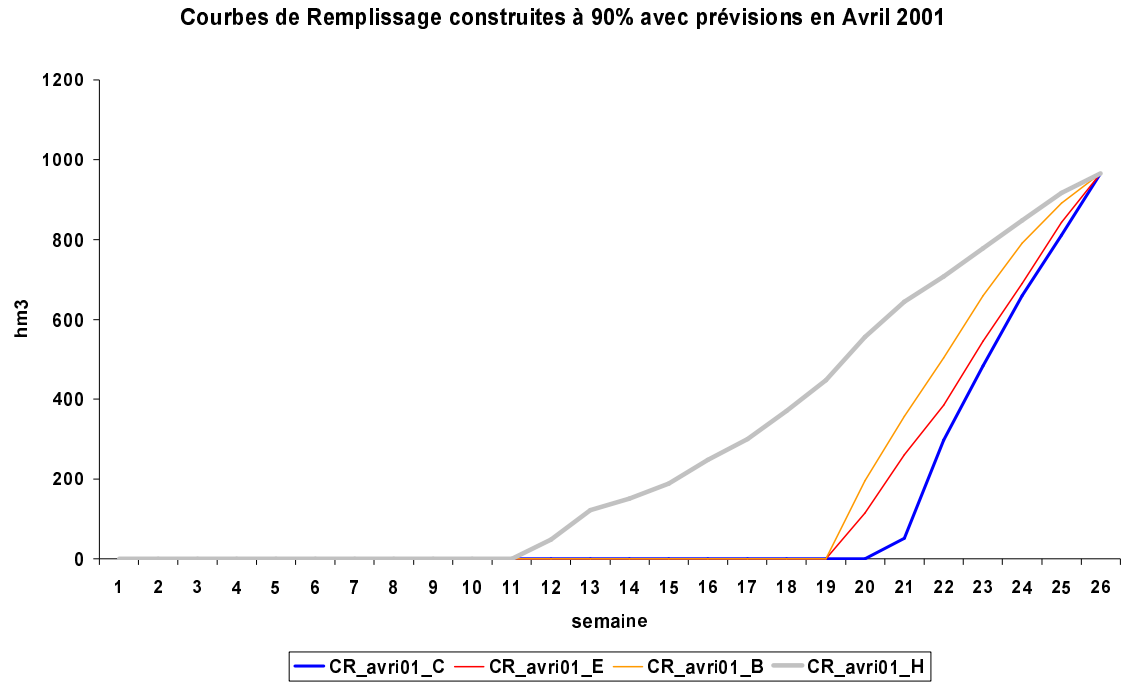



**Figure 10 : Courbes de remplissage établies du 1er janvier 2001 au 30 juin 2001 (semaine 1 à 26).**

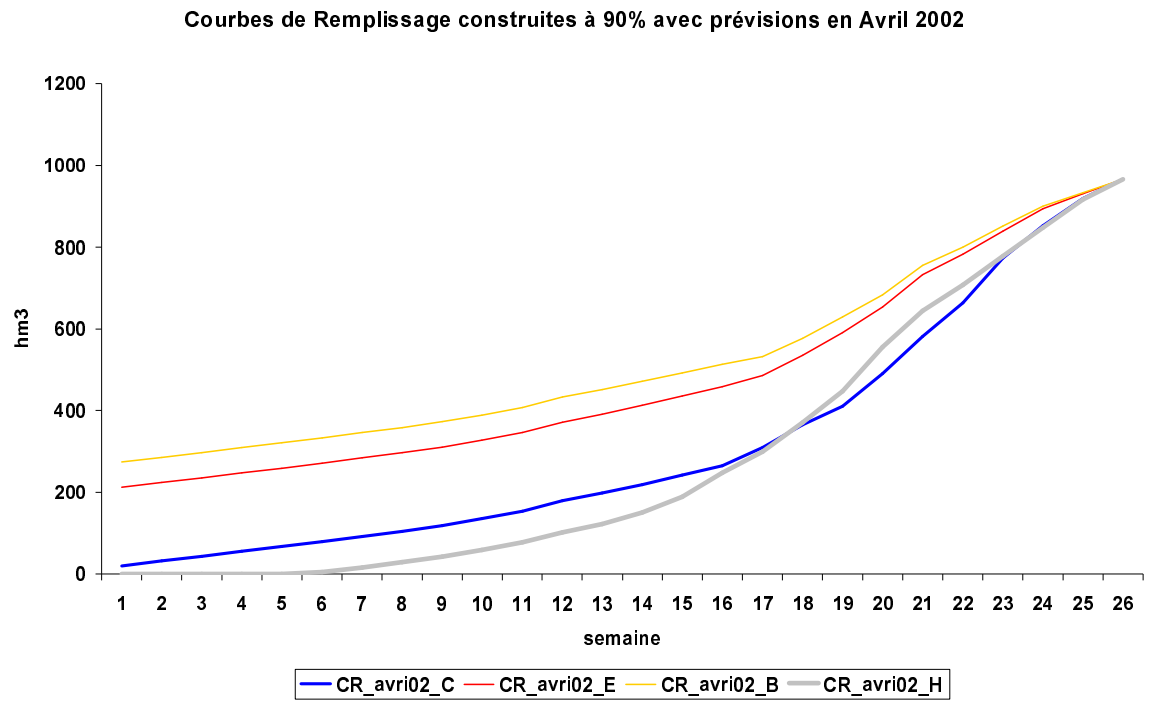



**Figure 11 : Courbes de remplissage établies du 1er janvier 2002 au 30 juin 2002 (semaine 1 à 26).**

### 3.3.4 La valeur de la prévision pour « une année à neige »

On reproduit ici la gestion de l'année 2001 sur la base de chacune des connaissances que peut avoir l'usager sur l'hydrologie à venir : en utilisant l'historique, les prévisions DTG brutes ou expertisées, puis en connaissant l'avenir.

La Figure 12 montre qu'en gestion libre (sans la contrainte touristique), la prise en compte de la prévision permet d'anticiper sur les forts apports printaniers de fusion nivale. L'utilisation de la prévision de DTG permet de se rapprocher de la gestion optimale en avenir certain, en creusant le stock en début de printemps, de manière à turbiner une partie des débits entrants pendant que les coûts marginaux sont encore suffisamment élevés.

Les 4 trajectoires obtenues restent cependant très proches. Les différences de gains qui en résultent sont de l'ordre de 1% (Tableau 7). Ces écarts, qui peuvent paraître faibles de prime abord, sont cependant associés à des gains très importants compte tenu de l'équipement hydroélectrique de la chaîne Durance et des débits turbinés.

Remarque : Sur toutes les figures, la cote touristique est représentée par un segment bleu représentant le niveau minimum acceptable pour les activités nautiques du 1er juillet au 31 août de l'année en cours. Lorsque la contrainte touristique est intégrée dans la règle de gestion (par exemple la Figure 13), l'objectif donné au simulateur PARSIFAL est uniquement d'atteindre la cote touristique au 1er juillet de l'année en cours. Le simulateur n'ayant aucune contrainte après le 1er juillet, on pourra observer des trajectoires passant sous la cote touristique durant la période estivale.

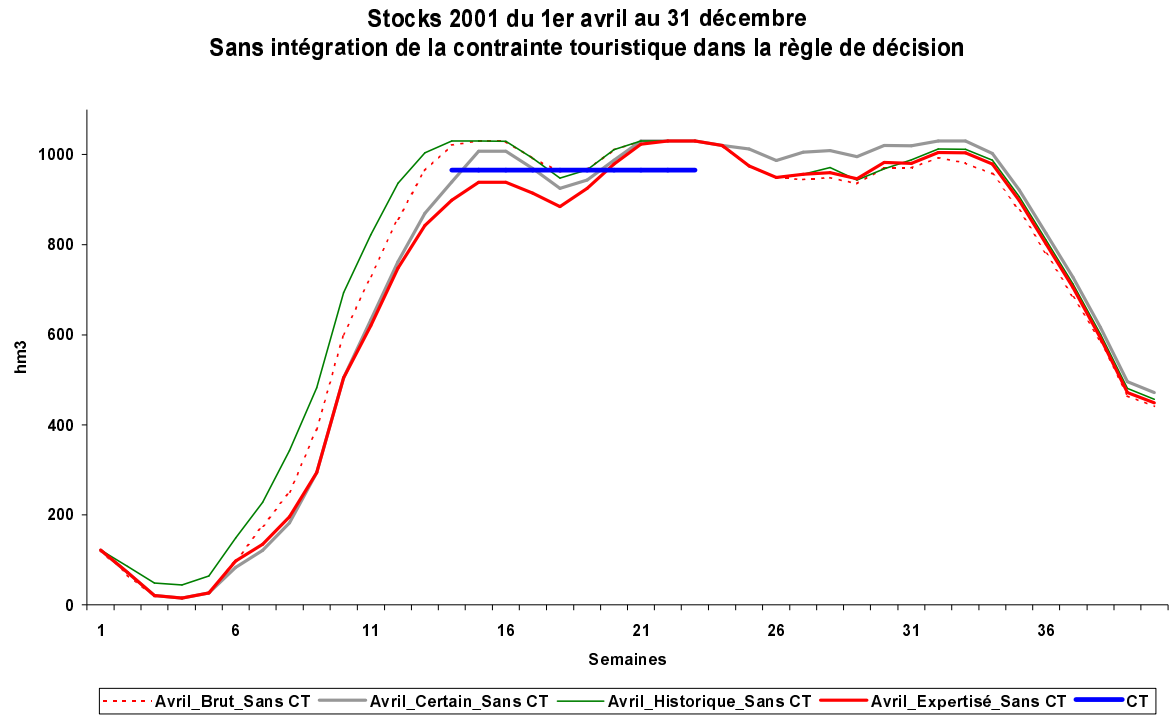



**Figure 12 : Trajectoires des stocks du 1er avril au 31 décembre 2001, sans l'intégration de la contrainte touristique dans la règle de gestion. Les 4 trajectoires résultent des différentes gestions réalisées : sur la base des connaissances historiques, grâce à une prévision brute ou expertisée, ou en avenir certain. Le stock initial est pris égal à 122 hm$^3$.**

Nous notons la remarquable valeur relative des prévisions de DTG (aux environs de 80%, Tableau 8) sur une échelle allant de la prévision naïve (0%) à la prévision parfaite (100%).

Avec intégration de la contrainte touristique, les trajectoires s'écartent très franchement (Figure 13). La gestion historique reste très « prudente » quant aux apports futurs. Cette fois, les gains de production sont de l'ordre de 8% (Tableau 7).

La gestion réalisée sur la base des prévisions de DTG permet de creuser beaucoup le stock, tout en respectant la contrainte, par une bonne connaissance des apports futurs. La prévision brute, et la prévision expertisée amènent à une gestion qui se confond presque avec la gestion en avenir certain. Les gains approchent 90% des gains maximum réalisables (Tableau 8).

NB : les coûts marginaux très faibles en juin invitent à ne pas turbiner à cette période. Cela explique que la retenue peut être remplie avec un peu d'avance pour la gestion avec contrainte touristique. On observe par exemple en Figure 14 que la trajectoire obtenue avec une gestion basée sur l'historique décolle franchement de sa courbe de remplissage durant le mois de juin.

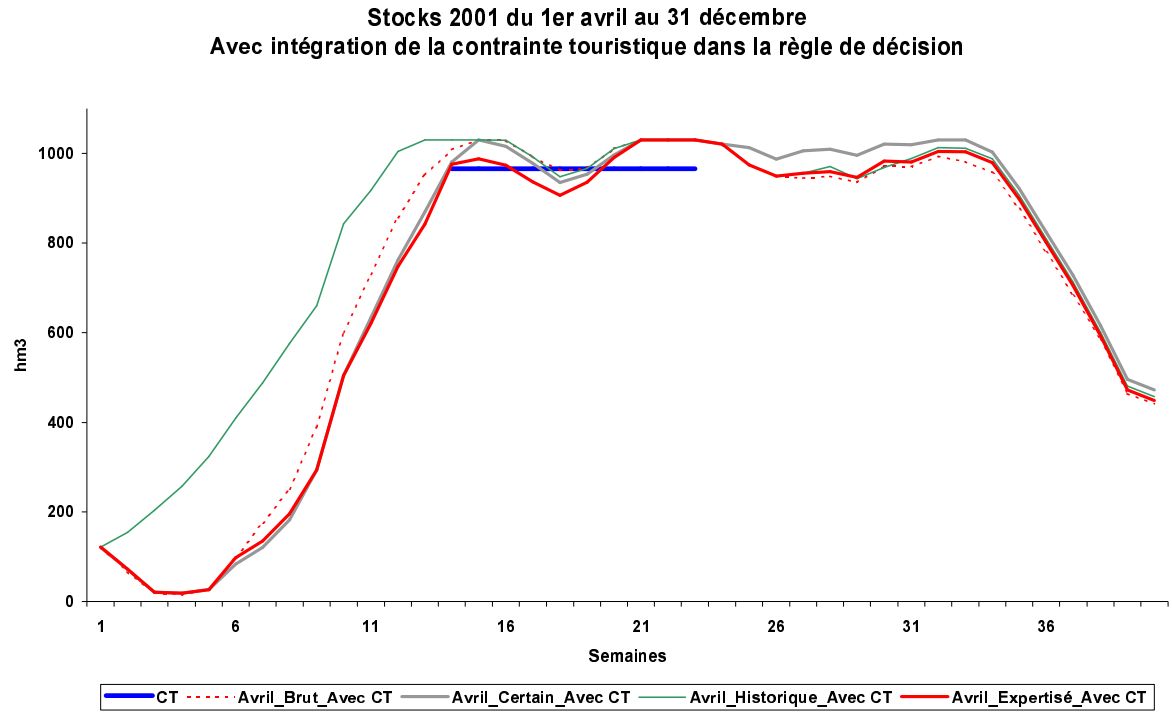



**Figure 13 : Trajectoires des stocks du 1er avril au 31 décembre 2001, avec intégration de la contrainte touristique dans la règle de gestion.**

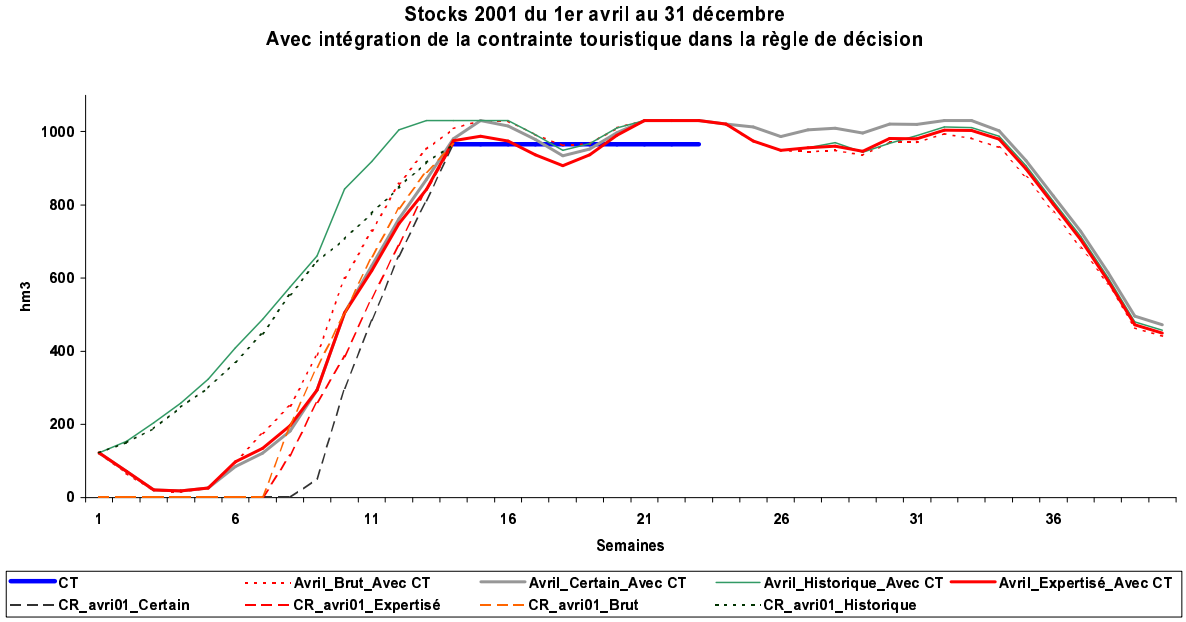



**Figure 14 : Trajectoires des stocks du 1er avril au 31 décembre 2001, avec contrainte touristique. Visualisation des courbes de remplissage.**

| Gestion à partir du 1er avril 2001 (année très humide) | Connaissances historiques | Prévision brute | Prévisions expertisées | Avenir certain |
|---|---|---|---|---|
| **Sans C Touristique Stock initial = 122 hm$^3$** | 100 | 101,1 | 101,2 | 101,4 |
| **Avec C Touristique Stock initial = 122 hm$^3$** | 100 | 108,5 | 108,0 | 108,6 |

**Tableau 7 : Gains (normalisés) obtenus avec la gestion réalisée à l'aide des connaissances historiques (base 100), de la prévision brute, de la prévision expertisée et en avenir certain.**

| Au 1er avril 2001 (Année humide) | Gain relatif obtenu avec la prévision expertisée | Gain relatif obtenu avec les sorties brutes de modèle |
|---|---|---|
| **Sans C Touristique SI = 122 hm$^3$** | 86% | 76% |
| **Avec C Touristique SI = 122 hm$^3$** | 93% | 98% |

**Tableau 8 : Valeur relative de la prévision experte et de la prévision brute : Entre connaissance historique (0%) et prévision parfaite (100%). Année 2001.**

En conclusion :

Lors d'une année à fort enneigement comme l'année 2001, la prévision permet une augmentation des gains de l'ordre de 1% en gestion libre. Avec la prise en compte de la contrainte touristique, les prévisions permettent de turbiner en début de période, lorsque les coûts marginaux sont encore élevés. Les gains sont alors particulièrement conséquents : de l'ordre de 8% (Tableau 7).

Si ce dernier résultat était pour le moins attendu, les résultats présentés au Tableau 8 montrent de surcroît la très grande valeur relative des prévisions de DTG, toujours proche du maximum réalisable (avenir certain), quel que soit le mode de gestion (libre ou sous contrainte), et des gains absolus générés.

### 3.3.5 La valeur de la prévision pour « une année sans neige »

L'utilisation de la prévision (et la réception du signal « apports de fusion nivale très faibles ») permet de gérer avec bien plus de parcimonie les débits entrants en début de période (Figure 15). La gestion ainsi réalisée est extrêmement proche de la gestion en avenir certain. L'usager réalise ainsi de meilleurs placements à l'automne[22]. Il obtient des gains conséquents par rapport à la gestion qu'il aurait effectuée sur la base de ses connaissances historiques (Tableau 9).

[22] Sur les graphiques, on rappelle que l'origine des abscisses est prise au 1er avril. L'automne débute donc aux alentours de la semaine 25 sur les graphiques.

Tout comme au paragraphe précédent, on note la remarquable performance relative de la prévision de DTG, en particulier de la prévision expertisée (plus de 95% de la performance maximum réalisable, Tableau 10).

| **Gestion à partir du 1er avril 2002 (année très sèche)** | **Connaissances historiques** | **Prévision brute** | **Prévisions expertisées** | **Avenir certain** |
|---|---|---|---|---|
| **Sans C Touristique Stock initial = 122 hm$^3$** | 100 | 105,1 | 105,6 | 105,8 |
| **Avec C Touristique Stock initial = 122 hm$^3$** | 100 | 100,6 | 100,7 | 100,8 |
| **Avec C Touristique Stock initial = 500 hm$^3$** | **100** | **95,2** | **95,7** | **100,6** |

**Tableau 9 : Gains (normalisés) obtenus avec la gestion réalisée à l'aide des connaissances historiques (base 100), de la prévision brute, de la prévision expertisée et en avenir certain.**

| **Au 1er avril 2002 (Année sèche)** | **Gain relatif obtenu avec la prévision expertisée** | **Gain relatif obtenu avec les sorties brutes de modèle** |
|---|---|---|
| **Sans C Touristique SI = 122 hm$^3$** | 96.1% | 87.4% |
| **Avec C Touristique SI = 122 hm$^3$** | Sans objet | Sans objet |
| **Avec C Touristique SI = 500 hm$^3$** | Sans objet | Sans objet |

**Tableau 10 : Valeur relative de la prévision experte et de la prévision brute : Entre connaissance historique (0%) et prévision parfaite (100%). Année 2002.**

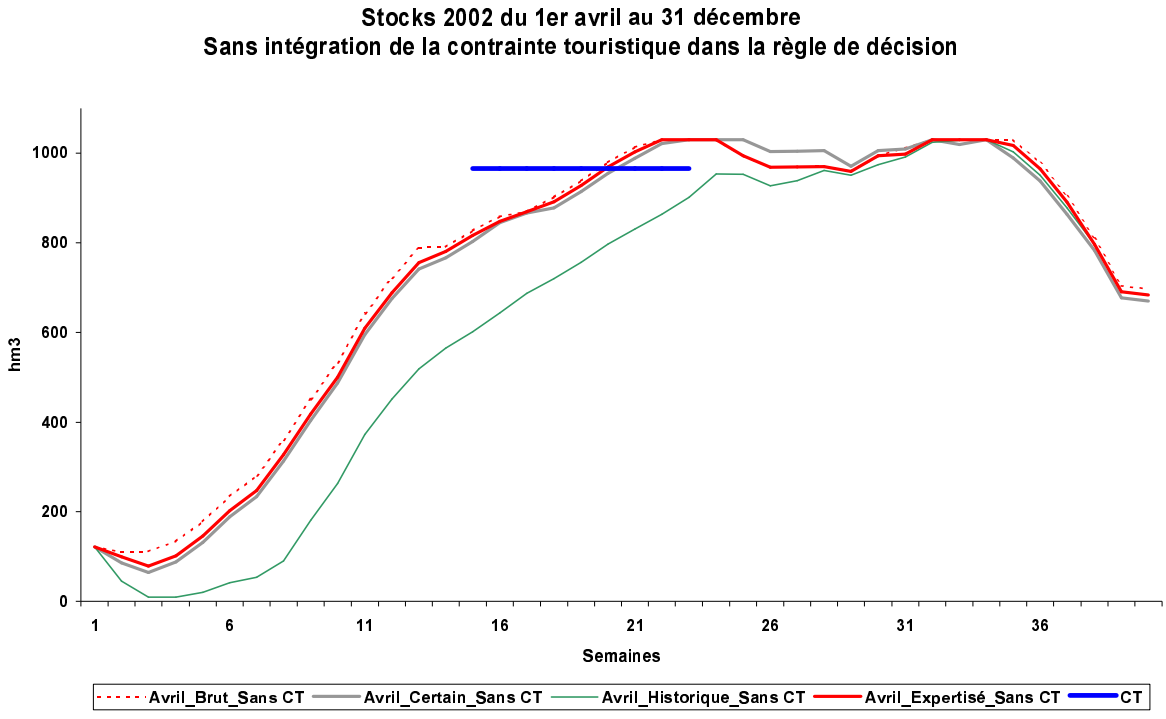


**Figure 15 : Trajectoires des stocks du 1er avril au 31 décembre 2002, sans l'intégration de la contrainte touristique dans la règle de gestion. Les 4 trajectoires résultent des différentes gestions réalisées : sur la base des connaissances historiques, grâce à une prévision brute ou expertisée, ou en avenir certain. Le stock initial est pris égal à 122 hm$^3$.**

Avec intégration de la contrainte touristique, la Figure 16 montre que le signal d'alarme est tiré trop tard. Même sans turbiner un m$^3$ d'eau dès le 1er avril, les apports sont suffisamment faibles pour ne pas permettre le remplissage de la retenue au 1er juillet.

Dans ce cas, toutes les stratégies de gestion sont les mêmes : interdire tout turbinage dès le 1er avril et jusqu'au 1er juillet pour tenter d'atteindre la cote touristique. Toutes les trajectoires étant confondues, la valeur relative des prévisions de DTG est nulle.

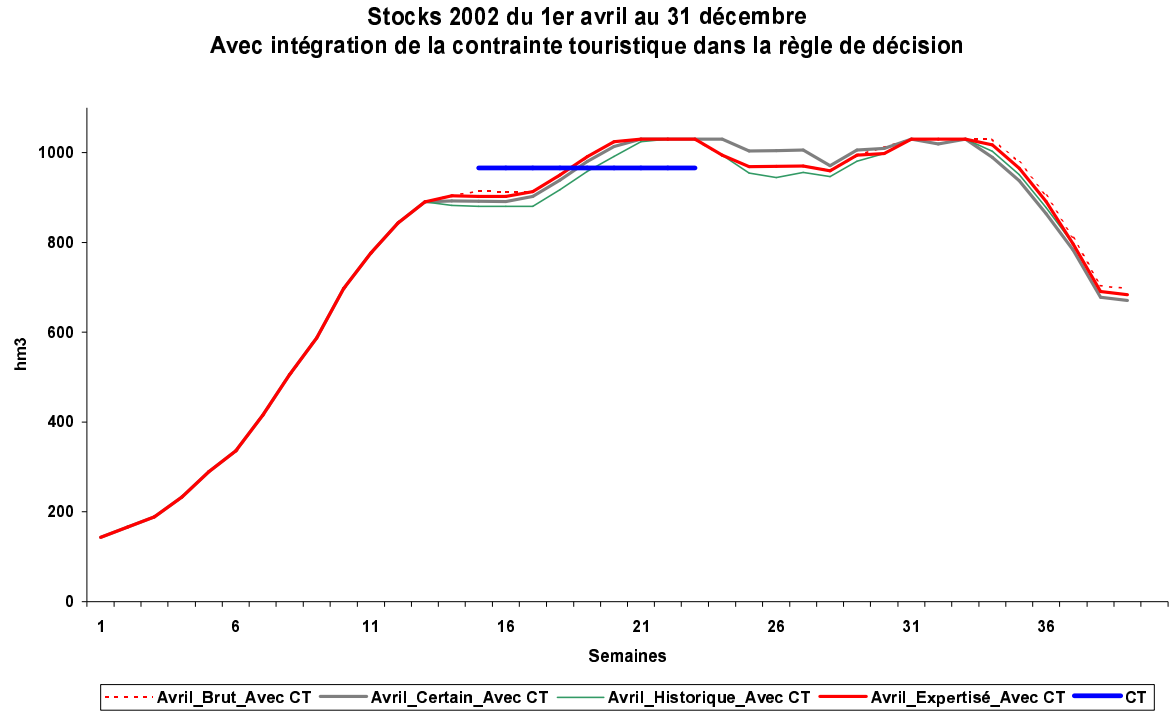



**Figure 16 : Trajectoires des stocks du 1er avril au 31 décembre 2002, avec intégration de la contrainte touristique dans la règle de gestion. Le stock initial est pris égal à 122 hm$^3$.**

Pour distinguer l'apport de la prévision de DTG, avons réalisé une deuxième simulation avec un stock initial plus élevé : 500hm$^3$ (Figure 17).

On observe Figure 17 que la gestion basée sur l'historique conduit en début de période à creuser plus qu'en avenir certain. Les prévisions brutes ou expertisées amènent à une gestion très prudente, creusant à peine en début de période par crainte de ne pas atteindre la cote touristique au 1er juillet. Ce qui est moins visible, c'est que la gestion basée sur l'historique est la seule à ne pas atteindre l'objectif de cote touristique au 1er juillet : à cette date, le creux enregistré par rapport à l'objectif est de 14 hm$^3$, ce qui correspond à une cote 50cm sous la cote touristique.

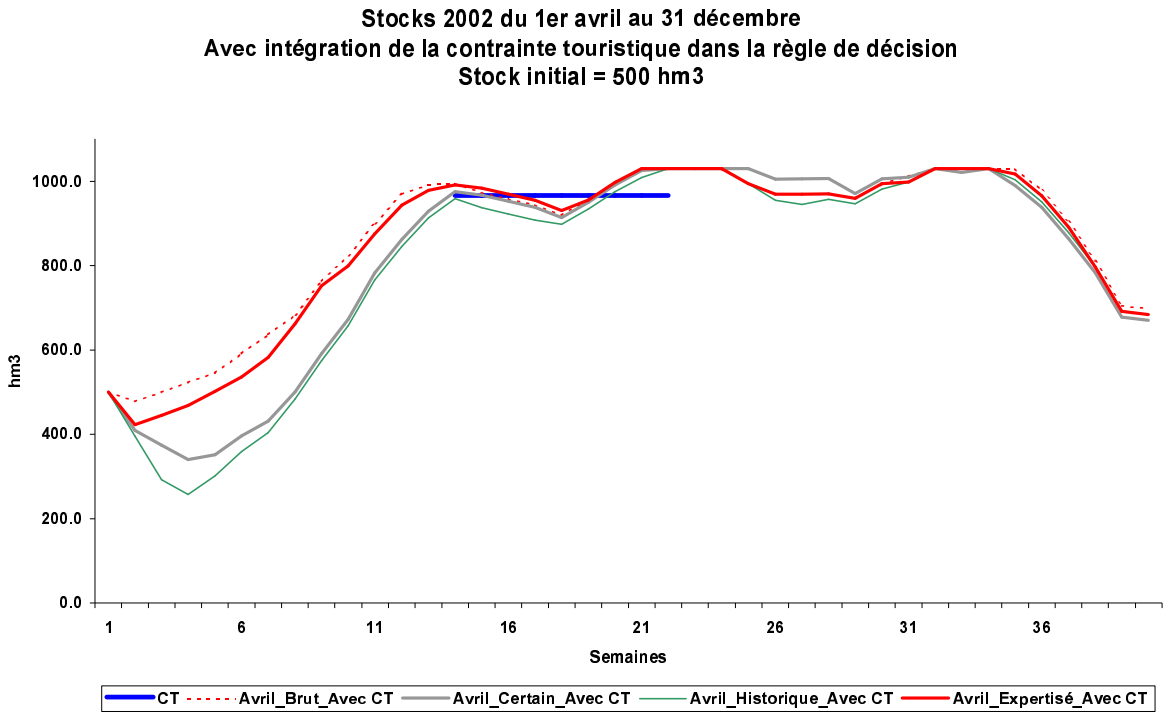



**Figure 17 : Trajectoires des stocks du 1er avril au 31 décembre 2002, avec intégration de la contrainte touristique dans la règle de gestion. Le stock initial est pris égal à 500 hm$^3$.**

La Figure 18 montre que la gestion basée sur l'historique, ignorant la faiblesse des apports à venir, conduit à beaucoup creuser en début de période jusqu'à croiser sa courbe de remplissage. Les apports de printemps sont en effet inférieurs en volume au quantile 10% sur l'historique (Tableau 11). La courbe de remplissage étant construite pour assurer l'atteinte de la cote touristique à 9 chances sur 10, les apports printaniers de l'année 2002 ne permettent pas de suivre cette courbe de remontée. La dissymétrie des apports printaniers permet toutefois d'éviter le pire : le mois de juin étant relativement pluvieux[23], la cote touristique est presque atteinte. Cette gestion apporte néanmoins des gains substantiels en valorisant des m$^3$ d'eau en début de période, au moment où les coûts marginaux sont encore élevés.

| Apports 4-6 en hm3 | Apports observés | Espérance des apports Prévision expertisée | Espérance des apports Prévision Brute | Quantile 10% apports historiques | Moyenne des apports historiques |
|---|---|---|---|---|---|
| Année 2002 | 771 | 733 | 688 | 837 | 1245 |

**Tableau 11 : Apports cumulés du 1er avril au 30 juin dans la retenue de Serre-Ponçon. Observés, espérance des prévisions, moyenne et quantile 10% des historiques.**

La gestion en avenir certain creuse un peu moins en début de période, non pas bien entendu par prudence vis à vis de l'objectif de remplissage, mais pour réserver des m$^3$ qui seront turbinés à l'automne, au moment où les coûts marginaux sont à nouveau élevés.

Les gestions utilisant les prévisions brute ou expertisée de DTG sont quant à elles extrêmement prudentes. Ces prévisions sont d'une part légèrement plus pessimistes (en

[23] En volume, les apports du mois de juin 2002 correspondent au quantile historique 21%. Si on est encore loin de la normale historique, notons tout de même que les apports du mois d'avril 2002 avaient été inférieurs aux plus faibles apports enregistrés sur 50 ans.

espérance) que les volumes observés (Tableau 11). D'autre part, la contrainte de gestion à 9 chances sur 10 entraîne avec ces nouvelles connaissances des courbes de remplissages très élevées (Figure 11).

Cet exemple illustre un « coup de chance » : la gestion basée sur des prévisions historiques conduit à une gestion énergétique très proche de la gestion en avenir certain, en manquant seulement de très peu l'atteinte de la cote touristique au 1er juillet.

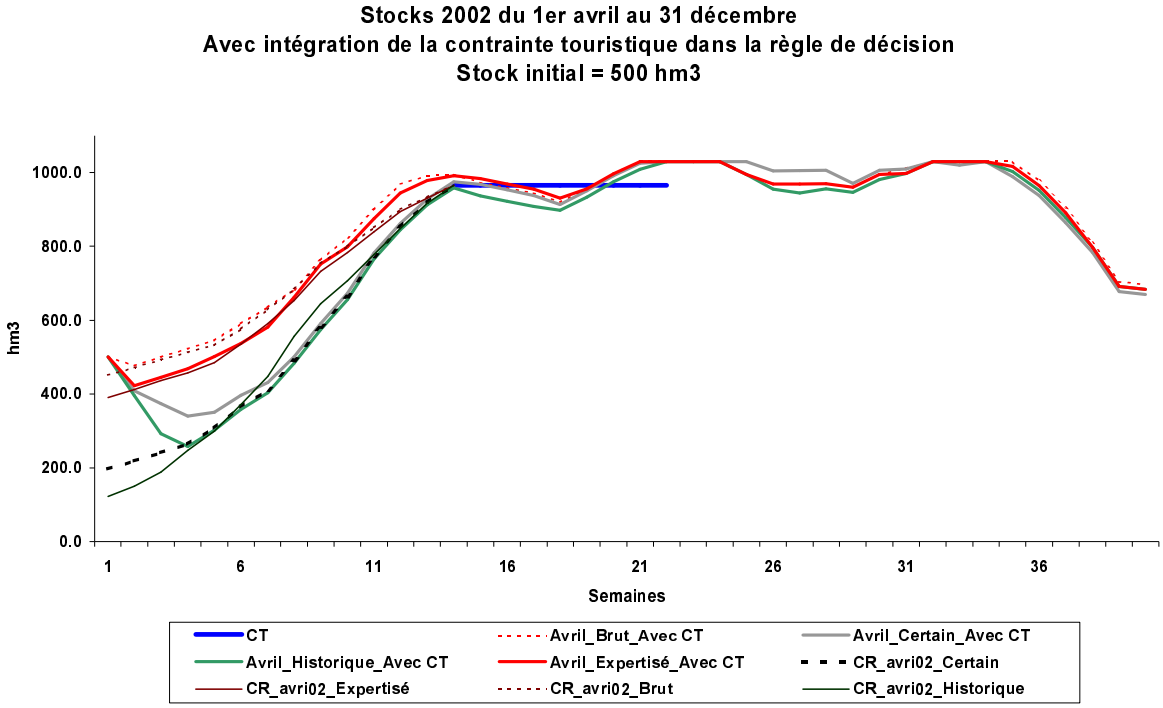



**Figure 18 : Trajectoires des stocks du 1er avril au 31 décembre 2002, avec contrainte touristique (stock initial 500 hm³). Superposition des courbes de remplissage.**

La gestion basée sur l'historique permet de profiter du début de la période pour beaucoup creuser, pendant que les coûts marginaux sont encore élevés. Cette gestion génère un gain très proche de celui de la gestion en avenir certain (Tableau 9), mais sans réussir à satisfaire la contrainte touristique.

La stratégie basée sur les prévisions DTG sacrifie de l'ordre de 4% de gain pour s'assurer le remplissage de la retenue à 9 chances sur 10, compte tenu de ces informations.

En conclusion :

La gestion d'un aménagement lors d'une année à faible enneigement comme l'année 2002 permet de montrer plusieurs attributs fondamentaux de la qualité d'une bonne prévision d'apport :

- En gestion sans contrainte, elle permet d'anticiper sur une période de vaches maigres, en stockant l'eau en début de période pour la valoriser à l'automne lorsque les coûts marginaux deviennent élevés.
- En gestion avec contrainte, lorsque la cote du lac est basse au cœur de l'hiver, la prévision peut permettre de tirer le signal d'alarme à temps (ce qui s'est effectivement passé en temps réel, cf. 3.3.6).

- En gestion avec contrainte, lorsque la cote du lac est moyenne au 1er avril, les prévisions permettent d'éviter des turbinages en début de période, qui auraient empêché l'atteinte de la cote touristique. Dans ce cas de figure, la gestion de Serre-Ponçon avec la prévision de DTG (expertisée ou brute) permet d'atteindre la cote touristique, contrairement à la gestion que l'usager aurait réalisée sur la base de ses connaissances historiques.

Dans le premier cas, la prévision de DTG permet d'augmenter les gains liés au turbinage de 5% environ. Les deux derniers cas montrent qu'en gestion sous contrainte, l'utilisation de la prévision diminue les risques de non-atteinte de la cote touristique. Ce type de valorisation (maintien de bonnes relations avec les pouvoirs publics et l'organisme de tutelle, consolidation de la bonne image d'EDF auprès du grand public et des collectivités locales, etc.) ne pourrait se mesurer qu'à l'aide d'une évaluation contingente. Les arguments évoqués en 3.1 ont cependant montré qu'une telle évaluation avait peu de chance d'aboutir à des résultats robustes.

### 3.3.6 La gestion réelle du lac de Serre-Ponçon en 2001 et 2002

En 2001, les prévisionnistes du CHA relèvent très tôt dans l'hiver que les stocks neigeux en Haute Durance sont particulièrement abondants pour la saison. Dès le début du mois de janvier 2001, ils en informent l'entité en charge d'optimiser la gestion de la ressource en eau sur ce bassin. Celle-ci entame alors une campagne importante de turbinage, pendant la saison durant laquelle la demande est très forte, et les coûts de production élevés (note n° 19, p 41). La cote du réservoir décroît donc rapidement… jusqu'au 1er mars, lorsque le bassin connaît une crue, et que les capacités de turbinage sont insuffisantes pour évacuer les débits entrants. Les prévisionnistes confirmant l'importance du stock neigeux, le gestionnaire turbine alors à plein régime jusqu'au 1er mai. A partir de cette date, les apports de fusion nivale sont suffisant pour atteindre la cote touristique au 1er juillet (Figure 19).

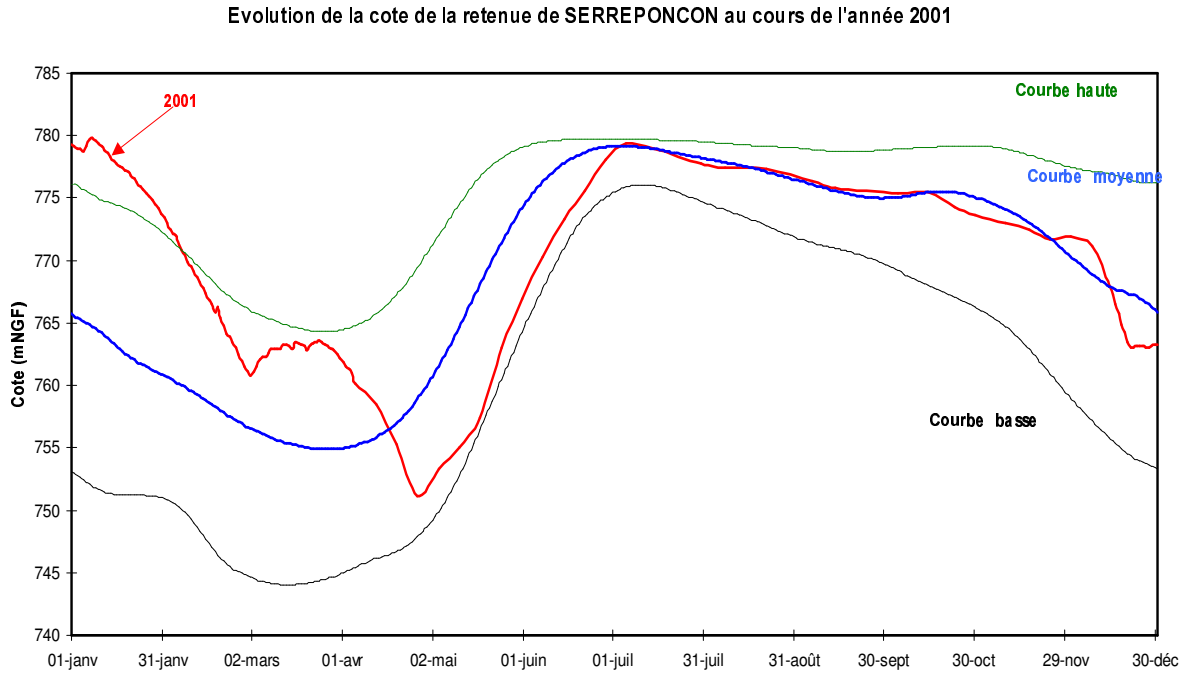



**Figure 19 : Trajectoire de la cote de la retenue de Serre-Ponçon au cours e l'année 2001. Courbe haute et courbe basse sont les maxima et minima historiques depuis la mise en eau (1961). La courbe moyenne est la moyenne des cotes historiques relevées.**

En 2002, les prévisionnistes décèlent très tôt un déficit important du stock de neige, et en informent le gestionnaire dès la première semaine de janvier. Celui-ci gère alors avec une extrême parcimonie tout turbinage, et la cote remonte très lentement. La cote touristique est finalement atteinte avec un mois d'anticipation, du fait des précipitations importantes qui se sont abattues sur le bassin durant la fin du printemps (Figure 20).

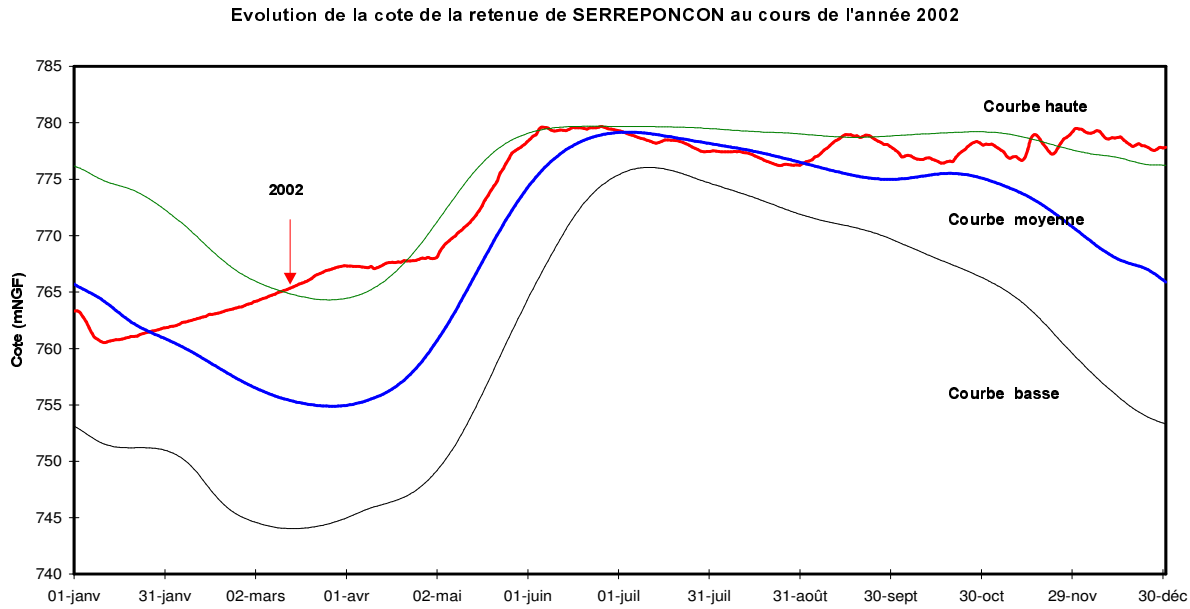



**Figure 20 : Trajectoire de la cote de la retenue de Serre-Ponçon au cours e l'année 2002.**

### 3.3.7 Discussion et perspectives

Cette étude nous fournit des ordres de grandeur sur la valeur de la prévision. Le choix de deux années contrastées conduit cependant à des estimateurs de la borne supérieure de cette valeur. Lors d'une année normale, il semble raisonnable de penser que la différence de gestion entre l'utilisation de la prévision de DTG et celle de l'historique sera bien plus ténue. Les résultats obtenus ici suggèrent que la valeur de la prévision varie entre 1% des gains pour une année normale[24] à 8% pour une année très contrastée.

Notons que l'estimation de la valeur de la prévision peut parfois être « bruitée » par la réalisation des pluies futures : l'estimation du stock neigeux peut-être excellente au 1er mars, et donner lieu à une prévision très "bonne", qui peut pourtant être invalidée par un mois de juin extrêmement pluvieux.

Pour compléter cette étude, il serait utile d'étudier la valeur de la prévision pour d'autres années que 2001 et 2002. Prendre l'espérance des 10 dernières années de prévision pourrait constituer un indicateur plus robuste de la valeur de celle-ci. Cependant, même une période de 10 années ne serait pas suffisante pour évaluer précisément la probabilité d'un échec du remplissage.

[24] Pour une gestion sans contrainte touristique, on peut supposer que la valeur relative de la prévision de DTG pour une année normale est supérieure à celle d'une année « à neige ». Dans le cas d'une année à neige, les débits entrants et turbinés sont suffisamment importants pour que les différences de placement énergétique entre les diverses gestions restent marginales (Figure 12).

Une modélisation plus fidèle à la gestion effective que réalise l'exploitant demanderait une actualisation bimensuelle des valeurs de l'eau calculées suite à la mise à jour de la prévision.

Nous suggérons également quelques raffinements dans la modélisation. Nous avons dû simplifier à outrance le modèle de la chaîne Durance qui avait déjà été construit pour d'autres études. En effet, le nombre de nos simulations était incompatible avec les temps de calcul associés au modèle global. De futures modélisations de la valeur de la prévision pourraient reprendre certaines caractéristiques fondamentales du système global de la chaîne Durance. En particulier, la modélisation des prélèvements agricoles à l'aval de la retenue Serre-Ponçon est une contrainte supplémentaire pour EDF qui doit turbiner un débit suffisant pour assurer un soutien à l'irrigation. La prise en compte, même grossière, de la chaîne aval permettrait également d'intégrer une autre contrainte majeure de la gestion opérationnelle : la saturation des débits turbinables dans les usines à l'aval de Serre-Ponçon. En effet, si le débit maximal turbinable à l'usine de Serre-Ponçon est de 380 $m^3$/s, l'usine de Curbans en aval ne peut turbiner que 210 $m^3$/s. La capacité du réservoir tampon étant très faible (6 $hm^3$, soit une saturation en quelques heures si les deux usines turbinent à plein régime), il importe ne pas turbiner des débits qui ne pourraient pas à nouveau être valorisés dans la chaîne de production en aval, et seraient déversés.

L'intercomparaison de la valeur des prévisions d'apport entre plusieurs bassins versants de montagne permettrait d'arbitrer les développements techniques à venir : mise en place ou amélioration d'un modèle hydrologique, optimisation du réseau de mesure (ajout ou suppression de pluviomètres, de nivomètres, etc.).

## 3.4 Conclusions sur la valorisation économique des prévisions opérationnelles

Même une information imparfaite et imprécise a une valeur économique substantielle. C'est en particulier le cas lorsque cette information donne des renseignements flous, mais à long terme.

Outre une meilleure connaissance de ce que la prévision rapporte effectivement à l'usager, la démarche développée dans ce Chapitre mérite d'être étendue en vue d'un "**pilotage par la performance" du système de prévision**. Autrement dit, adapter les moyens mis en place pour répondre aux enjeux de l'usager.

Ce dispositif permettrait alors d'arbitrer entre les différentes actions d'amélioration du système de prévision en fonction de leur efficacité pour l'usager. On peut aussi déceler d'éventuels gisements d'améliorations « gratuites », comme déterminer la portée optimale de la prévision de persistance par rapport à la distribution historique.

L'objectif est de cibler les actions d'amélioration du système de prévision les plus bénéfiques pour l'usager. Un travail d'intercomparaison permettrait de préciser sur quels bassins, avec quel délai d'anticipation, pour quelle portée, avec quelle fréquence d'actualisation, et dans quelles conditions hydrologiques, les prévisionnistes doivent concentrer leurs efforts. On pourrait également alléger leur charge de travail là où la différence de qualité entre prévision brute et prévision expertisée n'est pas significative pour l'usager.

Enfin, même à court terme, on pourrait mesurer la valeur ajoutée d'une prévision probabiliste par rapport à une prévision déterministe. Nous savons en effet que pour optimiser la gestion de la ressource en eau à l'aide des deux modèles SKIPPY et PARSIFAL, l'usager est en mesure de valoriser une information probabiliste, quelle qu'en soit la portée.

## 3.5 Bibliographie

Atger, F., 2001. Verification of intense precipitation forecasts from single models and ensemble prediction systems. Non Linear Processes in Geophysics, vol. 8, 401-417.

Atger, F., 2000. La prévision du temps à moyenne échéance en France, La Météorologie, vol. 8, 30, 61-86.

Bateman, I., J., Langford, I., H., Turner, R., K., Willis, K., G., Garrod, G., D., 1995. Elicitation and truncation effects in contingent valuation studies, Ecological Economics, vol. 12, 161-179.

Bonnaire, F., 2002, Retour d'expérience sur les prévisions issues du modèle MORDOR à 4 jours sur l'Ain à Vouglans (de septembre1999 à août 2001), EDF-DTG, rapport interne.

Bourgueil, J.-M., 1989. Une méthode de gestion et son logiciel associé pour les systèmes hydro-électriques à un ou deux lacs, Revue de l'Energie, n°410.

Brunswik, E., 1956. Perception and the representative design of psychological experiments (2nd Ed.), Berkley: University of California Press.

Davis, D., Duckstein, L., Krzysztofowicz, R., 1979. The worth of hydrologic data for nonoptimal decision making, Water Resources Research, vol. 15, 6, 1733-1742.

Garçon, R., 1996. Prévision opérationnelle des apports de la Durance à Serre-Ponçon à l'aide du modèle MORDOR, bilan de l'année 1994-1995. La Houille Blanche, vol.5.

Johnson, S., Holt, T., 1997. The value of weather information, *in* Economic value of weather and climate forecasts, Katz, R., Murphy, A. (Ed.), Cambridge University Press, 75-107.

Katz, R. Murphy A. (Ed.), 1997. Economic value of weather and climate forecasts, Cambridge University Press.

Krzysztofowicz, R., 2001. The case for probabilistic forecasting in hydrology, Journal of Hydrology, vol. 249, 2-9.

Lave, L., B., 1963. The value of better weather information to the raisin industry, Econometrica, vol. 31, 151-164.

McQuigg, J., 1971. Some attempts to estimate the economic response of weather information, Weather, vol. 26, 60-68.

Moclinda, A., Iana, G., Lévêque, F., 1998. Optimizing water power generation of the Olt river cascade using PARSIFAL software, Proceedings of the conference 'Modelling, Testing and Monitoring for Hydro Power plants-III, Aix en Provence (France).

Mroz, P., Raven, R., 1993. Levels of student understanding and reasoning associated with televised weather information, Bulletin of the American meteorological society, vol. 74, 425-438.

Paquet, E., Garçon, R., 2002. Caprices du climat et de l'hydrologie en Haute Durance : nos prévisions d'apports plurimensuels sont-elles encore fiables ?, La Houille Blanche, vol.8-2002.

Ringeisen, M., 2003. Fondements méthodologiques du logiciel Skippy, EDF, note interne.

Savage, L. J., 1954, The Foundations of statistics. Wiley.

Stanovich, K. E., 1999. Who is rational? Studies of individual differences in reasoning. Mahweh, NJ: Erlbaum.

Stewart, R., 1997. Forecast value: descriptive decision studies, *in* Economic value of weather and climate forecasts, Katz, R., Murphy, A. (Ed.), Cambridge University Press, 150-181.

# 4 Le pari de la prévision probabiliste

*« Dès que l'on parle de risque, il faut commencer par estimer celui de se tromper »*
*R. Garçon.*

*Eliciter des probabilités subjectives, c'est formuler, à l'aide de l'outil mathématique des probabilités, un jugement sur sa propre expertise.*

## 4.1 Comment éliciter des probabilités subjectives ?

### 4.1.1 Introduction : révéler le savoir tacite du prévisionniste

Le Chapitre 3 nous a permis d'estimer la valeur économique de l'information. Une telle modélisation suppose que toute la substance de l'information utilisée dans la prise de décision est contenue dans les chiffres communiqués :

- soit un vecteur de prévision, par exemple de débits ($Q_J$, $Q_{J+1}$, …, $Q_{J+N}$), où N est la portée (en jours) de la prévision dans le cas de la prévision déterministe (section 3.2),
- soit un ensemble de vecteurs de prévision équiprobables dans le cas de la prévision probabiliste (section 3.3).

En pratique, une telle hypothèse n'est pas toujours vérifiée, notamment lorsque la question de la Sûreté Hydraulique vient se mêler aux objectifs de production. Dans ce cas, l'usager se fait une idée du risque hydrométéorologique en consultant les indications qualitatives consignées dans le bulletin écrit. Si la situation devient plus tendue, l'usager s'enquiert plus avant du risque évoqué, par une conversation téléphonique informelle avec le prévisionniste. Sur la base de cette discussion, l'usager prend alors sa décision.

Dans ce cas de figure, force est de constater que :

1. les informations utilisées, qui concernent principalement l'incertitude de la prévision, ne sont ni quantifiées ni tracées ;
2. la règle de décision de l'usager n'est ni explicite, ni stationnaire ;
3. la prise de décision ne peut se modéliser simplement en supposant que l'usager est l'unique décideur : il n'est qu'un maillon d'une organisation qui réagit face à une crue, en fonction de son (dys)fonctionnement en période de crise, de sa capacité à transmettre les informations en situation tendue, etc. ;
4. la valeur de l'information est délicate à exprimer en terme économique, car il s'agit surtout de définir et de quantifier un niveau de sûreté.

Pour répondre aux points 2, 3 et 4, et comprendre comment est valorisée l'information dans de telles circonstances, nous abandonnerons le point de vue économique, pour adopter une démarche systémique, plus à même d'intégrer les facteurs organisationnels (Chapitre 5).

Dans ce Chapitre 4, nous nous focalisons sur le point 1. Notre objectif est de recueillir les informations qualitatives, orales et informelles transmises par téléphone en gestion de crise. Mieux, de les faire exprimer de manière systématique et formalisée, et de les conserver sur un support qui permette de réaliser a posteriori des retours d'expérience. Nous voulons obtenir au final une information quantitative, régulière (quotidienne) sur la vraisemblance de toutes les évolutions météorologiques ou hydrologiques possibles.

Fournir cette information, c'est précisément l'objet des probabilités. Leur usage permet en effet de recueillir et de tracer des informations sur l'incertitude liée à la prévision, c'est à dire entre autres, sur le risque d'occurrence d'un événement intense (crue, tempête, etc.). Les probabilités peuvent donc révéler des informations effectivement utilisées par l'usager dans sa prise de décision, mais qui jusque-là restaient insaisissables.

En outre, comme nous l'avons montré en 2.3, la prévision probabiliste a également pour vertu de clarifier la communication entre acteurs et de diminuer les risques de mauvaise interprétation.

### 4.1.2 Experts et machines en compétition sur le terrain des incertitudes

Il existe plusieurs méthodes pour produire une prévision hydrologique ou météorologique probabiliste, selon la "matière première" utilisée pour construire les fonctions de densité de probabilité (pdf) des variables à prévoir :

- **L'adaptation statistique en sortie d'un modèle physique de l'atmosphère**. Disposant d'un historique de prévisions et d'observations, on peut croiser dans une table de contingence les variables prévues et observées, par classe (par exemple pour les pluies 0-5mm, 5-10 mm, etc.). On déduit ensuite de cette table de contingence la distribution des précipitations attendues pour une journée donnée, conditionnellement à la prévision émise ce jour-là. Suite aux inondations de Chicoutimi en 1996, Marco Latraverse (Hydro-Québec, comm.pers.) a appliqué cette méthode au bassin versant du lac Kénogami[25].

- **La prévision d'ensemble** exploite un ensemble de sorties résultant de plusieurs runs d'un même modèle physique (déterministe). Chaque sortie correspond à une variation dans l'état initial du modèle ou son paramétrage. La prévision d'ensemble peut s'appliquer à des prévisions météorologiques issues d'un modèle atmosphérique, ou à des prévisions hydrologiques issues d'un modèle de bassin versant[26].

  Les météorologues américains utilisent des ensembles comprenant 15 membres. Sur le Vieux Continent, les prévisionnistes disposent des ensembles du CEPMMT comprenant 51 membres[27].

  Notons ici la particularité du modèle ANALOG qui construit les ensembles sur la base d'analogies entre la situation prévue et les champs de pression observés durant les 50 dernières années. Dans sa version actuelle à EDF, le modèle sélectionne les 50 meilleurs analogues. La prévision quantitative des précipitations est alors donnée par la distribution empirique des pluies associées à ces 50 situations historiques, qui se rapprochent le plus de la situation courante (Obled et al., 2002, Bontron, 2004).

- **Les probabilités subjectives** utilisent les indications quantitatives de l'expert sur l'incertitude de ses prévisions. Pour estimer la vraisemblance des divers scénarios qu'il envisage, ce dernier s'appuie sur les ensembles, mais aussi sur son expérience de la

---

[25] Le lac Kénogami est un plan d'eau de 50 km² situé à l'exutoire d'un bassin versant de 3390 km², qui reçoit en moyenne 1200 mm par année (ce qui en fait l'une des région les plus arrosées du Québec). Ce bassin versant, situé dans la région montagneuse des Laurentides (40% de la superficie du BV est à plus de 650 m d'altitude), est drainé par trois rivières principales : les rivières aux Ecorces, Pikauba et Cyriac. Des très fortes précipitations ont causé de grosses inondations en juillet 1996 particulièrement dans la ville de Chicoutimi.

[26] en y injectant différentes conditions initiales représentant des incertitudes comme l'état de saturation des sols, mais surtout divers scénarios météorologiques représentant les incertitudes, souvent bien plus importantes, sur les prévisions de précipitations.

[27] Pour faire face aux temps de calculs élevés, les européens utilisent un modèle atmosphérique plus sommaire que le modèle opérationnel "officiel" (troncature plus faible), tandis que les américains forment des ensembles à partir de 15 membres pris à des moments différents : certains sont obtenus à partir de conditions initiales dérivées des observations à l'instant t, d'autres à partir d'observations à $t+\Delta t$, ce qui pose un problème évident de cohérence (R. Krzysztofowicz, conférence à l'Engref, 24 sept. 2002).

fiabilité des modèles pour la situation en cours, et sur sa connaissance experte de la météorologie locale et de la réaction des bassins.

Les deux premières méthodes préconisent donc un système de prévision probabiliste automatique (calibré au besoin à l'aide d'adaptation statistique) qui tourne en parallèle avec le système de production de prévisions expertisées. De tels systèmes de prévision automatiques peuvent être mal perçus par les prévisionnistes opérationnels. C'est le cas à Météo-France, où ceux-ci voient leur expertise concurrencée par un nouveau système de prévision "qui n'a plus besoin d'eux". A partir des ensembles fournis par le CEPMMT, Météo-France produit en effet (à titre expérimental) des prévisions probabilistes[28] automatiques, sans contrôle humain. Les prévisionnistes eux-mêmes étant invités à éluder la question de l'incertitude de prévision (note n°8, p 20), cela accentue leur méfiance à l'égard des outils automatiques, qui en un sens, "font mieux qu'eux", en quantifiant leurs incertitudes a priori (F. Atger, comm. pers.).

De telles tensions existent aussi au sein des services météorologiques américains[29], qui testent la production de prévisions probabilistes de pluie (Schaake et al., 2001 et Krzysztofowicz, 2001) selon les deux écoles : prévision automatique et prévision expertisée.

La troisième méthode suppose que l'expert est le plus à même d'évaluer ses propres incertitudes de prévision. Avec cette méthode :

- le savoir tacite du prévisionniste est valorisé ;
- l'expertise humaine complète le traitement automatique ;
- le prévisionniste peut intégrer des informations variées et rattraper nombre de défaillances ; le processus d'élaboration de la prévision probabiliste est plus robuste en situation perturbée ou dégradée ;
- les retours d'expérience permettent au prévisionniste d'améliorer l'estimation de ses incertitudes. Une boucle de progrès peut être engagée, alors que la prévision probabiliste automatique suppose l'hypothèse de stationnarité, c'est à dire de non-amélioration des performances du modèle initial ;
- le modélisateur se fait complice des prévisionnistes, en les aidant à formuler au mieux l'essentiel de leur savoir, tandis que la machine les concurrence.

Pour ces raisons, cette dernière méthode nous semblait mieux à même de répondre à notre question de recherche (Chapitre 1). Nous avons donc choisi d'expérimenter à EDF l'élicitation des probabilités subjectives, dans le but d'améliorer la prévision opérationnelle et la communication des risques hydrométéorologiques.

---

[28] Prévisions probabilistes automatiques de température essentiellement, en mode semi-opérationnel en 2003 (F. Atger, comm. pers.). L'usager qui teste ce nouveau produit est EDF, pour la prévision de la consommation d'électricité.

[29] Bien plus qu'en Europe, en Amérique du Nord, et particulièrement aux Etats Unis, la culture du risque incite à quantifier les incertitudes. Quel qu'en soit le domaine, le citoyen et le politique américain ont toujours demandé aux experts d'évaluer des intervalles de confiance, dès qu'on savait qu'un chiffre était entaché d'un niveau d'incertitude suffisamment élevé pour avoir un impact important sur la société. Ainsi, lorsque les instituts de sondage français diffusent dans les media avant des élections « M. X 68% et M. Y 32% », sans aucune mention de l'incertitude, on trouve aux Etats Unis ou au Canada dans les journaux destinés au grand public : « Mr. X est crédité d'un score de 68% ± 4%, à 19 chances sur 20 » ; c'est à dire que l'institut de sondage estime à 0,95 la probabilité que Mr X. obtienne un score compris entre 64% et 72%. (V. Fortin, comm. pers.).

### 4.1.3 Comment amener le prévisionniste aux probabilités subjectives ?

Il existe plusieurs moyens de faire exprimer ses incertitudes à l'expert. En plus de sa prévision déterministe, celui-ci pourra déterminer :

- **un intervalle de crédibilité** dans lequel il estime que la variable prévue à x% de chance de tomber (Murphy et Winkler, 1974)[30] ;

  *"Mon intervalle de crédibilité à 80% sur le débit maximal de l'Arc à Modane durant la journée de demain est [20 $m^3/s$ ; 70 $m^3/s$]"*

- **un quantile de probabilité** sur la variable prévue (Krzysztofowicz *et al.*, 1993), c'est à dire un seuil en dessous duquel la variable prévue n'a qu'une probabilité fixée de rester[31] ;

  *"J'estime le quantile 90% à 70 $m^3/s$ pour le débit maximal de l'Arc à Modane dans la journée de demain"*

- **La probabilité de franchissement de divers seuils** par la variable prévue (Murphy et Dann, 1984).

  *"J'estime qu'il y a 1 chance sur 4 que le seuil des 30 $m^3/s$ sur l'Arc à Modane soit dépassé dans la journée de demain"*

La production de ces informations conduit le prévisionniste à dépasser la simple expertise de modèles probabilistes (qui serait cousine de la critique habituelle de modèles hydrologiques ou météorologiques) pour s'impliquer personnellement en prenant des paris sur l'avenir.

Les deux premières méthodes consistent à estimer une grandeur associée à une probabilité fixée. Elles semblent êtres plus intuitives que la troisième, les prévisions obtenues montrant une meilleure fiabilité (Murphy et Winkler, 1974). En effet, les prévisionnistes préfèrent manipuler leurs grandeurs habituelles (des pluies, des débits, etc.) plutôt que des probabilités (un chiffre abstrait entre 0 et 1).

Nous retiendrons ces deux premières méthodes pour expérimenter l'élicitation des probabilités subjectives avec les prévisionnistes d'EDF.

### 4.1.4 Le cas particulier des prévisions de précipitation : des pdf mixtes

Quel que soit le prévisionniste, le bassin choisi et la journée de prévision, la loi de probabilité subjective est mixte : la probabilité qu'il pleuve 0mm n'est presque jamais réduite à 0, alors que la probabilité qu'il pleuve exactement une hauteur h non nulle est toujours réduite à 0. La loi de probabilité sur une prévision de pluie se modélise donc généralement par un Dirac en 0, associé à une loi continue sur [0 ; +∞[, par exemple une loi de Weibull (Krzysztofowicz, 2002) de la forme $[x] = c.x^{c-1}.\exp[(-x)^c]$.

En conséquence, Krzysztofowicz (2002) demande aux prévisionnistes

- de déterminer dans un premier temps la probabilité subjective d'occurrence de pluie significative (> 0,1mm) $\nu$ ;

[30] Couramment cet intervalle de crédibilité est donné via quantiles de probabilité (un IC à 80% est donné à l'aide des quantiles 10% et 90%)

[31] Ainsi, le quantile 90 est le seuil que la variable n'a qu'1 chance sur 10 de dépasser.

- de donner dans un deuxième temps les trois quantiles ($x_{25}$, $x_{50}$, $x_{75}$) conditionnellement à l'occurrence de pluie significative sur le bassin.

Pour simplifier le travail du prévisionniste, nous choisissons quant à nous de ne parler qu'en probabilité totale. Ainsi, lorsque le prévisionniste est convaincu qu'il ne va pas pleuvoir à 3 chances sur 4, ils mettra les quantiles 25, 50 et 75 à 0[32] Nous lui évitons ainsi la contrariété suivante : s'il prévoit avec une bonne confiance qu'il ne pleuvra pas (comme la plupart du temps), on ne lui répliquera pas aussitôt : "Et si jamais il pleut, quelle est votre prévision centrale, et votre intervalle de confiance à 50% ?"[33].

Notre approche rendra en revanche plus délicate l'évaluation a posteriori de la fiabilité des prévisions émises (cf Annexe 6).

Une voie d'amélioration serait de leur demander de s'exprimer en probabilités totales, et de les interroger en plus sur leur évaluation de la probabilité de pluie significative (>0,1 mm) (Vincent Fortin, comm. pers).

Enfin, pour sensibiliser le prévisionniste à l'impossibilité de parler de risque sans évaluer celui de se tromper, nous lui demandons en outre de fournir systématiquement un quantile élevé : le quantile 90.

## 4.2 L'introduction des probabilités subjectives en salle de prévision

### 4.2.1 Objectifs

Nous avons débattu en 4.1 de l'intérêt des probabilités pour faire exprimer au prévisionniste son incertitude sur ses prévisions. Les probabilités subjectives nous ont paru les plus à même de contribuer à l'amélioration de la prévision opérationnelle et de la communication des phénomènes à risques. Nous avons identifié les méthodes les plus intuitives pour initier le prévisionniste à leur utilisation :

- la détermination d'intervalles de crédibilité et de quantiles de probabilité ;
- l'utilisation de probabilités totales plutôt que de probabilités mixtes.

Notre objectif était d'appliquer ces pratiques en salle de prévision à EDF, de manière à étudier leur pertinence pour l'amélioration de la performance du système de prévision hydrométéorologique de l'entreprise. Il nous fallait pour cela mettre en place un protocole d'expérimentation.

---

[32] dans ce cas, seule la fourniture des quantiles 90, voire 99 peut nous informer sur les risques de pluies significative.

[33] Notons, bien entendu, que l'information fournie à l'aide des probabilités totales est moins riche que celle fournie à l'aide des probabilités conditionnelles : 3 paramètres ($x_{25}$, $x_{50}$, $x_{75}$) avec les probabilités totales contre 4 avec les probabilités conditionnelles ($\nu$, $x_{25}$, $x_{50}$, $x_{75}$).
Ajoutons que Krzysztofowicz et Sigrest (1999) avaient commencé par travailler avec les prévisionnistes en probabilités totales, avant que leur expérience ne les fasse changer de méthodologie (R. Krzysztofowicz, comm. pers.). Ces deux auteurs se sont en effet aperçus que les prévisionnistes raisonnaient très souvent de manière conditionnelle (par exemple, en plein été, une situation convective peut produire de violents orages si la convection est suffisante, rien sinon). Les probabilités conditionnelles leur semblaient donc l'outil mathématique adéquat pour traduire le jugement des prévisionnistes.

### 4.2.2 Le stage de sensibilisation à l'usage des probabilités en prévision opérationnelle

A l'automne 2001, nous avons proposé un premier stage de formation et de sensibilisation des prévisionnistes à la quantification des incertitudes en prévision opérationnelle (description en Annexe 2).

Ce stage, qui a eu lieu dans chaque centre de prévision (CHA et CHPMC), avait pour principal objectif de sensibiliser les prévisionnistes à l'intérêt de communiquer l'incertitude liée à chaque prévision météorologique ou hydrologique.

Nous leur avons proposé de fournir une estimation de leurs propres incertitudes sur les précipitations attendues sur six bassins, dans deux situations météorologiques différentes (Tableau 12 et Figure 21).

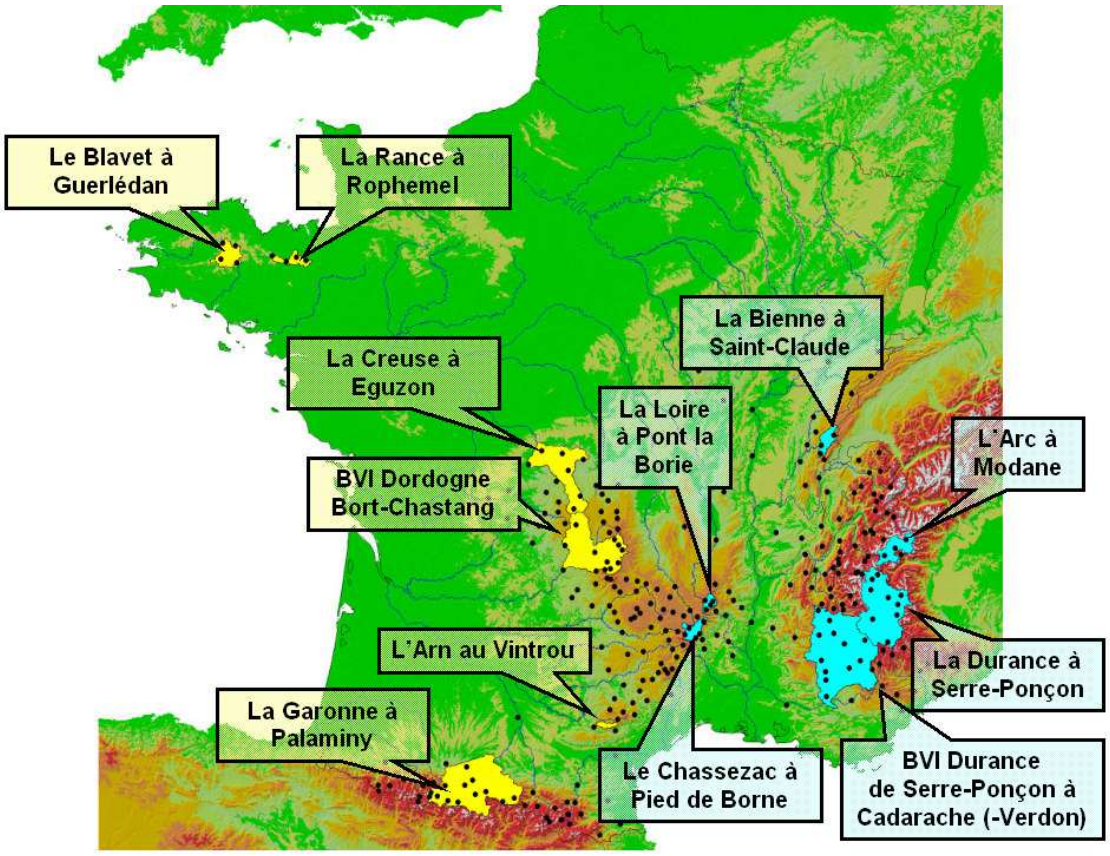



**Figure 21 : Localisation des 6 bassins versants du CHPMC (en jaune) et des 6 du CHA (en bleu). Prévision probabiliste de précipitation. Stage de sensibilisation-formation à EDF (les points localisent les pluviomètres).**

| Equipe de prévision | Bassin versant | Surface (km²) |
|---|---|---|
| CHA | La Durance à Serre-Ponçon | 3580 |
| | La Bienne à Saint-Claude | 450 |
| | L'Arc à Modane | 840 |
| | La Loire à Pont dela Borie | 230 |
| | Le Chassezac à Pied de Borne | 360 |
| | BVI Durance de Serre-Ponçon à Cadarache (sauf Verdon) | 6000 |

| Equipe de prévision | Bassin versant | Surface (km²) |
|---|---|---|
| **CHPMC** | Le Blavet à Guerlédan | 680 |
| | La Rance à Rophemel | 380 |
| | La Garonne à Palaminy | 4500 |
| | La Creuse à Eguzon | 2400 |
| | L'Arn au Vintrou | 130 |
| | BVI Dordogne de Bort à Chastang | 3100 |

**Tableau 12 : Surface des 6 bassins versants du CHA et des 6 du CHPMC.**
**Prévision probabiliste de précipitation. Stage de sensibilisation-formation à EDF.**

Entre les deux situations météorologiques présentées, nous avons à chaque fois réalisé un débriefing collectif sur la fiabilité de l'estimation que les prévisionnistes venaient de faire de leurs propres incertitudes. Pour alimenter ce temps de discussion, nous leur avons également proposé le questionnaire de culture générale inspiré par Biais et al. (2002, voir 8.2.1), avant d'en analyser ensemble les résultats.

Expérience menée au CHA

Les prévisionnistes ont directement estimé les quatre quantiles 25, 50, 75 et 90 sur leur prévision de la lame d'eau spatiale sur 6 bassins versants.

*Contrôle de cohérence* : Les prévisionnistes ont ensuite été amenés à parier sur ces quantiles, pour contrôler la conformité des chiffres fournis avec leur jugement intime de la situation ("type one goodness", Murphy, 1993). On a ainsi vérifié s'ils étaient effectivement prêts à prendre des paris en cohérence avec les chiffres annoncés pour quantifier leurs incertitudes.

Les paris proposés étaient de la forme suivante :

> *Donnez vos préférences entre le jeu A et le jeu B. Pesez bien ce que chacun des jeux signifie, et donnez vos préférences en vous mettant vraiment dans la peau du parieur que vous seriez en telle situation.*
>
> **Jeu A** : dans une urne, il y a 6 boules rouges, et une boule verte. On tire une boule au hasard.
>
> Si c'est la boule verte qui sort, vous perdez 100 FF ; si c'est une rouge, vous gagnez 30 FF.
>
> **Jeu B** : Si la lame d'eau spatiale sur l'Arc à Modane pour les 24h à venir dépasse L90 (tel que vous l'avez vous-même défini), vous perdez 100 FF ; sinon, vous gagnez 30 FF.
>
> ***A quel jeu préférez-vous jouer ?***

Ils avaient ainsi le choix de prendre soit un **pari en univers risqué** (tirage d'une boule dans une urne), soit un **pari en univers incertain** (la prévision météorologique).

Expérience menée au CHPMC

Les prévisionnistes ont tout d'abord été amenés à prendre des paris sur la lame d'eau spatiale prévue sur chacun des 6 bassins.

Les paris proposés étaient de la forme suivante :

*Par la suite, on va vous demander de nous donner vos préférences entre plusieurs jeux. Pesez bien ce que chacun des jeux signifie, et donnez vos préférences en vous mettant vraiment dans la peau du parieur que vous seriez en telle situation.*

***Choix n°-1***

**Jeu A** : dans une urne, il y a 2 boules rouges, et 3 boules bleues. On tire une boule au hasard.

Si c'est une boule bleue qui sort, vous empochez 100 FF ; si c'est une rouge, vous êtes quitte.

**Jeu B** : Si la lame d'eau spatiale sur les 24h à venir (cumul 8h-8h) dépasse 5 mm sur le Blavet à Guerledan, vous empochez 100 FF ; sinon, vous êtes quitte.

***A quel jeu préférez-vous jouer ?***

*Contrôle de cohérence* : On a ensuite demandé aux prévisionnistes de fournir plusieurs quantiles de probabilité (25, 50, 75 et 90) sur la lame d'eau attendue sur chacun des 6 bassins. Cela nous a permis de contrôler la cohérence entre paris pris dans un 1er temps et quantiles élicités dans un 2nd temps.

### 4.2.3 La prévision probabiliste au fil de l'eau

Suite à ce stage de sensibilisation, l'équipe des prévisionnistes du CHPMC a mis en place un banc d'essai pour estimer, au quotidien, ses incertitudes de prévision. Il s'agit donc d'une expérience réalisée "au fil de l'eau" pour produire des prévisions probabilistes.

Ils ont eux-même choisi les variables à prévoir :

- Le vent à Bordeaux, en rafales (vent instantané maximum) ;
- Les débits moyens journaliers de la Creuse à Eguzon, du Tarn à Millau, et de la Garonne à Palaminy ;
- Les lames d'eau spatiales sur les quatre bassins versants suivants : la Creuse à Eguzon (2400 km²), le Tarn à Millau (2170 km²), la Garonne à Palaminy (4400km²) et l'Arn au Vintrou (130 km²) (Figure 22).

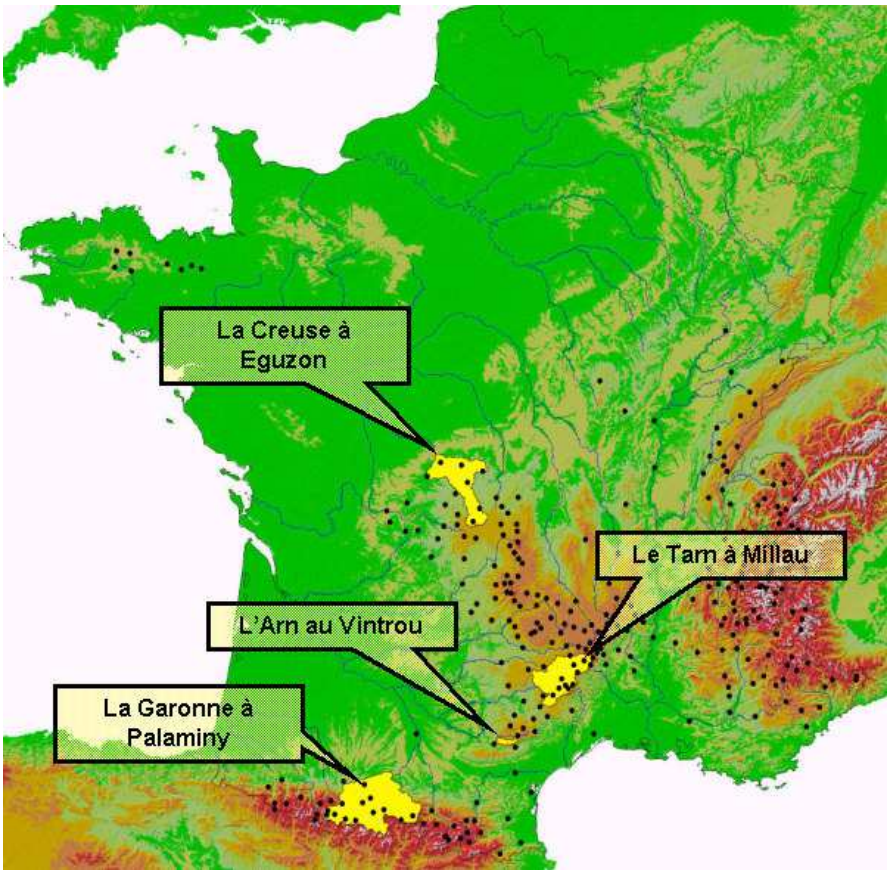



**Figure 22 : Localisation des bassins versants pour la prévision probabiliste de précipitation. Expérience semi-opérationnelle à EDF.**

De juin à décembre 2002, les 8 prévisionnistes du CHPMC ont alors produit des prévisions probabilistes aux horizons J et J+1 sur ces paramètres hydrométéorologiques :

Pour caractériser leurs incertitudes sur la prévision courante, il leur était demandé de fournir pour chaque prévision les quantiles 10%, 50% et 90%. En d'autres termes, on leur demandait de fournir, en plus de leur prévision déterministe (tendance centrale ou scénario hydrométéorologique le plus probable), un intervalle de crédibilité à 80% centré en probabilité autour de cette prévision déterministe[34].

[34] R. Krzysztofowicz (comm.pers.) souligne que dans la littérature, **la définition d'une prévision déterministe** n'est nulle part donnée :

- est-ce "la réalisation future" ? ce serait une prédiction (et non une prévision), ne laissant aucune place aux incertitudes (Krzysztofowicz, 2001) ; pourtant, en hydrométéorologie, personne ne peut ignorer l'existence d'incertitudes conséquentes.
- est-ce "la réalisation la plus probable aux yeux du prévisionniste" ? cela reviendrait à prévoir une pluie nulle pour toutes les situations convectives (pleuvoir précisément 0mm étant toujours plus probable que pleuvoir précisément une hauteur X strictement positive)

R. Krzysztofowicz a pour sa part choisi de définir la prévision déterministe comme l'estimation ponctuelle telle que la réalisation a autant de chance de tomber au-dessus qu'au-dessous (selon le jugement du prévisionniste). **La prévision déterministe est donc définie comme la médiane de la fonction de probabilité tirée du jugement du prévisionniste.** C'est cette définition que nous avons également tâché de faire adopter aux prévisionnistes de DTG.

Pour les lames d'eau, les prévisionnistes ont fourni uniquement les quantiles 50% et 90% : le quantile 10% aurait été presque toujours fixé à 0, et n'aurait donc présenté que très peu d'intérêt.

En ne demandant que trois chiffres pour caractériser la prévision d'un paramètre et l'incertitude qui lui est associée, notre démarche était volontairement minimaliste. Il s'agissait avant tout de trouver une manière très intuitive de quantifier l'incertitude courante de prévision, et de rester aussi fidèle que possible au jugement du prévisionniste concernant la situation en cours.

Compte tenu des données manquantes (en prévisions ou en observations), nous avons finalement recueilli 115 journées de prévision, pour deux horizons (J et J+1), et 8 paramètres.

## 4.3 Conclusion

Convaincus de la pertinence des probabilités subjectives pour la prévision opérationnelle, nous les avons introduites en salle de prévision, sans trop bouleverser les habitudes de travail des prévisionnistes.

Nous nous sommes ainsi appuyés sur l'existant : le prévisionniste avec ses outils, mais aussi ses doutes, ses convictions, ses avis. Celui-ci, avec les seules armes de son jugement, a été amené à prendre des paris sur l'avenir, paris qui se traduisent en probabilités subjectives.

Les résultats de ces expériences ont constitué une matière première qui nous a permis de :

- démontrer, pour le cas concret d'EDF la valeur ajoutée de la prévision probabiliste (Chapitre 6) ;
- étudier la perception des incertitudes en salle de prévision, et d'en déduire les précautions à prendre lors du déploiement opérationnel de la prévision opérationnelle probabiliste (Chapitres 7, 8 et 9).

## 4.4 Discussion : ce que nous croyons, ce que nous connaissons, et ce que nous croyons connaître…

Que cherchons-nous à éliciter au juste, les croyances de l'expert ou ses connaissances ? Les mots sont en effet révélateurs de deux conceptions différentes du comportement d'un individu sous incertitude :

- *« Nous formulons des **croyances** parce que c'est pratique pour prendre efficacement de bonnes décisions : peu importe si ces croyances sont le reflet d'une quelconque réalité, ni même si cette réalité existe »*. C'est un point de vue constructiviste, qui consiste à penser que toute donnée ou observation est une réponse à un questionnement, et n'a de sens qu'en référence à ce questionnement. Cette position a été défendue par Bachelard (1934), ou encore par Paul Valéry, auteur de la célèbre formule « il n'y a pas de données ; il n'y a que des modèles. ». On peut résumer ce courant de pensée par l'assertion suivante :
  *« J'y crois parce que c'est pratique »*
- *« Nos **connaissances** sont les bribes d'une réalité unique et universelle qui nous est partiellement cachée par l'étendue de notre ignorance. C'est sur la base de nos connaissances que nous prenons des décisions ».* C'est bien l'assertion suivante qui

est à la base de tout :
*« J'y crois parce que c'est vrai »*

Dans un cas comme dans l'autre, on s'accorde à dire que la probabilité n'est pas un attribut de l'événement considéré, comme sa propension à survenir, mais bien un construit du sujet connaissant, (ou croyant connaître…) et pensant. « L'incertitude est une notion relative à l'individu, qui ne peut être séparée de son corpus de connaissances, d'intuitions et de culture » (Bernier et al., 2000, p. 157).

Nous ne ferons pas par la suite d'hypothèse ontologique sur une possible réalité unique et universelle. Cette hypothèse nous semble de surcroît très forte dans le cas de la prévision hydrométéorologique. Elle supposerait en effet un déterminisme (voire une « prédestination météorologique »), plus ou moins caché à nous, pauvres mortels, et qui le serait moins pour les prévisionnistes, encore moins pour ceux qui parmi eux sont les plus expérimentés. Une telle hypothèse est inutile dans le cas qui nous intéresse (puisque Savage (1954) a montré que les seules préférences permettent de construire des probabilités, à la condition que l'on respecte certains axiomes comportementaux).

Nous nous attacherons donc à éliciter les croyances de l'expert, révélées au travers de ses préférences entre différents paris. Ces croyances nous révèleront à la fois le comportement de l'expert en univers incertain (Chapitre 7), les biais qui peuvent apparaître dans son jugement (Chapitre 8), même si la méthode peut bouleverser les habitudes des prévisionnistes (Chapitre 9). Nous verrons au Chapitre 6 que la langue des probabilités - qui découle de l'élicitation des croyances de l'expert – révèle et améliore la performance du système de prévision face à l'enjeu de Sûreté Hydraulique.

Pourtant, les chercheurs de diverses sciences humaines sont loin d'être unanimes sur le bien-fondé de l'usage des probabilités pour la communication entre l'expert et le décideur (Chapitre 5). Il convient donc de bien définir les différentes postures épistémologiques afin de mieux positionner notre recherche.

## 4.5 Bibliographie

Bachelard, G., 1934, Le nouvel esprit scientifique, Presses universitaires de France (6e Ed.).

Bernier, J., Parent, E., Boreux J.-J., 2000. Statistiques pour l'environnement, traitement bayésien de l'incertitude, Tec et Doc, Lavoisier.

Bontron, G., 2004. Prévision quantitative des précipitations : adaptation probabiliste par recherche d'analogues. Utilisation des réanalyses NCEP/NCAR et application aux précipitations du sud-est de la France. Thèse de l'INPG, janvier2004.

Golding, B., W., 2000. Quantitative precipitation forecasting in the UK, Journal of Hydrology, 239, 286-305.

Krzysztofowicz, R., 2002. Bayesian system for probabilistic river stage forecasting, Journal of Hydrology, Vol. 268, 1-4, 16-40.

Krzysztofowicz R., 2001. Hydrologic uncertainty processor for probabilistic river stage forecasting: precipitation-dependent model, Journal of Hydrology, Vol. 249, 1-4, 46-68.

Krzysztofowicz R., 2001. Integrator of uncertainties for probabilistic river stage forecasting: precipitation-dependent model, Journal of Hydrology, Vol. 249, 1-4, 69-85.

Krzysztofowicz, R., 2001. The case for probabilistic forecasting in hydrology, Journal of Hydrology, Vol. 249, 2-9.

Krzysztofowicz R., 2001. Reply to comment by Schaake, J.C., Welles, E., Graziano, T., on "Bayesian theory of probabilistic forecasting via deterministic hydrologic model", Water Resources Research, Vol. 37, n° 2, 441-442.

Krzysztofowicz, R., Drzal, W., J., Drake, T., R., Weyman, J., C., Giordano, L., A., 1993. Probabilistic quantitative precipitation forecasts for river basins. Weather and Forecasting, vol. 8., n°4, 424-439.

Krzysztofowicz, R., Sigrest, A., 1999. Calibration of probabilistic quantitative precipitation forecasts. Weather and Forecasting, 14(3), 427-442.

Murphy, A. H., 1993. What is a good forecast ? An essay on the nature of goodness in weather forecasting. Weather and forecasting, vol. 8, no 2, 281-293.

Murphy, A., H., Winkler, R., L., 1974. Credible interval temperature forecasting: some experimental results. Monthly Weather Review, vol. 102, 784-794.

Savage, L. J., 1954, The Foundations of statistics. Wiley.

Schaake, J.C., Welles, E., Graziano, T., 2001. Comment on "Bayesian theory of probabilistic forecasting via deterministic hydrologic model". Water Resources Research, Vol. 37, n° 2, 439.

# 5 L’usage des probabilités est sujet à controverses

*On sait les problèmes de la prévision déterministe, et les difficultés pour s’en dépêtrer.*

*On sait aussi les atouts de la prévision probabiliste.*

*Pourtant, les théoriciens de diverses disciplines en sciences économiques et sociales ne sont pas d’accord sur l’usage de cette langue pour parler du risque.*

## 5.1 Le point de vue des systémiciens et sociologues des organisations

***Passer au quantitatif pour formaliser les incertitudes n'est pas intuitif pour le prévisionniste ni limpide pour l'usager ; l'information de prévision perd de sa richesse.***

Les systémiciens et les sociologues des organisations tendent à privilégier les aspects qualitatifs des informations véhiculées. Ces chercheurs préfèrent les raisonnements intuitifs à la logique discursive.

Le paradigme de la complexité développé par Morin (La Méthode, tomes I à V, 1977 à 2001)[35] invite le chercheur à modéliser tout système complexe en se concentrant sur les interactions croisées entre les différents composants du système, et sur le phénomène de récursivité[36] qui en affecte le fonctionnement.

Appliquée à notre question de recherche[37], cette option consiste à considérer comme système global l'ensemble des réseaux d'acteurs qui interviennent dans la gestion des ouvrages hydroélectriques (à tous les niveaux hiérarchiques et fonctionnels, depuis la conception jusqu'au pilotage opérationnel des installations).

La modélisation du système consiste à décrire son fonctionnement global sur la base des observations que l'on peut en faire lorsque celui-ci est soumis à divers éclairages et à diverses contraintes. Cette modélisation intègre les aspects techniques, humains et organisationnels dans le fonctionnement du système.

La valeur de l'information de prévision ne peut alors s'évaluer qu'à l'aune de la qualité du fonctionnement global du système, en particulier de sa résilience[38] lors d'événements extrêmes. Dans cette perspective, on cherche à mettre en lumière des modes de communication entre les protagonistes qui soient intuitifs pour chacun d'eux. Plutôt que de passer par l'artifice d'un outil mathématique qui nécessite un travail de cryptage (élicitation des probabilités subjectives) puis de décryptage (compréhension et interprétation des quantiles fournis dans la prévision), le modélisateur s'attache à observer le domaine de recouvrement maximal entre le savoir tacite détenu par l'usager et les connaissances du prévisionniste lors de l'exercice de son expertise hydrométéorologique. L'information est d'autant plus valorisée qu'elle remonte suffisamment dans les niveaux d'expertise du prévisionniste sans sortir du domaine de compétence de l'usager.

Deux types d'information n'ont donc aucune valeur pour l'usager : les tautologies pour lesquelles l'usager n'apprend rien, et les informations nécessitant un niveau d'expertise trop élevé pour lesquelles l'usager ne comprend rien. Entre ces deux extrêmes existent toutes les

---

[35] Ce paradigme est repris par Le Moigne (1992) à propos de "la modélisation des systèmes complexes".

[36] D'une part, en fonctionnant, les composants d'un système se transforment, en nature et en structure (de nouveaux modes d'organisation apparaissent). D'autre part le modélisateur et le système modélisé -c'est à dire l'organisation- interagissent ; l'intentionnalité du chercheur guide son action de modélisation, qui influe sur le système modélisé. Les signaux perçus par le modélisateur sur le comportement du système et ses évolutions orientent également sa manière de modéliser. Si l'on accepte de prendre en compte cette récursivité, on ne peut donc plus parler d'objet d'étude mais de projet de modélisation.

[37] Notre question de recherche peut se résumer ainsi : Comment formuler puis communiquer de manière opérationnelle des prévisions hydrométéorologiques entachées d'incertitude pour que ces informations soient valorisées au mieux par l'usager ?

[38] La résilience est l'aptitude d'un système complexe à rester sûr lorsqu'il est soumis à des perturbations (Wybo, 2001) (6.1.6).

prévisions, incertaines et imparfaites, pour lesquelles tout l'art du modélisateur est de trouver l'ensemble des notions techniques pour lesquelles le prévisionniste et l'usager sont au diapason, la sémantique qui est leur terrain d'entente, tous ces éléments qui permettent au prévisionniste de faire de l'usager un complice de son expertise et d'améliorer la résilience du système.

La mesure de la valeur de l'information ne peut se faire que qualitativement, grâce à une métrologie dépendante du système. A défaut de pouvoir répondre à la question "combien vaut l'information ?", on peut néanmoins répondre aux questions "en quoi l'information est utile au bon fonctionnement global du système ?", puis "comment améliorer le fonctionnement du système et en accroître les synergies ?".

Considérons le système de gestion de crue, c'est à dire l'organisation et les moyens techniques dont elle s'est dotée pour gérer les ouvrages hydroélectriques en crue et assurer la Sûreté Hydraulique. La valeur de la prévision délivrée par le centre hydrométéorologique s'apprécie en comparant le niveau de sûreté atteint par le système de gestion de crue avec et sans la présence d'un système de prévision (0). Pour mesurer ces divers niveaux de sûreté différentes méthodes peuvent être utilisées : l'évaluation des défenses en profondeur du système (6.1.4), l'estimation de ses marges de manœuvre (6.1.5), ou la description de la résilience (6.1.6),

Ces méthodes permettent de mettre en lumière les types d'organisation résilients, et de décrire en quoi la prévision contribue à l'amélioration de la Sûreté Hydraulique. Elles ne permettent cependant pas de placer le décideur devant les choix suivants :

- quel est le risque acceptable auquel il accepte de se soumettre ?
- quelle est la hauteur des moyens à mettre en œuvre pour atteindre ce niveau de risque acceptable ?
- comment arbitrer entre différents moyens qui concourent à la Sûreté Hydraulique (maintenance préventive, surdimensionnements lors de la conception des installations, personnel de quart, prévision hydrométéorologique, etc.) ?

C'est notamment pour aider le décideur à faire face à ces arbitrages qu'une approche quantitative est nécessaire ; une telle approche permet de mesurer d'une part les incertitudes de prévision, et d'autre part la valeur de l'information produite. C'est l'approche qui est prônée par les économistes.

## 5.2 Le point de vue des économistes

***Utiliser les probabilités pour formaliser les incertitudes de prévision est indispensable pour que le modélisateur mesure la performance du système de prévision, que le prévisionniste apprenne à mieux se connaître, et que l'usager valorise au mieux la prévision.***

Quantifier les incertitudes et utiliser l'outil mathématique nécessite d'adopter un cadre de travail reposant sur plusieurs hypothèses opposées à celles du sociologue des organisations. En particulier, les économistes supposent que le modélisateur n'influe pas sur son objet d'étude : l'organisation des acteurs qui interviennent dans la gestion des ouvrages hydroélectriques. Cela revient à supposer qu'il n'y a pas de récursivité dans le système, ni entre le modélisateur et le système.

A la différence des systémiciens, les économistes se placent ici dans le cadre de la méthode analytique[39] . Ils découpent le système étudié en éléments simples : le prévisionniste et l'usager, qui échangent de l'information. Pour modéliser chacun d'eux comme un agent économique, ils découpent ensuite ses actions, ses préférences et ses croyances en des ensembles discrets :

- pour le prévisionniste, l'élicitation des probabilités subjectives révèlera ses croyances sur l'état futur de la nature (décrit par diverses variables hydrométéorologiques) ;
- pour l'usager, ses actions avec et sans l'information de prévision, et ses préférences sur les conséquences de ces actions révèleront la valeur de l'information de prévision.

Pour pouvoir modéliser les incertitudes de prévision par des probabilités subjectives, il faut supposer la rationalité du prévisionniste au sens des quatre axiomes qui fondent la statistique décisionnelle (Savage, 1954) :

A1 : invariance vis-à-vis de l'achat ou de la vente d'un pari ;

A2 : existence d'une métrique pour laquelle il y a insensibilité aux effets d'échelle lors de l'évaluation du prix d'achat d'un pari ;

A3 : refus de transactions systématiquement déficitaires ;

A4 : acceptation du raisonnement conditionnel.

D'autre part, pour modéliser le comportement d'un agent économique tel que l'usager de la prévision aux prises avec ses enjeux opérationnels, les économistes adoptent certaines hypothèses, bien que celles-ci soient rarement explicitées. On rappelle ici les plus fondamentales (Walliser, 2000) :

H1 : les croyances d'un acteur sont de même nature que celles du modélisateur ;

$H1_{bis}$ : le modélisateur est doté d'un modèle quasi-parfait du système, alors que l'acteur n'en possède qu'un modèle partiel et bruité, défini par référence au précédent ;

H2 : les croyances de l'acteur se nourrissent essentiellement d'informations externes, et non de savoirs tacites et inexprimés, ancrés dans son esprit ;

$H2_{bis}$ : à partir des informations qui lui sont données, il y a révision et non reconstruction des croyances de l'acteur (principe d'économie) ;

H3 : l'acteur cherche toujours à optimiser ses actions en fonctions de leurs conséquences (rationalité de l'acteur);

$H3_{bis}$ : l'acteur a au préalable simulé mentalement les conséquences de toutes les stratégies possibles.

La modélisation systémique est descriptive, alors que les principes qui guident la modélisation analytique sont **normatifs** : l'agent est supposé agir de manière sous-optimale par rapport aux normes comportementales édictées[40], du fait de ses défaut cognitifs ou de normes sociales ou

[39] Etymologiquement, le mot analyse traduit l'idée de décomposition. Selon la définition du Petit Robert, l'analyse est une méthode comportant un examen discursif en vue de discerner les éléments. En substance, l'analyse présuppose de découper l'objet de recherche en autant d'éléments qu'il faut pour que chacun d'eux puisse être accessible à notre faculté de raisonner (méthode cartésienne).

[40] Pour l'usager, ces normes prescrivent d'envisager l'ensemble des actions possibles, leurs conséquences, puis d'optimiser. Le prévisionniste quant à lui est supposé envisager l'ensemble des évolutions météorologiques possibles, leur vraisemblance, leurs conséquences en terme de débit au droit d'un aménagement, puis construire la fonction de répartition du débit prévu.

professionnelles inadaptées à certains types de situation. En conséquence, les économistes cherchent à améliorer la rationalité des agents en montrant quels pourraient être les gains si chacun était parfaitement rationnel.

En ce sens, on retrouve en psychologie cognitive des hypothèses (Kahneman, Slovic, Tversky, 1982, noté par la suite KST) qui font parfaitement écho à celles énoncées plus haut à propos de l'économie (Tableau 13). Notons que pour désigner un individu, on emploie le terme de sujet en psychologie, d'acteur, ou d'agent en économie, et de décideur en statistique décisionnelle.

| **Economie** | **Psychologie cognitive** |
|---|---|
| Von Neuman et Morgenstern (1943), Savage (1954), Walliser (2000), etc. | Kahneman, Slovic et Tversky (1982), Palmarini (1995), etc. |
| H1 : Les croyances d'un acteur sont de même nature que celles de l'économiste. | H1 : Le psychologue et le sujet utilisent les mêmes concepts (vraisemblance, probabilités, etc.). |
| H1bis : L'économiste est doté d'un modèle quasi-parfait du système, alors que l'acteur n'en possède qu'un modèle partiel et bruité, défini par référence au précédent. | $H1_{bis}$ : "Le sujet utilise un nombre limité de principes heuristiques qui l'amènent parfois à des erreurs sévères et systématiques" (KST). Le psychologue quant à lui manie correctement les concepts et les règles de calcul des probabilités. |
| H2 : Les croyances de l'acteur se nourrissent essentiellement d'informations externes, et non de savoirs tacites et inexprimés, ancrés dans l'esprit de l'acteur. | H2 : Les croyances de l'acteur se nourrissent essentiellement d'informations externes, et non de savoirs tacites et inexprimés, ancrés dans l'esprit de l'acteur.<br>Exemple du choix des priors égaux à la proportion des taxis verts et des taxis bleus dans le fameux problème du même nom[41]. On suppose que le sujet n'utilise pas de savoir tacite comme celui-ci : "la compagnie la plus grande est plus sérieuse dans le choix de ses conducteurs" (KST). |

**Tableau 13 : Synoptique des postulats qui fondent l'économie cognitive et la psychologie cognitive.**

[41] Lyon et Slovic, 1976, Tversky et Kahneman, 1982. Le problème posé est le suivant "Une nuit, après avoir causé un accident, un taxi a pris la fuite. Il n'existe que deux compagnies de taxi dans cette ville, les Taxis Verts et les Taxis Bleus. 85% des taxis de la ville sont Verts ; les 15% restants sont Bleus. Un témoin de l'accident rapporte qu'il s"agissait d'un Taxi Bleu. Le juge a fait tester la fiabilité de ce témoin dans les mêmes circonstances que celles de la nuit de l'accident. L'étude a conclu que le témoin nommait "taxi bleu" 80% des taxis Bleus, mais nommait 20% de ces taxis Bleus "taxi vert". Le témoin appelait aussi "taxi vert" 80% des taxis Verts, mais appelait "taxi bleu" 20% des taxis Verts. Quelle est la probabilité que le taxi impliqué dans l'accident soit Bleu ?"

## 5.3 Le point de vue avancé dans cette thèse

***L'introduction des probabilités permet d'acquérir un nouvel outil approprié à la gestion collective du risque hydrométéorologique, et répond à trois objectifs :***
***1. analyser la perception des incertitudes lors de l'élaboration de la prévision ;***
***2. mesurer la valeur économique de la prévision pour l'usager ;***
***3. évaluer puis améliorer la performance du système de prévision vis-à-vis de l'enjeu de Sûreté Hydraulique.***

Dans un premier temps, nous procédons suivant **la méthode analytique**, pour découper le système de gestion des ressources en eau en trois acteurs principaux : le prévisionniste, l'usager en charge de l'optimisation des moyens de production hydroélectriques (symbolisé par Max, en section 2.1.3), et l'usager en charge de la Sûreté Hydraulique sur les ouvrages (symbolisé par Jéo, en section 2.1.3).

### 5.3.1 Des probabilités pour analyser la perception des incertitudes lors de l'élaboration de la prévision

Poser les hypothèses A1 à A4 énoncées en 5.2 nous permet d'utiliser l'outil mathématique des probabilités pour analyser quantitativement le comportement d'un expert prévisionniste en situation d'incertitude. Nous pouvons ainsi rendre compte de la connaissance qu'il a de sa propre connaissance, et évaluer sa perception du risque hydrométéorologique en diverses situations (Chapitres 7, 8 et 9).

Lors de l'analyse du comportement du prévisionniste, deux logiques peuvent s'affronter (Stanovich, 1999)[42] :

- **La logique des Panglossiens** (de Maître Pangloss, dans Voltaire (1759), pour qui "tout va pour le mieux dans le meilleur des mondes possibles") : ceux-ci se situent dans un courant néo-darwinien, et pensent que les individus s'adaptent à leur environnement, notamment en développant (à force d'apprentissage de type essai-erreur) des heuristiques qui leur permettent de prendre globalement des décisions satisfaisantes. Tous les biais et défauts présents dans leurs raisonnements proviennent du fait que l'expérimentateur les a arrachés à leur environnement naturel, et que dans ces conditions, ils ne disposent plus de la globalité des outils, informations et heuristiques auxquels ils ont l'habitude de faire appel.
  Hilton (1995) montre par exemple que les règles sociales régissant la communication entre les sujets font partie intégrante du contexte social dans lequel celui qui écoute doit se forger une opinion. Ces règles exigent du sujet qui écoute qu'il aille plus loin que la seule information qui lui est délivrée, et qu'il intègre dans son jugement sa perception de celui qui lui parle[43]. Ne pas prendre en compte cette hypothèse pourrait conduire à attribuer à tort des défauts cognitifs à une personne qui aurait pourtant pris une décision rationnelle compte tenu du contexte social dans lequel s'est effectuée la prise de décision.

- **La logique des Mélioristes** : ceux-ci pensent non seulement que les sujets ont des défauts cognitifs, mais ils prétendent en plus pouvoir les corriger à l'aide d'exercices plus ou moins académiques.
  Savage (1954) postule que « le rôle de la théorie mathématique des probabilités est de

[42] Les résultats présentés aux Chapitres 7 et 8 nous permettront d'illustrer les interprétations très différentes que font panglossiens et mélioristes à propos de phénomènes observés en psychologie expérimentale.

[43] Hilton (1995) nomme cette hypothèse "conversational inference". Selon celle-ci, celui qui écoute infère de la conversation certaines qualités de son interlocuteur, qui orienteront ensuite l'interprétation du message reçu.

> permettre à la personne qui l'utilise de détecter les irrationalités de son propre comportement, que celui-ci soit réel ou simplement envisagé. Il est également entendu que, ayant détecté une irrationalité, la personne pourra s'en défaire. » (p 57).
> Selon Krzysztofowicz et Sigrest (1999), "la calibration des prévisionnistes révèle certaines faiblesses dans leur jugement" et il faudrait concevoir "un entraînement individualisé des prévisionnistes pour améliorer leur calibration" (p 439).

Nous retenons donc certains arguments des Panglossiens pour réaliser les expériences dans des conditions aussi "écologiques" que possibles (méthodes interactives et intuitives d'élicitation des probabilités subjectives, expérimentation réalisée au fil de l'eau, dans les conditions de travail opérationnel, etc.). Le point de vue panglossien peut également d'enrichir l'interprétation de certains résultats obtenus (surconfiance, Chapitre 7, biais, Chapitre 8, ou aversion à l'incertitude, Chapitre 9) en intégrant des éléments du contexte social dans lequel s'élabore la prévision.

L'école mélioriste oriente quant à elle nos recherches sur la perception des incertitudes (et notamment la surconfiance), pour aider les experts à améliorer la connaissance de leur propre connaissance.

Nous répondons ainsi à l'objectif 1 dans les Chapitres 7, 8 et 9 : nous analysons la perception des incertitudes lors de l'élaboration de la prévision, puis nous mesurons la fiabilité des prévisionnistes dans l'estimation de leurs incertitudes.

### 5.3.2 Des probabilités pour mesurer la valeur de la prévision pour l'usager en charge de l'optimisation de la production

Pour caractériser le comportement de l'usager en charge de l'optimisation de la production nous nous plaçons dans le cadre de l'analyse économique, en adoptant notamment les hypothèses H1 à H3. Nous obtenons ainsi la valeur de l'information pour un usager qui sait optimiser ses actions en fonction de la prévision fournie, qu'elle soit déterministe (3.2) ou probabiliste (3.3).

Comme dans sa gestion opérationnelle l'usager utilise un logiciel d'optimisation de la production, il ne peut valoriser que des informations quantitatives sur l'aléa hydrologique. L'utilisation de prévisions probabilistes lui permet de disposer d'une information riche sur l'avenir à long terme, et donc d'améliorer substantiellement sa gestion de la ressource en eau, même si cette information peut sembler floue de prime abord (3.3).

Tout laisse à penser que sur le court-terme également, la prévision probabiliste améliore l'optimisation de la production par rapport à l'utilisation d'une donnée déterministe. Pour des raisons pratiques, nous n'avons cependant pas pu tester cette hypothèse.

Nous répondons donc en partie à l'objectif 2 dans le Chapitre 3 : mesurer la valeur économique de la prévision pour l'usager.

### 5.3.3 Les probabilités, un nouvel outil pour une meilleure gestion du risque lié à la Sûreté Hydraulique

La quantification des incertitudes de prévision permet également de mesurer et de tracer la performance du système de prévision (Chapitre 6). C'est par l'usage systématique des probabilités par les prévisionnistes que nous pouvons formaliser leur savoir sur le niveau

d'incertitude de la prévision, savoir qui à ce jour est transmis ou non à l'usager au gré des sollicitations.

Comme nous l'avons évoqué en 5.1, la modélisation systémique est particulièrement appropriée pour montrer la contribution de la prévision à l'amélioration de la Sûreté Hydraulique. C'est même l'unique moyen de décrire la valorisation de l'information, dans la mesure où les décideurs responsables de la Sûreté Hydraulique à EDF n'ont pas su se prononcer sur leur consentement à payer pour atteindre tel ou tel niveau de sûreté (6.2.2).

Nous ne pouvons donc pas quantifier la valeur de l'information fournie en terme d'amélioration de la Sûreté Hydraulique. L'usage des probabilités permet cependant de révéler une part du savoir tacite du prévisionniste concernant les risques à venir, information qui peut être particulièrement utile à l'usager (2.3). En ce sens, l'usage des probabilités révèle en partie la performance du système de prévision (6.2).

L'usage des probabilités fournit de plus des éléments quantitatifs d'évaluation de cette performance vis-à-vis de la gestion des crues ; un élément majeur d'évaluation de la performance de la prévision est le critère de **détectabilité** (6.2).

Enfin, l'usage des probabilités améliore la performance du système de prévision face aux enjeux opérationnels. On montre par exemple que le quantile 90 de la prévision probabiliste est bien plus adéquat pour gérer le risque de crue que la prévision déterministe (6.2).

### 5.3.4 Conclusions

La systémique s'oppose à l'approche analytique sur des hypothèses fondamentales, qui permettent de retrouver la dualité Panglossiens-Mélioristes (Tableau 14).

Ces approches duales sont pourtant complémentaires. La systémique nous permettra de mieux explorer le système global de gestion de crue, et de rendre intelligibles certains mécanismes complexes qui contribuent à la Sûreté Hydraulique (0). Une telle approche rend cependant très délicate toute tentative de quantification.

L'approche analytique permet quant à elle d'estimer la valeur de l'information vis-à-vis de l'enjeu d'optimisation de la production (Chapitre 3). Cette approche analytique est également nécessaire pour modéliser les incertitudes de prévision par des probabilités subjectives (adoption des axiomes A1 à A4, 5.2). Produire des prévisions sous forme probabiliste (Chapitre 4) nous permet alors :

- de mieux mettre en lumière la perception des incertitudes lors de l'élaboration de la prévision (Chapitres 7, 8 et 9)
- d'apporter de nouveaux outils qui permettent de mesurer certains facteurs-clefs pour la Sûreté Hydraulique, comme la détectabilité des pluies intenses (6.2).

| **Economie et psychologie cognitives**<br>**Mélioristes** | **Systémique**<br>**Panglossiens** |
|---|---|
| Von Neuman et Morgenstern (1943), Savage (1954), Walliser (2000), Kahneman, Slovic, Tversky (1982), Palmarini (1995), etc. | Le Moigne (1990), Morin (1977, 2000), Piaget (1967), Lagadec (2000), Brunswik (1956), Gigerenzer et al. (1991), Juslin (1994), etc. |
| H1 : Les croyances d'un acteur sont de même nature que celles de l'économiste.<br><br>*Ces croyances sont jugées sur leur **validité empirique** (critère épistémologique).* | H1 : Les croyances d'un sujet sont essentiellement tacites, et font appel à son intuition et à son expérience ; celles du modélisateur sont formalisées, conceptualisées, axiomatisées pour former un espace déductivement clos.<br><br>*Les croyances du sujet sont jugées sur leur **efficacité opérationnelle** (critère praxéologique).* |
| $H1_{bis}$ : Le modélisateur est doté d'un modèle quasi-parfait du système, alors que l'acteur n'en possède qu'un modèle partiel et bruité, défini par référence au précédent. | $H1_{bis}$ : L' expérience et les capacités d'adaptation du sujet lui ont permis d'adopter un comportement adapté aux situations qu'il rencontre ; le modélisateur ne peut élaborer qu'un modèle partiel et bruité de ce comportement. |
| H2 : Les croyances de l'acteur se nourrissent essentiellement d'informations externes, et non de savoirs tacites et inexprimés, ancrés dans son esprit. | H2 : Les croyances de l'acteur se nourrissent essentiellement de savoirs tacites, ou bien d'informations externes dont le caractère informel, incertain et évolutif laisse une large part à l'interprétation. |
| $H2_{bis}$ : A partir des informations qui lui sont données, il y a révision et non reconstruction des croyances de l'acteur (principe d'économie). | $H2_{bis}$ :Il y a reconstruction des croyances pour surmonter des dissonances cognitives (Festinger, 1957, note n°88).<br>Ainsi des ruptures créatrices (Lagadec 2000) en période de crise chez un individu ou dans une organisation. |
| *Théories normatives* | *Théories explicatives et descriptives* |

**Tableau 14 : Oppositions entre économie et psychologie cognitives (école mélioriste) et systémique (école panglossienne). Comparatif des postulats qui fondent chacune de ces disciplines.**

## 5.4 Bibliographie

Bachelard, G., 1934, Le nouvel esprit scientifique, Presses universitaires de France (6e Ed.)

Brunswik, E., 1956. Perception and the representative design of psychological experiments (2nd Ed.), Berkley: University of California Press.

Gigerenzer, G., Hoffrage, U., Kleinbölting, H., 1991. Probabilistic mental models: a Brunswikian theory of confidence, Psychological Review, Vol. 98, 4, 506-528.

Festinger, L.A., 1957. A Theory of Cognitive Dissonance, Stanford University Press.

Hilton, D., 1995. The social context of reasoning: conversational inference and rational judgment. Psychological Bulletin, vol. 118, 2, pp 248-271.

Juslin, P., 1994. The Overconfidence Phenomenon as a Consequence of Informal Experimenter.Guided Selection of Almanac Items, Organizational Behavior and Human Decision Processes, Vol. 57, 2.

Kahneman, D., Slovic, P., Tversky, A., 1982. Judgment under uncertainty: Heuristics and biases. Cambridge University Press.

Krzysztofowicz, R., Sigrest, A., 1999. Calibration of probabilistic quantitative precipitation forecasts. Weather and Forecasting, 14(3), 427-442.

Lagadec, P., 2000, Ruptures créatrices, Editions d'Organisation.

Le Moigne, 1990. La modélisation des systèmes complexes, Dunod, Paris.

Lyon, D., Slovic, P., 1976. Dominance of accuracy information and neglect of base rates in probability estimation, Acta Psychologica, 40, 287-298.

Morin, E., 1977, La Méthode, tome 1, La nature de la nature, Editions du Seuil.

Morin, E., 1980, La Méthode, tome 2, La vie de la vie, Editions du Seuil.

Morin, E., 1987, La Méthode, tome 3, La connaissance de la connaissance, Editions du Seuil.

Morin, E., 1991, La Méthode, tome 4, Les idées, leur habitat, leur vie, leur mœurs, leur organisation, Editions du Seuil.

Morin, E., 2000, La Méthode, tome 5, L'identité humaine, Editions du Seuil.

Palmarini, M., P., 1995, La réforme du jugement, ou comment ne plus se tromper, Odile Jacob (trad. De L'Illusione di sapere, Arnoldo Editore, 1993).

Piaget, J., 1967. Biologie et connaissance. Gallimard, Paris.

Savage, L. J., 1954, The Foundations of statistics. Wiley.

Stanovich, K. E., 1999. Who is rational? Studies of individual differences in reasoning. Mahweh, NJ: Erlbaum.

Tversky, A., Kahneman, D., 1982. Evidential impact of base rates. In Kahneman, D., Slovic, P., Tversky, A., Judgment under uncertainty: Heuristics and biases. Cambridge University Press, 153-160.

Voltaire, 1759 : Candide, ou l'optimisme. Traduit de l'allemand de M. le docteur Ralph , avec les Additions qu'on a trouvées dans la poche du Docteur, lorsqu'il mourut à Minden, l'An de Grâce 1759. Réédition 1975, Hatier, Paris.

Von Neumann, J., Morgenstern, O., 1943. Theory of Games and Economic Behavior, Princeton University Press.

Walliser, 2000. L'Economie Cognitive, Odile Jacob.

Wybo, J.-L., 2001, Risky business, Tomorrow, June.

# 6 Les probabilités : une langue qui révèle et améliore la performance du système de prévision face aux enjeux de Sûreté

*L'information probabiliste permet de mieux se comprendre dès lors qu'on parle de risque. Mais comment en mesurer la valeur ajoutée par rapport à une information de type déterministe (simple chiffre fourni) ou catégorique (il y a risque oui ou non) ?*

*La systémique permet de décrire, sans le quantifier, l'apport des prévisions opérationnelles à la Sûreté Hydraulique.*

*L'analyse du cas d'école d'un usager utilisant une prévision probabiliste permet quant à elle de mesurer la valeur ajoutée de cette dernière face aux événements extrêmes.*

## 6.1 Sait-on mesurer la performance du système de prévision en matière de Sûreté Hydraulique?

### 6.1.1 La vulnérabilité des installations hydroélectriques au risque de crue

Pour l'exploitant, "rater une crue", c'est ne pas réussir à atteindre l'un des "deux objectifs vitaux", selon le texte de doctrine X-EX-03-05[44] :

- empêcher la submersion de l'ouvrage afin de maintenir son intégrité ;
- assurer la transparence de l'ouvrage, ou ne pas aggraver les conséquences de la crue, c'est à dire ne pas provoquer de sur-débits à l'aval de l'ouvrage, et limiter les sur-gradients à 50% de plus que les gradients des débits entrants (et pour les barrages mobiles en rivière, empêcher la submersion des berges en amont immédiat).

Qu'en coûterait-il à EDF si de tels risques de sur-débits se concrétisaient ? La doctrine X-EX-03-05 décline les conséquences à l'externe[45] d'une crue mal gérée :

- les conséquences économiques (responsabilité civile) : réparation des dommages ;
- les conséquences institutionnelles (responsabilité administrative) : annulation du titre par l'autorité administrative, remise en cause de la mission de gestion de la ressource hydraulique nationale, voire remise en cause de la pérennité de l'Etablissement ;
- les conséquences judiciaires (responsabilité pénale) : pour les personnes morales aussi bien que physiques.

Et puisque "les aménagements électriques, répartis sur une grande partie du territoire, sont un vecteur important de l'image d'EDF vis-à-vis de la collectivité" (X-EX-02-00), on peut ajouter à ces conséquences la perte de confiance du public, des collectivités locales, voire éventuellement des Etats qui ont accordé une concession à EDF ou à l'une de ses filiales.

Jean-Yves Delacoux, responsable d'EDF délégué à la production hydraulique, le rappelait lors de la 3e Convention Nationale pour la Sûreté Hydraulique, le 15 mai 2003 : « *La survenue d'un accident grave compromettrait fortement l'avenir de l'exploitation hydraulique et les avantages attachés à ce mode de production, que sont sa souplesse et sa rentabilité.* »

D'où "l'importance d'assurer une gestion de crue irréprochable" (X-EX-03-05)… Mais quels sont les moyens à mettre en œuvre dans la gestion des crues pour pouvoir la qualifier d'irréprochable ?

Pour répondre à cette question, il faudrait savoir mesurer formellement le niveau de Sûreté Hydraulique en gestion de crue. A ce jour, la modélisation systémique constitue le cadre le plus approprié pour intégrer tous les facteurs techniques, humains et organisationnels qui interfèrent dans le système complexe de gestion de crue (5.1).

### 6.1.2 Mesure de la Sûreté : l'approche systémique

Comme nous l'avons évoqué en 5.3, la mesure de la performance du système de prévision amène à deux postures épistémologiques antagonistes. Pour modéliser la contribution de la prévision à la Sûreté Hydraulique, la première attitude, panglossienne (Stanovich, 1999) et

---

[44] Les éléments référencés ainsi sont des textes de doctrine qui régissent l'exploitation des ouvrages de production à EDF.

[45] Les conséquences internes EDF d'une crue mal gérée ont été décrites en 6.1.5 : engravement des retenues, des prises d'eau, voire des conduites d'adduction, usure des machines, etc.

systémique (Piaget, 1967, Le Moigne, 1990), s'impose si l'on admet l'idée suivante : **la situation de danger est une situation complexe.**

La définition des trois mots clefs que sont « situation », « danger », et « complexe » permet de mieux comprendre l'attitude épistémologique du systémicien (définitions toutes trois inspirées du dictionnaire Petit Robert) :

- **situation** : ensemble des relations concrètes qui, à un moment donné, unissent un sujet ou un groupe au milieu et aux circonstances dans lesquels il doit vivre et agir.

  Cette définition sous-tend les notions d'interactions croisées et de récursivité présentées au Chapitre 5 (5.1 et note n° 36). Elle suppose également qu'on définisse les acteurs et groupes d'acteurs concernés par la situation de danger, de la prévision hydrométéorologique aux commandes des organes d'évacuation sur un ouvrage hydroélectrique ; de l'ingénierie et la conception à la maintenance ; des décideurs stratégiques aux pilotes opérationnels.

- **danger** : ce qui menace ou compromet la sûreté des biens et des personnes.

- **complexe** : on utilisera ce mot sous la même acception que les chercheurs en psychologie : ensemble perçu globalement, sans analyse de ses parties composantes.

  Une autre définition pourrait compléter et légitimer cette première : une situation complexe est une situation qui dépasse les capacités d'analyse d'un individu ou d'un groupe d'individus.

  Cette dernière définition s'adapte bien à la situation de danger.

« **La systémique est la discipline dont le projet est l'élaboration et le développement des méthodes de modélisation des phénomènes perçus ou conçus complexes comme et par un système général.** » Piaget (1968). Les concepts clefs de cette approche appliquée au champ des sciences du danger (JL Wybo, comm. pers.) sont observations, perturbations, et comportements :

- **observations** : de nombreuses observations d'un système complexe, au travers d'expériences vécues puis formulées par les hommes qui le composent, permettent de rendre ce système plus intelligible ;
- **perturbations** : plutôt que de découper un système pour en faire une analyse détaillée et méthodique qui en permettra la compréhension, la systémique nous amène à adopter une attitude de recherche résolument différente. On observe comment le système réagit dans toute sa complexité à des perturbations qu'il subit, ou que l'expérimentateur lui fait subir ;
- **comportements** : on déduit des observations du système ayant subi diverses perturbations son comportement sous telle ou telle contrainte.

Pour les événements extrêmes[46], le choix de l'approche systémique est légitimé par la complexité de tous les éléments qui interagissent pour donner lieu à une catastrophe.

[46] La notion d'événement extrême n'est utilisée ici que par opposition à événement courant. On désigne par là un événement hydrométéorologique remarquable (d'une période de retour de 10 à 50 ans), mais qui, associé à des conditions d'exploitation dégradées peut provoquer des accidents.

NB : Pour de tels événements, il est en effet difficile de raisonner en fréquence (tout au plus peut-on évaluer une probabilité subjective d'occurrence[47]), d'en mesurer les effets a priori[48], d'en comprendre et d'en prévoir les causes.

Une telle difficulté trouve sa source dans :

- **l'incertitude** sur l'évaluation des facteurs connus dans les mécanismes conduisant à des défaillances du système,
- **l'incomplétude** : existence de facteurs inconnus influant dans ces mécanismes,
- **l'ignorance** : mécanismes inconnus que seul un retour d'expérience post-accidentel permettra de cerner en partie,

de toute personne confrontée à une situation de danger soit en situation réelle[49], soit au cours d'un retour d'expérience.

### 6.1.3 Application de l'approche systémique à la gestion d'un ouvrage en crue

On se place dans le cas d'événements rares : les crues importantes voire extrêmes, d'une période de retour de 10 ans ou plus. Lors de tels événements, l'expérience montre que la probabilité :

- d'une défaillance technique (automatisme de contrôle, organes d'évacuation, moyens de communication),
- humaine (erreur de l'opérateur),
- organisationnelle (mauvaise circulation de l'information, manque de clarté sur le rôle et la responsabilité de chaque agent, insuffisance de l'apprentissage individuel ou collectif de la gestion de crue),

est elle-même très faible[50].

Le risque qui nous intéresse est donc associé à des probabilités très faibles, ce qui explique qu'on ne puisse pas le mesurer en terme de fréquence d'ESSH (événement significatif

---

[47] La fourniture de prévisions probabilistes donnant la tendance centrale et les quantiles élevés sur un débit prévu ne fait que quantifier le risque de crue (ce qui est déjà bien). Cependant, celui-ci n'est qu'un élément qui interagit avec d'autres pour conduire ou non à la catastrophe. Des accidents qui ont marqué l'histoire hydroélectrique à EDF, tels celui du Drac (1992), montrent qu'en dehors de tout aléa hydrologique, le système de production hydroélectrique peut avoir des défaillances.

[48] le guide d'analyse des ESSH (événement significatif concernant la sûreté hydraulique) ne propose qu'une mesure a posteriori d'un incident ou d'un accident.

[49] dans ce cas, il faut rajouter à l'incertitude, l'incomplétude et l'ignorance de l'agent, l'urgence de la situation dans laquelle il se trouve, qui diminue ses capacités cognitives et intellectuelles.

[50] Encore qu'il faille considérer la **dépendance** entre différents événements dommageables pour le parc de production hydroélectrique. En cas de crue importante, les éléments hydrométéorologiques sont réunis pour qu'il y ait **concomitance** de plusieurs événements naturels dommageables : pluies intenses pouvant provoquer des glissements de terrain (qui peuvent couper l'accès au barrage) et créer de gros embâcles dans la retenue, foudre pouvant endommager les lignes de communication ou de télégestion des ouvrages ainsi que certains automatismes électroniques, vents violents pouvant créer des amorçages sur les lignes de transport d'électricité et faire décrocher les groupes hydroélectriques, vents violents pouvant également créer des chablis et couper certains axes de communication (routes, téléphone, etc.).
En cas de crue importante, les éléments sont également réunis pour que chaque agent, ou l'organisation dans son ensemble commette des erreurs : urgence de la situation, stress, contraintes hiérarchiques ou réglementaires, manque d'information sur la situation en cours, de moyens de contrôle, etc.

concernant la Sûreté Hydraulique)[51]. Comme on ne peut pas construire de mesures se fondant sur les réalisations passées (soit que celles-ci n'ont jamais eu lieu, soit qu'elles ont eu lieu en des périodes et des endroits trop différents pour ne pas être inter-comparables), les mesures de la fiabilité des systèmes complexes sont fondées sur l'idée de « distance qui nous sépare de la catastrophe ».

Pour quantifier la sûreté d'un système, trois types de mesure sont proposées ci-dessous. Les deux premières sont surtout techniques, même si elles peuvent intégrer dans certaines limites des facteurs humains (probabilité d'erreur humaine). La troisième intègre non seulement les facteurs humains, mais aussi les facteurs organisationnels.

La Sûreté Hydraulique, notion bien plus récente que la Sûreté Nucléaire, a peu été modélisée à ce jour. C'est pourquoi on prendra soin d'illustrer, pour chacune des trois méthodes de mesure de fiabilité des systèmes complexes, leur application possible à la Sûreté Hydraulique.

### 6.1.4 Mesure de la Sûreté par le « potentiel de danger »

Ce type de mesure, et la gestion du risque qui en découle, s'est développé dans le début des années 1980 en France, dans le domaine de la Sûreté Nucléaire. L'émergence de la notion de défense en profondeur et la mise en place des premiers ESS (événements significatifs pour la Sûreté Nucléaire) a conduit à l'identification :

1. des événements redoutés (types de vulnérabilité du système) ;
2. des parades associées, appelées aussi « boucles de rattrapages » (redondances, automatismes de contrôle, mécanismes d'alerte de l'opérateur, etc.) ;
3. des trajectoires qui mèneraient à ces événements si toutes les parades étaient successivement mises en défaut.

Le niveau de sûreté se mesure alors en temps réel par le nombre de boucles de rattrapage qui séparent la situation courante de telle ou telle catastrophe.

Application possible à la mesure de la valeur de la prévision de crue pour la Sûreté Hydraulique :

La prévision peut constituer une barrière de plus dans la défense en profondeur, tout autant qu'une vulnérabilité du système, si celui-ci est sensible à une information fausse ou mal interprétée.

L'approche par le potentiel de danger fait intervenir trop d'inconnues (circonstances d'exploitation, efficacité du vecteur de communication, fiabilité de l'interprétation de l'usager, etc.) pour que l'on puisse quantifier le gain (ou la perte) de Sûreté apporté par la prévision.

### 6.1.5 Mesure de la Sûreté par la « manœuvrabilité »

Cette approche se fonde sur l'évaluation de l'ensemble des paramètres de contrôle possibles. Pour chacun d'eux, on examine :

[51] Les ESSH permettent cependant un retour et un partage d'expérience, qui seront les moteurs d'un apprentissage individuel et collectif à la gestion des risques et des crises d'une plus grande ampleur.

1. l'amplitude de manœuvre actuelle ;
2. la position du point de fonctionnement par rapport aux limites de manœuvre ;
3. le coût du franchissement d'une des limites.

Cette méthode a l'avantage de pouvoir être calculée automatiquement et de focaliser l'attention des opérateurs sur les commandes sensibles.

Application à la gestion d'un ouvrage hydroélectrique en période de crue

On perçoit les applications immédiates de cette approche dans le domaine de l'exploitation d'un ouvrage en période de crue. La question de la marge de manœuvre de l'exploitant se pose pour :

- assurer le bon fonctionnement des installations hydroélectriques dans des conditions de sécurité satisfaisantes pour les tiers ;
- assurer la sûreté de l'ouvrage, en maintenant le niveau sous la cote de retenue normale d'exploitation (cote du plus haut niveau admis pour l'exploitation du réservoir dans des conditions normales d'exploitation), sous la cote des plus hautes eaux (cote qui correspond à l'évacuation de la crue de projet) ou sous la cote de sûreté (cote qui correspond au niveau d'eau au delà duquel la sécurité de l'ouvrage n'est plus garantie) ;
- turbiner l'équivalent du débit entrant de manière à maintenir le niveau d'eau dans la retenue à la cote nominale d'exploitation ;
- abaisser le plan d'eau depuis la cote nominale d'exploitation jusqu'à celle qui permet le passage de l'ouvrage en régime torrentiel[52]. ;
- passer en régime torrentiel (effacement du barrage) grâce à l'utilisation des différentes vannes (vannes de fond, et vannes au niveau du barrage), sans aggraver les conséquences de la crue par rapport à ce qui se passerait en l'absence de barrage : en particulier, ne pas provoquer de sur-débits artificiels ;
- fermer les prises d'eau des différentes adductions de la retenue, pour limiter les apports ;
- pomper l'eau vers un autre réservoir dans un état moins critique (niveau de remplissage suffisamment bas, et localisation à l'écart de la zone d'extension de la crue).

Certaines limites sont particulièrement critiques. Ainsi en est-il de l'instant limite au-delà duquel l'abaissement du plan d'eau par turbinage n'est plus possible *(passé cet instant, les débits entrants sont plus élevés que les capacités de turbinage)*. Dans ces conditions :

- Soit on provoque des sur-débits importants à l'aval en passant trop brutalement l'ouvrage en régime torrentiel (contraire au cahier des charges et aux obligations réglementaires).

  Le coût du franchissement de cette limite est alors l'éventuelle mise en cause de la responsabilité civile, pénale et administrative de l'entreprise, ainsi que l'impact en terme

[52] Le régime torrentiel correspond à l'effacement du barrage (vannes grandes ouvertes), c'est à dire que l'écoulement est similaire à ce qui se passerait en l'absence de l'ouvrage.

d'image pour l'entreprise si la sécurité des personnes ou des biens à l'aval de l'ouvrage est menacée.

- Soit on n'ouvre que partiellement les vannes pour maintenir la cote constante tout en maîtrisant les variations de débit à l'aval.

Dans ce cas, le coût de franchissement de la limite résulte de l'engravement de la retenue : diminution de la capacité utile de la retenue, et éventuellement coût du désengravement de la retenue par une entreprise prestataire.

De même, pour un barrage de haute montagne, l'instant limite qui permet d'envoyer une équipe fermer les prises d'eau *(passé cet instant, l'orage gronde déjà, et le chemin d'accès qui permet d'ordinaire d'accéder en 4x4 aux prises d'eau en deux heures est probablement déjà impraticable, du fait de coulées de boue par exemple)* :

- Soit les capacités de turbinage permettent de turbiner les entrants dans l'usine sur le cours d'eau principal en contrebas. Dans ce cas là, le coût est faible.
- Soit le turbinage à plein régime est insuffisant, voire impossible : par exemple lorsque le cours d'eau en aval est en crue, le canal de fuite de l'usine en aval peut être en charge. L'orage a également pu faire tomber les lignes, et dans ce cas les groupes électriques « décrochent » : faute de pouvoir évacuer l'énergie produite, les turbines s'arrêtent en urgence.

Dans ce cas, les débits qui ne peuvent plus être turbinés sont déversés. Le coût associé peut être très élevé, surtout si les conditions suivantes sont réunies :

- La nature, qui n'a pas vu une goutte d'eau dans le lit du torrent depuis la construction du barrage en amont, a envahi le lit d'arbres et de végétation
- L'homme, qui n'est plus habitué à voir couler de l'eau, a sous-dimensionné les buses d'écoulement sous les routes qui croisent le lit à sec, et a construit aux environs des habitations et infrastructures (de loisir notamment, liées à l'attrait touristique de ces régions de montagne).

Le coût résulte alors des dégâts provoqués sur ces habitations et infrastructures par les flots déversés, qui débordent du lit et charrient divers corps solides.

D'où, un fois encore, un coût en terme de responsabilité civile, pénale et administrative, ainsi qu'en terme de perte d'image de l'entreprise, particulièrement auprès de trois groupes d'acteurs :

- l'autorité administrative de tutelle, les collectivités locales, et les autres acteurs locaux de l'eau, à l'heure où EDF négocie de nombreux renouvellements de titre d'autorisation ou de concession d'ouvrages hydroélectriques ;
- d'éventuels futurs actionnaires, à l'heure où l'entreprise se prépare à ouvrir son capital ;
- les clients français et européens, qui, avec l'ouverture progressive du marché de l'électricité ont le libre choix de leur fournisseur d'électricité.

Dans tous les cas, il faut ajouter le coût de désengravement des prises d'eau, voire également des conduites d'adduction.

Application possible à la mesure de la valeur de la prévision de crue pour la Sûreté Hydraulique :

La valeur de la prévision se traduit le plus souvent par l'augmentation des marges de manœuvre relatives à la sécurité:

- les délais d'anticipation sur l'événement dont dispose l'exploitant sont améliorés : alors que celui-ci n'avait connaissance que des débits entrants dans la retenue, les prévisionnistes l'informent des débits mesurés en amont, des pluies mesurées au réseau sol (ou visualisées grâce aux images Radar et satellitaires), et des prévisions météorologiques ;
- l'incertitude sur la situation vue par l'exploitant diminue : le prévisionniste lui fournit une vision globale qui l'aide à prendre de la hauteur sur les orages locaux, et de mieux d'estimer le risque de crue et les évolutions de celle-ci en temps réel.

Dans ce cas encore, il est difficile de chiffrer l'amélioration de la Sécurité liée à l'augmentation des marges de manœuvre.

### 6.1.6 Mesure de la Sûreté à l'aide de la notion de résilience

En théorie des organisations, la résilience est définie comme l'aptitude d'un système complexe à rester sûr lorsqu'il est soumis à des perturbations (Wybo, 2001).

Ici, le système complexe désigne l'organisation composée de l'ensemble des acteurs impliqués dans la gestion de crue, depuis l'ingénierie jusqu'au pilotage opérationnel des installations et les moyens matériels dont elle s'est dotée pour cette mission (organes de contrôle, réseau d'observation, etc.).

La résilience est donc l'aptitude d'un système à résister (en subissant d'éventuelles déformations) lorsqu'il est soumis à des contraintes de natures et intensités diverses.

On modélise le système avec des composants caractérisés par leur nature et leurs propriétés (élastique ou plastique[53], avec divers coefficients d'élasticité, de plasticité, et de résistance à la rupture) et possédant entre eux des liaisons (lâches, souples, rigides, etc.). Ces liaisons reportent de manière homogène ou hétérogène les contraintes globales au niveau de chaque composant du système complexe. Chaque composant subit ou non des déformations (élastiques ou plastiques). Le produit de ces déformations est la réponse globale du système aux contraintes qu'on lui a imposées.

Ainsi, le jeu des liaisons entre les différents éléments et celui des propriétés de chaque élément déterminent la résilience du système.

Application à la gestion d'un ouvrage hydroélectrique en période de crue

Après avoir défini les acteurs ou réseaux d'acteurs impliqués dans la gestion d'un ouvrage en crue, on modélise leurs interactions à partir d'observations du comportement de l'organisation lors d'incidents ou d'accidents. On appréhende ainsi le comportement de l'organisation

[53] **élasticité** : aptitude d'un objet à se déformer sous l'effet d'une contrainte, et à retrouver sa forme initiale une fois que la contrainte a cessé.
**plasticité** : aptitude d'un objet à se déformer sous l'effet d'une contrainte, et à conserver sa forme modifiée une fois que la contrainte a cessé.

soumise à de fortes contraintes, comme elle le sera en période de crue (urgence et gravité de la situation, contraintes réglementaires, etc.).

Dans l'entreprise EDF, les liaisons sont très rigides entre les agents qui gèrent une crue : le rôle et la responsabilité de chacun dans la gestion de l'ouvrage en période de crue sont précisément édictés :

- les rapports entre les prévisionnistes de DTG[54], les agents en charge de la programmation de la production au COOP, les agents en charge de l'exploitation et de la maintenance des installations hydroélectriques au GEH, et enfin, ceux qui gèrent la téléconduite des installations en temps réel au PHV sont codifiés de manière contractuelle ;
- au sein des entités, le rôle de chacun est décrit ; par exemple :
    - ◊ à DTG : prévisionniste d'astreinte, prévisionniste de soutien, équipe réseau (pluviomètres et stations hydrométriques), etc.
    - ◊ chez l'exploitant : directeur du GEH, ingénieur exploitation, chef de quart, etc.
- la conduite à tenir est explicitée dans les éléments de doctrine, les consignes de crue, les instructions permanentes.

Les pressions exercées sont les **pressions réglementaires** (avec mise en cause possible de la responsabilité civile, pénale et administrative de l'exploitant), les **pressions hiérarchiques**, et **l'urgence de la situation**. Etant donné la rigidité de la structure, ces contraintes peuvent se retrouver au niveau des agents. Ces derniers les absorbent de diverses manières :

- certains utilisent placidement leur expérience pour réagir au mieux dans la situation dans laquelle ils se trouvent ;
- d'autres agissent mal sous l'effet du stress et de l'urgence ;
- d'autres appliquent méthodiquement et mécaniquement les consignes de crue ;
- d'autres se reposent sur les autres, etc…

Admettons qu'il existe un défaut structurel dans ce dispositif de gestion de la crue (vecteur de communication peu efficace, consignes de crue mal établies, ou lacunaires, etc.). Il est fort probable que l'organisation y ait remédié spontanément par une boucle de rattrapage. La Sûreté Hydraulique repose alors sur un agent particulier qui connaît la personne à contacter pour avoir une information complémentaire, ou a l'expérience d'une situation similaire qui n'a été prévue par aucune consigne de crue, et connaît la chose à faire en pareille situation.

Pour de multiples crues mineures, cette boucle de rattrapage peut rester invisible au yeux de tous, mais en cas de « gros coup », c'est au niveau de cet élément que tout risque de s'effondrer. Toute la rigidité alentours aura fait que l'ensemble des contraintes subies par la structure est reprise par l'élément le plus malléable (celui qui a toujours « encaissé » et « pris sur lui » pour rattraper la situation) ; c'est celui qui rompra le premier. De plus, ce sera à lui qu'incombera la faute, au yeux de la hiérarchie et peut-être aussi de l'administration.

[54] On rappelle la signification des sigles : DTG (Division technique générale), COOP (Centre opérationnel d'optimisation de la production), GEH (Groupement d'exploitation hydraulique), PHV (Poste hydraulique de vallée). Voir définitions associées dans le glossaire, en page 9.

Les limites de la résilience sont, dans ce cas, la trop grande plasticité d'un composant du système combinée avec la rigidité du reste du dispositif.

Généralisation :

On peut imaginer d'autres cas où l'ensemble des composants sont trop « mous », et pour lesquels, de dérive en laisser-aller, on aboutit à une catastrophe (pas à EDF !).

On peut aussi envisager une structure dans laquelle les liaisons sont tout aussi rigides que dans le cas énoncé ci-dessus, mais où chacun des agents demeure « inflexible » : strict respect des missions qui lui sont confiées, sans aucune souplesse ni esprit d'initiative. Dans ce cas, passé un seuil limite de contrainte acceptable, le système casse comme du verre.

De nombreux exemples montrent ainsi que la résilience est le produit d'un subtil agencement d'éléments plus ou moins plastiques ou élastiques, reliés entre eux par des liaisons plus ou moins rigides.

Application possible à la mesure de la valeur de la prévision de crue pour la Sûreté Hydraulique :

La prévision améliore la résilience grâce à un partage d'information entre prévisionniste et usager :

- remontée d'informations de terrain, sur les conditions hydrométéorologiques locales et sur les contraintes d'exploitation ;
- fourniture d'informations sur la situation météorologique à une échelle plus globale, sur les évolutions attendues, et d'éventuels risques à venir (crue, tempête).

La notion de résilience permet d'illustrer la valorisation de l'information de prévision, sans proposer de quantification.

### 6.1.7 Conclusion

Partant d'une approche systémique, trois types de mesures permettent de mieux cerner en quoi le système de prévision des crues contribue à la Sûreté Hydraulique. Chaque mesure aide à tracer dans le temps les évolutions d'un système et à pointer d'éventuels dysfonctionnements techniques ou organisationnels pouvant nuire à la Sûreté Hydraulique en gestion de crue.

Cependant, faute d'être quantitatives, de telles mesures ne permettent pas de réaliser des arbitrages entre diverses actions d'amélioration de la Sûreté Hydraulique (5.1).

## 6.2 Amélioration du système de prévision face aux événements extrêmes

### 6.2.1 En quoi et comment la prévision des crues contribue-t-elle à la Sûreté Hydraulique ?

C'est la question que nous avons posée aux décideurs d'EDF dans le domaine de la Sûreté Hydraulique lors d'une réunion qui s'est tenue le 6 juin 2003 à Grenoble dans les locaux de DTG avec Michel Poupart, Chargé de Mission Sûreté Hydraulique, et Jacques Lemarchand, Inspecteur Sûreté Hydraulique d'EDF.

La question posée est délicate dans la mesure où très peu de consignes de crue font directement référence aux prévisions de DTG, les manœuvres d'exploitation des ouvrages hydrauliques en crue ne reposant que sur les observations.

Nous avons montré l'intérêt, pour améliorer la Sûreté Hydraulique (SH), de disposer d'une prévision hydrométéorologique fiable. L'anticipation sur l'événement attendu aide à s'y préparer dans de bonnes conditions, et à gérer plus sereinement la crue : mise en place d'équipes de quart, vérification du bon fonctionnement des organes d'évacuation (vannes de fond ou de surface) et des outils de communication et d'observation (radio, téléphone, fax, transmission des données d'exploitation et des observations du réseau sol, etc.).

Nous avons confronté les décideurs d'EDF à la question centrale de notre recherche : Comment piloter par la performance les moyens de prévision hydrométéorologique mis en œuvre pour la SH ?

Concrètement, nos questions étaient les suivantes :

- Quel est le consentement à payer d'EDF pour améliorer la Sûreté Hydraulique de ses ouvrages en période de crue ? *Des exemples d'amélioration étaient décrits à l'aide des notions de défense en profondeur, de manœuvrabilité, ou de résilience.*
- Les moyens dont EDF s'est dotée pour assurer une gestion sûre de ses ouvrages en crue sont-ils à la mesure des enjeux évalués en 6.1.1 ?
- Si ces questions paraissent trop abruptes, quelles sont les dépenses qu'on entreprend sur d'autres axes contribuant à la Sûreté Hydraulique, soit pour mesurer le risque (remise à jour des études de crue extrême, essais et mesures des variations de débit lors de lâchers d'eau, etc.), soit pour le réduire (redimensionnement des déversoirs de crue, emploi de saisonniers pour sensibiliser les estivants à proximité des ouvrages, etc.) ?

Les réflexions engagées ont soulevé de nouvelles questions :

- Les moyens actuels sont-ils adaptés aux enjeux opérationnels ?
- Permettent-ils d'aboutir, in fine, à un risque acceptable pour l'entreprise EDF ?
- Les moyens mis en œuvre pour la Sûreté Hydraulique sont-ils pilotés par leur performance ?
- Sinon, comment se fait l'arbitrage entre les différents moyens qui contribuent à la Sûreté Hydraulique des ouvrages en période de crue : personnel de quart, nombre de prévisionnistes, redondances sur l'ouvrage, outils météo (Radar par exemple), niveau de maintenance préventive, etc. ?

En cas d'incapacité à répondre à toutes ces questions, doit-on se fixer pour objectif de tout faire pour être « exemplaire » (en référence à l'état de l'art), sans se préoccuper des coûts associés ?

### 6.2.2 Les réponses des décideurs, et les conséquences pour notre étude

A ce jour à EDF, les décideurs ne sont pas en mesure de répondre à de telles questions.

En conséquence, les Mélioristes (Stanovich, 1999) concluraient que les arbitrages se font "au petit bonheur", entre les divers moyens d'assurer la Sûreté Hydraulique : emploi de personnel saisonnier pour sensibiliser les estivants aux dangers liés au fonctionnement des centrales, maintenance préventive sur les organes de sécurité des barrages, maintenance du réseau de mesure au sol (pluviomètres, stations hydrométriques, etc.), fiabilisation des automates de contrôle, etc.

Les Panglossiens (Stanovich, 1999) fidèles à eux-mêmes, concluraient que l'organisation s'est parfaitement adaptée à ses objectifs de Sûreté Hydraulique. Une telle adaptation se serait faite en dehors de toute démarche codifiée, mais plutôt par la créativité et l'initiative de chacun, les partages d'expérience à différents niveaux de l'organisation, la capitalisation des analyses d'incidents d'exploitation, les divers retours d'expérience, et les groupes de travail qui rassemblent périodiquement les exploitants de l'Hydraulique. En ce sens, les moyens s'adaptent "d'eux-mêmes" aux objectifs : dès qu'un manque se fait sentir, l'organisation s'adapte pour y remédier.

Gardons-nous de tomber dans l'un ou l'autre de ces excès, et tenons-nous en à la question spécifique de la valorisation de la prévision des crues.

A l'aide des notions de défense en profondeur (6.1.4), de manœuvrabilité (6.1.5), et de résilience (6.1.6), nous avons constaté que la prévision de crue contribue effectivement à la Sûreté Hydraulique. Il nous a cependant été impossible de quantifier la part de cette contribution. Nous savons simplement que la prévision de crue est valorisée par l'exploitant, dans la mesure où elle lui permet une meilleure anticipation de l'événement.

Faute de mieux, nous allons étudier la performance du système de prévision, dans sa capacité à détecter les crues avec suffisamment d'anticipation sur leur occurrence. En effet, ce qui intéresse l'usager, c'est le franchissement d'un seuil de débit pour lequel il doit engager des manœuvres d'exploitation particulière, en vue d'assurer la Sûreté Hydraulique sur les installations dont il a la charge.

### 6.2.3 La performance de la prévision des épisodes de pluie intense

Nous avons choisi avec les prévisionnistes quatre bassins pour lesquels ils sont régulièrement contactés par l'usager, particulièrement sensible au risque de crue : ce sont les bassins de l'Arn, du Tarn, de la Creuse, et de la Garonne.

Nous étudierons la performance des prévisions de pluie (plutôt que de débit), pour des raisons pratiques : durant cette période, nous avons eu bien plus d'épisodes de pluie intenses (lame d'eau spatiale supérieure à 20mm sur le bassin) que de crues (car les stocks hydriques des bassins étaient rarement saturés).

Les prévisionnistes ont produit chaque jour durant 6 mois des prévisions probabilistes sur les lames d'eau attendues à J et à J+1, en fournissant deux chiffres :

- Le quantile 50 : nous l'avons appelé la "tendance centrale" de la prévision. C'est le chiffre qu'ils associent naturellement à la prévision déterministe qu'ils ont l'habitude de produire (cf. note n° 34 p 70).
- Le quantile 90 : c'est la "prévision du risque". Il s'agit pour eux d'un exercice nouveau : fixer une hauteur de précipitation qui n'a qu'une chance sur dix d'être dépassée.

On suppose l'usager particulièrement sensible aux événements manqués, c'est à dire aux crues qui n'avaient pas été prévues dans un temps suffisant pour lui permettre d'agir. Il est en somme prêt à assumer un certain taux de fausses alertes (par exemple 80%) pour obtenir un taux d'événements manqués aussi bas que possible. L'arbitrage entre taux de fausses alertes et taux d'événements manqués s'illustre ainsi : sur un ouvrage hydroélectrique, une fausse alerte engendre des pertes de production (l'abaissement du plan d'eau entraîne une diminution du rendement des groupes, et les $m^3$ déversés sont des kWh en moins), alors qu'un événement manqué mettra en cause la sécurité des personnes et des biens à l'aval de l'ouvrage.

Imaginer que l'usager fonde sa règle de décision sur le quantile 90, c'est supposer que le coût d'un événement manqué est dix fois supérieur à celui d'une fausse alerte. Dès que la probabilité excède 0,1, l'usager préfère agir comme si la crue devait arriver.

On suppose que l'usager est sensible au franchissement du seuil 20mm, et qu'il a besoin d'un délai d'anticipation de 24h pour entreprendre des opérations de protection. Son critère a posteriori d'appréciation du système de prévision est le suivant : tant que le taux de fausses alertes n'excède pas 80%, il est d'autant plus satisfait que le taux d'événements manqués est faible[55].

NB : On postule que cet usager est sensible à un seuil unique (que l'on suppose égal à 20mm). Si on lui annonce une prévision de lame d'eau à 30mm (en quantile 50% ou 90%, suivant son critère de décision), et qu'il pleut ensuite 300mm, on considèrera que l'usager est satisfait : il a pu se protéger contre le risque annoncé.

NB : A quel niveau de référence peut-on comparer la performance du système de prévision opérationnel ?

La comparaison par rapport aux prévisions naïves (Chapitre 3) n'a ici aucun sens : la prévision de persistance ne pourra, par construction, détecter de franchissement de seuil qu'après-coup. Quant à la prévision fondée sur l'historique climatologique ou hydrologique, elle ne ferait que rappeler, chaque jour, la valeur de la crue biennale, décennale, centennale, etc.

La comparaison par rapport à la prévision parfaite sera immédiate : celle-ci ne donnera jamais lieu ni à une fausse alerte, ni à un événement manqué.

En conséquence, pour avoir tout de même un étalon par rapport auquel on puisse mesurer la performance des prévisionnistes dans la détection des pluies intenses, on comparera les performances du modèle opérationnel ANALOG (Duband, 1982, Obled et al., 2002) qui délivre chaque jour des prévisions probabilistes de pluie, avec les prévisions probabilistes qui émanent du jugement expert de ces prévisionnistes.

[55] La fonction d'utilité associée est dite dactylographique : une condition sine qua non sur un critère, puis seulement une satisfaction linéaire par rapport à un autre critère. De telles fonctions d'utilité semblent effectivement exister « dans la nature » : passé un seuil de taux de fausses alertes, le système de prévision perd toute crédibilité aux yeux de l'usager… à force d'avoir crié au loup, plus personne ne prend au sérieux les messages d'alerte.

Nous pourrions cependant imaginer qu'une approche plus fine serait de pondérer les deux événements Fausse Alerte et Evénement Manqué suivant leurs désagrément respectifs pour l'usager, et de les encoder dans une seule et même fonction d'utilité.

### 6.2.4 L'aptitude des experts à anticiper les épisodes de pluie intense

Le Tableau 15 résume, pour les 120 prévisions émises à J+1, le nombre de fois où le prévisionniste a annoncé un franchissement de seuil, en "tendance centrale" ou en "prévision du risque" (i.e. avec le quantile 50% ou 90%).

| | **Creuse** | | **Tarn** | | **Arn** | | **Garonne** | |
|---|---|---|---|---|---|---|---|---|
| **Quantiles** | **q50** | **q90** | **q50** | **q90** | **q50** | **q90** | **q50** | **q90** |
| Nombre d'événements prévus | 6 | 30 | 4 | 13 | 4 | 11 | 2 | 18 |
| Nombre d'événements observés | 6 | | 7 | | 7 | | 5 | |
| Nombres d'événements observés correctement prévus | 2 | 5 | 2 | 5 | 2 | 3 | 2 | 3 |

**Tableau 15 : Table de contingence de la prévision experte de pluie intense.**
**Seuil 20mm sur tous les bassins sauf celui de la Creuse (15mm).**

NB : La Creuse ayant été particulièrement pauvre en épisode de pluie intense sur la période considérée, nous avons dû prendre un seuil à 15mm pour avoir des observations de franchissement de seuil.

Ce tableau nous permet déjà d'observer que le nombre d'alertes augmente considérablement suivant qu'on utilise le quantile 90% plutôt que le quantile 50%, mais qu'en contrepartie, surtout sur le Tarn et sur la Creuse, le nombre d'événements bien anticipés augmente aussi considérablement.

Taux de fausses alertes et taux d'événements manqués

| **Précipitations intenses** | **Observées** | **Non observées** |
|---|---|---|
| **Prévues** | H | FA |
| **Non Prévues** | M | CR |

**Tableau 16 : Tables de contingence de la prévision de pluie intense**
**H: nombre d'événements correctement prévus. FA: nombre de fausses alertes.**
**M: nombre d'événements manqués. CR: nombre de rejets corrects.**

Comme dans les autres Chapitres de cette thèse, on définit ici TEM (Taux d'événements manqués) et TFA (Taux de fausses alertes) de la manière suivante[56] :

[56] Une autre définition du taux de fausses alertes est utilisée habituellement dans la littérature : TFA=FA/(FA+CR), en particulier pour tracer le diagramme ROC qui possède certaines propriétés intéressantes (Mason 1982). Cette définition a l'avantage d'être homogène avec celle du TEM : dans les deux cas, on conditionne par rapport aux observations (H+M, et FA+CR). Autre avantage : à elles deux, les définitions de TFA et TEM font intervenir les 4 nombres du tableau de contingence (H, M, FA et CR), alors que notre définition ne fait pas intervenir le nombre de rejets corrects CR.

L'usager, et en particulier Jéo (cf. 2.1.3), a cependant coutume de conditionner aux observations lorsqu'il s'agit de parler d'événements manqués ("Si on fait le compte des 4 crues de l'année dernière, il y en a 2 que vous n'aviez pas prévues à temps"), et de conditionner aux prévisions lorsqu'il s'agit de fausses alertes ("Quand on regarde les 5 fois où vous nous avez mis en alerte, on n'a eu que 2 crues"). Cette manière de juger de la qualité du système de prévision est sous-optimale (un usager rationnel a besoin de connaître le nombre de rejets corrects CR pour construire sa règle de décision optimale). Nous avons néanmoins choisi de la conserver pour placer le prévisionniste au plus près des préoccupations de Jéo, à qui il a l'habitude de fournir des prévisions.

**TEM=M/(H+M), et TFA=FA/(FA+H).**

Exemple de calcul du couple (TFA, TEM)

On s'intéresse par exemple à la prévision de la lame d'eau à J+1 sur le Tarn ; et on considère le quantile 90% :

- il y a eu 7 franchissements de seuil observés ;
- 13 fois le prévisionniste avait placé son quantile 90% à 20mm ou plus ;
- sur ces 13 « alertes », il y a donc eu 5 détections effectives d'un franchissement de seuil, et 8 « fausses alertes » ;
- sur les 7 franchissements de seuil observés, 5 ont été détectés, et 2 sont des « événements manqués ».

On en déduit ainsi le taux de fausses alertes et le taux d'événements manqués :

**TFA** = (13-5)/13 = **62%**

**TEM** = (7-5)/7 = **29%**

On fait les mêmes calculs pour le quantile 50%, et pour les autres paramètres. Les résultats sont reportés en Figure 23, avec en abscisse le TFA, et en ordonnée le TEM.

Positionnement des prévisions de chacun des paramètres dans l'espace (TFA, TEM) :

Le tableau de contingence permet de calculer chaque taux de fausses alertes et chaque taux d'événements manqués selon le bassin et le quantile choisi. On représente les points obtenus dans un espace où le TFA est en abscisses, et le TEM en ordonnée. Les quantiles 50 et 90 de chaque prévision sont reliés par un segment (sur chaque bassin, le quantile 50 présente systématiquement un TEM plus élevé et un TFA plus faible que le quantile 90).

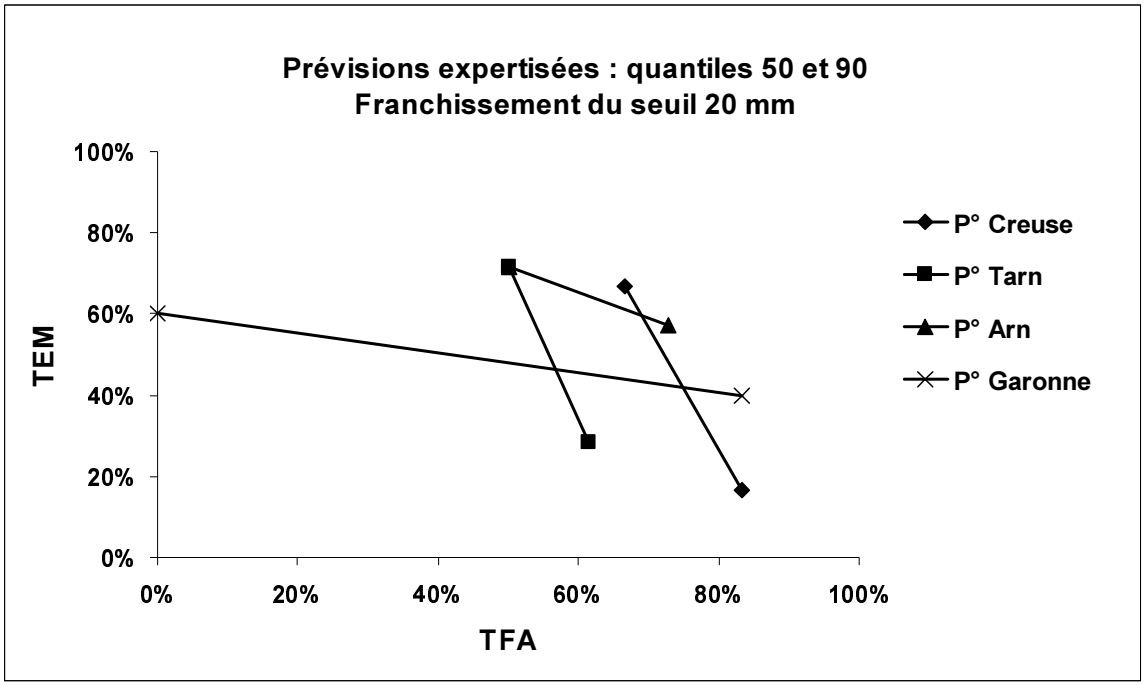



**Figure 23 : Détection par les experts des pluies intenses (>20mm) avec 24h d'anticipation.**
**Règle de décision basée soit sur le quantile 50 soit sur le quantile 90**
**En abscisses, le taux de fausses alertes (TFA), en ordonnée, le taux d'événements manqués (TEM)**

Toutes les prévisions ont un TFA du même ordre de grandeur (60 à 80%[57]), que l'on considère comme acceptable pour l'usager.

Le critère de performance du système de prévision tient donc dans sa seule détectabilité de pluies intenses : plus le TEM est bas, meilleur est le système de prévision.

On peut retenir de la Figure 23 que sur les bassins du Tarn et de la Creuse l'usager a un net avantage à utiliser le quantile 90% de la prévision probabiliste : le TEM décroît considérablement (de 70% pour le quantile 50 à 20% pour le quantile 90) alors que le TFA reste dans des limites raisonnables.

Sur les bassins de l'Arn et de la Garonne en revanche, même en utilisant le quantile 90 des prévisions émises, le TEM reste supérieur à 40%, pour un TFA de 80%.

Remarque :

Intéressons-nous aux prévisions sur le Tarn et sur la Creuse, qui sont toutes deux performantes. Elles traduisent en fait un compromis différent : la prévision sur la Creuse produit un TFA bien plus élevé que celle sur le Tarn, et un TEM bien plus faible. On montre en Annexe 6 que cet effet résulte du comportement du prévisionniste : plus craintif de rater un épisode de pluie intense sur la Creuse que sur le Tarn, il donne des prévisions biaisées positivement pour se couvrir. On dit dans le jargon qu'il « sort le parapluie ». C'est le biais bien connu des météorologistes sur les bassins à enjeu, nommé "overforecasting bias" (Krzysztofowicz et Sigrest, 1999, Murphy et Daan, 1984).

### 6.2.5 Comparaison à la performance des analogues

Comparons à présent la performance de ces prévisions expertisées avec les sorties du modèle ANALOG sur ces quatre bassins (Figure 24).

Avec le quantile 50% (c'est à dire pour la tendance centrale de la prévision), les prévisionnistes détectent mieux que le modèle ANALOG l'occurrence de pluies intenses. Ils disposent en effet d'autres outils (observations en temps réel, imagerie RADAR et satellitaire, modèle ARPEGE ou ALADIN, etc.) qui leur permettent de bien anticiper la situation à risque.

En revanche, un usager qui serait vraiment sensible au risque de précipitation intense (pour qui le coût d'un événement manqué serait dix fois plus lourd que celui d'une fausse alerte) aurait gagné à utiliser les prévisions issues des analogues durant cette période sur le Tarn, l'Arn et la Garonne. Les taux d'événement manqués pour le modèle ANALOG sont bien plus bas que pour les prévisions expertes sur ces trois bassins.

Il n'y a que sur le bassin de la Creuse que la « prévision du risque » (c'est à dire la détection grâce au quantile 90% du franchissement de seuil) est bien meilleure chez les experts que pour le modèle ANALOG.

Pourtant, les prévisionnistes qui ont produit au cours de ces 6 mois d'expérimentation des prévisions probabilistes disposaient tous les jours des sorties du modèle ANALOG, en plus de tous leurs autres outils de prévision et d'observation.

Il semble paradoxal que les prévisionnistes n'aient pas plus utilisé le modèle ANALOG, vu la performance de ce modèle à détecter des épisodes de pluie intense : sur deux des quatre bassins, l'utilisation du quantile 90 du modèle aurait permis de ne rater aucun événement,

[57] à l'exception du quantile 50% de la prévision de la Garonne pour laquelle les deux seules fois où le prévisionniste annonçait un franchissement de seuil, celui-ci s'est réalisé. Le prévisionniste aura néanmoins « raté » sur ce bassin 3 épisodes sur 5 en tendance centrale (c'est à dire en fixant son quantile 50%).

pour un taux de fausses alertes identique à celui produit par l'utilisation du quantile 90 produit par les prévisionnistes (un TFA d'environ 80%).

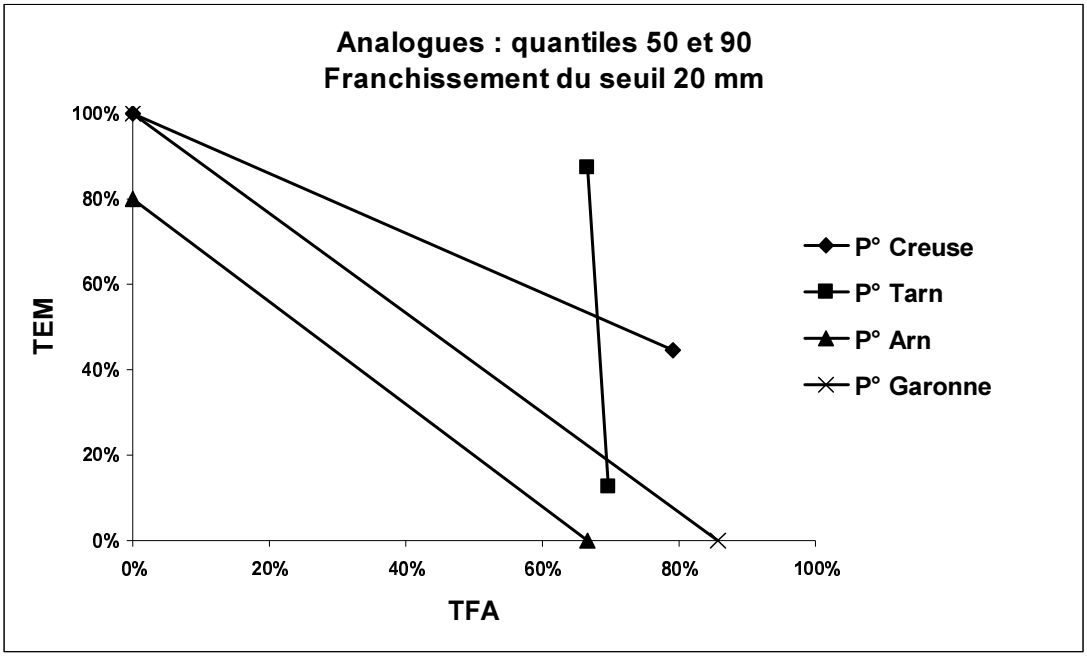



**Figure 24 : Détection par le modèle des analogues des pluies intenses (>20mm) avec 24h d'anticipation. Règle de décision basée soit sur le quantile 50 soit sur le quantile 90 En abscisses, le taux de fausses alertes (TFA), en ordonnée, le taux d'événements manqués (TEM)**

Nous avançons l'explication suivante pour expliquer ce paradoxe. En "prévision centrale", le modèle ANALOG est très peu performant pour détecter des épisodes de pluie intense : plus de 80 à 100% d'événements manqués ! Les prévisionnistes sont donc tentés de se concentrer sur toutes leurs autres sources d'information pour évaluer les risques météorologiques, en délaissant le modèle ANALOG.

### 6.2.6 Conclusion

La prévision probabiliste apporte un plus : elle permet de s'adapter aux enjeux des usagers en leur proposant un système de prévision ayant une meilleure détectabilité (en échange d'un taux de fausses alertes plus élevé). Cette adaptation se fait en toute transparence : lorsque le prévisionniste annonce un quantile 90 à 30 mm, il affiche qu'il y a 9 chances sur 10 pour que cette valeur ne soit pas dépassée.

A l'opposé, lorsque le prévisionniste avait l'habitude de fournir un chiffre déterministe, et qu'il pressentait un risque de pluie intense sur un bassin comportant des aménagements vulnérables (en chantier par exemple), il annonçait volontairement un chiffre plus élevé que ce qu'il prévoyait effectivement en tendance centrale. Ce processus était parfaitement opaque pour l'usager, qui ne pouvait pas savoir a priori quel était le risque (i.e. la probabilité) que le chiffre annoncé soit atteint voire dépassé.

Il reste cependant aux prévisionnistes du chemin à parcourir pour améliorer la qualité de ce nouveau type de prévision. D'une part, en terme de calibration, nous avons pu observer que subsistaient d'anciens réflexes de surestimation systématique des lames d'eau attendues, par crainte de rater un épisode de pluie intense (comportement "parapluie").

Les prévisionnistes ont également de grandes marges de progrès pour améliorer leur capacité à déceler un faible risque de pluie intense. Sur plusieurs bassins ils sont, en prévision du

risque (c'est à dire en utilisant le quantile 90 pour détecter les pluies intenses), moins performants que le modèle ANALOG, alors qu'ils disposent chaque jour des sorties de ce modèle. Si les prévisionnistes sont meilleurs que les sorties du modèle ANALOG pour détecter une pluie intense "en tendance centrale", ils doivent en revanche davantage tenir compte des sorties du modèle ANALOG pour rester vigilant aux événements extrêmes qui s'éloignent de ce qui leur semble prévisible.

## 6.3 Discussion

Les trois mesures de la sûreté d'un système complexe (6.1.4, 6.1.5 et 6.1.6) permettent de mieux comprendre comment en pratique la prévision des crues contribue à la Sûreté Hydraulique.

Il nous faut une cependant une mesure quantitative pour pouvoir **piloter le système de prévision par sa performance**, comme proposé en 3.4 pour l'optimisation de la production hydroélectrique. Une telle mesure nous aurait ici permis de positionner le système de prévision hydrométéorologique par rapport à d'autres éléments qui contribuent à la Sûreté Hydraulique : maintenance préventive, redondance des organes d'ultime secours du barrage, surdimensionnements lors de la conception des installations, présence de personnel de quart, etc. (5.1).

Nous n'avons cependant pas pu quantifier la valeur de la prévision vis-à-vis de l'enjeu de sûreté comme nous l'avions fait au chapitre 3 vis-à-vis des enjeux économiques. Une telle quantification impose en effet l'utilisation du concept de **consentement à payer** (3.1). Or force est de constater que les décideurs ne savent pas (ou ne veulent pas) donner leur consentement à payer pour une information météorologique, que celle-ci leur soit utile pour améliorer leur compétitivité économique (voir l'expérience malheureuse et non publiée de Brown et Murphy, décrite en 3.1), ou que celle-ci leur permette d'améliorer leur niveau de sûreté[58] (6.2.2).

La méthode analytique nous a cependant permis de travailler sur un concept-clef pour la gestion de crue : la détection avec suffisamment d'anticipation des épisodes de pluie intense (6.2). Vis-à-vis de la détection de ces phénomènes, le quantile 90 de la prévision est bien mieux adapté qu'une simple prévision déterministe (6.2.3) et permet en théorie d'améliorer la performance de la prévision (6.2.4)[59].

Pour qu'en pratique la prévision probabiliste produise effectivement une amélioration de la Sûreté Hydraulique en gestion de crue, notons cependant qu'il faut prendre quelques précautions pour ne pas bouleverser le système actuel de gestion de crue, un système qui fonctionne déjà "plutôt bien"[60].

Pour ne pas bouleverser les habitudes de communication entre acteurs (ce qui diminuerait la résilience de l'organisation, 6.1.6), nous préconisons de **ne pas remplacer** un mode de

---

[58] Les décideurs d'EDF veulent simplement « être inattaquables » en montrant qu'ils ont été exemplaires en matière de Sûreté Hydraulique. On montre cependant en Annexe 8 que les décideurs ne sont jamais à l'abri du risque juridique, et que prendre des décisions en fonction d'éventuelles conséquences juridiques mène à une impasse.

[59] Nous avons cependant pu constater qu'en matière de "prévision du risque" les prévisionnistes avaient encore une bonne marge de progrès (6.2.6).

[60] Durant les deux crues du Tarn de novembre et décembre 2003, malgré leur importance (toutes deux décennales) et leur soudaineté (des gradients de montée de débit de 250 $m^3/s$ à l'heure, le débit étant passé pour la 2$^e$ crue de 80 $m^3/s$ à 2000 $m^3/s$ dans la soirée et le début de nuit), aucun incident significatif n'a été à déplorer.

communication des incertitudes par un autre, mais d'**ajouter** un nouvel élément d'information dans le bulletin écrit ou dans les discussions téléphoniques entre prévisionniste et usager : le quantile 90 sur les débits à venir.

## 6.4 Bibliographie

Duband, D., 1981. Prévision spatiale des hauteurs de précipitations journalières, La Houille Blanche, n° 7/8.

Krzysztofowicz, R., Sigrest, A., 1999. Calibration of probabilistic quantitative precipitation forecasts. Weather and Forecasting, 14(3), 427-442.

Le Moigne, 2000, Science des systèmes, Encyclopædia Universalis

Murphy, A. H., Daan, H., 1984. Impacts of feedback and experience on the quality of subjective probability forecast : Comparison of results from the first and second years of the Zierikzee Experiment, Monthly Weather Review, vol. 112, 413-423.

Piaget, J., 1967. Biologie et connaissance. Gallimard, Paris, 1345 p.

Obled, C., Bontron, G., Garçon, R., 2002. Quantitative precipitation forecasts: a statistical adaptation of model outputs through an analogues sorting approach, Atmospheric Research, vol. 63.

Stanovich, K. E., 1999. Who is rational? Studies of individual differences in reasoning. Mahweh, NJ: Erlbaum.

Wybo, J.-L., 2001, Risky business, Tomorrow, June.

X-EX-03-05 : La Gestion des aménagements hydroélectriques lors d'épisodes de crue, juin 1998

X-EX-02-00 : La maintenance des installations hydroélectriques, mai 1997

X-EX-04-06 : Guide d'analyse des ESSH, juin 2002

# 7 Une langue qui révèle la sous-estimation des incertitudes de l'expert

*L'usage de cette langue trahit chez celui qui la parle certains traits de son caractère.*
*Une trop grande confiance en lui par exemple.*
*Mais la prise de conscience de ce type de défaut est salutaire pour l'expert.*
*Elle lui permet de mieux répondre à la question qu'il se pose souvent :*
*Comment se faire comprendre lorsqu'on n'est pas sûr de ce que l'on avance ?*

## 7.1 Introduction

La prévision hydrométéorologique est un outil indispensable à la gestion des ressources en eau. C'est particulièrement le cas pour des exploitants hydroélectriques, tels que Hydro-Québec qui contrôle 8 systèmes fluviaux (sur une surface totale d'environ 700 000 km²), et gère 49 usines hydroélectriques (pour une puissance totale installée de 37 GW) ; ou encore pour EDF, qui contrôle 75% des eaux de surface françaises, et gère plus de 450 usines hydroélectriques (puissance totale installée : 20 GW).

Dans le Chapitre 2 nous avons présenté le fonctionnement des deux centres de prévision opérationnels d'EDF à Grenoble et à Toulouse, en réponse aux nombreux enjeux opérationnels des usagers de la prévision. Nous avons également souligné les limites du mode essentiellement déterministe sur lequel se communique la prévision. Nous avons donné un argumentaire en faveur de la prévision probabiliste, qui permettrait au prévisionniste de s'exprimer plus finement, et à l'usager de disposer d'une information plus riche, tout en évitant l'arbitraire lié au flou qualitatif et informel.

Rappelons tout d'abord que si la majorité des centres de prévision opérationnels ne délivrent à ce jour que des prévisions déterministes, celles-ci sont généralement complétées par des indications qualitatives portant principalement sur l'évaluation des incertitudes associées.

La question qui se pose alors aux prévisionnistes est de quantifier explicitement leurs propres incertitudes. Atger (1999, 2001) donne des exemples de prévision probabiliste en météorologie où la fonction de densité de probabilité (pdf) est estimée en bruitant les données d'entrée du modèle numérique de prévision d'ensemble du CEPMMT (Centre Européen pour les Prévisions Météorologiques à Moyen Terme).

Cependant, la prévision opérationnelle repose généralement sur l'expertise de prévisionnistes expérimentés. Ceux-ci sont en effet capables, en particulier dans l'urgence d'une situation de crise, de traiter d'importants volumes de données d'origine hétérogène, par exemple : divers produits issus de différents modèles de prévision météorologique (déterministes ou ensemblistes), données télé-transmises depuis le réseau de mesure au sol (précipitation, température, hauteur d'eau dans les rivières, etc.), images Radar et satellitaires. Même si les modèles numériques sont de plus en plus largement utilisés du fait de leur performance sans cesse croissante, l'interprétation par un expert prévisionniste est toujours, en pratique, le moyen le plus courant pour produire des prévisions opérationnelles.

Dès lors, on comprend bien l'intérêt de s'en remettre au prévisionniste pour lui demander de quantifier ses incertitudes[61], sachant qu'il a déjà l'habitude de les qualifier. Il est en effet

[61] On pourrait en effet faire une simple adaptation statistique à partir de séries de prévisions expertisées et de données observées et construire des pdf conditionnelles aux prévisions émises. D'une part, ce serait évincer l'expert de son propre domaine d'expertise (l'évaluation des risques hydrométéorologiques), ce qui rendrait la démarche peu acceptable dans la communauté des prévisionnistes. D'autre part, comme toute adaptation statistique suppose d'avoir à faire à des processus stationnaires (ni changement climatique, ni changement de modèle météorologique), cette démarche conduirait tout simplement à interdire aux prévisionnistes tout progrès. Faire une adaptation statistique sur les prévisions expertisées irait donc à l'encontre des méthodes de travail et d'apprentissage du prévisionniste, réalisant chaque jour un retour d'expérience à partir des prévisions qu'il avait émises la veille, analysant et mémorisant ses erreurs et ses succès, pour affiner ses modèles, concepts, et heuristiques. Savage (1971) illustre non sans humour l'abîme dans lequel une telle adaptation statistique nous ferait tomber : "You might discover with experience that your expert is optimistic or pessimistic in some respect and therefore temper his judgments. Should he suspect you of this, however, you and he may well be on the escalator to perdition" (p. 796).

habitué à répondre à des sollicitations téléphoniques pour préciser son niveau d'incertitude, la vraisemblance de telle ou telle évolution météorologique, son appréciation du risque de dépassement d'un seuil d'alerte sur un débit.

Plusieurs moyens peuvent être utilisés pour quantifier ce niveau d'incertitude (4.1.3). On peut par exemple demander au prévisionniste de fournir un intervalle de crédibilité qui reflète son degré de conviction que la variable prévue tombe dans l'intervalle fourni (Murphy et Winkler, 1974), un quantile de probabilité sur la variable prévue (Krzysztofowicz *et al.*, 1993), c'est à dire un niveau que la variable prévue n'aura qu'une probabilité fixée de dépasser, ou enfin la probabilité que la variable prévue dépasse divers seuils (Murphy et Dann, 1984).

Cependant, la quantification experte des incertitudes souffre de plusieurs limites. Parmi elles, Murphy (1993) a identifié l'inconsistance. On définit ici la consistance comme le degré d'adéquation entre le jugement interne du prévisionniste cherchant à comprendre la situation météorologique et à en prévoir les évolutions, et l'information de prévision qu'il diffuse à l'usager[62](Murphy, 1993).

Le prévisionniste doit bien souvent prendre en compte des éléments extérieurs à son appréciation experte des évolutions météorologiques, comme la vulnérabilité des régions concernées par un risque d'orage, ou plus généralement, l'impact de ses prévisions sur les activités des usagers. Selon Murphy (1991) et Murphy et Winkler (1974), cette attitude amène le prévisionniste à surestimer le risque d'occurrence d'événements qui ont de lourdes conséquences pour les usagers.

Le biais de surestimation des prévisions[63], bien connu dans la communauté des météorologistes (Murphy et Dann, 1984), pourrait donc être une manifestation de l'inconsistance. Ce biais peut se traduire par la surestimation de la variable à prévoir (Doswell et Fueck, 1989), de quantiles de probabilités sur celle-ci (Krzysztofowicz et Sigrest, 1999), ou des probabilités que certains seuils soient dépassés (Murphy et Daan, 1984). Murphy (1991) a cependant montré que ce biais de surestimation était plus faible pour des prévisions probabilistes. L'auteur considère en effet que le vecteur de communication probabiliste est plus adapté à traduire la complexité du jugement du prévisionniste, et donc que les distorsions entre ce qu'il pense et ce qu'il communique sont moindres que s'il était obligé de fournir une prévision déterministe (communiquer un seul chiffre), voire une prévision catégorique (dire si il y a ou non un risque).

Krzysztofowicz et al. (1993) a expérimenté des moyens pratiques pour réduire cette inconsistance. Après avoir fourni divers quantiles de probabilité, le prévisionniste est invité à considérer à nouveau l'ensemble des chiffres qu'il a fourni, et à s'assurer notamment qu'il considère bien qu'il y a équiprobabilité entre plusieurs événements :

- autant de chances que la hauteur d'eau mesurée tombera à l'intérieur de l'intervalle de crédibilité compris entre les quantiles 25% et 75% qu'à l'extérieur
- autant de chance que la hauteur d'eau mesurée tombera en dessous du quantile 25% qu'au dessus du quantile 75%

---

[62] Murphy (1993), définit la consistance comme la première qualité d'une prevision, i.e. "the correspondence between forecaster's judgments and their forecasts (type 1 goodness, or consistency)".

[63] R. Garçon (communication orale) remarque que ce biais de surestimation est lié à l'existence d'un risque associé à de fortes valeurs de la variable à prévoir (risque de précipitations intenses, de crue, de vent violent, etc.). A l'inverse, l'expert peut être amené à se prononcer sur un risque d'étiage sévère comme ce fut le cas sur la Loire à l'été 2003 (source froide des réacteurs nucléaires installés en bordure du fleuve). Dans ce cas, l'expert qui voudrait se couvrir face à ce risque serait enclin à sous-estimer les débits prévus.

Nous examinerons spécifiquement la question de la consistance en section 9.1. Nous faisons ici (dans les sections 7 puis 8) l'hypothèse que le prévisionniste est consistant, c'est à dire qu'il fournit des prévisions dont la qualité première est d'être conforme à son jugement (Murphy, 1993).

Cette hypothèse étant posée, demeure la question de la fiabilité des prévisions probabilistes. Nous entendons ici par fiabilité le degré de correspondance entre les probabilités prévues et les fréquences observées (Murphy, 1973). La fiabilité requiert un niveau d'exigence plus élevé que la simple consistance. On demande au prévisionniste non seulement de donner sa prévision sans être influencé par des facteurs externes ; celui-ci doit en plus quantifier le niveau d'incertitude associé à ses prévisions. La fiabilité des prévisions probabilistes est discutable, du fait de l'existence de deux biais.

Le premier biais que nous venons d'évoquer, à savoir le biais de surestimation des prévisions demeure même lorsqu'on permet au prévisionniste de se servir du langage des probabilités, « le langage intérieur des prévisionnistes » (Sanders, 1963), qui évite par conséquent « de consacrer du temps et des efforts à traduire des jugements probabilistes en prévisions catégoriques [ou déterministes] » (Murphy, 1991). Le défaut de fiabilité que constitue le biais de surestimation ne s'explique pas uniquement par l'inconsistance des prévisionnistes, obligés d'utiliser un langage qui n'est pas le leur (la langue déterministe ou catégorique), et d'intégrer des préoccupations qui ne sont pas les leurs (enjeux des usagers). Nous venons de voir que même en utilisant le langage probabiliste, censé abstraire le prévisionniste de la prise en compte des enjeux de l'usager (Murphy, 1991), le biais de surestimation demeure (Krzysztofowicz et Sigrest, 1999, Murphy et Daan, 1984).

La sous-estimation des incertitudes (surconfiance) est un deuxième biais qui a rarement été étudié dans le contexte de la prévision hydrométéorologique opérationnelle. Ce biais consiste pour le prévisionniste à surestimer ses capacités à fournir une prévision précise. Krzysztofowicz et Sigrest (1999) ont décelé un tel biais chez des prévisionnistes en hydrométéorologie, en notant que l'intervalle de confiance à 50% (c'est à dire l'intervalle compris entre les quantiles 25% et 75%) était trop étroit en moyenne.

La surconfiance peut également se traduire par une surestimation systématique des probabilités élevées (associées à des événements courants), et une sous-estimation systématique des probabilités faibles (associées à des événements rares et sévères). Cela signifie que le prévisionniste est trop confiant dans ses capacités à prévoir qu'un phénomène se réalisera ou non. Murphy et Dann (1984) (noté par la suite M&D) ont étudié de manière extensive le biais de surestimation durant deux années d'expérimentation, sans toutefois relever de biais de surconfiance. Pourtant, l'un des prévisionnistes qui a pris part à l'expérience de M&D était manifestement surconfiant. La figure 5 de M&D montre en effet que durant l'année 1981-1982, le prévisionniste C surestimait systématiquement les probabilités élevées, et sous-estimait les probabilités faibles. Il est vrai que pour tous les autres prévisionnistes de l'expérience de M&D, le biais le plus flagrant était le simple biais de surestimation des probabilités. Pourtant, les résultats présentés par M&D laissent suggérer une légère surconfiance, associée à ce biais de surestimation des probabilités : toutes les probabilités sont surestimées, mais les probabilités les plus élevées sont plus surestimées que les probabilités les plus faibles.

Depuis plusieurs décennies de nombreuses études en psychologie expérimentale ont identifié la surconfiance comme l'un des principaux obstacles à l'évaluation correcte des probabilités subjectives (Kahneman et al., 1982, et références incluses). Cependant, les auteurs ne s'accordent pas sur l'interprétation qu'on peut faire de ce phénomène.

Certains auteurs interprètent la surconfiance comme une déficience cognitive, convaincus que le jugement des individus est caractérisé par certaines irrationalités et autres défauts systématiques (les théories normatives de Savage, 1954, pour le cadre général, et Krzysztofowicz, 1983, pour une application au cas de la prévision hydrométéorologique opérationnelle). Ces théoriciens reçurent l'appellation de "Mélioristes" par Stanovich (1999) du fait de leur espoir de voir se développer des méthodes d'entraînement pour réduire ces défauts cognitifs.

D'autres auteurs interprètent plutôt la surconfiance comme une adaptation adéquate et naturelle de l'esprit humain à son environnement social (Cosmides et Tooby, 1996). Ceux-ci furent baptisés "Panglossiens" en référence à Maître Pangloss qui considère que "tout va pour le mieux dans le meilleur des mondes possibles" (Voltaire, 1759).

Fischhoff (1982) illustre bien chacune des interprétations que l'on peut faire de la surconfiance : "Lorsque les médecins surestiment la vraisemblance d'une maladie, c'est peut-être parce qu'ils ont une mauvaise connaissance de leurs connaissances, mais peut-être aussi par peur de se voir accuser de négligence, ou encore par avidité des retombées financières que pourraient apporter des tests supplémentaires."

Nous étudions dans ce Chapitre 7 le phénomène de surconfiance dans les centres de prévision opérationnels d'EDF et d'Hydro-Québec. Nous présentons un exemple de surconfiance dans un centre de prévision opérationnelle (7.2). Nous analysons l'effet d'un feed-back aux prévisionnistes sur la fiabilité de l'estimation de leurs incertitudes (7.3). Nous examinons ensuite les moyens d'assurer la durabilité à long terme d'une réduction de la surconfiance à la suite d'un tel feedback (7.4). Les résultats obtenus nous amèneront à formuler quelques remarques conclusives (7.5). Nous émettons enfin des propositions pour une implémentation opérationnelle d'un système de prévision probabiliste expertisée qui prenne en compte le phénomène de surconfiance (7.6).

## 7.2 Un exemple de surconfiance dans un centre de prévision opérationnelle

### 7.2.1 Contexte

Depuis plus d'une décennie, il est demandé aux météorologues d'Hydro-Québec de fournir leurs prévisions opérationnelles sous forme probabiliste. Les prévisionnistes fournissent ainsi des intervalles de crédibilité à 70% sur 14 villes québécoises pour les cumuls quotidiens de précipitations prévues pour les échéances J à J+3. Ces données sont alors injectées dans un interpolateur qui renvoie des intervalles de crédibilité à 70% sur les lames d'eau spatiales, pour chacun des 89 bassins gérés par Hydro-Québec.

Ces lames d'eau par bassin servent à alimenter les modèles hydrologiques de prévisions d'apport, selon trois scénarios : la borne inférieure de l'intervalle, sa borne supérieure, et son milieu. L'hydrologue les considère alors comme les quantiles 15%, 85%, et la médiane.

On dispose pour cette étude de deux ans de données, provenant d'une population bien restreinte puisqu'il s'agit d'un seul prévisionniste de Trans-Energie[64] (et de son adjoint en cas d'absence). Nous avons décidé de ne nous intéresser qu'aux mois d'été (juin à septembre), pour exclure les épisodes neigeux. A la sortie de l'interpolateur, on dispose alors de prévisions

[64] Trans-Energie est le gestionnaire du réseau de transport d'électricité du Québec.

sur 89 bassins versants, émises sur 140 journées différentes, et à chaque fois pour 4 échéances (de J à J+3), soit près de 50 000 prévisions.

### 7.2.2 Résultats

On considère que l'usager a un modèle simple d'interprétation des intervalles de crédibilité fournis : pour une échéance donnée, il associe à chaque prévision la même probabilité qu'on note θ que l'observé tombe dans l'intervalle fourni.

Cela revient à admettre deux hypothèses :

- **l'hypothèse d'homogénéité** : pour une échéance de prévision donnée, tous les intervalles fournis ont la même chance de recueillir la réalisation
- **l'hypothèse d'indépendance** : pour une échéance de prévision donnée, la probabilité qu'un intervalle de crédibilité recueille la réalisation ne dépend pas de la journée de prévision, ni de l'auteur de la prévision, ni des autres paramètres à prévoir.

θ est donc la probabilité effective qu'une réalisation tombe à l'intérieur de l'intervalle de crédibilité fourni. Cette probabilité effective signifie "valeur de la probabilité telle que doit l'interpréter l'usager" ; on la compare à la probabilité "demandée au prévisionniste", qui elle vaut 0,7.

Le Tableau 17 présente les estimations de θ suivant les échéances à l'aide des 50 000 prévisions analysées. Par une estimation Bayésienne de l'incertitude sur θ (Annexe 1) on présente également un intervalle de confiance à 95% sur ce paramètre compte tenu des hypothèses d'indépendance et d'homogénéité.

| Echéance | Estimateur de θ | $IC_{95\%}$ [θ\|n,k] |
|---|---|---|
| J | 0,53 | [0,52 ; 0,54] |
| J+1 | 0,51 | [0,50 ; 0,52] |
| J+2 | 0,47 | [0,46 ; 0,48] |
| J+3 | 0,43 | [0,42 ; 0,44] |

**Tableau 17 : Calibration des intervalles de crédibilité à 70% fournis par les prévisionnistes d'Hydro-Québec. Estimateur et intervalle de confiance à 95% de θ.**

L'hypothèse d'indépendance que peut faire l'usager est discutable[65] dans la mesure où chaque jour les prévisions de précipitation sur les 89 bassins versants sont réalisées à partir des intervalles de crédibilité fournis par le prévisionniste pour 14 villes québécoises. Le prévisionniste peut de surcroît lui-même produire des prévisions interdépendantes : il peut par exemple surestimer les précipitations sur plusieurs des 14 villes. Si l'interpolateur ajoute de la dépendance entre les prévisions de précipitation émises, notons qu'un prévisionniste aurait lui-même très certainement fourni des prévisions très interdépendantes si on lui avait demandé de fournir quotidiennement 89 intervalles de crédibilité.

[65] Ces hypothèses qui peuvent paraître très fortes sont cependant celles qu'est en droit de se faire l'usager. Lorsqu'il utilise les prévisions probabilistes qui lui sont fournies, il ne se préoccupe pas de savoir qui était l'auteur de la prévision, ni de tisser des relations entre l'estimation des incertitudes de prévision courante avec la prévision de la veille, etc.

Malgré la faiblesse d'une telle hypothèse, qui conduit à obtenir des intervalles de confiance sur θ trop étroits, la constance observée dans ces résultats n'en demeure pas moins troublante :

- 0,70 est bien au-dessus de tous les intervalles de confiance sur θ quelle que soit l'échéance de prévision
- θ décroît strictement avec l'échéance de la prévision

On peut considérer les échéances successives de prévision comme des niveaux de difficulté croissants dans la question à laquelle le prévisionniste se trouve confronté. La vérification des prévisions météorologiques confirme ce jugement intuitif : les prévisions à long terme sont moins performantes en moyenne que celles à plus court terme[66] (Atger, 2000).

Le premier résultat permet de conclure que le prévisionniste d'Hydro-Québec est surconfiant. Alors qu'on lui demande de produire des prévisions sous forme d'intervalles de crédibilité à 70%, seulement 50% des intervalles fournis contiennent effectivement la lame d'eau observée. Tout comme Krzysztofowicz et Sigrest (1999), nous montrons ici que les intervalles de crédibilité fournis sont trop étroits en moyenne.

Le second résultat met également en avant un autre fait intéressant : plus une question est difficile, plus le prévisionniste est surconfiant. Ce phénomène qui fait l'objet d'une large littérature en psychologie a été baptisé "hard-easy effect"[67] (Lichtenstein, Fischhoff, 1977, Suantak et al., 1996, Juslin et al. 2000).

Le Tableau 18 montre en effet qu'en moyenne les prévisionnistes n'élargissent pas leurs intervalles de crédibilité lorsque l'échéance de la prévision croît. A l'échéance J+3, le prévisionniste affiche même en moyenne des incertitudes plus faibles qu'à l'échéance J[68].

---

[66] Il ne fait aucun conteste qu'en moyenne la difficulté de la prévision s'accroît avec l'échéance de la prévision. Néanmoins, il existe certaines configurations particulières pour lesquelles cette assertion n'est pas vraie. R. Garçon (comm. pers.) donne l'exemple de la prévision hydrologique sur la Garonne au 8 septembre 2003. Le fleuve vient de subir une première crue (les sols sont donc bien saturés), et le prévisionniste attend pour la journée à venir 25 à 60mm de pluie sur ce bassin. Son intervalle de crédibilité à 80% sur le débit de la Garonne à J+1 aurait donc été de [100 $m^3/s$ ; 500 $m^3/s$]. La situation météorologique à 1 semaine étant anticyclonique, le prévisionniste s'attend avec bien moins d'incertitude un retour au débit de base (70 $m^3/s$), et aurait pu donner comme $IC_{80\%}$ : [60 $m^3/s$ ; 100 $m^3/s$].

[67] Nous proposons deux explications techniques pour éclairer ce Hard-easy effect dans le cadre de la prévision météo.

La première, suggérée par J. Shanteau (comm. pers.) est que les prévisionnistes ont bien plus de feed-back sur leurs prévisions à J pour les échéances plus lointaines. Ils ont en effet l'habitude de débuter leur travail de prévision par un court retour d'expérience sur leurs prévisions de la veille. Ce retour d'expérience étant bien plus rare pour les prévisions émises deux ou trois jours avant, il leur est plus difficile de se corriger leurs défauts de sous-estimation des incertitudes des prévisions associées à ces échéances plus lointaines.

R. Garçon (comm. pers.) suggère quant à lui que c'est la similarité des méthodes utilisées pour produire des prévisions quelle qu'en soit l'échéance qui peut être à l'origine de cette non prise en compte de l'augmentation de l'incertitude de prévision. En effet, quelle que soit l'échéance de la prévision (de J à J+3), le prévisionniste dispose globalement des mêmes outils météo qu'il utilise de la même façon pour traduire les phénomène identifiés en manifestations sensibles (précipitations). Ces outils étant déterministes (la prévision météo d'ensemble n'est pas encore utilisée quotidiennement par les prévisionnistes de DTG), le prévisionniste attribue la même incertitude aux prévisions de précipitation à J qu'à J+3.

[68] Nous proposerons des éléments d'explication au paragraphe 7.2.3.

| J | J+1 | J+2 | J+3 |
|---|---|---|---|
| 2.36 | 2.41 | 2.25 | 2.13 |

**Tableau 18 : Moyenne de la largeur des intervalles selon l'échéance (en mm). Intervalles de crédibilité à 70% fournis par les prévisionnistes d'Hydro-Québec.**

Le Tableau 19 donne un élément d'explication : le prévisionniste n'augmente qu'une fois sur trois la largeur de l'intervalle fourni en passant d'une échéance à la suivante. La majeure partie du temps, le prévisionniste conserve la même largeur de l'intervalle (par effet d'ancrage), voire la diminue.

| | De J à J+1 | De J+1 à J+2 | De J+2 à J+3 | En moyenne, d'une échéance à la suivante |
|---|---|---|---|---|
| Proportion des intervalles dont la **largeur augmente** | 35% | 33% | 31% | **33%** |
| Proportion des intervalles dont la **largeur diminue** | 29% | 26% | 32% | **29%** |
| Proportion des intervalles dont la **largeur reste constante** | 36% | 42% | 38% | **39%** |

**Tableau 19 : Evolution de la largeur des intervalles avec l'échéance. Intervalles de crédibilité à 70% fournis par les prévisionnistes d'Hydro-Québec.**

### 7.2.3 Interprétation et discussion

Si l'on s'en réfère à l'approche décrite en 7.1 comme mélioriste, ces résultats éclairent le fait que les prévisionnistes ne savent pas estimer correctement le niveau de leur ignorance lorsqu'ils émettent une prévision. Une telle surconfiance pourrait être le fruit de la pratique traditionnelle en prévision opérationnelle. Cette pratique (Atger, 2000) consiste à identifier des phénomènes météorologiques d'échelle généralement large (plusieurs milliers de kilomètres), prévoir leur évolution, et traduire ces évolutions en terme de variables auxquelles l'usager est sensible (ici, la lame d'eau spatiale sur un bassin versant). Une telle tradition pourrait avoir induit dans l'esprit des prévisionnistes l'idée que les manifestations atmosphériques sont strictement déterministes[69].

[69] Certains prévisionnistes me semblent envisager les évolutions atmosphériques futures avec l'idée d'un déterminisme absolu, malgré les difficultés à mesurer l'état initial de l'atmosphère et à modéliser finement les différents mécanismes météorologiques. On trouve par exemple sur le site de Météo-France des explications pédagogiques sur la prévision de tempête : "les dépressions sont un moteur naturel et provisoire qui transforme en vent «l'essence» contenue dans le rail des dépressions" (www.meteo.fr). L'auteur de la formule reconnaît lui-même que l'analogie est simpliste, notamment du fait de "la notion de variabilité interne sans cause externe qui ne trouve guère son équivalent dans le système très rigide que constitue une voiture". Cependant, des discussions informelles avec les prévisionnistes d'EDF ont montré que cette idée mécaniste était tenace dans leurs esprits : avec une bonne connaissance à la fois de l'état initial de l'atmosphère et des "mécanismes qui en décrivent le comportement", on peut parfaitement prévoir les évolutions atmosphériques. Popper (1984) montre comment une telle idée mécaniste peut s'apparenter à une vision du "monde-horloge", et aboutir à la notion de déterminisme. Plus généralement, Frédéric Atger (comm. pers.) me faisait également remarquer que cette

Des discussions informelles avec certains prévisionnistes opérationnels d'EDF ont montré qu'à leurs yeux, la meilleure prévision qu'ils puissent produire consistait à fournir à l'usager leur meilleure estimation du paramètre à prévoir, à l'aide de tous leurs outils de modélisation et d'observation, et grâce à leur propre expérience. Il ne reste donc plus aucune place à un énoncé probabiliste, que ce soit dans la compréhension des phénomènes météorologiques ou dans l'information fournie à l'usager.

Du point de vue mélioriste, la conclusion serait donc que les prévisionnistes ne sont pas correctement calibrés dans l'estimation de leurs incertitudes parce qu'on ne leur a jamais enseigné à l'être. Nous examinerons en 7.3 la question de la formation et de l'entraînement des prévisionnistes pour améliorer leur faculté d'évaluer correctement l'incertitude liée aux prévisions qu'ils émettent.

Les Panglossiens quant à eux (7.1) affirmeraient que les prévisionnistes se sont parfaitement adaptés aux attentes des usagers. Ceux-ci attendent en effet des prévisionnistes qu'ils leur fournissent des éléments qui leur permettront de prendre rapidement et efficacement une décision. Comme cette décision finira bien par être déterministe (et ce, même si tout le processus cognitif a intégré des éléments probabilistes), le moyen le plus rapide, le plus simple et le plus efficace d'aboutir à une décision est d'utiliser des informations déterministes.

D'autre part, Fischhoff (1982, p 439) a montré que certaines motivations pouvaient expliquer la surconfiance chez les experts : « Bien entendu, si on fait appel à de tels experts (et si on les paye) à la mesure de la confiance qu'ils inspirent, ils pourraient être tentés de masquer l'étendue de leur ignorance ». Shanteau et al. (2002) a proposé un indicateur (nommé CWS, pour Cochran-Weiss-Shanteau) pour exclure ce type d'experts. Selon leur critère, un expert doit être à la fois fiable (au sens de Murphy, 1973, 7.1) et discriminant, c'est à dire ayant « la capacité à déceler de subtiles différences imperceptibles pour le grand public ». La fiabilité des prévisions opérationnelles peut être mesurée a posteriori, en comparant probabilités prévues par chacun des prévisionnistes et fréquences observées. La qualité de discrimination semble en revanche être beaucoup plus délicate à mesurer puisque cette mesure est basée sur l'inter comparaison de prévisions produites par différents prévisionnistes à partir d'une même situation météorologique, et que les prévisionnistes sont très rarement en compétition lors de l'élaboration de la prévision opérationnelle.

Cela nous amène à prendre en compte l'effet des normes sociales tant chez les experts que chez les décideurs. Il me semble que chez les décideurs (ici, les usagers), celui qui sait prendre rapidement des décisions (même si l'on s'aperçoit après coup qu'elles étaient mauvaises) est bien plus valorisé que celui qui mesure, analyse, intègre les diverses sources d'incertitudes, pèse le pour et le contre, et au final montre de la mesure dans ses arguments et de la lenteur à prendre des décisions[70].

De telles normes sociales existent également dans le milieu de la prévision opérationnelle. Soll (1996) a par exemple montré à partir de l'expérience de Murphy et Dann (1984) que « les prévisionnistes en météo sont évalués et récompensés selon la précision et la calibration de

---

manière de concevoir la prévision météorologique découle également de la longue tradition de la physique qui a mené au mécanisme (« mode de fonctionnement de ce qu'on associe à une machine » selon la définition du dictionnaire Petit Robert).

[70] Nous pouvons supposer qu'avec l'émergence de la notion de « principe de précaution » ce type de norme sociale va peu à peu évoluer, en valorisant mieux la prise en compte des incertitudes.

leurs prévisions[71] ». Nous suggérons ici que les prévisionnistes d'Hydro-Québec décrits dans ce Chapitre et qui ont manifesté une telle surconfiance peuvent avoir été soumis à une norme sociale récompensant bien plus la précision d'une prévision que sa calibration. En ce sens, ces experts pourraient se rapprocher de ceux décrits par Fischhoff (1982).

Si les prévisionnistes sont effectivement soumis à de telles normes sociales, on peut se demander s'il existe un quelconque moyen de réduire leur surconfiance. Serait-ce finalement la norme sociale elle-même qu'il faudrait changer ? Nous apportons des éléments de réponse dans les paragraphes 7.3 et 7.4.

## 7.3 Sensibiliser et former les prévisionnistes pour réduire la surconfiance

### 7.3.1 Contexte

Nous avons vu au Chapitre 2 que les centres hydrométéorologiques opérationnels d'EDF ont toujours délivré des prévisions sous forme déterministe[72].

Lors d'une journée de sensibilisation des prévisionnistes à la quantification de leurs incertitudes de prévision (Chapitre 4), il leur a été demandé de fournir, pour une situation météorologique donnée, leur estimation des intervalles de crédibilité à 50% de la lame d'eau spatiale attendue dans 6 bassins versants pour la journée à venir[73]. Les prévisionnistes disposaient pour cela de leurs outils de prévision habituels (Annexe 3), et travaillaient en binôme, comme dans leur pratique opérationnelle de la prévision. Les binômes ont été placés dans des salles séparées afin qu'ils ne s'influencent pas mutuellement.

Globalement, les 8 prévisionnistes de Toulouse, et les 6 de Grenoble ont chacun fourni 6 prévisions de lame d'eau spatiale pour une seule échéance, et pour deux situations météorologiques différentes. Soit un total de 162 prévisions (6 données manquantes).

[71] Nous devrions rajouter également la vitesse à laquelle la prévision est émise, c'est à dire l'anticipation qu'elle permet sur les phénomènes annoncés. Un tel attribut de qualité est souvent omis, peut être du fait que les exigences de l'usager dans ce domaine sont souvent satisfaites : il reçoit bien en temps et en heure la prévision pour les échéances souhaitées.

[72] A l'exception des prévisions d'apport de remplissage des grands réservoirs saisonniers.

[73] C'est à dire la précipitation prévue intégrée sur l'ensemble du bassin versant, et sur les 24 heures à venir (de 8h à 8h, temps universel).

### 7.3.2 Résultats

Avec les mêmes hypothèses et les mêmes notations qu'en 7.2, le Tableau 20 présente les résultats obtenus dans chaque expérience (au CHA et au CHPMC).

| | | **Avant feedback** | **Après feedback** |
|---|---|---|---|
| **Prévisionnistes du CHA (Grenoble)** | **Estimateur de θ** | **0,20** | **0,47** |
| | $IC_{95\%}$ [θln,k] | [0,09 , 0,37] | [0,32 , 0,63] |
| **Prévisionnistes du CHPMC (Toulouse)** | **Estimateur de θ** | **0,21** | **0,54** |
| | $IC_{95\%}$ [θln,k] | [0,11 , 0,34] | [0,40 , 0,68] |

**Tableau 20 : Calibration des intervalles de crédibilité à 50% fournis par les prévisionnistes d'EDF. Estimateur et intervalle de confiance à 95% de θ.**

Tout comme en 7.2.2, l'hypothèse d'indépendance est discutable, dans la mesure où chacun des prévisionnistes était appelé à donner sa prévision de la lame d'eau spatiale attendue sur 6 bassins versants, pour la même situation météorologique. Dans ces conditions, si le modèle météorologique reproduit mal la situation atmosphérique, cela peut conduire tous les prévisionnistes aux mêmes erreurs de prévision, et avoir le même effet sur plusieurs bassins.

Malgré la faiblesse de cette hypothèse, la similarité des comportements observés dans chacun des centres hydrométéorologiques a de quoi surprendre :

- Avant sensibilisation et feed-back, θ vaut environ 0,2 ; cela indique une sous-estimation marquée des incertitudes
- Après feed-back sur leur estimation des incertitudes, 0,5 est très proche du maximum de vraisemblance sur θ, ce qui suggère une absence de surconfiance.

### 7.3.3 Interprétation et discussion

Le feed-back que nous avons donné aux prévisionnistes concernait leur fiabilité, c'est à dire la calibration de leurs prévisions probabilistes. L'efficacité d'un tel feed-back pour améliorer les qualités de calibration suggérée ici a déjà été montrée par Lichtenstein et Fischhoff (1980).

On peut dès lors se poser la question de l'interprétation d'une telle amélioration. Les prévisionnistes ont-ils amélioré leurs facultés cognitives ? Est-ce la norme sociale qui a changé ?

L'interprétation mélioriste (7.1) selon laquelle les prévisionnistes auraient amélioré leurs facultés cognitives nous amène à considérer les limites d'une telle expérience. Les prévisionnistes peuvent en effet avoir réduit leur surconfiance dans un but purement académique, c'est à dire pour se conformer aux exigences d'un tel exercice. Ils ont en effet pu considérer cette expérience comme un jeu pour lequel on ne leur a pleinement dévoilé les règles qu'à la fin d'une première partie, et durant le débriefing et le feed-back. Ils auraient alors adapté leurs réponses au second tour pour tenter de satisfaire aux consignes de ce qu'ils ont pu considérer comme un examen. Dans cette hypothèse, l'expérience n'aurait rien changé à leur perception et à leur formulation des incertitudes lors de l'élaboration quotidienne de la prévision.

Cette expérience a cependant pu faire prendre conscience aux prévisionnistes de leur incapacité à estimer correctement leurs incertitudes. Dans ce cas, ils ont pu réajuster leur perception des limites de leur savoir. Si tel a été effectivement le cas, nous pourrions conclure de la même façon que Lichtenstein et Fischhoff (1980) : un stage intensif de sensibilisation et de formation à l'estimation des incertitudes, avec exercices et feed-back, améliore très sensiblement la fiabilité de l'estimation des incertitudes. Les prévisionnistes seraient tout du moins mis sur la bonne route, entrant dans le cercle vertueux de prise de conscience de leur incertitude, formulation de celle-ci, puis retour d'expérience quotidien et apprentissage, pour améliorer à terme la fiabilité de leur estimation courante des incertitudes de prévision.

Nous pourrions également interpréter cette disparition de surconfiance avec le point de vue panglossien : c'est peut-être la norme sociale à laquelle se conformaient les prévisionnistes qui a changé. Il nous semble cependant curieux qu'une norme sociale puisse être modifiée de manière si soudaine[74].

Connaître la nature de ce changement de comportement dans l'affichage des incertitudes est fondamental pour assurer la pérennité d'une telle amélioration. S'il s'avère que le point de vue mélioriste est le plus perspicace, il conviendra d'axer notre travail avec les prévisionnistes sur la pratique d'exercices académiques. Si ce sont les Panglossiens qui ont la meilleure analyse, il nous faudra dans ce cas travailler sur la norme sociale.

Quelle que soit la raison de cette disparition de surconfiance, la question de la durabilité d'une telle amélioration demeure. C'est la question à laquelle nous allons maintenant tâcher de répondre.

## 7.4 Assurer la pérennité de la fiabilité de l'estimation des incertitudes par les prévisionnistes

### 7.4.1 Contexte

Si le stage de sensibilisation et de formation des prévisionnistes a permis la disparition du phénomène de surconfiance lors de l'élaboration de la prévision, nous nous posions la question de la durabilité de tels résultats. Nous doutions en effet qu'une simple journée d'exercice puisse avoir des effets à long terme. La calibration des prévisions probabilistes est d'autant plus difficile à atteindre que la simple calibration déterministe[75] est déjà délicate à obtenir chez des prévisionnistes opérationnels (Olson et al., 1995).

Pour tester la durabilité à long terme de cette absence de surconfiance, une quantification au jour le jour des incertitudes de prévision est nécessaire. Pourtant, ce ne fut pas la raison qui nous conduisit à faire perdurer l'expérience décrite au Chapitre 4 par une expérimentation quotidienne et semi-opérationnelle. Notre principal objectif était en effet d'"améliorer la performance du système de prévision", en particulier face au risque de crue (ou d'autres événements météorologiques rares et dommageables pour le système de production ou de distribution d'électricité). Nous voulions donc travailler conjointement avec les

[74] V. Fortin (comm. pers.) remarque que la norme professionnelle peut rester la même (« être performant pour l'entreprise »), l'indicateur de performance étant périodiquement redéfini par la direction de l'entreprise. En Amérique du Nord, une part significative des revenus des employés étant significativement rattachée à l'amélioration d'un indicateur de performance, ceux-ci sont habitués à changer très rapidement de comportement pour s'adapter fréquemment aux nouveaux critères sur lesquels ils sont pécuniairement récompensés.

[75] La calibration probabiliste, ou fiabilité, est le degré de correspondance entre les probabilités prévues et les fréquences observées (Murphy, 1973). La calibration déterministe est l'adéquation entre la moyenne des prévus et des observés.

prévisionnistes pour mettre au point un système de prévision qui soit plus adapté à la gestion des risques[76]. C'est donc pour améliorer la détectabilité du système de prévision aux phénomènes météorologiques à risques que nous avons proposé aux prévisionnistes de fournir systématiquement le quantile 90% de leurs prévisions[77]. Les prévisionnistes ont eux-mêmes choisi les variables à prévoir associées à des enjeux de sécurité importants pour l'usager. Ils ont ainsi retenu une prévision de vent en rafale, 3 prévisions de débit et 4 prévisions de précipitation (Chapitre 4).

Les prévisionnistes ont donc fourni quotidiennement, en plus de leur prévision déterministe (médiane de leur prévision probabiliste, cf. note n° 34 p 70), les quantiles 10% et 90% sur la variable à prévoir. En d'autres termes, ils ont fourni en plus de leur prévision déterministe habituelle un intervalle de crédibilité à 80% centré en probabilité autour de cette prévision.

Pour les lames d'eau, les prévisionnistes ont choisi de fournir uniquement les quantiles 50% et 90%. D'une part le quantile 10% aurait été presque toujours fixé à 0, et n'aurait donc présenté que très peu d'intérêt ; d'autre part, le quantile 90% leur semblait bien plus primordial pour faire face aux enjeux de sûreté[78]. On appellera par la suite ce doublet ($q_{50\%}$ ; $q_{90\%}$) « demi intervalle de confiance à 80% ». Une bonne estimation des incertitudes signifierait que 40% des réalisations tombent dans ces demi-intervalles de confiance à 80%.

Nous analysons ici les prévisions probabilistes fournies quotidiennement durant les 6 premiers mois de l'expérience. Nous avons collecté durant cette période 945 prévisions de lame d'eau, 732 prévisions de débit, et 187 prévisions de vent.

### 7.4.2 Résultats

Avec les mêmes hypothèses et les mêmes notations qu'en 7.2, et 7.3, le Tableau 21 présente les résultats obtenus pour la prévision probabiliste des lames d'eau spatiales. Pour les deux échéances, $\theta$ est largement supérieur à 0,4. Ces résultats permettent donc de conclure à une sous-confiance de la part des prévisionnistes, qui ont fourni des intervalles trop larges par rapport à leur niveau de connaissance.

| Echéance | Estimateur de $\theta$ | $IC_{95\%}$ $[\theta\|n,k]$ |
|---|---|---|
| J | 0,61 | [0,57 , 0,65] |
| J+1 | 0,64 | [0,60 , 0,68] |

**Tableau 21 : Calibration des demi-intervalles de crédibilité à 80% fournis par les prévisionnistes d'EDF. Lames d'eau spatiales. Estimateur et intervalle de confiance à 95% de $\theta$.**

[76] Le mot risque est ici prix dans son sens courant d' « événement dommageable plus ou moins prévisible ».

[77] Notre objectif était donc d'adapter le couple (TFA, TEM) à des usagers très sensibles aux événements manqués. Pour ces usagers, le coût d'un événement manqué est largement supérieur à celui d'une fausse alerte. Nous avons retenu, pour fixer un cas d'école, que l'usager avait un coût d'événement manqué 9 fois plus élevé que celui d'une fausse alerte (dès que la probabilité de franchissement de seuil serait supérieur à 0,1, il engage des actions de protection ; à chaque fois, il sait pourtant qu'il y a 9 chances sur 10 pour que le seuil ne soit pas franchi, contre 1 chance sur 10 pour que celui-ci soit franchi). On se rapproche ainsi bien plus des enjeux opérationnels de l'usager en charge de la Sûreté Hydraulique qu'en sanctionnant de la même façon un événement manqué et une fausse alerte, comme la prévision déterministe nous amènerait à le faire (Chapitre 6).

[78] Pour cette raison, le quantile 50% (qui désigne également la prévision déterministe, cf. note n° 34) est parfois désigné dans le langage courant des prévisionnistes, par "prévision en tendance centrale", "prévision classique". Le quantile 90% est quant à lui parfois appelé "prévision risquée".

Le diagramme de fiabilité (Figure 25) montre que pour les deux échéances J et J+1, le quantile 90 est parfaitement calibré alors que le quantile 50 présente un biais négatif très marqué : il est dépassé trois fois sur quatre.

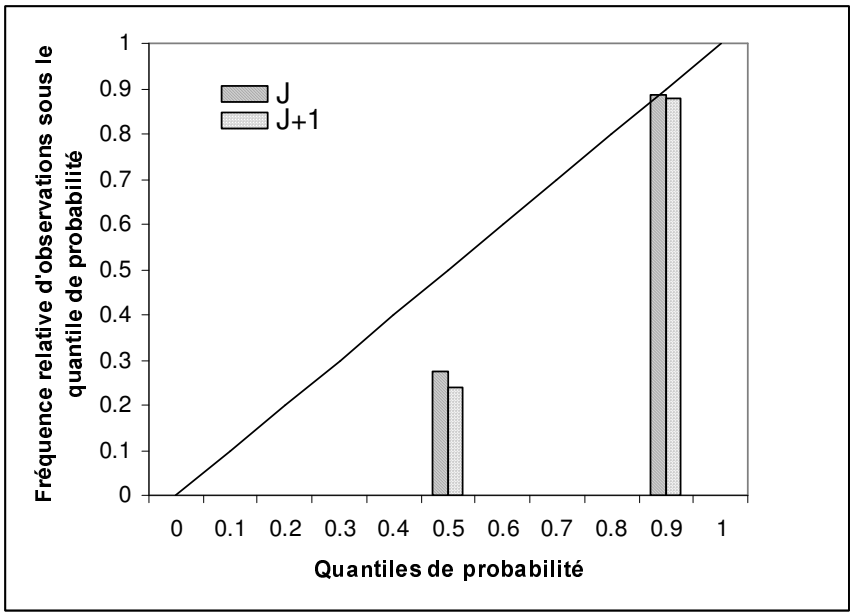



**Figure 25 : Diagramme de fiabilité pour les quantiles 50% et 90% fournis par les prévisionnistes d'EDF. Lames d'eau spatiales aux échéances J et J+1.**

Le quantile 50 peut être considéré comme la prévision déterministe qu'avaient l'habitude de produire les prévisionnistes (cf. note n° 34). Le biais en fréquence que nous venons d'évoquer (quantile 50 dépassé plus fréquemment qu'une fois sur deux) se traduit alors par un biais quantitatif de la prévision déterministe (la moyenne des prévus plus faible que la moyenne des observés). C'est ce qu'indique le Tableau 22. Ce biais est en partie dû à une trop grande fréquence de prévision de l'événement « aucune pluie significative sur le bassin versant »[79] (Tableau 23).

| Précipitations | Moyenne |
|---|---|
| Quantile 50 de la prévision Echéance J | 2,9 mm |
| Quantile 50 de la prévision Echéance J+1 | 2,9 mm |
| Observé | 3,4 mm |

**Tableau 22 : Moyenne des précipitations prévues et observées (expérience semi-opérationnelle à EDF).**

[79] R. Garçon (comm. pers.) suggère que cette surévaluation de l'événement « pluie nulle sur le bassin » peut refléter une tendance pour les prévisionnistes à fournir le mode plutôt que la médiane de la loi de probabilité sur la prévision de pluie. En effet, quel que soit le prévisionniste, le bassin choisi et la journée de prévision, la loi de probabilité subjective est mixte (Chapitre 4) : la probabilité qu'il pleuve 0mm n'est presque jamais réduite à 0, alors que la probabilité qu'il pleuve exactement une hauteur h non nulle est toujours réduite à 0. Le mode de telles lois mixtes étant à 0 (valeur toujours la plus probable), la confusion du prévisionniste entre mode et médiane entraîne une tendance à mettre trop souvent la médiane à 0.

| Evénement | Fréquence |
|---|---|
| La médiane de la prévision à J est nulle | 56% |
| La médiane de la prévision à J+1 est nulle | 57% |
| Aucune pluie observée | 44% |

**Tableau 23 : Fréquence des précipitations prévues et observées égales à 0. (expérience semi-opérationnelle à EDF).**

En conclusion, s'il y a un défaut dans l'estimation des incertitudes (Tableau 21), celui-ci n'est pas dû à la sous-estimation du risque de fortes pluies, mais à la surestimation des chances qu'il ne tombe pas une goutte d'eau. Cela suggère que les prévisionnistes sont au fait des limites de leur savoir pour prévoir les événements météorologiques rares et sévères, mais qu'en revanche, ils sous-estiment largement les pluies attendues en tendance centrale.

### 7.4.3 Interprétation

Deux interprétations différentes peuvent être avancées pour expliquer ces différences de calibration entre les quantiles 50 et 90.

Il se peut tout d'abord que nous ayons rompu la symétrie dans la tâche qui consiste à encadrer la meilleure estimation du paramètre à prévoir par un intervalle de crédibilité. Nous avons évoqué en 7.4.1 les raisons qui nous ont conduit à ne demander aux prévisionnistes que les quantiles 50 et 90., c'est à dire des ½ intervalles de crédibilité. L'asymétrie de cette tâche a cependant pu provoquer un biais de sous-estimation de la médiane. Les prévisionnistes ont en effet pu restaurer la symétrie de la tâche et rendre celle-ci plus intuitive en fournissant une sorte de borne inférieure qu'ils estimeront vraisemblablement dépassée, plutôt que la médiane.

Nous pourrions également fournir une toute autre interprétation, en invoquant l'idée que les processus cognitifs impliqués dans l'élicitation de chacun des deux quantiles sont différents. D'un côté, lorsqu'un prévisionniste doit fournir la médiane de sa prévision probabiliste, il peut donner sa meilleure estimation de la variable à prévoir, et fournir comme il en a l'habitude une prévision déterministe (note n°78). Comme toute prévision déterministe, cette estimation peut être quantitativement biaisée. Cependant, un tel biais est indépendant de l'estimation par le prévisionniste de ses incertitudes.

Le fait que ce biais soit rigoureusement identique pour chaque échéance de prévision (Tableau 22) confirme cette idée. Alors que le prévisionniste perçoit et gère des incertitudes croissantes lorsque l'échéance de la prévision passe de J à J+1, son évaluation moyenne de la lame d'eau la plus probable ne varie pas. Ce résultat suggère donc que le processus d'évaluation du scénario le plus probable n'interfère pas avec celui d'évaluation des incertitudes associées à la prévision. Lorsqu'on demande au prévisionniste de fournir un quantile 50, celui-ci s'attachera à fournir sa meilleure prévision possible avec les moyens dont il dispose, et en faisant abstraction des incertitudes associées.

D'un autre côté, lorsqu'on demande au prévisionniste de fournir un quantile de probabilité à 90%, il raisonnera à l'aide de scénarios, de leur fréquence climatologique associée, et de leur vraisemblance à ses yeux au moment où il élabore sa prévision. Il devra exclure les événements les plus rares et les moins probables pour ne conserver que ceux dont il estime la

probabilité à plus d'une chance sur dix. Exclure trop rapidement certains événements rares, c'est à dire sous-estimer la vraisemblance des événements rares est une forme de surconfiance (7.1).

Les feed-back réalisés ont été axés sur la calibration des quantiles élevés : savoir parler du risque en intégrant celui de se tromper, savoir évaluer, au quotidien, la vraisemblance de petits ou grands « scénarios-catastrophes » parmi diverses évolutions atmosphériques, etc. Les prévisionnistes ont donc pu réajuster correctement leur quantile 90 grâce à l'attention quotidienne qu'ils ont portée à la fiabilité de l'estimation de leurs incertitudes. Dans le même temps, il se peut tout à fait qu'ils n'aient pas porté attention aux biais qui pouvaient exister dans leur prévision déterministe.

Les processus cognitifs par lesquels les prévisionnistes améliorent chacune des calibrations (déterministe et probabiliste, note n°75) de leurs prévisions pourraient être indépendants.

La bonne calibration du quantile 90% reflèterait alors une bonne estimation des incertitudes par les prévisionnistes, tandis que le défaut de calibration du quantile 50% ne révèlerait qu'un biais déterministe.

Cette bonne estimation des incertitudes par les prévisionnistes est confirmée par l'analyse de leurs prévisions probabilistes de débit durant la même période. Les résultats présentés dans le Tableau 24 suggèrent en effet une bonne calibration globale des prévisions de débit. Une légère sous-confiance est même suggérée pour l'échéance J, même si celle-ci n'est pas significative compte-tenu de l'échantillon analysé. On décèle en revanche le phénomène du "hard-easy effect", qui consiste à sous-estimer d'autant plus ses incertitudes que la question posée est difficile (7.2.2)[80].

| **Echéance** | **Estimateur de θ** | **$IC_{95\%}$ [θ\|n,k]** |
|---|---|---|
| J | 0,82 | [0,78 , 0,85] |
| J+1 | 0,74 | [0,69 , 0,78] |

**Tableau 24 : Calibration des intervalles de crédibilité à 80% fournis par les prévisionnistes d'EDF. Prévisions des débits. Estimateur et intervalle de confiance à 95% de θ.**

Le diagramme de fiabilité présenté en Figure 26 montre un biais constant de surestimation sur chaque quantile, qui conduit à décaler les intervalles de crédibilité de même que la prévision déterministe vers le haut tout en gardant une bonne estimation des incertitudes elles-mêmes. En effet, les intervalles de crédibilité à 80% recueillent bien environ 80% des réalisations. Ce résultat semble donc montrer que les prévisionnistes peuvent à la fois surestimer les débits tout en ayant une estimation correcte de leurs incertitudes.

[80] On montre en Annexe 4 comment le biais d'ancrage sur une même largeur d'intervalle pour les prévisions à J et J+1 contribue à ce phénomène de hard-easy effect.

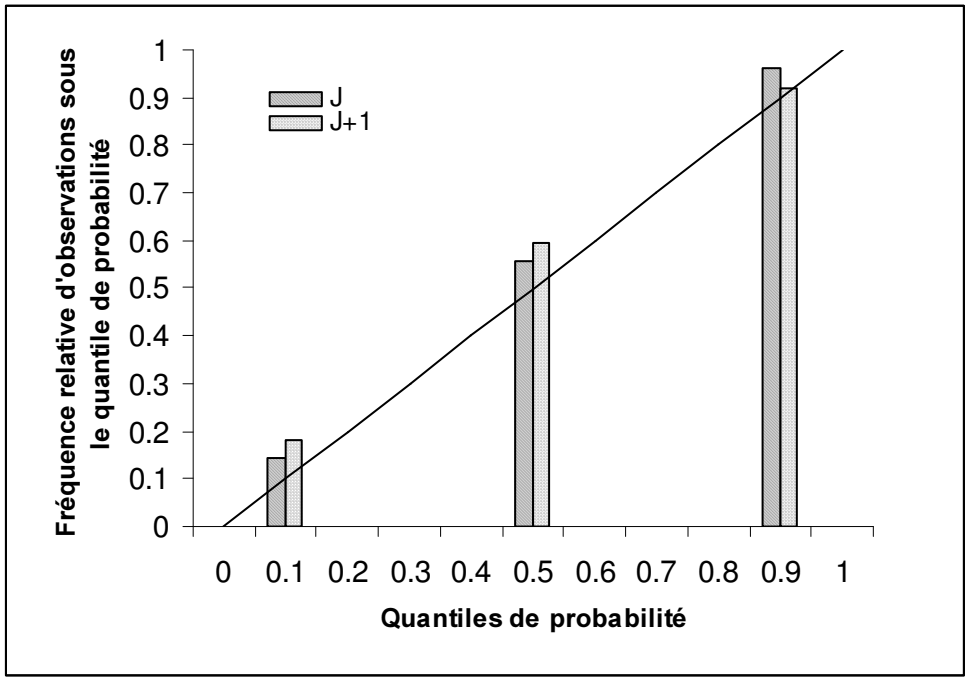



**Figure 26 : Diagramme de fiabilité pour les quantiles 10%, 50% et 90% fournis par les prévisionnistes d'EDF. Débits prévus aux échéances J et J+1.**

On peut donc conclure qu'aussi bien pour les prévisions de lame d'eau que pour celles des débits, les prévisionnistes ont une estimation fiable de leurs incertitudes[81]. Ce fait n'empêche pas que les prévisions soient entachées d'un biais quantitatif, négatif dans le cas des prévisions de pluie, positif dans le cas des prévisions de débit.

## 7.5 Discussion : Quand le prévisionniste se met à la place du décideur…

Si les prévisionnistes avaient vécu l'expérience au fil des jours comme un exercice académique, leur jugement aurait progressé lentement et de manière asymptotique vers une absence de surconfiance. Selon Fischhoff (1982), débiaiser les gens est en effet une tâche lente et difficile. Nous avons vu qu'au contraire les prévisionnistes sont brutalement passés d'une surconfiance marquée à une estimation correcte de leurs incertitudes.

Nous avançons l'idée qu'un tel changement a été motivé par l'adaptation à une nouvelle norme sociale. Le critère de performance à l'aune duquel les prévisionnistes avaient l'habitude d'être jugés (et de se juger eux-même) était la précision de leur prévision. Ce critère de performance se mesurait par la distance entre les prévus et les observés.

Avant de commencer cette nouvelle expérience au fil de l'eau, nous avons défini ensemble un nouveau critère qui nous semblait plus adapté à la prévision des phénomènes à risque pour l'usager, tels que les crues ou les pluies intenses. Nous avons supposé un seuil d'alerte (20mm en 24h) pour lequel l'usager demande à être prévenu 24h à l'avance. Il s'agit alors pour le

[81] Cette fiabilité globale peut néanmoins masquer des marques de surconfiance plus ponctuelles, par exemple lorsque le prévisionniste affiche un niveau d'incertitude très faible (Annexe 5).

prévisionniste d'avoir un taux d'événement manqué (TEM) aussi bas que possible, tout en conservant un taux de fausses alertes (TFA)[82] dans des limites raisonnables[83].

En choisissant ce nouveau critère, le couple (TFA, TEM), nous avons donc internalisé les enjeux de l'usager ayant la responsabilité de la sûreté d'installations hydroélectriques. Il ne s'agissait plus désormais de fournir une prévision la plus précise possible, mais de permettre à l'usager d'anticiper au mieux l'occurrence de phénomènes à risque.

Les résultats (Tableau 25) montrent que sur les 20 franchissements de seuil observés, seulement 6 avaient été prévus par la prévision déterministe, alors que la quantile 90% amène à en détecter 11 (Tableau 26). Le prix à payer est que le nombre de fausses alertes passe de 7 pour la prévision déterministe à 51 avec le quantile 90%.

| **Quantile 50%,** (échéance J+1 seuil 20 mm/24h) | **Observés** | **Non observés** |
|---|---|---|
| **Prévus** | H=6 | FA=7 |
| **Non Prévus** | M=14 | CR=432 |

**Tableau 25 : Tables de contingence pour le quantiles 50%.**
**Anticipation du franchissement du seuil de 20 mm en 24h sur la lame d'eau spatiale à J+1.**
**H: nombre d'événements correctement prévus. FA: nombre de fausses alarmes.**
**M: nombre d'événements manqués. CR: nombre de rejets corrects.**

| **Quantile 90%,** (échéance J+1 seuil 20 mm/24h) | **Observés** | **Non observés** |
|---|---|---|
| **Prévus** | 11 | 51 |
| **Non Prévus** | 9 | 388 |

**Tableau 26 : Tables de contingence pour le quantile 90%.**
**Anticipation du franchissement du seuil de 20 mm en 24h sur la lame d'eau spatiale à J+1.**

L'objectif que s'étaient fixé les prévisionnistes d'améliorer le système de prévision en terme de détection des phénomènes à risque a donc été partiellement atteint. Ils ont en effet très largement amélioré leurs qualités de détectabilité, même si le taux d'événements manqués reste encore supérieur à 50% en utilisant le quantile 90% (Tableau 27).

[82] On rappelle que le taux de fausses alertes (TFA) est la proportion du nombre d'alertes données à tort sur le nombre total d'alertes communiquées. Le taux d'événements manqués (TEM) est la proportion du nombre d'événements observés qui n'avaient pas été anticipés sur le nombre total d'événements observés. Cf. Note n°56.

[83] Pour fixer ce niveau de taux de fausses alertes raisonnable, seule une révélation des préférences du décideur est possible. Comme en la matière, le dialogue avec l'usager est pour le moins délicat (il ne veut bien souvent entendre parler ni de fausses alertes, ni d'événements manqués), on se fixe un cas d'école qui nous semble avoir les bons ordres de grandeur : un taux de fausses alertes inférieur à 80% pour une prévision de pluie à J+1 nous semble raisonnable.

| Echéance J+1 Seuil 20mm/24h | Quantile 50% | Quantile 90% |
|---|---|---|
| **TEM** | 70% | 45% |
| **TFA** | 54% | 82% |

**Tableau 27 : Taux d'événements manqués et Taux de fausses alertes pour les quantiles 50% et 90%. Anticipation du franchissement du seuil de 20mm en 24h sur la lame d'eau spatiale à J+1.**

Ces résultats montrent donc l'extrême difficulté de prévoir plus de 24h à l'avance l'occurrence de pluies intenses avec les outils actuels, en particulier sur des bassins versants de moyenne montagne[84]. Atger (2001) montre que même avec une interprétation ingénieuse des modèles déterministes ou ensemblistes[85], un taux d'événement manqué faible ne peut être obtenu qu'au prix d'un taux de fausses alertes élevé.

Au delà de ce constat, ces résultats suggèrent que lorsqu'on place les prévisionnistes dans le domaine décisionnel, leur comportement évolue vers une meilleure fiabilité de l'estimation de leurs incertitudes. Dans le même temps, leur faculté de détectabilité des pluies intenses s'améliore : le TEM réduit fortement, alors que le TFA s'accroît de manière raisonnable (et reste sous la barre des 80%).

Ces résultats sont donc compatibles avec l'hypothèse d'un changement de norme sociale à l'origine de cette disparition de surconfiance. Nos conclusions font donc la part belle aux Panglossiens (7.1). Il semble en effet que les prévisionnistes fournissaient des prévisions surconfiantes en réponse à une norme sociale qui les poussait à agir ainsi, les jugeant et les récompensant sur un critère de précision. Lorsque cette norme sociale est abandonnée pour une autre (comme ici, la détection des phénomènes à risque), le phénomène de surconfiance disparaît.

Le syndrome de surestimation systématique révélé par Murphy et Dann (1984) n'a touché dans nos expérimentations que les prévisions de débit. Ce résultat peut suggérer que pour ces prévisions, qui sont des "produits finis" intéressant directement l'usager, la norme sociale influe plus fortement que pour les prévisions de pluie, qui sont avant tout une variable de contrôle interne au système de prévision. La nouvelle norme sociale a donc pu à la fois mettre fin au phénomène de surconfiance (par prise de conscience des incertitudes), et induire un biais de surestimation dans les prévisions de débit (par excès de prudence).

## 7.6 Résumé, conclusions et perspectives

Les prévisionnistes opérationnels peuvent être significativement surconfiants (7.2). Un simple stage d'une journée de sensibilisation et de formation à la quantification des incertitudes peut

[84] Les bassins versants retenus étaient situés dans les massifs montagneux des Pyrénées, du Massif Central, et des Cévennes. Figure 22, p 70.

[85] La règle de décision la plus simple et la plus courante est celle pour laquelle l'alerte est donnée si le modèle déterministe prévoit le franchissement du seuil qui intéresse l'usager. Atger (2001) manipule cette règle de décision à l'aide d'autres leviers (utilisés de manière séparée ou conjointe) : alerter lorsque les sorties du modèle dépassent un certain niveau (niveau que l'on fait varier autour du seuil qui intéresse l'usager), alerter si le modèle prévoit des précipitations intenses à une certaine distance du point considéré (et faire varier cette distance). Le modèle ensembliste permet l'utilisation d'un levier de plus : faire varier le seuil de probabilité sur lequel on décide ou non d'alerter.

faire disparaître cette surconfiance (7.3). Pour faire perdurer une telle amélioration de la fiabilité des prévisions, les Mélioristes auraient préconisé la poursuite d'exercices académiques visant à corriger les défauts cognitifs des prévisionnistes. Nous avons pour notre part expérimenté le point de vue panglossien, considérant que la surconfiance était le résultat d'une adaptation des prévisionnistes à leur environnement social, et à la norme récompensant la précision. Des discussions avec les prévisionnistes à propos de voies d'amélioration du système de prévision pour l'usager confronté à des enjeux de sûreté nous ont conduit à adopter un nouveau critère, mettant en avant les qualités de détectabilité des phénomènes météorologiques à risque (7.4). L'adoption de ce nouveau critère, et de la nouvelle norme sociale qu'il a produit, semble avoir à la fois amélioré leur performance dans la détection de ces phénomènes, et fait disparaître le phénomène de surconfiance (7.5). Cette nouvelle norme a cependant pu induire un biais de surestimation dans les prévisions (7.5).

Ces résultats nous permettent de faire quelques suggestions pour une implémentation opérationnelle d'un système de prévision probabiliste expertisée qui soit épargné par le phénomène de surconfiance :

- Montrer aux prévisionnistes l'intérêt de produire des prévisions sous cette nouvelle forme. Gagner leur adhésion à ce nouveau moyen de communication des incertitudes est le préliminaire indispensable à toute expérimentation semi-opérationnelle de prévision probabiliste.
- Sensibiliser les prévisionnistes au phénomène de surconfiance. On pourra leur montrer les conditions dans lesquelles celui-ci apparaît, et les différentes interprétations qu'on peut avoir de ce phénomène.
- Choisir conjointement avec les prévisionnistes des variables et les sites présentant de forts enjeux pour l'usager. Cela permettra d'une part de réduire le biais académique de l'exercice[86]. D'autre part, le temps passé à sonder leurs incertitudes sur ces prévisions et l'élicitation des probabilités associées leur permettra de disposer d'une meilleure assise pour répondre aux éventuelles sollicitations téléphoniques des usagers, désirant obtenir des précisions sur le risque attendu[87].
- Développer des méthodes interactives et intuitives d'élicitation des probabilités subjectives. Pour éviter le caractère fastidieux de l'exercice, on tâchera de les interroger sur quelques variables intégrées dans l'espace et dans le temps (par exemple une lame d'eau sur un bassin versant, ou le débit à son exutoire) plutôt que de nombreuses cartes où ils devraient tracer les contours d'isoprobabilité de dépassement de seuil, pour plusieurs probabilités, et plusieurs seuils, à des pas de temps rapprochés.
- Faire des feed-back réguliers basés sur la calibration des prévisions émises. Ces feed-back doivent servir de support de discussion avec les prévisionnistes et entre les prévisionnistes pour mieux comprendre leur perception de leurs incertitudes suivant la situation en cours, les conditions particulières de prévision (indisponibilité de certains outils de prévision ou d'observation, défaillance technique, etc.).

[86] Si les prévisionnistes considèrent l'élicitation des probabilités subjectives comme une sorte d'examen, ou d'exercice déconnecté de leurs enjeux opérationnels, ils pourront montrer de l'inconsistance dans leur réponses.

[87] V. Fortin (comm.pers.) propose de vérifier la calibration des prévisionnistes sur également d'autres sites témoins pour s'assurer qu'une amélioration de la performance sur les sites choisis ne s'accompagne pas d'une détérioration de la performance sur certains autres sites. Une telle détérioration pourrait se produire soit parce que les prévisionnistes ne consacrent plus assez de temps à la prévision sur ces sites moins importants, soit parce qu'ils transposent des heuristiques adaptées aux sites témoins, mais inadéquates ailleurs.

- Montrer aux prévisionnistes la valeur ajoutée de ce nouveau type de communication des incertitudes. Ce n'est qu'en comprenant l'intérêt d'une telle expérimentation semi-opérationnelle que les prévisionnistes pourront échapper au biais académique inhérent à toute expérimentation en sciences humaines (9.1.7).

Comme nous l'avons suggéré au Chapitre 3, la quantification des incertitudes pourrait permettre d'améliorer le dialogue entre prévisionnistes et usagers, particulièrement lorsqu'ils parlent de risque. Ainsi, en toute transparence, ceux-ci pourraient communiquer sur un ou plusieurs quantiles élevés à propos d'une prévision. D'une part, cela éviterait au prévisionniste d'avoir besoin de surestimer ses prévisions pour "se couvrir". D'autre part, cela placerait l'usager devant les limites du système de prévision. L'usager serait alors invité à arbitrer sur le couple (TFA, TEM) qui lui est optimal, parmi tous les couples possibles pour une performance donnée du système de prévision.

Cependant, comme le souligne Atger (2001), un tel changement nécessite d'éduquer l'usager à cette nouvelle forme d'expression des incertitudes, et au maniement de ces nouveaux concepts (probabilité, vraisemblance, etc.). Il semble que nous n'ayons pas encore atteint ce « meilleur des mondes possibles » (Voltaire, 1759). Il est encore difficile, à l'heure actuelle, de demander à l'usager de se prononcer sur le couple (TFA, TEM).

## 7.7 Bibliographie

Atger, F., 2001. Verification of intense precipitation forecasts from single models and ensemble prediction systems. Non Linear Processes in Geophysics, 8, 401-417.

Atger, F., 2000, La prévision du temps à moyenne échéance en France, La Météorologie, 8$^{e}$ série, n°30.

Atger, F., 1999a. Tubing : an alternative to clustering for the classification of ensemble forecasts. Weather and forecasting, vol. 14.

Atger, F., 1999b. The skill of Ensemble Prediction Systems. Monthly Weather Review, vol. 127.

Biais, B., Hilton, D., Mazurier, K., Pouget, S., 2002. Psychological Dispositions and Trading Behaviour. Les cahiers recherche du Centre de Recherche en Gestion de Toulouse, vol. 145, (http://www.univ-tlse1.fr/iae/310-crgcahiers.php).

Cooke, W. E., 1906. Forecasts and verifications in western Australia. Monthly Weather Review, vol. 34, 23-24.

Cosmides, L., Tooby, J., 1996. Are humans good intuitive after all? Rethinking some conclusions from the literature on judgment under uncertainty. Cognition, 58, 1-73.

Doswell, C. A., Flueck, J. A., 1989. Forecasting and verifying in a filed research project: DOPLIGHT '87. Weather and Forecasting, 4, 97-109.

Fischhoff, B., 1982, Debiasing, in Kahneman, D., Slovic, P., Tversky, A., 1982. Judgment under uncertainty: Heuristics and biases. Cambridge University Press, 422-444.

Hilton, D., 1995. The social context of reasoning: conversational inference and rational judgment. Psychological Bulletin, vol. 118, 2, pp 248-271.

Juslin, P., Winman, A., Olsson, H., 2000. Naive empiricism and dogmatism in confidence research: a critical examination of the hard-easy effect, Psychological Review, vol. 107, no 2, 384-396

Kahneman, D., Slovic, P., Tversky, A., 1982. Judgment under uncertainty: Heuristics and biases. Cambridge University Press.

Krzysztofowicz, R., 2002. Bayesian system for probabilistic river stage forecasting, Journal of hydrology, vol. 268.

Krzysztofowicz, R., 1983. Why should a forecaster and a decision maker use Bayes theorem, Water Ressources Research, vol.19, 2, 327-336.

Krzysztofowicz, R., Sigrest, A., 1999. Calibration of probabilistic quantitative precipitation forecasts. Weather and Forecasting, 14(3), 427-442.

Krzysztofowicz, R., Sigrest, A., 1997. Local climatic guidance for probabilistic quantitative precipitation forecasting. Monthly Weather Review, 125, 3, 305-316.

Krzysztofowicz, R., Drzal, W., J., Drake, T., R., Weyman, J., C., Giordano, L., A., 1993. Probabilistic quantitative precipitation forecasts for river basins. Weather and Forecasting, vol. 8.

Lichtenstein, S., Fischhoff, B., 1980, Training for calibration. Organizational Behavior and Human Performance, vol. 26, 149-171.

Lichtenstein, S., Fischhoff, B., 1977. Do those who know more also know more about how much they know? Organizational behavior and human decision process, vol. 39, 98-114.

Mason, S., J., 1982. A model for assessment of weather forecasts, Australian Meteorological Magazine, vol. 30, 291-303.

Murphy, A. H., 1993. What is a good forecast ? An essay on the nature of goodness in weather forecasting. Weather and forecasting, vol. 8, no 2, 281-293.

Murphy, A. H., 1991. Probabilities, odds, and forecast of rare events, Weather and Forecasting, vol. 6, no 2, 302-307.

Murphy, A. H., 1973. A new vector partition of the probability score. Journal of Applied Meteororology, 12, 595-600

Murphy, A. H., Daan, H., 1984. Impacts of feedback and experience on the quality of subjective probability forecast : Comparison of results from the first and second years of the Zierikzee Experiment, Monthly Weather Review, vol. 112, 413-423.

Murphy, A., H., Winkler, R., L., 1974. Credible interval temperature forecasting: some experimental results. Monthly Weather Review, vol. 102, 784-794.

Olson, D. A., Junker, N.W., Korty, B., 1995. Evaluation of 33 years of quantitative precipitation forecasting at the National Meteorological Center, Weather and Forecasting, vol. 10.

Popper, K., L'Univers irrésolu, plaidoyer pour l'indéterminisme, (tr. fr. R. Bouveresse) Hermann, Paris, 1984.

Sanders, F.,1963. On subjective probability forecasting, Journal of Applied Meteorology, vol. 2, 2, 191-201.

Savage, L. J., 1954, The Foundations of statistics. Wiley.

Savage , L., J., 1971. Elicitation of personal probabilities and expectations, Journal of the American Statistical Association, Vol. 66, 336, 783-801.

Shanteau, J., Weiss, D. J., Thomas, R. P., Pounds, J. C., 2002. Performance-based assessment of expertise: How to decide if someone is an expert or not, European Journal of Operational Research, vol. 136, no. 2, 253-263.

Soll, J., B., 1996. Determinants of overconfidence and miscalibration: the roles of random error and ecological structure, Organizational behavior and human decision processes, vol. 65, no. 2, 117-137.

Stanovich, K. E., 1999. Who is rational? Studies of individual differences in reasoning. Mahweh, NJ: Erlbaum.

Suantak, L., Bolger, F., Ferrell, W. R., 1996. The Hard-Easy Effect in Subjective Probability Calibration, Organizational Behavior and Human Decision Processes, vol. 67,no 2, 201-221.

Voltaire, 1759 : Candide, ou l'optimisme. Traduit de l'allemand de M. le docteur Ralph , avec les Additions qu'on a trouvées dans la poche du Docteur, lorsqu'il mourut à Minden, l'An de Grâce 1759. Réédition 1975, Hatier, Paris.

# 8 La relation entre surconfiance et biais : résultats expérimentaux

*Mais au fond, les gens sous-estiment-ils vraiment leurs incertitudes ?*

*La surconfiance ne serait-elle pas plutôt le reflet d'un autre biais de jugement ?*

*Un biais quantitatif, qui traduirait les idées fausses qui trottent dans la tête de tous les gens d'une même communauté, comme celle des prévisionnistes.*

*Un biais quantitatif, qui serait produit par un manque de retour d'expérience.*

## 8.1 Introduction

Les prévisionnistes en hydrométéorologie peuvent être surconfiants (Chapitre 7). Nous avons également évoqué que ce phénomène de sous-estimation des incertitudes est très général, et concerne une très large catégorie de personnes et d'activités. Les personnes exerçant une activité d'expert semblent être tout particulièrement enclines à ce phénomène. Les médecins cliniciens par exemple, ont révélé au travers de nombreuses études leur tendance à afficher des incertitudes bien trop faibles en regard de leur niveau de connaissance. (Friedman et al., 2001, et Fischhoff, 1982, pour une revue de la littérature sur la question).

Nous avons également présenté (7) la dualité entre les Panglossiens et les Mélioristes (Stanovich, 1999) pour interpréter ce phénomène. Pour les premiers, la surconfiance permet de prendre efficacement des décisions, avec un minimum de temps et d'énergie. Ce phénomène peut également résulter d'une adaptation à une norme sociale ou professionnelle privilégiant la précision d'une information au détriment de sa fiabilité (7.5), ou l'aplomb avec lequel on délivre cette information au détriment de la modestie (Fischhoff, 1982). Blanton et al. (2001) suggèrent également que la surconfiance puisse être motivée par la volonté de réduire les dissonances cognitives (Festinger, 1957)[88]. Kramer et al. (1993) montrent que des facteurs affectifs, tels que le désir de voir son ego savant et compétent, peuvent contribuer au phénomène de surconfiance.

Les Mélioristes quant à eux considèrent la surconfiance comme une déficience cognitive. Cette approche a souvent reçu l'appellation « heuristic and biases », expression restée célèbre suite à la publication de "Judgment under uncertainty : Heuristics and biases" (Kahneman et al., 1982). Ainsi apprend-on en introduction que les sujets utilisent "des heuristiques en nombre limité, [qui,] bien qu'utiles en général, mènent parfois à des erreurs graves et systématiques" (Kahneman et al., 1982, p 3). Du fait de la finitude des capacités d'assimilation et de traitement de l'information (empan cognitif), l'information disponible est mal utilisée. Les sujets sont en somme considérés comme des "processeurs d'informations à capacité limitée" (Tetlock, 1991). En particulier, les sujets ne portent pas assez d'attention à la qualité ou à l'importance des indications qui leur sont données (Griffin et Tversky, 1992). Ils pondèrent mal les facteurs, estimant parfois la probabilité d'une concomitance d'événements supérieure à la probabilité d'occurrence de chaque événement pris séparément (Palmarini, 1995)[89]. Les Mélioristes ont également noté une certaine "résistance" à l'information. Les informations complémentaires ne sont utilisées que si elles renforcent la position acquise (Koriat et al., 1980). L'accroissement d'information n'améliore donc pas la fiabilité de l'expert mais augmente sa confiance en son propre jugement.

---

[88] La théorie de la dissonance cognitive (Festinger, 1957) explique que, dans l'inconfort dû à une incompatibilité entre certaines croyances et actions d'un sujet, celui-ci tente de réduire cette "dissonance cognitive". Le sujet utilise alors l'un des 3 moyens suivants : minimiser l'importance de la croyance en conflit, adopter une nouvelle croyance qui rétablit l'équilibre, ou cesser le comportement contradictoire. Selon Blanton et al. (2001), la surconfiance serait motivée chez un sujet par la réduction de la dissonance entre son amour-propre et la perception de son niveau d'ignorance, ou entre le sentiment de contrôle de la situation et sa perception de l'incertitude.

[89] Palmarini (1995) donne l'exemple de sujets confrontés à la description suivante (p 88) : « Linda, 31 ans, est célibataire, extravertie et extrêmement brillante. Elle a une maîtrise de philosophie. Lorsqu'elle était étudiante, elle était très intéressée par les problèmes de discrimination raciale et d'injustice sociale et participait aux manifestations anti-nucléaires. » A partir de cette description, les sujets ont pour la majorité fourni une probabilité plus faibles à chacune des assertions « Linda est militante féministe » (cas A) et « Linda est employée de banque » (cas B) qu'à l'assertion « Linda est employée de banque et militante féministe active » (cas A et B).

Nous avons émis l'idée (7.4.3), sans toutefois nous y être trop attardés, que le biais quantitatif sur une variable (une surestimation systématique par exemple) peut être déconnecté de la surconfiance associée à cette variable. On a ainsi relevé qu'un prévisionniste pouvait être très fiable dans l'estimation de ses incertitudes tout en produisant des prévisions quantitativement biaisées.

Dans cette partie 8, nous investiguons plus en profondeur les relations qui peuvent exister entre biais et estimation des incertitudes.

Nous utilisons les mêmes données qu'en 7, c'est à dire les prévisions d'Hydro-Québec, les résultats de l'expérience de culture générale, et ceux des deux expériences en laboratoire avec les Centres Hydrométéorologiques.

## 8.2 Surconfiance et biais dans les réponses à un questionnaire de culture générale : pièges dans les questions ou biais cognitif ?

### 8.2.1 Contexte

Nous voulions initialement proposer le test de culture générale, inspiré par Biais et al. (2002), dans un but pédagogique : nous voulions montrer aux personnes interrogées le phénomène de surconfiance. Pour cela, quoi de plus convaincant que de l'illustrer sur leur propre comportement ?

L'expérience a été réalisée avec quatre populations, pour un total de 96 personnes :

- les participants au colloque de Rochebrune 2002 : 19 personnes
- des élèves en première année de l'Engref : 26 personnes
- les participants aux journées des doctorants de l'Ecole Doctorale Géoscience et Ressources Naturelles : 39 personnes
- les prévisionnistes en hydrologie d'Hydro-Québec (5 personnes), et les gestionnaires de réservoirs d'Environnement Québec (7 personnes).

Notre objectif était donc de sensibiliser les personnes rencontrées (notamment les statisticiens Bayésiens du colloque de statistique décisionnelle de Rochebrune 2002[90]) à la difficulté d'avoir une bonne connaissance de l'étendue de leurs connaissances.

Dans le même temps, nous nous posions également certaines questions :

- Comment se comporte un expert lorsqu'il est confronté aux limites de son savoir ?
- L'expert connaît-il mieux que l'homme de la rue les limites de son savoir ?
- Quelle perception ont les sujets de leur corpus de connaissance ?

S'agissant tout particulièrement de statisticiens Bayésiens, nous nous demandions quelle était leur perception de l'incertitude et leur usage des probabilités pour la quantifier.

De manière plus générale, nous cherchions à comprendre dans quelle mesure les probabilités subjectives pouvaient correctement traduire les connaissances imparfaites d'une personne.

Le questionnaire était formulé de la sorte :

[90] Les actes du colloque de Rochebrune 2002 (comité scientifique E. Parent, J. Bernier, J.J. Boreux) sont disponibles sur le site : http://www.engref.fr/grese/rochebrune.html

*Pour les 10 questions ci-dessous, vous devez indiquer les bornes inférieure et supérieure de l'intervalle dans lequel vous estimez que la réponse a 80% de chance de se trouver. Vous n'avez pas besoin de connaître la bonne réponse, mais seulement d'évaluer le niveau de votre ignorance. Si votre évaluation est correcte, 8 de vos intervalles contiendront effectivement la bonne réponse*[91].

*1) Age de Martin Luther King à sa mort*
*2) Longueur du Nil (km)*
*3) Nombre de pays membres de l'OPEP*
*4) Nombre de livres dans le nouveau testament*
*5) Nombre d'habitants au Texas (millions)*
*6) Année de naissance de J.-S. Bach*
*7) Nombre de juges à la cour internationale de La Haye*
*8) Diamètre de la Lune (km)*
*9) Hauteur des chutes du Niagara (m)*
*10) Profondeur maximale des océans (m)*

(Réponses : 39, 6671, 13, 27, 17, 1685, 15, 3476, 51, 11520)

Seuls les élèves de l'Engref ont eu un questionnaire légèrement différent : à la place de l'année de naissance de Bach, on leur a demandé l'année de naissance de Diderot. On traitera donc les réponses obtenues comme des réponses à un questionnaire à 11 questions, sachant que chaque population n'a répondu qu'à dix d'entre-elles.

### 8.2.2 On observe de la surconfiance dans les réponses au questionnaire

On observe (Figure 27) que quelle que soit la population, le nombre moyen (par individu) d'intervalles comprenant la bonne réponse est bien en deçà de 8.

[91] Si nous avions été plus rigoureux, nous aurions rajouté « à la dispersion d'échantillonnage près ». L'objectif de ce préambule était cependant de fournir à la personne interrogée, en plus d'une représentation probabiliste, une représentation ensembliste du problème posé. Ce 2e type de représentation est en général plus intuitif (Cosmides et Tooby,, 1996).

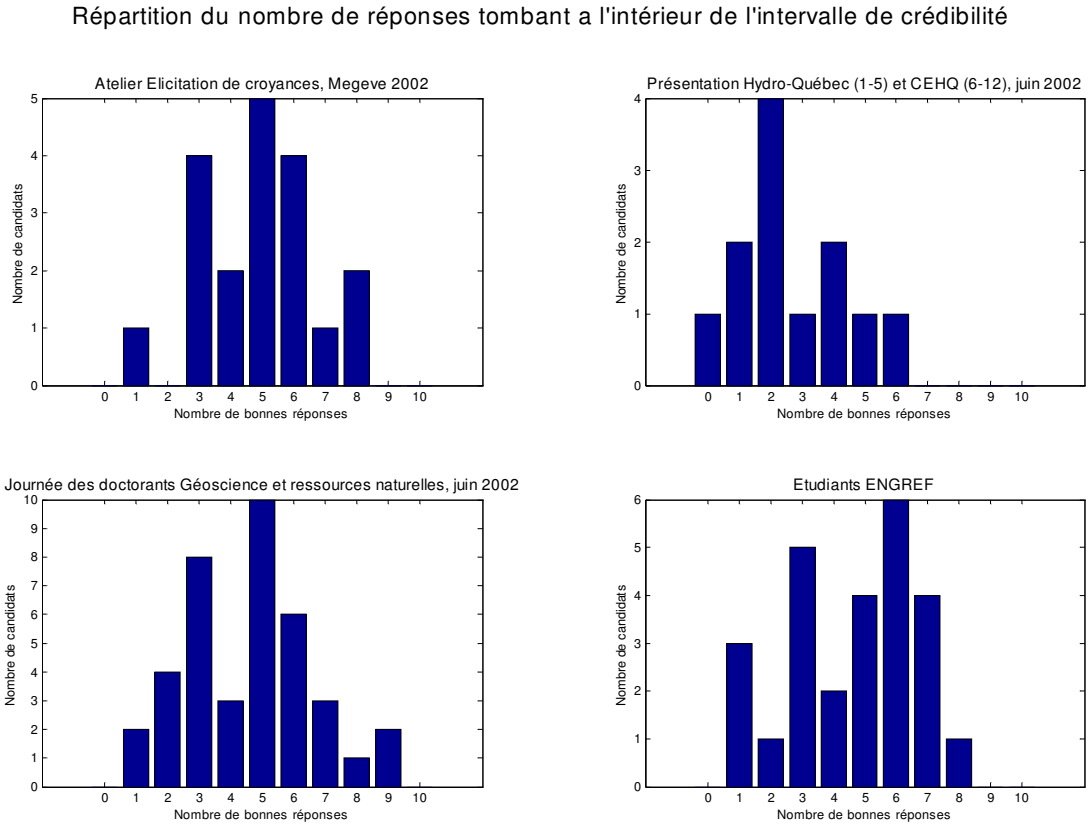



**Figure 27 : Résultat du test de culture générale pour chacune des quatre populations interrogées.**

Comme en 7.2.2, nous faisons l'hypothèse suivante :

> **La probabilité que l'intervalle fourni contienne la bonne réponse est la même pour chaque question, et pour chaque individu.**
>
> **On la note θ.**

Dans le cadre de cette expérience, cela signifie que :

- les questions portent sur des domaines suffisamment **variés** pour qu'une erreur à l'une n'entraîne pas nécessairement une erreur à une autre. On peut donc les considérer **indépendantes** les unes des autres.
- les participants sont suffisamment **honnêtes** pour qu'une erreur chez l'un n'entraîne pas nécessairement une erreur chez son voisin[92].

[92] Gigerenzer et al. (1991) et Juslin et al. (1994) notent que l'indépendance des réponses de plusieurs individus est loin de se limiter à la question de l'honnêteté. Leurs modèles interprétatifs, dits "modèles écologiques" reposent sur l'hypothèse que les individus partagent les mêmes heuristiques (mêmes clefs de comparaison, stratégies similaires de recherche d'informations pertinentes, etc.), et une culture commune. Ces auteurs réfutent donc l'hypothèse selon laquelle N individus répondant à la même question de culture générale peuvent être considérés comme autant de réalisations indépendantes de la même expérience. Nous poserons cependant cette hypothèse d'indépendance, et analyserons si les données recueillies nous permettent de confirmer ou d'infirmer cette hypothèse.

- les populations (et les questions) sont suffisamment **homogènes**, de sorte qu'on puisse considérer que pour chaque question, et pour l'un quelconque des participants dans une population donnée, la probabilité de succès est la même : $\theta$.
- les participants « acceptent de **jouer le jeu** », et répondent honnêtement, pour chaque question, par un intervalle sur lequel ils seraient prêts à parier à 4 contre 1. On exclut donc en particulier les participants adoptant la stratégie suivante : mettre un intervalle $[0 ; +\infty[$ à 8 questions, et saborder les deux dernières en mettant par exemple des intervalles inclus dans $]-\infty ; 0]$.
  *On suppose donc la bonne foi de l'expérimentateur, qui n'essaie pas de piéger le candidat par des questions à tiroir, et la bonne volonté du candidat, qui n'essaie pas de contourner les règles du jeu pour obtenir le résultat attendu.*

Avec ces hypothèses, et selon la méthode utilisée en 7 (décrite en Annexe 1), les réponses fournies au questionnaire par la population globale nous permettent d'obtenir un intervalle de crédibilité à 95% sur $\theta$ en plus du simple estimateur $\overline{\theta}$ :

$\overline{\theta} = 0{,}45$, et $IC_{95\%}[\theta|n,k] = [0{,}42 ; 0{,}48]$

*avec* *n=939 intervalles de crédibilité à 80% fournis*

*et* *k=424 bonnes réponses comprises dans ces intervalles de crédibilité*

0,8 n'est donc pas inclus dans l'intervalle de confiance à 95% de la distribution a posteriori de $\theta$, compte tenu des résultats obtenus pour la population totale[93].

On en conclut donc que les individus sont surconfiants ; c'est à dire qu'ils ont une large tendance à sous-estimer leurs incertitudes.

### 8.2.3 A-t-on choisi des questions-pièges ?

L'expérience proposée a donc parfaitement atteint son « but pédagogique ». Elle nous a permis de montrer le phénomène de surconfiance à et sur chaque population interrogée. Notre conclusion a donc toujours recueilli l'approbation de tous : manifestement, les sujets sont (très) surconfiants.

De tous ? Pas tout à fait… sitôt donnés les résultats du questionnaire au séminaire de Rochebrune 2002, un participant nous a accusés d'avoir incidemment fourni des « questions biaisées », autrement dit, des questions pour lesquelles la majorité des sujets a des idées fausses. Les réponses à de telles questions seraient donc biaisées, c'est à dire que tout le monde aurait tendance à se tromper dans le même sens (sous-estimation ou surestimation). Afin d'éviter toute équivoque avec les biais cognitifs que peuvent avoir les sujets, nous nommerons par la suite de telles questions des questions-piège.

Remarquons tout de suite que le questionnaire a pu comporter des questions-pièges sans que son concepteur ait eu les mauvaises intentions dont l'a accablé ce participant. Juslin (1994) a en effet réalisé l'expérience suivante : il a demandé à des sujets de choisir de manière informelle des questions qui permettront par la suite de mesurer les connaissances des sujets.

---

[93] Pour mémoire, si l'on ne tient compte que de la population française (on verra par la suite qu'elle a eu un comportement homogène face au questionnaire), on obtient les résultats suivants :, $\overline{\theta}$ =0,48 et $IC_{90\%}[\theta|n,k] = [0{,}44 ; 0{,}51]$

Ses résultats montrent que les sujets ont tous été amenés à sélectionner des questions-piège[94]. Ainsi, le simple choix de questions – choix que l'expérimentateur pouvait juger innocent - peut l'amener de fait à sélectionner des questions-piège.

Pour analyser si le questionnaire proposé était réellement piégeant, on commence par définir le biais dans les réponses fournies. Une question sera dite piège lorsqu'elle aura produit des réponses biaisées.

Pour une question donnée,

- Soit n1 le nombre d'individus qui surestiment la réponse,
  *C'est à dire que la bonne réponse est inférieure à la borne inférieure de leur intervalle de confiance.*
- Soit n2 le nombre d'individus qui sous-estiment la réponse
  *C'est à dire que la bonne réponse est supérieure à la borne supérieure de leur intervalle de confiance.*

Le **biais positif b+** est alors défini par n1/(n1+n2) ; le **biais négatif b-** par n2/(n1+n2)

Notons que le biais négatif est par construction le complémentaire à 1 du biais positif, et qu'une question n'est pas biaisée si b+ (ou b-) = 0,5

Le **biais absolu** b est le minimum du biais positif et négatif ; il est compris entre 0 et 0,5, et vaut 0,5 si les réponses ne sont pas biaisées

Une telle définition a le mérite d'être indépendante de l'ordre de grandeur des réponses, ou de l'ordre de grandeur des incertitudes sur la réponse, ou même de la difficulté des questions : elle ne prend pas en compte le nombre de personnes qui ont obtenu la bonne réponse dans leur intervalle de confiance.

On s'intéresse désormais au biais positif. On utilise à nouveau la méthode développée en Annexe 1. Nous affectons à b+ le prior (non informatif) de Jeffrey sur [0,1]. Après mise à jour avec les données recueillies, on obtient pour b+ un posterior dont la couverture à 95% est représentée Figure 28.

[94] Des discussions informelles ont permis de faire ressortir que les questions étaient sélectionnées par les sujets lorsque ceux-ci les jugeaient « intéressantes » (Juslin, 1994, p236). Le mot « intéressantes » peut suggérer une interprétation des résultats de cette expérience : peut-être que la conception du questionnaire a été considérée comme un jeu, où, comme dans tous les jeux, il y aurait des gagnants et des perdants. Dans ce cas-là, l' « intérêt » des concepteurs du questionnaire est de choisir des questions pour lesquelles les gens interrogés se tromperaient.

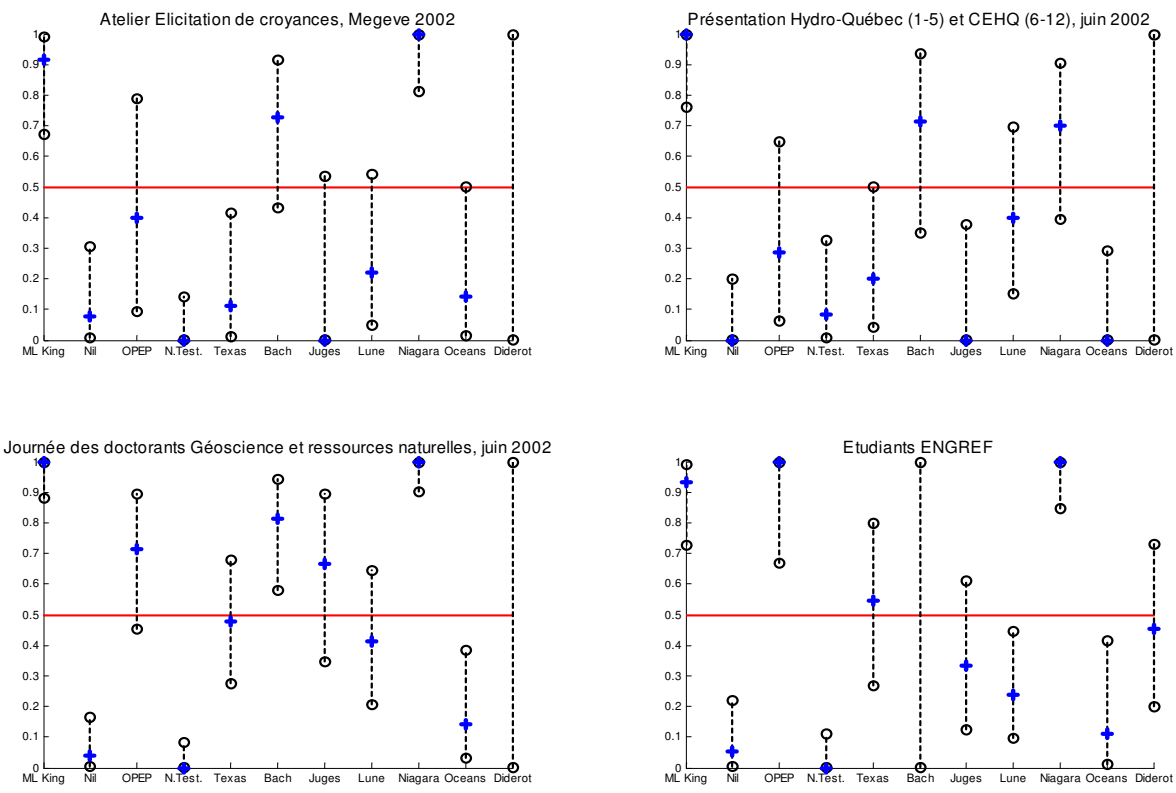



**Figure 28 : Biais dans les réponses au questionnaire de culture générale, pour les quatre populations. Estimation Bayésienne et des incertitudes associées (intervalle de confiance à 95% sur b+).**

Les trois populations françaises se comportent de manière très similaire aux questions qui leurs sont posées : les biais vont tous dans le même sens (positif ou négatif), et sont du même ordre de grandeur. Dans la suite, nous agrègerons donc ces trois populations pour étudier cette question du biais.

Ce test nous invite donc à rejeter pour chaque population au moins six questions sur dix. Les cinq questions qui ont été systématiquement biaisées pour les trois populations françaises sont les suivantes :

| | | |
|---|---|---|
| 1) Age de Martin Luther King à sa mort | *présente un biais* | + |
| 2) Longueur du Nil (km) | | - |
| 4) Nombre de livres dans le nouveau testament | | - |
| 9) Hauteur des chutes du Niagara (m) | | + |
| 10) Profondeur maximale des océans (m) | | - |

Remarque :

*Si l'on ajoute la question « 6) Année de naissance de J.-S. Bach », qui présente un biais positif, mais qui n'est significatif que pour les doctorants de Géoscience et Ressources Naturelles, on trouve que le test est bien balancé entre les questions induisant des biais positifs, et celles induisant des biais négatifs.*

*Cela explique qu'un tel questionnaire biaisé ait pu passer inaperçu : chacun se trompe tantôt dans un sens, tantôt dans un autre, et tant que collectivement, les sujets n'ont pas décidé de confronter les résultats, question par question, il leur est bien difficile de s'en rendre compte… sur les quatre populations interrogées, une seule personne (au colloque de Rochebrune) aura eu la perspicacité (voire l'audace) d'attribuer la source de la surconfiance aux "pièges" dans les questions.*

La Figure 29 présente les résultats obtenus pour la population française agrégée (83 membres). Ces résultats nous conduisent à rejeter huit questions, dont les réponses sont jugées biaisées avec une confiance de 95%.

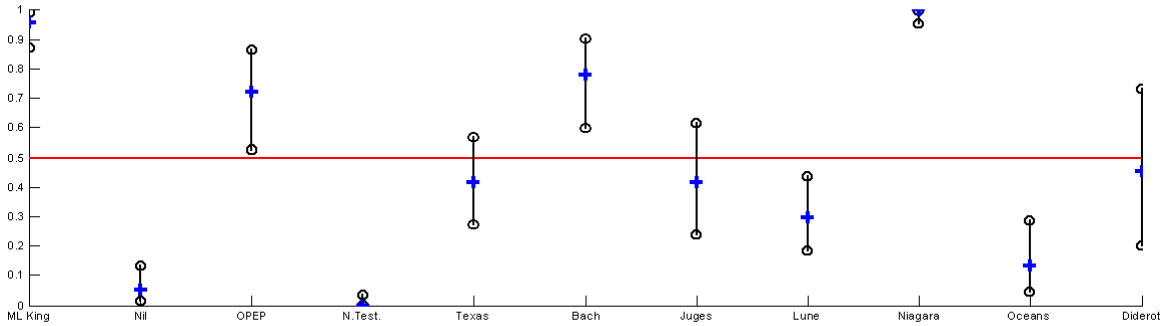



**Figure 29 : Biais dans les réponses au questionnaire de culture générale, pour la population française agrégée. Estimation Bayésienne et des incertitudes associées (intervalle de confiance à 95% sur b+).**

On en conclut donc avec une confiance de 95% que dans les 10 questions auxquelles chaque candidat de la population française était confronté se trouvaient 8 questions-piège.

### 8.2.4 Pièges dressés par l'expérimentateur ou surconfiance des sujets ?

Il est donc manifeste que le questionnaire proposé comportait une large majorité de questions pièges, qui ont su se servir d'idées fausses largement répandues pour induire les candidats en erreur.

Il n'en demeure pas moins que les sujets ont également pu montrer une large surconfiance, due à leur caractère, aux normes auxquelles ils sont soumis, ou à d'éventuelles déficiences cognitives. Comment faire la part entre la surconfiance obtenue en sélectionnant des questions-piège, et la part de la surconfiance qui découlerait du comportement intrinsèque des individus sous incertitude ? Pour répondre à cette question, on estime à nouveau θ en retirant les questions-piège.

Dans un premier temps, on choisit de rejeter les questions dont on est sûr à 99,9% qu'elles entraînent des réponses biaisées dans la population française. Avec un prior de Jeffrey pour le biais positif b+, une couverture à 99,9% du posterior de b+ obtenu avec les résultats de la population française nous conduit à rejeter 5 questions sur 10 :

1) Age de Martin Luther King à sa mort
2) Longueur du Nil (km)
4) Nombre de livres dans le nouveau testament
9) Hauteur des chutes du Niagara (m)
10) Profondeur maximale des océans (m)

On obtient alors l'intervalle de crédibilité à 95% sur θ et l'estimateur $\overline{\theta}$ :

$\overline{\theta} = 0{,}57$ et $IC_{95\%}\ [\theta|n,k] = [0{,}52\ ;\ 0{,}62]$

Avec toutes les questions, nous avions obtenu $\overline{\theta} = 0{,}48$ et $IC_{90\%}\ [\theta|n,k] = [0{,}44\ ;\ 0{,}51]$ (note n°93).

$\theta$ augmente donc significativement lorsqu'on retire les questions "très" piégeantes (exclues avec 99,9% de confiance). Conformément à l'intuition qu'on pouvait avoir, les questions-pièges ont bien accru l'effet de surconfiance chez les individus.

Cependant, 0,8 n'est donc toujours pas inclus dans l'intervalle de confiance à 95% avec le questionnaire privé des cinq questions pièges. Même en privant le questionnaire des 5 questions produisant les plus forts biais, les individus présentent encore une sous-estimation des incertitudes bien marquée.

Retirons à présent toutes les questions provoquant un biais dans les réponses, avec un seuil de rejet à 95%. Nous sommes ainsi amenés à ne conserver plus que les questions suivantes :

5) Nombre d'habitants au Texas (millions)
7) Nombre de juges à la cour internationale de La Haye
11) Année de naissance Diderot

On obtient alors les résultats suivants : $\overline{\theta}=0,60$ et $IC_{95\%}[\theta|n,k] = [0,53 ; 0,67]$

Si $\overline{\theta}$ augmente encore légèrement, $IC_{95\%}[\theta|n,k]$ n'inclut toujours pas 0,8 ; loin s'en faut. Ainsi, même en retirant toutes les questions-pièges, les individus montrent encore une sous-estimation marquée de leurs incertitudes.

Les résultats présentés en 8.2.3 ont montré une très large présence de questions-piège, entraînant les individus à sous-estimer leurs incertitudes. On montre ici que ces pièges n'expliquent pas à eux-seuls la surconfiance décelée dans les réponses au questionnaire de culture générale. Face à des questions neutres, les individus ont tout de même sous-estimé leurs incertitudes. Le tort est donc partagé entre l'expérimentateur qui a sélectionné des questions fallacieuses, et les sujets qui se sont montré surconfiants même pour des questions neutres.

### 8.2.5 Conclusions sur le test de culture générale

Les résultats énoncés ci-dessus permettent d'édicter les conclusions suivantes (sous les hypothèses rappelées en 8.2.2) :

- Lorsqu'on pose une question à quelqu'un, celui-ci a généralement tendance à surestimer ses connaissances et sous-estimer son ignorance sur cette question. Il croit mieux savoir qu'il ne sait effectivement.
- Cette sous estimation des incertitudes est d'autant plus importante que la question est piègeante, mais reste marquée même pour des questions "neutres" (c'est à dire des questions qui n'entraînent pas de réponses biaisées).

Tâchons à présent de répondre à l'objection de Juslin et al. (1994) et Juslin (1994) sur l'invalidité de l'hypothèse d'indépendance entre les réponses de plusieurs individus à une question de culture générale sélectionnée pour induire les gens en erreur. Nous avons montré que sous l'hypothèse d'indépendance, les réponses obtenues étaient manifestement surconfiantes, et que cette surconfiance était très bien corrélée avec le biais des questions.

D'autre part, nous avons pu affirmer avec une confiance de 95% que 8 questions sur 10 présentaient un biais, ce qui nous a conforté dans l'idée de Juslin (1994) que l'expérimentateur a incidemment sélectionné des questions qui faisaient appel à des heuristiques largement répandues n'induisant les gens en erreur que dans certains rares cas (ceux-là mêmes qui ont été choisis pour le test). Nous en concluons donc que le biais est une source de surconfiance,

qui peut être utilisée à dessein par l'expérimentateur pour, sinon produire, du moins exagérer le phénomène de surconfiance.

Si les réponses obtenues présentent donc un tel biais, c'est qu'elles ne peuvent être considérées comme indépendantes : si à l'extrême, 20 individus utilisent les mêmes heuristiques et disposent de la même culture, cela ressemble à "20 sujets prenant les paris sur un seul lancer d'une pièce de monnaie" : soit tout le monde gagne, soit tout le monde perd (Juslin, 1994).

En raisonnant par l'absurde[95] on peut donc conclure que les réponses obtenues au questionnaire de culture générale sont manifestement dépendantes, et que cette dépendance résulte d'une sélection de questions fallacieuses.

Cependant, nous pouvons objecter à Juslin (1994) d'une part que ces questions-pièges ne font qu'accentuer le phénomène de surconfiance, qui reste malgré tout présent pour des questions neutres (8.2.4).

D'autre part, nous émettons également une réserve quant à l'hypothèse faite par Juslin (1994) sur la supposée absence de biais concernant le jugement des individus dans des "conditions écologiques" (i.e. dans leur univers habituel, notamment professionnel). Peut-on en effet affirmer que les réponses aux questions de la vie de tous les jours, personnelle ou professionnelle, ne sont pas biaisées ?

C'est à cette question que les deux tests suivants vont nous permettre de répondre.

## 8.3 Détection d'un biais dans le jugement des prévisionnistes. Doit-on y voir la source de la surconfiance ?

*On reprend ici les résultats des expériences de prévision probabiliste à EDF (Chapitre 4) et à Hydro-Québec (7.2).*

### 8.3.1 Résultats à EDF

On calcule de manière analogue le biais mis en lumière par le test de culture générale.

On pose la même question à tous les prévisionnistes (quelle est la lame d'eau attendue sur un bassin versant donné pour les prochaines 24h ?), et on observe si toutes les erreurs des prévisionnistes vont dans le même sens. Pour une question donnée, on recense les prévisions pour lesquelles l'observé tombe à l'intérieur de l'intervalle de crédibilité fourni, ce qui nous permet de calculer l'estimateur et l'$IC_{95\%}$ sur $\theta$. On compte le nombre de surestimations et le nombre de sous-estimations, puis on calcule l'$IC_{95\%}$ sur b+ par la méthode exposée en 8.2. Si 0,5 n'appartient pas à cet intervalle de confiance sur b+, on considère que la question induit une réponse biaisée.

Lorsque l'on dit que la question induit une réponse biaisée, il s'agit bien du biais décelé en interrogeant une seule fois n personnes (de même qu'en 8.2). Ce biais révèle donc des idées

[95] Supposer l'indépendance des réponses aux questions nous a amenés à mettre en lumière leur biais, et donc leur dépendance.

fausses largement répandues dans la population interrogée. En ce sens, le biais défini ici et en 8.2 s'apparente à un indicateur de la difficulté intrinsèque d'une question[96].

Par exemple, pour les prévisions du CHPMC sur la Rance, pour la première situation météorologique qui était proposée (i.e. avant feed-back), personne n'avait surestimé la lame d'eau spatiale. Deux prévisionnistes ont fourni un intervalle de crédibilité dans lequel l'observé est tombé, les 6 autres prévisionnistes ayant sous-estimé la lame d'eau attendue. On peut alors estimer le biais dans les réponses à cette question par l'intervalle de confiance suivant : $IC_{95\%}$ $[b+|n,k]= [8.10^{-5} ; 0,33]$. On conclut avec une confiance de 95% que les réponses à cette question étaient effectivement biaisées.

On fait de même avec chacune des 6 questions (i.e. les prévisions sur les 6 bassins versants), et on comptabilise celles qui ont produit des réponses biaisées.

Les résultats obtenus durant chaque journée de formation respectivement au CHA et au CHPMC sont contenus dans le Tableau 28 et le Tableau 29.

| | Avant la sensibilisation à la sous-estimation des incertitudes | Après la sensibilisation |
|---|---|---|
| **Estimateur de θ** | 0,20 | 0,45 |
| **$IC_{95\%}$ [θ\|n,k]** | [0,09; 0,37] | [0,31; 0,60] |
| **Questions produisant des réponses biaisées $0,5 \notin$ IC95% [b+\|n,k]** | 4 sur 6 (67%) | 3 sur 6 (50%) |

**Tableau 28 : Calibration des intervalles de crédibilité à 50% fournis par les prévisionnistes du CHA (EDF), et pourcentage de prévisions biaisées.**

[96] Selon Gigerenzer et al. (1991), ce biais traduit plutôt le caractère « piégeant » des questions, c'est à dire leur caractère non représentatif (au sens de Brunswik, 1956). Gigerenzer et al. (1991) construisent ainsi leur « probabilistic mental model (PMM) » qui permet de rendre compte de la différence entre question difficile et question piège. Dans la vie courante, les individus utilisent des heuristiques pour répondre à un type de question, et connaissent la validité de leurs heuristiques (i.e. le nombre de fois où elles amènent à une réponse correcte dans la vie courante, « validité écologique »). L'expérimentateur sélectionne, consciemment ou non (Juslin, 1994), des questions où précisément ces heuristiques conduisent à un résultat incorrect (questions pièges). Les individus donneront donc des réponses fausses, dont ils évalueront pourtant la validité comme étant égale à la « validité écologique ». C'est ainsi que Gigerenzer et al. (1991) expliquent à la fois la surconfiance dans les réponses à un questionnaire de culture générale, et le hard-easy effect (plus les gens se trompent à une question, plus l'expérimentateur la définit comme difficile, alors qu'elle est simplement « piégeante »). Quant à mettre une échelle de difficulté sur les questions, ces auteurs affirment que c'est une tâche hautement subjective, qui, hors cas particulier, est vouée à l'échec.

En 8.5, nous apporterons deux objections à ce PMM : d'une part, en l'absence de question piège (c'est à dire de biais dans les réponses), les individus ont montré une surconfiance marquée (8.2.5). D'autre part, l'exemple de la prévision météorologique nous permet de construire aisément une échelle légitime de difficulté des questions : en moyenne, la difficulté de la prévision s'accroît lorsque son échéance augmente (7.2.2, et en particulier la note n°66). A l'aide de cette échelle qui nous permet de nous abstraire du problème d'éventuels pièges dressés par l'expérimentateur, le phénomène du hard-easy effect reste marqué (7.2.2).

| | Avant la sensibilisation à la sous-estimation des incertitudes | Après la sensibilisation |
|---|---|---|
| **Estimateur de θ** | 0,21 | 0,54 |
| **$IC_{95\%}$ [θ\|n,k]** | [0,11; 0,34] | [0,40 ; 0,68] |
| **Questions produisant des réponses biaisées 0,5∉ IC95% [b+\|n,k]** | 5 sur 6 (83%) | 3 sur 6 (50%) |

**Tableau 29 : Calibration des intervalles de crédibilité à 50% fournis par les prévisionnistes du CHPMC (EDF), et pourcentage de prévisions biaisées.**

En 8.2, le Tableau 20 nous présentait déjà les résultats quant à la calibration des prévisions fournies, ainsi que l'effet du feed-back sur celle-ci. Les pourcentages de réponses biaisées nous amènent à deux nouvelles conclusions :

- Bien que les situations météorologiques et les bassins versants aient été choisis pour leur intérêt vis-à-vis des enjeux hydroélectriques d'EDF, et sans la mauvaise intention d'induire les candidats en erreur, les experts ont eu un jugement biaisé pour la majorité des questions qui leur étaient posées.
- Un feed-back sur la sous-estimation des incertitudes a le double effet de réduire cette sous-estimation, et de diminuer le biais du jugement des individus.

Les prévisionnistes ont ainsi pu réajuster leur jugement suite à une double prise de conscience : non seulement ils sont surconfiants, mais ils sont tous « programmés » pour faire les mêmes erreurs sur les mêmes bassins pour les mêmes situations météorologiques. Ils se sont donc réajustés pour produire des prévisions probabilistes reflétant correctement leur degré d'incertitude modulo le fait qu'il leur arrive d'être piégés. Autrement dit, ils ont pris une « marge de sécurité » dans l'estimation de leurs incertitudes de prévision. Pour les questions qui ne se sont a posteriori pas révélées piégeantes (3 questions dans chaque expérience), on décèle en effet une très large sous-confiance (Tableau 30).

| | Après sensibilisation au CHA | Après sensibilisation au CHPMC |
|---|---|---|
| **Estimateur de θ pour les bassins non biaisés** | 0,76 | 0,88 |
| **$IC_{95\%}$ [θ\|n,k]** | [0,55; 0,90] | [0,70 ; 0,96] |

**Tableau 30 : Calibration des intervalles de crédibilité à 50% fournis par les prévisionnistes du CHA et du CHPMC (EDF), pour les bassins non biaisés.**

Ainsi, si les prévisionnistes n'ont pas su éviter de tomber dans les mêmes pièges y compris après feed-back (le nombre de questions qui les a piégés est passé de 4 ou 5 à 3), ils ont néanmoins réajusté leur appréciation générale des incertitudes, de sorte à fournir des prévisions parfaitement calibrées en terme d'estimation des incertitudes.

### 8.3.2 Résultats à Hydro-Québec

On reprend la même analyse que ci-dessus avec les prévisions d'Hydro-Québec.

Notons que contrairement au biais défini en 8.2 et 8.3, le biais est décelé ici en posant n questions à une même personne. Il révèle non plus « des idées fausses largement répandues dans une population donnée », mais « les mauvaises habitudes d'un prévisionniste », qui sous-estimerait ou surestimerait beaucoup trop souvent sa prévision pour un bassin et une échéance donnés.

Nous avons par exemple obtenu les résultats suivants pour deux bassins (échéance J+1), pour lesquels la prévision présentait un biais :

Sur le bassin "Hart Jaune" : 25 surestimations, 46 succès, et 51 sous-estimations. On peut alors estimer le biais de cette question : $IC_{95\%}$ [b+|n,k]= [0,23 ; 0,44], et conclure à un biais négatif (biais de sous-estimation).

Sur le bassin "Bersimis 2", en obtient en revanche 40 surestimations, pour 62 succès, et 16 sous-estimations. Ainsi, $IC_{95\%}$ [b+|n,k]= [0,59 ; 0,82], ce qui nous amène à conclure à un biais positif (biais de surestimation).

Le Tableau 31 ajoute au Tableau 17, présenté précédemment, le pourcentage de réponses biaisées pour chaque échéance.

| **Echéance** | **Questions produisant des réponses biaisées 0,5∉ IC95% [b+\|n,k]** | **Estimateur de θ** | **$IC_{95\%}$ [θ\|n,k]** |
|---|---|---|---|
| J | 9,8% | 0,53 | [0,52 ; 0,54] |
| J+1 | 14,1% | 0,51 | [0,50 ; 0,52] |
| J+2 | 12,0% | 0,47 | [0,46 ; 0,48] |
| J+3 | 18,5% | 0,43 | [0,42 ; 0,44] |

**Tableau 31 : Calibration des intervalles de crédibilité à 50% fournis par les prévisionnistes d'Hydro-Québec, et pourcentage de prévisions biaisées par échance.**

Ces résultats qui nous avaient permis de présenter le "hard-easy effect" en 7.2.2 nous montrent également que plus une question est difficile, plus elle entraîne un biais important dans les réponses.

Cette fois, il n'y a plus aucune ambiguïté entre question difficile et question piégeante. Toutes les questions posées étaient neutres, car posées de manière opérationnelle par l'usager pour 140 journées différentes, sans même que le prévisionniste soit au courant que ses prévisions aient été archivées.

De même qu'il y a un lien entre difficulté des questions et surconfiance dans les réponses, il semble y a avoir un lien entre difficulté dans les questions et biais dans les réponses. Notons toutefois que les biais révélés dans ces prévisions ne sont pas du même ordre de grandeur que ceux obtenus avec les questions de culture générale : ici, moins de 20% des questions entraînent des réponses biaisées, alors qu'avec le même critère de rejet, la proportion était de 80% pour le questionnaire de culture générale.

Qu'en est-il en retirant a posteriori les bassins pour lesquels les prévisions étaient biaisées ? Les valeurs de θ varient de façon non-significative : il faut aller jusqu'à la 3e décimale pour déceler une augmentation (Tableau 32). Nous n'avons pas présenté les intervalles de confiances qui restent strictement identiques (à la 2e décimale) que l'estimation soit faite sur l'ensemble des bassins ou sur les bassins « neutres ».

| Echéance | Estimateur de θ<br>Sur l'ensemble des bassins | Estimateur de θ<br>pour les bassins non biaisés |
|---|---|---|
| J | 0.529 | 0.532 |
| J+1 | 0.507 | 0.511 |
| J+2 | 0.474 | 0.475 |
| J+3 | 0.428 | 0.429 |

**Tableau 32 : Calibration des intervalles de crédibilité à 50% fournis par les prévisionnistes d'Hydro-Québec, sur l'ensemble des bassins, et en retirant ceux produisant des prévisions biaisées.**

Ces résultats nous permettent de conclure que si la surconfiance et le biais augmentent tous deux avec l'échéance de la prévision, ce n'est pas l'augmentation du nombre de bassins entraînant des prévisions biaisées qui entraîne une augmentation de la surconfiance.

## 8.4 Interprétation

Nous avons montré que le biais n'est pas seulement dû à l'intentionnalité de l'expérimentateur ou le fait du hasard ou de mécanismes plus ou moins inconscients (Juslin, 1994) dans le choix des questions. Le biais est également inhérent au candidat et à sa culture (les sujets se trompent tous de la même façon sur des questions de culture générale, ou dans la formulation d'une prévision météorologique). Nous avions suggéré au Chapitre 7 que les prévisionnistes pouvaient avoir été amenés à produire des prévisions surconfiantes pour s'adapter à leur environnement social. Les résultats obtenus ici aussi bien avec les prévisions hydrométéorologiques, qu'avec le questionnaire de culture générale pourraient suggérer de la même façon que les individus ont pu être éduqués à avoir un jugement biaisé.

Le biais, ainsi décelé a pu être de deux natures :

- idées fausses de nombreuses personnes sur une même question de culture générale
- mauvaise habitude d'un prévisionniste face à une question posée un grand nombre de fois : la prévision de la lame d'eau sur un bassin et pour une échéance donnés[97].

Quelle qu'en soit sa nature, le biais peut contribuer à la surconfiance, par un effet d'ancrage loin de la valeur correcte.

C'est bien entendu le cas pour les réponses au questionnaire de culture générale (8.2.5). L'expérience avec les prévisionnistes d'EDF a confirmé que les biais peut contribuer à la surconfiance : les prévisionnistes devaient en être eux-mêmes convaincus pour fournir des estimations si fiables de l'incertitude, se sachant biaisés.

[97] Cette mauvaise habitude peut être partagée par toute une équipe de prévisionnistes, voire par la communauté des prévisionnistes (le biais de surestimation, Krzysztofowicz et Sigrest, 1999, et Murphy et Daan, 1984). En cela, les deux natures de biais peuvent se rejoindre : mauvaise habitude et idée fausse largement répandue.

Comme nous l'avions évoqué en note n° 92 le biais décelé dans les réponses au questionnaire de culture générale peut provenir du fait que les individus ne donnent pas des réponses indépendantes à ce type de questions car ils partagent une culture commune. Or la culture[98] induit chez l'individu des biais d'ancrage par rapport aux choses qui lui sont familières : de mauvaises associations d'idées ou des analogies inadaptées entraîneront des raisonnements erronés. En prenant le cas des trois questions les plus biaisées, on peut imaginer les raisonnements suivants pour un candidat bénéficiant d'une culture française moyenne :

- Longueur du Nil : Les fleuves qui sont familiers à ce type de candidat sont les grands fleuves français comme la Loire. "*Un grand fleuve comme la Loire doit faire 1000 km ; le Nil doit être bien plus grand*", sans imaginer une telle disproportion entre la taille des deux fleuves.
- Nombre de livres dans le Nouveau Testament : Le candidat sait qu'il y a 4 évangiles ; il suppose qu'il doit exister quelques autres livres dans le Nouveau Testament, mais l'association d'idée évangiles-Nouveau Testament est tellement forte qu'il ne peut envisager que les évangiles soient en nombre si faible par rapport à l'ensemble des livres du Nouveau Testament.
- Hauteur des chutes du Niagara : "*Puisque c'est très connu, ce doit être très impressionnant, donc très haut*". Habituellement, le caractère impressionnant pour une chute d'eau (classiquement, une cascade) aux yeux du candidat, c'est sa hauteur (et en second lieu son débit). Ici, c'est la largeur qui est impressionnante.

Nos explications confirment en partie les conclusions de Juslin (1994) : c'est surtout parce que les individus ont un jugement biaisé sur les questions qu'on leur pose qu'ils sous-estiment leurs incertitudes.

Cependant, l'auteur affirme que le biais est artificiellement introduit en sélectionnant des questions de culture générale pour lesquelles la plupart des individus utilisaient des heuristiques qui, si elles fonctionnent de manière satisfaisante dans la plupart des cas, se trouvent mises en défaut dans les cas sélectionnés par l'expérimentateur. Juslin (1994) postule donc qu'en choisissant au hasard une sélection de questions, et en plaçant l'individu dans des conditions aussi proches que possible de ses conditions « naturelles », celui-ci ne sera plus surconfiant (le modèle "écologique" de Juslin et al., 1994).

Or, si nous avons montré que des questions-pièges ont été délibérément introduites dans le questionnaire de culture générale et que ces pièges entraînent une aggravation de la surconfiance, trois conclusions viennent infirmer le postulat de Juslin.

- D'une part, même pour des questions de culture générale qui n'entraînaient aucun biais dans les réponses, nous avons observé la présence de surconfiance.
- D'autre part, nous avons également constaté - certes dans une moindre mesure que pour le questionnaire de culture générale - la présence de réponses biaisées de la part du prévisionniste de Trans-Energie, pour des questions on ne peut plus "neutres", alors même que celui-ci était placé dans son "univers naturel". Celui-ci ignore même que ses prévisions sont stockées dans une base de données, et font l'objet d'analyses statistiques ; lui-même ne conserve aucun historique des prévisions qu'il a émises.

[98] On utilise ce mot sous l'acception suivante : Ensemble des formes acquises de comportement dans les sociétés humaines (Dictionnaire Petit Robert)

- Enfin, même en retirant les bassins sur lesquels le prévisionniste se fait habituellement piéger, la surconfiance du prévisionniste ne varie pas.

Ce dernier résultat montre que même s'il existe un biais dans le jugement de l'expert, l'effet d'ancrage ne suffit pas à lui seul à expliquer l'effet de surconfiance.

## 8.5 Résumé, et conséquences pratiques en psychologie expérimentale.

Dans le Chapitre 7 nous nous étions intéressés au phénomène de surconfiance chez les prévisionnistes opérationnels, et aux moyens à mettre en œuvre pour réduire cette surconfiance. Les résultats obtenus suggéraient que le biais et la surconfiance pouvaient être des phénomènes en partie indépendants (7.4.3). Malgré leurs éventuelles idées fausses sur la lame d'eau à attendre en tendance centrale de la prévision, les prévisionnistes avaient une bonne perception du risque météorologique (c'est à dire de l'événement grave qui n'a qu'une probabilité faible d'occurrence), et évaluaient correctement l'ampleur de leurs incertitudes.

Dans ce Chapitre 8, une investigation plus approfondie sur le lien entre le biais et l'estimation des incertitudes a permis de nuancer ce résultat. Des idées fausses sur une question, si elles sont suffisamment importantes, peuvent créer un effet d'ancrage suffisamment fort pour que la bonne réponse se situe en dehors des intervalles de crédibilité fournis (8.2). C'est ainsi que le biais (c'est à dire les mêmes idées fausses partagées par de nombreuses personnes) peut être l'une des sources de la surconfiance mesurée chez les individus.

Partant, on imagine bien qu'on puisse concevoir un questionnaire de culture générale qui accentue la surconfiance mesurée. Il suffit pour ce faire de sélectionner soigneusement des questions fallacieuses, en cherchant des questions qui vont à l'encontre des idées répandues, de l'intuition générale, des ordres de grandeur courants. La chose est d'autant plus facile à réaliser qu'on peut même concevoir un tel questionnaire en toute bonne foi, avec la simple intention de retenir des questions « intéressantes » (Juslin, 1994).

On ne peut pas pour autant conclure de manière aussi radicale que les Panglossiens, comme Juslin (1994) et Gigerenzer et al. (1991) que la surconfiance n'est due qu'au biais de l'expérimentation. Nous avons pu en effet remarquer qu'il restait une surconfiance marquée une fois ôtées toutes les questions pièges.

Néanmoins, il convient d'être extrêmement attentif aux dépendances qui peuvent exister dans les réponses à un questionnaire de culture générale, du fait que les individus utilisent bien souvent les mêmes heuristiques. En ce sens, nos explications avancées en 8.4 s'inspirent du "probabilistic mental model" (Gigerenzer et al., 1991) qui suppose que les sujets peuvent partager les mêmes heuristiques (même mécanismes cognitifs, mêmes clefs de comparaison, stratégies similaires de recherche d'informations pertinentes, etc.).

Les deux conclusions majeures que les résultats contenus dans cette section 8 nous permettent de tirer sont les suivantes :

- **la surconfiance subsiste même lorsque le biais a disparu (8.2.4).**
- **le biais subsiste même lorsque les questions ne comportent pas de piège (i.e. quand l'individu est dans son environnement "écologique"). Ce biais "naturel" peut accentuer le phénomène de surconfiance (8.4).**

Ces résultats nous invitent également à nuancer les interprétations données par les Mélioristes (Kahneman et al., 1982), qui semblent exclure que leurs résultats puissent être faussés par des artéfacts dû à l'expérimentation : la sélection de questions-pièges, le fait qu'en laboratoire les sujets ne disposent pas de la richesse et de la complexité des informations avec laquelle ils ont

l'habitude de construire leur jugement, etc. Dans de telles conditions on peut observer une augmentation de la surconfiance des sujets (8.2.3).

Nos résultats nous ont également permis d'investiguer le "hard-easy effect". En premier lieu, ceux-ci nous permettent de souligner l'ambiguité qui existe entre question difficile et question-piège. Si l'on mesure la difficulté d'une question au nombre de personnes qui ont donné des réponses correctes, on risque fort de mesurer à quel point cette question est piégeante (8.2). Cependant, toute autre mesure de la difficulté des questions posées aura indubitablement une importante composante subjective. Ce n'est que dans des rares cas comme celui de la prévision opérationnelle qu'une mesure de la difficulté des questions est incontestable : la difficulté est en moyenne croissante avec l'échéance de la prévision. Disposant alors d'une telle relation d'ordre, on a pu étudier indépendamment comment variaient le biais et la surconfiance en fonction de la difficulté des questions. Les résultats obtenus ont alors suggéré qu'un accroissement de la difficulté dans les questions entraînait conjointement une augmentation du biais et de la surconfiance. Plutôt que d'accepter d'avoir une idée vague sur des questions difficiles, le prévisionniste préfère avoir une idée précise, pour laquelle il dispose de théorie, concepts, et modèles grâce auxquels il peut exercer son expertise et communiquer avec les autres prévisionnistes… au risque que cette idée soit fausse (7.5).

Ces conclusions nous amènent à formuler quelques suggestions pratiques en psychologie expérimentale :

Il convient tout d'abord de prendre garde aux « pièges cachés » de l'expérimentation (comme la sélection de questions-pièges) et de l'exploitation des résultats (comme l'hypothèse souvent tacite que deux variables dont on ne sait rien sont indépendantes). Nous avons par exemple montré qu'une question prétendument difficile pouvait en fait être tout simplement une question-piège.

Nous proposons deux types de solutions pour explorer la phénomène de surconfiance en intégrant la question du biais.

Une première solution de type expérimental consiste à manipuler les sujets avant de leur poser des questions. Une telle manipulation pourrait être réalisée de plusieurs manières :

- une sensibilisation générale à ces deux types de problèmes que sont le biais et la surconfiance. Un court stage permettrait de sensibiliser des individus aux caractéristiques majeures de ces deux phénomènes. Ce type de stage de sensibilisation pourrait être similaire à celui présenté en 7.3, en intégrant toutefois des explications sur le phénomène de biais (i.e. les idées fausses).
- un feed-back sur un premier questionnaire mettant en évidence chacun des phénomènes. Ces feed-back, souvent utilisés en psychologie expérimentale, ont prouvé leur efficacité (Stone, Opel, 2000).
- des instructions préalables à l'expérimentation, telles que celles précédents le questionnaire de culture générale que nous avons proposé (8.2.1). Ces instructions invitent à rester prudent et modeste dans l'estimation de ses incertitudes pour limiter la surconfiance. Celles-ci peuvent également mettre en garde chaque individu sur le fait que des idées fausses circulent souvent sur bien des questions. De telles instructions ont cependant montré une faible efficacité (Lichtenstein et al., 1982).

Une deuxième solution, de type statistique, pourrait consister à produire un questionnaire sans aucune question piège. Pour obtenir une liste de questions neutres, V. Fortin (comm. pers.)

suggère une élimination séquentielle Bayésienne de questions pièges : pour toutes les questions, les réponses sont affectées d'un prior non informatif sur leur biais. A chaque expérimentation, on remet à jour la distribution de probabilité du biais pour les réponses à chaque question. Lorsque pour l'une d'entre elles les réponses deviennent significativement biaisées (0,5 sort de l'intervalle de recouvrement à 80% par exemple), on remplace la question par une autre, choisie au hasard. Le questionnaire serait ainsi purgé au fil des expérimentations de ses questions-pièges.

## 8.6 Bibliographie

Biais, B., Hilton, D., Mazurier, K., Pouget, S., 2002. Psychological Dispositions and Trading Behaviour. Les cahiers recherche du Centre de Recherche en Gestion de Toulouse, vol. 145, (http://www.univ-tlse1.fr/iae/310-crgcahiers.php).

Blanton H., Pelham, B., Dehart, T., Carvallo, M., 2001. Overconfidence as Dissonance Reduction, Journal of Experimental Social Psychology, Vol. 37**,** 373–385.

Brunswik, E., 1956. Perception and the representative design of psychological experiments (2nd Ed.), Berkley: University of California Press.

Cosmides, L., Tooby, J., 1996. Are humans good intuitive after all? Rethinking some conclusions from the literature on judgment under uncertainty. Cognition, 58, 1-73.

Festinger, L.A., 1957. A Theory of Cognitive Dissonance, Stanford University Press.

Fischhoff, B., 1982, Debiasing, in Kahneman, D., Slovic, P., Tversky, A., 1982. Judgment under uncertainty: Heuristics and biases. Cambridge University Press, 422-444.

Friedman, C., Gatti, G., Elstein, A., Franz, T., Murphy, G., Wolf, F., 2001. Are clinitians correct when they believe they are correct? Implications for medical decision support, Medinfo, Vol. 10, 1, 454-458.

Gigerenzer, G., Hoffrage, U., Kleinbölting, H., 1991. Probabilistic mental models: a Brunswikian theory of confidence, Psychological Review, Vol. 98, 4, 506-528.

Griffin, D., Tversky, A., 1992. The weighing of evidence and the determinants of confidence, Cognitive Psychology, Vol. 24, 411-435.

Juslin, P., Winman, A., Persson, T., 1994. Can overconfidence be used as an indicator of reconstructive rather than retrieval processes ?, Cognition, Vol. 54, 99-130.

Juslin, P., 1994. The Overconfidence Phenomenon as a Consequence of Informal Experimenter.Guided Selection of Almanac Items, Organizational Behavior and Human Decision Processes, Vol. 57, 2.

Kahneman, D., Slovic, P., Tversky, A., 1982. Judgment under uncertainty: Heuristics and biases. Cambridge University Press.

Koriat, A., Lichtenstein, S., Fischhoff, B., 1980, Reasons for confidence. Journal of experimental psychology: Human learning and memory, vol.6, 107-118.

Kramer, R., Newton, E., Pommerenke, P., 1993. Self-enhancement biases and negociator judgment: effects of self-esteem and mood, Organizational Behavior and human decision processes, Vol. 56, 1, 110-133.

Krzysztofowicz, R., Sigrest, A., 1999. Calibration of probabilistic quantitative precipitation forecasts. Weather and Forecasting, 14(3), 427-442.

Murphy, A. H., Daan, H., 1984. Impacts of feedback and experience on the quality of subjective probability forecast : Comparison of results from the first and second years of the Zierikzee Experiment, Monthly Weather Review, vol. 112, 413-423.

Lichtenstein, S., Fischhoff, B., Philips, L., 1982, Calibration of probabilities: the state of art to 1980, in Kahneman, D., Slovic, P., Tversky, A., 1982. Judgment under uncertainty: Heuristics and biases. Cambridge University Press, 422-444.

Stanovich, K. E., 1999. Who is rational? Studies of individual differences in reasoning. Mahweh, NJ: Erlbaum.

Stone, E., Opel, R., 2000. Training to improve calibration and discrimination: the effects of performance and environmental feedback, Organizational Behavior and human decision processes, Vol. 83, 2, 282-309.

Tetlock, P., E., 1991, An alternative metaphor in the study of judgment and choice: People as politicians, Theory and psychology, Vol.1, 4, 451-475.

Palmarini, M., P., 1995, La réforme du jugement, ou comment ne plus se tromper, Odile Jacob (trad. De L'Illusione di sapere, Arnoldo Editore, 1993).

# 9 Une langue adaptée au contexte opérationnel ?

*En plongeant le prévisionniste dans cet univers de tirages d'urne et de lancement de dés,*
*univers tellement étranger à celui qui est un familier des vents et du soleil,*
*des bruines et des tempêtes,*
*ne risque-t-on pas de lui faire perdre la boule ?*

## 9.1 Le prévisionniste confronté à ce nouveau langage : premiers balbutiements.

### 9.1.1 Introduction

Murphy (1993) rappelle que la première qualité d'une prévision est sa consistance ("type one goodness"), c'est à dire sa conformité avec le jugement personnel du prévisionniste devant la situation météorologique.

Selon Murphy (1993), la principale source d'inconsistance provient de ce que le prévisionniste s'adapte aux enjeux des usagers (en particulier à leur vulnérabilité) plutôt que de ne se concentrer que sur les évolutions atmosphériques. L'internalisation des enjeux de l'usager peut alors entraîner un biais de surestimation systématique[99], par excès de prudence de la part des prévisionnistes (Doswell et Flueck, 1989, Murphy, 1991, et la section 7.1).

On peut se demander cependant si, en dehors de toute internalisation des enjeux de l'usager, la simple méthode d'élicitation des probabilités peut éventuellement entraîner une forme d'inconsistance dans les prévisions émises.

Murphy et Winkler (1974) montrent par exemple que les prévisionnistes à qui l'on demande de fournir des intervalles de crédibilité sont plus fiables que ceux à qui l'on demande d'estimer la probabilité que l'observé tombe dans un intervalle de largeur fixe.

D'où la question plus générale : comment réagissent les prévisionnistes à des questions abstraites de jeu de hasard, décalées par rapport aux questions qu'ils se posent habituellement, ou faisant appel à des concepts qui leur sont peu familiers ?

L'inconsistance pourrait alors se traduire par des contradictions logiques entre plusieurs réponses fournies par le même prévisionniste.

Nous analyserons dans cette partie 9.1 dans quelle mesure le mode opératoire selon lequel nous avons recueilli des probabilités subjectives durant la journée de stage proposée aux prévisionnistes a pu perturber leur jugement et produire diverses formes d'incohérence.

### 9.1.2 Contexte

Protocole suivi au CHA

Durant la journée de stage au CHA, nous avions tout d'abord demandé aux prévisionnistes d'exprimer 4 quantiles de probabilité sur la prévision de la lame d'eau spatiale sur 6 bassins versants (Chapitre 4 etAnnexe 2). La signification de chacun de ces quantiles était expliquée en détail.

Ensuite, sur 2 bassins, on leur demandait de choisir entre un pari risqué (tirage d'une boule dans une urne dont on connaît la composition) et un pari incertain (lame d'eau spatiale supérieure à un quantile qu'ils venaient d'émettre). Le lecteur trouvera le descriptif du protocole en Annexe 2.

Tous les paris, incertains ou risqués avaient les mêmes enjeux : dans les deux premiers, soit on gagnait 100 FF, soit on était quitte, et dans les deux derniers, soit on perdait 100 FF, soit on gagnait 50FF. Tous les paris avaient une espérance de gain positive, et seuls les deux derniers paris associaient le franchissement d'un seuil élevé à une perte pour le prévisionniste. Cela permettait de se rapprocher de la norme sociale selon laquelle le prévisionniste est

[99] Dans le cas où le risque est lié à de fortes valeurs de la variable à prévoir (note n°79).

sévèrement jugé lorsque sa prévision a été largement dépassée, et qu'il n'a pas vu venir l'occurrence d'une pluie intense. D'avis d'experts en prévision opérationnelle (Frédéric Atger et Rémy Garçon), et de chercheur en sciences du danger (Jean-Luc Wybo), il semblait contre-intuitif d'associer des conséquences positives à l'occurrence de pluies intenses.

Les paris ont été construits de sorte à situer les quantiles $L_{40}$, $L_{60}$, puis $L_{86}$, et $L_{92}$ par rapport à leurs quantiles $L_{50}$, et $L_{90}$. A chacun des 4 choix de pari, il y avait une seule réponse correcte, l'autre impliquant une incohérence. Sur les quatre choix de pari, il y avait donc 4 possibilités d'erreur : $L_{90} \geq L_{92}$ , $L_{90} \leq L_{86}$ , $L_{50} \geq L_{60}$ , et $L_{50} \leq L_{40}$.

Nos questions sur le choix de paris n'avaient ainsi qu'un objectif de "contrôle de cohérence". De même que dans l'expérience de Krzysztofowicz et Sigrest (1999), il s'agissait avant tout de bien faire peser aux prévisionnistes la signification de chaque quantile, et de s'assurer que ceux-ci correspondaient bien à leur jugement ("type one goodness", Murphy, 1993). Notre objectif était simplement de sensibiliser les prévisionnistes à l'idée qu'estimer un quantile sur une prévision, c'est conditionner une prise de pari sur l'avenir (pari que prendra l'usager sur la base de cette information, Chapitre 6). Nous avons insisté sur ce point lors de l'introduction de notre expérimentation et des discussions que nous avions eues avec les prévisionnistes en préparation de cette journée de stage.

Avec 5 prévisionnistes avant feed-back[100], 6 après feed-back, et 2 bassins pour lesquels on posait à chaque fois 4 questions de choix de paris, nous avons posé un total de 5*4*2= 40 puis 6*4*2=48 questions.

Chaque question portant sur un choix de paris entraînait une réponse binaire (A ou B), juste ou fausse. Il y avait donc autant de possibilités d'incohérence que de réponses à fournir.

Protocole suivi au CHPMC

Au CHPMC, nous avons tout d'abord fait parier les prévisionnistes sur diverses hauteurs de précipitation que nous leur présentions sur chacun des 6 bassins (Annexe 2.). Les paris ont été construits de sorte à situer les quantiles $L_{40}$, $L_{60}$, puis $L_{86}$, et $L_{92}$ par rapport à deux valeurs a et b (a<b)qu'on leur fournissait. Les deux valeurs a et b variaient suivant la situation météorologique choisie, le bassin versant, mais également le questionnaire fourni à chacun des prévisionnistes. Afin de s'abstraire d'un éventuel biais d'ancrage sur des valeurs fixées, nous avions produit suivant les questionnaires trois jeux de valeurs (a,b), distribués aléatoirement à chacun des prévisionnistes.

Dans un second temps, nous avons demandé aux prévisionnistes d'estimer directement les quantiles $L_{25}$, $L_{50}$, $L_{75}$, et $L_{90}$ sur sa prévision de lame d'eau spatiale.

Nous avons pu ainsi contrôler la cohérence entre les 4 paris choisis, et nous assurer également que les quatre quantiles fournis étaient bien croissants.

Ce test nous a surtout permis de contrôler la cohérence des paris pris dans un premier temps, avec les quantiles émis dans un second temps. Il s'agissait de voir si les préférences révélées par le choix des paris n'étaient pas en contradiction avec les quantiles directement estimés.

Supposons ici que les quantiles émis soient strictement croissants, et que les paris pris soient cohérents entre eux.

La compatibilité entre les quantiles émis et les paris pris se traduit alors par exemple, pour le quantile $L_{90}$, par la condition : $L_{92} \leq L_{90}$ et $L_{90} \leq L_{86}$. Or on ne dispose que d'inégalités sur $L_{92}$

[100] Un prévisionniste a dû s'absenter en cours d'analyse de la 1ère situation météo. Il a en revanche participé à l'introduction, à la première analyse météo, et au feed-back qui l'a suivie.

et $L_{86}$ pour s'assurer que cette condition est bien respectée. Il existe donc 2 possibilités d'erreur pour $L_{90}$ : ($L_{92} \leq$ b et $L_{90} \geq$ b) ou ($L_{86} \geq$ b et $L_{90} \leq$ b)

Voici pour tous les quantiles toutes les possibilités d'incohérence entre quantiles émis et paris pris que notre questionnaire permettait. Elles sont au nombre de 7, pour un bassin et un prévisionniste :

Quantile 90 incompatible avec les paris pris : ($L_{92} \leq$ b et $L_{90} \geq$ b) ou ($L_{86} \geq$ b et $L_{90} \leq$ b)

Quantile 75 incompatible avec les paris pris : ($L_{86} \leq$ b et $L_{75} \geq$ b) ou ($L_{60} \geq$ a et $L_{75} \leq$ a)

Quantile 50 incompatible avec les paris pris : ($L_{60} \leq$ a et $L_{50} \geq$ a) ou ($L_{40} \geq$ a et $L_{50} \leq$ a)

Quantile 25 incompatible avec les paris pris : $L_{40} \leq$ a et $L_{25} \geq$ a

Cela fait donc au total, pour 6 bassins, 8 prévisionnistes, et 7 possibilités d'incohérence : 6*8*7=336 possibilités d'erreur.

L'inconsistance produite par le mode opératoire

Pour chacune des deux expériences effectuées au CHA et au CHPMC, si le protocole n'est pas suffisamment intuitif, on s'attend à ce que les prévisionnistes émettent des quantiles incompatibles avec leurs choix de paris. Ces incohérences seraient le produit de l'inconsistance du prévisionniste, inapte à faire part de son jugement en réponse à des questions trop pointues, incomprises, ou mal interprétées.

### 9.1.3 Résultats

Nous avons tout d'abord relevé, tant au CHA et qu'au CHPMC, que tous les quantiles directement élicités étaient strictement croissants. De même, dans chaque expérience, aucun prévisionniste n'a choisi des paris incompatibles entre eux.

Concernant le nombre de réponses manquantes, les prévisionnistes du CHA comme du CHPMC ont systématiquement fourni une estimation de chacun des 4 quantiles sur tous les bassins, et pour chaque situation considérée. Cependant, tous les prévisionnistes du CHA ne se sont pas prononcés sur chacun des paris : 36 réponses sur 40 nous ont été fournies avant feed-back, et 32 sur 48 après feed-back. Les prévisionnistes du CHPMC ont quant à eux répondu à tous les choix de pari.

Dans les deux expériences, l'inconsistance s'est manifestée par des incohérences entre quantiles émis et paris pris (Tableau 33 et Tableau 34). Le feed-back ne semble pas faire décroître la proportion de ces erreurs.

Ces incohérences entre quantiles émis et paris pris sont particulièrement nombreuses au CHA.

| Stage de formation au CHA | Avant feed-back | Après feed-back |
|---|---|---|
| N, nombre de possibilités d'erreur | 36 | 32 |
| n, nombre d'erreurs commises | 13 | 17 |
| Estimateur de la probabilité **ε** de commettre une erreur | 0,36 | 0,53 |
| $IC_{95\%}$ [ε\|N,n] | [0,21 ; 0,52] | [0,36 ; 0,70] |

**Tableau 33 : Probabilité qu'un prévisionniste commette une erreur (incohérence entre quantiles émis et paris pris). Stage de formation au CHA.**

| Stage de formation au CHPMC | Avant feed-back | Après feed-back |
|---|---|---|
| N, nombre de possibilités d'erreur | 336 | 336 |
| n, nombre d'erreurs commises | 29 | 25 |
| Estimateur de la probabilité **ε** de commettre une erreur | 0,09 | 0,07 |
| $IC_{95\%}$ [ε\|N,n] | [0,06 ; 0,12] | [0,05 ; 0,10] |

**Tableau 34 : Probabilité qu'un prévisionniste commette une erreur (incohérence entre quantiles émis et paris pris). Stage de formation au CHPMC.**

### 9.1.4 Interprétation

Un protocole expérimental peu représentatif des conditions de travail habituelles des prévisionnistes

Pour expliquer de telles erreurs, les Mélioristes mettraient en avant des défauts cognitifs : les prévisionnistes manqueraient en particulier d'exercice sur la manipulation des probabilités.

Une interprétation panglossienne de ces erreurs serait que notre protocole n'est pas assez intuitif, ni représentatif (au sens de Brunswik[101], 1956) des conditions de travail habituelles des prévisionnistes.

On relève donc trois entraves majeures à la bonne représentativité de notre expérience :

Tout d'abord, il était assez inhabituel pour les prévisionnistes d'être placés face à des hauteurs d'eau dont on demandait la probabilité de dépassement. Le processus mental habituel conduit le prévisionniste à converger vers un scénario qui lui semble le plus probable, puis à en considérer les manifestations du point de vue de l'usager. On comprend donc la difficulté pour un prévisionniste à faire le chemin inverse : à partir d'une lame d'eau qu'on lui soumet, remonter au scénario générateur, puis imaginer les évolutions atmosphériques qui ont produit une telle situation météorologique. Le prévisionniste risque alors de se raccrocher aux

[101] La représentativité d'une expérience de psychologie expérimentale est la prise en compte aussi exhaustive que possible de tous les éléments de contexte sur lesquels le sujet s'appuie dans la vie de tous les jours lors de ses prises de décision courantes (Brunswik, 1956).

situations déjà rencontrées ou théoriques qui pourraient amener à un tel résultat, et ne réfléchir plus que sur des considérations analogiques et fréquentielles[102], en s'éloignant du même coup de son analyse physique des évolutions atmosphériques.

Il était également inhabituel pour le prévisionniste de formuler son jugement en utilisant le concept de probabilité subjective. Dans ses considérations analogiques et fréquentielles, il a pu avoir une mauvaise interprétation des probabilités subjectives. En effet, certains prévisionnistes n'ont pas interprété les probabilités comme *un type de formalisation de l'idée que chacun se fait de sa propre incertitude de prévision dans la situation qui est la sienne*, mais comme une vérité qui lui serait cachée, et qu'il essaierait de deviner, par exemple en analysant une population de situations analogues (sortie d'une base de données, ou de sa propre expérience de situations vécues)[103].

Il était enfin inhabituel pour les prévisionnistes de ne pas disposer d'un historique récent sur la situation en cours. Ils ne pouvaient donc pas connaître le comportement des différents modèles dans la situation présente, en comparant les sorties de la veille au matin, avec les observés des dernières 24h. Ils n'avaient donc aucun élément pour apprécier la fiabilité des modèles dans la situation en cours.

Ces trois défauts de représentativité ont sorti le prévisionniste de ses conditions habituelles ("écologiques") de travail. D'où le risque de rupture causale[104], c'est à dire que la personne interrogée ne sait plus pourquoi elle répond ceci ou cela. Elle ne peut à la fois prendre en compte un contexte qui lui est étranger (le jeu, le pari, les urnes), tout en restant concentrée et méthodique dans l'analyse de la situation météorologique déjà suffisamment complexe[105].

<u>Des défauts de représentativité plus importants au CHA qu'au CHPMC</u>

La probabilité de commettre une erreur a été bien plus forte durant la journée d'expérimentation au CHA que celle au CHPMC. Les prévisionnistes ont pourtant sensiblement la même expérience, les mêmes outils, et les mêmes conditions de travail quotidiennes dans les deux centres de prévision.

---

[102] Cosmides et Tooby (1996) montrent que le sujet aborde bien plus aisément les probabilités lorsqu'il formule celles-ci en termes d'ensemble ou de fréquence qu'en termes de vraisemblance ou de pari.

[103] C. Colardelle de l'ENSMP, (comm. pers) a souligné qu'il n'était pas clair pour tous les prévisionnistes que l'exercice n'était pas de calculer la « probabilité pure ». A mon sens, ce concept n'existe pas, la probabilité étant un construit intellectuel, intrinsèquement subjectif, et donc loin de l'idée d'un chiffre unique et universel qui nous serait caché (et qu'on tenterait d'approcher par la loi des grands nombres par exemple). On modélise ici les paramètres hydrométéorologiques à prévoir, tels une lame d'eau spatiale sur un bassin versant ou un débit en un point d'un cours d'eau comme des variables aléatoires. C'est le choix du modélisateur qui cherche à quantifier au mieux une variable incertaine, et non celui de Dieu qui aurait assigné à chaque événement hydrométéorologique une loi de probabilité dont l'expert serait le mieux placé pour "deviner" la distribution. Je reprends ici une réflexion de Frédéric Atger (correspondance) à propos de la prévision probabiliste en général, et de la prévision d'ensemble en particulier : "I do not believe that there exist anything like an underlying pdf that would be "hidden" (by God ?) and that we should guess from clues left behind by angels-like ensembles !".

L'approche défendue ici est celle d'une partie de la communauté scientifique bayésienne, également soutenue par les épistémologues constructivistes tels Bachelard (1934) ou Le Moigne (1990).

[104] J.-L. Wybo (comm. pers.).

[105] C. Colardelle (comm. pers.) insiste également sur le sentiment d'insécurité face à ce changement de mode de travail, qui risque de modifier les habitudes des prévisionnistes. Les actuels détenteurs du pouvoir craignent également que le statu quo dont ils bénéficient ne s'affaiblissent, et que change la répartition du pouvoir au sein de l'équipe. Ces craintes peuvent également exacerber le sentiment de « perte de repères » face au nouveau problème qui leur est posé.

Les résultats obtenus suggèrent donc que le protocole adopté était particulièrement peu intuitif au CHA. Les incompréhensions dues à la nouveauté de l'utilisation des probabilités étaient particulièrement marquées au CHA (2e défaut de représentativité décrit ci-dessus). Plusieurs éléments de contexte permettent d'en donner une explication :

Les paris étaient placés après l'élicitation directe des quantiles, et ne devaient servir qu'à inviter les prévisionnistes à considérer à nouveau les quantiles qu'ils venaient d'émettre. Chacun de ces quantiles n'admettait à chaque fois qu'une réponse correcte, rationnelle. Plusieurs indications écrites et orales venaient appuyer l'idée que "rationnellement", "logiquement" "si votre évaluation des quantiles a été correcte", les quantiles émis doivent se traduire "de manière cohérente" par des paris. Aussi avions-nous appelé cette deuxième phase "test de cohérence", chacun des choix de paris "Test".

Tous les éléments étaient donc rassemblés pour que les prévisionnistes prennent cette phase d'expérimentation comme un examen. Un examen facultatif qui n'avait d'autre vertu qu'illustrative, pour montrer que « les prévisions sont le reflet de paris sur l'avenir que prend le prévisionniste ». Par conséquent, soit les prévisionnistes ne voulaient pas se faire juger, et n'ont pas répondu à ces questions (ce qui explique le nombre de réponses manquantes), soit ils voulaient passer avec succès ce test, et ils ont méthodiquement calculé les probabilités en jeu[106], soit ils ont répondu en pesant leurs préférences entre pari risqué et pari incertain[107]. (Annexe 3 pour visualiser les différences inter-individuelles).

Quoi qu'il en soit, de telles conditions étaient pour le moins inhabituelles pour les prévisionnistes. En pratique, lorsqu'on fait appel à eux, c'est pour qu'ils réduisent les incertitudes sur la vision du futur qu'a l'usager (passer d'une prévision naïve, climatologique, à la prévision fournie par les experts). Quand on leur pose une question, il n'existe donc jamais de « bonne réponse » qui permette de conclure que les prévisionnistes sont ou non rationnels. Notre protocole était ainsi très scolaire, et donc éloigné du contexte habituel dans lequel s'exerce l'expertise du prévisionniste[108].

Les résultats obtenus suggèrent en revanche que le protocole adopté au CHPMC était plus intuitif, plus représentatif des conditions naturelles, "écologiques", dans lequel l'expert a coutume d'exercer son métier de prévisionniste. L'expérience réalisée au CHPMC a en effet eu lieu après celle réalisée au CHA, ce qui nous a permis de modifier plusieurs éléments pour rendre le protocole d'élicitation plus intuitif :

- Tout d'abord, nous avions plus longuement explicité aux prévisionnistes du CHPMC notre approche des probabilités : une simple mesure de notre incapacité à connaître, observer, ou prévoir avec précision des événements. Cette mesure est donc par nature liée à chaque prévisionniste, à ses connaissances, à sa confiance en lui, et à son appréciation subjective des risques météorologiques.
- De plus, les éléments de contexte permettaient aux prévisionnistes une mise en confiance. Les choix de paris, placés avant l'élicitation directe des quantiles,

[106] Ces calculs ont parfois été laissés en marge, dans les questionnaires que les prévisionnistes nous ont rendus.

[107] Ils pourraient dans ce cas montrer de l'aversion à l'incertitude par rapport au risque (Knight, 1921). Cette question sera traitée en 9.3.

[108] Durant le cursus scolaire (jusqu'au second cycle universitaire), la plupart des problèmes soumis à l'élève ont déjà été formulés, conceptualisés et résolus avant celui-ci, et admettent une unique solution. Pour l'atteindre, l'élève devra appliquer les outils et méthodes qui lui ont été inculqués. En ce qui concerne le métier d'expert prévisionniste, les problèmes qu'on lui soumet n'admettent a priori aucune bonne solution. L'usager cherche simplement à réduire son ignorance de l'avenir.

pouvaient constituer un « galop d'essai », et permettre aux prévisionnistes de se jauger eux-mêmes par rapport à des idées a priori que pouvait avoir l'usager (vraisemblance que la lame d'eau dépasse 5mm, 10mm, etc.). Les chiffres choisis étaient neutres, au sens où ils ne venaient pas remettre en question des estimations qu'avait pu formuler chacun des prévisionnistes.

Ainsi, nous avons pu adapter des éléments de contexte de notre expérimentation pour que les choix de paris constituent une aide à la formulation des probabilités subjectives plutôt qu'un contrôle de cohérence qui s'apparenterait à un examen. Les conditions de l'expérience se sont donc un peu plus rapprochées des conditions habituelles des prévisionnistes.

Cependant, même au CHPMC, aucun des trois défauts de représentativité de l'expérience (0) n'était tout à fait absent. Cela peut contribuer à expliquer qu'il subsiste des incohérences malgré toutes les précautions prises.

### 9.1.5 Conclusion

Le stage proposé aux deux équipes de prévisionnistes avait pour objectif de les sensibiliser à l'estimation de leurs incertitudes et de les former à la prévision probabiliste. Nous voulions leur monter que les probabilités subjectives n'étaient qu'une traduction en termes quantitatifs de notions qui leur étaient familières : incertitude, vraisemblance, perception du risque, etc.

Cette "traduction" se faisait grâce à la notion de quantiles de probabilité et de prise de paris, qui étaient censés révéler leur degré de conviction ou d'incertitude sur leur prévision.

Cependant, comme toute traduction, celle-ci a pu mettre à mal "la pensée de l'auteur". Le prévisionniste a pu alors montrer des incohérences suivant la manière dont on lui a fait aborder l'estimation de ses incertitudes de prévision. En témoignent les incohérences entre quantiles émis et paris pris, particulièrement lorsque le contexte leur donnait l'impression de « passer un examen ».

Pourtant, lorsque dans sa pratique quotidienne le prévisionniste est appelé à donner des indications qualitatives sur son niveau d'incertitude, on peut penser que d'une sollicitation à l'autre, ses différentes réponses ne sont pas incohérentes entre elles.

Le protocole lui-même d'élicitation des probabilités subjectives, "traducteur de ses incertitudes en termes quantitatifs", a pu être à l'origine de certaines inconsistances, révélant ainsi les limites de notre protocole pour traduire le jugement du prévisionniste en situation courante.

Remarque : surconfiance et inconsistance

A la lecture particulière de chaque prévision, nous nous sommes questionnés sur le caractère « réaliste » de certaines fonctions de probabilités fournies : des Dirac par endroits (ailleurs que pour l'événement « pluie nulle sur le bassin » ; lorsque $L_{25} = L_{50} = 3$mm, par exemple), ou des quantiles $L_{75}$ et $L_{90}$ très rapprochés. Bien que de telles anomalies rendent les réponses peu crédibles[109], elles ne violent en rien la logique ni les règles des probabilités.

On peut supposer que l'inconsistance due au protocole expérimental peut contribuer au rapprochement, voire à l'égalité de ces quantiles. Le prévisionniste distinguant mal les niveaux de probabilité proches de 0 (1 chance sur 4, 1 chance sur 10, 1 chance sur 100), il ne

[109] Affirmer qu'il y a plus d'une chance sur quatre qu'il pleuve *exactement* 3mm sur tel bassin durant les prochaines 24h est une proposition mathématiquement correcte, mais très peu crédible.

dispose pas de suffisamment d'éléments[110] pour garnir la queue de distribution des pdf. C'est ainsi qu'il peut être amené à confondre les quantiles associés à des probabilités faibles[111].

Le resserrement voire l'égalité de certains quantiles est une forme de surconfiance (Krzysztofowicz *et al.*, 1993). Au final, ce comportement devrait conduire à fournir des intervalles de crédibilité trop étroits (surconfiance dans ses capacités à fournir une prévision précise), ou un quantile 90% trop bas (surconfiance dans ses capacités à exclure certaines valeurs élevées de la variable à prévoir).

La surconfiance a été étudiée en 7 et 8, et interprétée en supposant qu'elle révélait bien les convictions profondes du prévisionniste (ses croyances). Les résultats présentés dans cette section 9.1 suggèrent que l'inconsistance due au protocole expérimental peut également accroître le phénomène de surconfiance.

### 9.1.6 Discussion : faut-il placer le prévisionniste hors du monde ?

Il y a à nouveau antagonisme entre les deux écoles mélioriste et panglossienne (Stanovich, 1999) pour interpréter puis gérer ce problème d'inconsistance, c'est à dire de décalage entre ce que communique le prévisionniste et ce qu'il pense.

Une première idée, ancrée de longue date chez des météorologistes qu'on pourrait qualifier de Mélioristes (Sanders, 1963, Murphy, 1985, Murphy, 1991) et évoquée en section 7.1, est que les prévisionnistes auraient un jugement "interne" intrinsèquement probabiliste. « Ce jugement est fondé sur l'estimation de la vraisemblance de l'occurrence d'événements qui intéressent l'usager (...). Comme ce jugement représente les "croyances authentiques" du prévisionniste, il est intrinsèquement probabiliste[112]. » (Murphy, 1991).

L'élicitation des probabilités subjectives ne serait qu'un procédé qui permettrait de révéler ces pensées tout en douceur, sans les altérer. Cette élicitation ne ferait au contraire que libérer le prévisionniste de la tâche fastidieuse et délicate de « traduire des jugements probabilistes en prévisions catégoriques [ou déterministes] qui intègrent des considérations relatives à l'usager[113] » (Murphy, 1991).

Cependant, même dans la langue des probabilités, qui devrait permettre de formuler les incertitudes de prévision sans être influencé par les enjeux de l'usager (Murphy, 1985), le prévisionniste reste influencé par ceux-ci en montrant un biais de surestimation ("overforecasting", Krzysztofowicz et Sigrest, 1999, Murphy et Daan, 1984).

---

[110] Ces éléments peuvent être des notions ou concepts (risque, vraisemblance, probabilité, paris, etc.), des informations (visualisation des « tubes » (Atger, 1999) de la prévision d'ensemble lui permettant d'imaginer des scénarios extrêmes), ou sa simple expérience de prévision de phénomènes rares.

[111] On peut cependant trouver un terrain d'entente en associant des probabilités « parlantes » (une chance sur deux, une chance sur dix) à des grandeurs « parlantes » (seuils de l'exploitant qui retiennent la vigilance des prévisionnistes, valeurs historiques, ou valeurs « rondes » : 200 $m^3/s$, 50mm, 80km/h, etc.). D'où ce biais d'ancrage, qui provient donc tout autant de la nature de l'expérimentation que d'éventuels défauts des prévisionnistes.

[112] "In the context of subjective weather forecasting, forecasters formulate judgments regarding the likelihood of occurrence of the events of interest and then translate these judgments into forecasts. Since the judgments represent the forecasters' true beliefs, they are inherently probabilistic", (Murphy, 1991) p 303.

[113] "It would no longer be necessary for forecasters to devote any time or effort to the process of translating probabilistic judgment into categorical forecasts –a non trivial exercise involving (*inter alia*) user-related considerations", (Murphy, 1991) p 303.

La seule solution pour obtenir des prévisions fiables semble alors être l'enfermement du prévisionniste dans sa salle météo, avec interdiction absolue de tout contact avec l'usager[114].

Les Panglossiens quant à eux argumenteraient au contraire que ce qui pousse le prévisionniste à se lever (très tôt) le matin pour aller faire son travail de prévision, ce qui guide son analyse de la situation et oriente son attention, c'est sa conviction que sa prévision sera utile à l'usager. Une telle conviction est l'aiguillon qui guide son analyse experte de la situation météorologique et l'aide à anticiper sur ses évolutions.

Priver le prévisionniste de cette dimension téléologique essentielle à son métier d'expert en lui demandant de s'abstraire des enjeux des usagers (et de ce fait, de l'utilité de sa prévision), ne serait pas simplement impopulaire. Cela pourrait également priver le prévisionniste de certains repères, ne sachant notamment plus quel critère de performance il doit se définir pour s'améliorer[115].

Le critère de performance fait en effet partie intégrante du contexte social dans lequel le prévisionniste exerçait son métier en formulant puis en communiquant des prévisions. Celui-ci est notamment porté par une norme professionnelle qui régit le rapport prévisionniste-usager (et les rapports entre les prévisionnistes eux-mêmes).

Nous avons montré au Chapitre 7 deux exemples de norme professionnelle amenant à des résultats antagonistes concernant la fiabilité de l'estimation des incertitudes de l'expert :

- Soit le prévisionniste (tel le médecin spécialiste) doit passer pour celui qui sait, qui a un avis tranché, et afficher des incertitudes très faibles pour assurer ainsi son statut d'expert (Friedman et al. 2001). Il sera jugé sur la précision de ses prévisions.
- Soit le prévisionniste est reconnu comme un maillon indispensable à la Sûreté Hydraulique, en ce qu'avant tout le monde, il peut alerter sur le risque d'occurrence de phénomènes intenses. Dans ce cas, il sera jugé sur la détectabilité (Chapitre 7).

Privé de contexte social, ou refusant le caractère arbitraire d'un critère de performance qui ne soit pas fondé sur l'intérêt de la prévision pour l'usager, le prévisionniste pourrait se montrer indécis et versatile dans la production de ses prévisions.

Ce type d'inconsistance, est, à l'inverse, produit par l'absence de contact avec l'usager. Dans ce cas, l'expérimentation mise en place ne serait pas représentative (au sens de Brunswik, 1956) des conditions naturelles, "écologiques", dans lequel l'expert a coutume d'exercer son métier de prévisionniste.

Nous avons donc présenté les deux faits antagonistes. D'un côté, c'est un fait éprouvé que les prévisionnistes montrent de l'inconsistance lorsqu'ils intègrent les enjeux de l'usager dans la communication de leurs prévisions (Krzysztofowicz et Sigrest, 1999, Murphy et Daan, 1984). Cette inconsistance se traduit par un défaut de fiabilité des prévisions (biais de surestimation). Cependant, à vouloir monter un protocole de prévision "neutre" en coupant de manière radicale les prévisionnistes de toute considération des enjeux de l'usager, on peut voir

---

[114] Interdiction également de sortir de sa salle météo pour quelque motif que ce soit, afin qu'il n'ait pas connaissance des vulnérabilités spatiales du territoire sur lequel il se doit de produire des prévisions "neutres". Les feed-back des usagers seraient en particulier proscrits, puisqu'ils pourraient influencer le prévisionniste par la prise en compte d'enjeux ponctuels de ceux-ci, indépendants de la qualité "intrinsèque" de la prévision (critère du type écart quadratique moyen entre le champ des prévus et celui des observés). Le canal d'information entre prévisionniste et usager devrait donc être impérativement à sens unique.

[115] Seul un prévisionniste autant passionné de météo que de jeux mathématiques et logiques pourrait considérer cette tentative d'abstraction des conséquences de sa prévision comme un nouvel exercice intellectuel.

apparaître un "biais académique". Les prévisionnistes ne percevant plus l'intérêt de leur prévision pour l'usager, ils risquent de perdre eux-mêmes leur source de motivation et de se montrer inconsistants dans leurs prévisions. En réponse à des questions trop abstraites, et décalées de leurs préoccupations quotidiennes, les prévisionnistes pourraient alors se montrer incohérents (contradictions logiques entre leurs différentes réponses).

Ainsi, un mode opératoire en décalage avec l'environnement social habituel du prévisionniste pourrait entraîner une autre forme d'inconsistance que celle habituellement décrite dans la littérature.

### 9.1.7 De la nécessité d'intégrer les enjeux de l'usager dans le processus d'élaboration de la prévision

Selon Piaget (1967), « connaître ne consiste pas à copier le réel, mais à agir sur lui et à le transformer (en apparence ou en réalité), de manière à le comprendre en fonction des systèmes de transformation auxquels sont liées ces actions. ».

Toute connaissance répond à un questionnement ; il n'y a pas de connaissance neutre, absolue. Ce questionnement qu'a le prévisionniste face à une situation météo est alimenté par l'intérêt qu'a l'usager pour la prévision qui lui sera fournie. Le « que me demande-t-on ? » amène le « quelles sont les questions à me poser ? » ; puis « quels sont les scénarios météo amenant à des situations pouvant avoir un impact sur les enjeux de l'usager ? »

Sans cette dimension téléologique, le savoir de l'expert risque d'être inepte, et sa pensée décousue[116]. En ce sens, l'existence de « prévisions internes, isolées par définition de considérations relatives à leur utilisation » défendue par Murphy (1991)[117] est discutable.

Il me semble donc important d'intégrer dans le processus de décision de l'expert cette dimension téléologique, quitte à agir sur elle : fixer ensemble de nouveaux objectifs d'amélioration du système de prévision, pour précisément le placer au plus près des enjeux des usagers. Nous avons vu en section 7 que changer cet objectif pourrait du même coup transformer la norme sociale à laquelle sont soumis les prévisionnistes.

Intégrer les enjeux de l'usager pourrait donc amener un double bénéfice : un système de prévision plus performant pour répondre à ses enjeux, et des prévisions plus fiables (meilleure estimation de l'incertitude), comme nous l'avons montré en section 7.

---

[116] A quel paramètre faut-il s'intéresser ? Pourquoi ? Dans ce cas, dois-je regarder le scénario le plus probable, ou dois-je porter mon attention sur la possible occurrence de phénomènes à risque. Ces questions montrent que l'objectif précède la construction du jugement, et lui est nécessaire.

[117] "Since internal forecasts by definition are isolated from considerations related to their use, little if any justification exists for not making these forecasts correspond to the respective judgments" (Murphy, 1991).

## 9.2 Interlude : peut-on séparer parole et pensée ?

Le prévisionniste est-il en désaccord avec ce qu'il communique à l'extérieur du centre de prévision (Murphy, 1991) ?

Si tel est le cas, comment percevoir le jugement intime du prévisionniste, avant qu'il ne le travestisse en prévisions adaptées aux enjeux des usagers ? Comment peut-on dissocier parole et pensée ? Et tout d'abord, cette disjonction entre parole et pensée est-elle légitime ?

Pour que le prévisionniste soit effectivement en désaccord avec ce qu'il communique, il faut que sa pensée ait été préalablement formulée, explicitée et construite. Il peut alors se rendre compte que celle-ci ne correspond pas aux attentes des usagers. C'est ce que suggère Murphy (1991), faisant certainement part de sa propre expérience de prévisionniste qui n'a parfois pas communiqué des prévisions en accord avec ses convictions profondes, afin de « se couvrir ».

C'est également ce que suggère R. Garçon (comm. pers.), en donnant un exemple récent. Durant l'été 2003, l'expert sus-nommé fut sollicité par la direction de l'entreprise préoccupée de la canicule, en particulier par la température des fleuves au droit des centrales nucléaires (et qui en sont la source froide). Dans leur dialogue avec les ministères, les décideurs d'EDF voulaient identifier le scénario le plus pessimiste, pour préparer les parades associées[118]. Les décideurs d'EDF, associant intuitivement ce scénario pessimiste à une poursuite de la canicule, demandent alors explicitement au prévisionniste de fournir des scénarios de température d'eau « dans l'hypothèse d'un maintien de la canicule ». Le prévisionniste, intimement convaincu qu'un scénario extrêmement pessimiste est une diminution progressive et modérée de la température de l'air se voit obligé par la nature de la demande de l'usager de fournir un scénario trop peu vraisemblable à ses yeux (c'est à dire un scénario dont la probabilité est très voisine de 0). Le prévisionniste fournit finalement la prévision demandée, basée sur le scénario trop peu vraisemblable, plus une prévision basée sur un scénario qu'il juge pessimiste, mais suffisamment vraisemblable pour qu'on puisse en tenir compte.

Les interrelations entre parole et pensée ont particulièrement été étudiées en psychologie du développement (Piaget, 1947, Gelman et Bymes, 1992), et la question de l'antécédence de l'une par rapport à l'autre a toujours fait débat. Selon Chomsky (2000) - qui insiste sur le fait qu'on ne peut trouver la réponse que par introspection - les idées précèdent les mots[119]. Pour Vygotsky (1962) au contraire, le langage fonde la pensée. Selon cet auteur, le développement cognitif résulte d'une appropriation du langage, les opérations mentales prenant corps dans la structure du langage. Entre ces deux extrêmes, Piaget (1947) adopte une position plus

---

[118] En théorie de la décision, cette stratégie correspond à celle bien connue du Minimax : identifier toutes les scénarios possibles, leurs conséquences associées pour le décideur, et engager toutes les actions possibles pour réduire ces conséquences. C'est la stratégie dite de l'assuré : « se prémunir contre tous les accidents de la vie ». Une telle politique appliquée strictement conduirait rapidement à rassembler tous les moyens disponibles pour se protéger contre une éventuelle attaque d'extra-terrestres qui voudraient anéantir la planète Terre. C'est pourquoi le décideur ne considère usuellement que les scénarios « suffisamment vraisemblables ». Néanmoins, dans la mesure où une telle stratégie ne tient pas compte des probabilités associées aux divers scénarios, elle reste bien entendu irrationnelle.

[119] "That is the traditional question: can you have thought without language? If you ask how much we know about that topic, the answer is "not much". What we know is by introspection. Now what seems to me obvious by introspection is that I can think without language. In fact, very often, I seem to be thinking and find it hard to articulate what I am thinking. It is a very common experience at least for me and I suppose for everybody to try to express something, to say it and to realize that is not what I meant and then to try to say it some other way and maybe come closer to what you mean; then somebody helps you out and you say it in yet another way. That is a fairly common experience and it is pretty hard to make sense of that experience without assuming that you think without language." Chomsky (2000), p 76.

nuancée : langage et pensée se nourrissent mutuellement, le langage nécessitant la pensée pour le développer, et réciproquement.

La position de Murphy (1991), qui rejoint donc celle de Chomsky (2000) est discutable dans la mesure où elle suppose l'existence de pensées préétablies, substantielles, qui planent dans la tête du prévisionniste en attendant, en cas de sollicitation, de trouver des mots adéquats pour les traduire.

L'idée que la pensée précède la parole peut ainsi être remise en cause. La formulation, même intérieure d'une idée, pensée, remarque, réflexion, commentaire, les font exister, les font éclore, leur donne vie. Notre pensée ne précède pas le mot, elle naît au moment même où nous la formulons. « Nos pensées naissent toutes habillées » (Oscar Wilde).

S'il l'on admet que la pensée et la parole jaillissent d'un même élan, il apparaît manifeste que tout processus d'élicitation n'est jamais neutre. La parole structure la pensée, et la manière de faire naître la parole influence la formulation de la pensée.

- D'une part, le processus d'élicitation oriente forcément le questionnement du sujet, qui ne peut manier que des concepts auxquels il a donné du sens (Piaget, 1967). D'où l'importance de « faire comprendre » aux prévisionnistes ces nouveaux concepts et notions et que sont le pari, la vraisemblance, la probabilité. Non par une démarche mélioriste d'exercices de perfectionnement, mais en leur montrant l'intérêt de ces nouveaux outils, l'objectif auquel ils permettent de tendre. Objectif qui est aussi le leur : améliorer le système de prévision.
- D'autre part, ce processus structure la pensée en faisant naître la parole de manière logique et méthodique. D'où l'importance d'éduquer l'usager pour qu'il puisse recevoir cette nouvelle parole. Cette éducation se fera également en partant de considérations téléologiques afin de montrer comment l'outil est construit pour répondre à des objectifs essentiels pour l'usager. Le « à quoi ça sert » guidera le « comment ça marche ».

Les résultats présentés en 7 puis en 9.1 ainsi que ces quelques considérations auront peut-être contribué à écorner l'image d'une prévision "absolue", "neutre", dans un monde idéal où le prévisionniste pourrait livrer son jugement expert sur l'évolution météorologique, enfin débarrassé des contingences liées à l'utilisation de la prévision.

Au contraire, la prise en compte de l'intérêt de la prévision pour l'usager pourrait permettre au prévisionniste et à l'usager de construire ensemble de nouveaux outils plus appropriés à la gestion du risque.

En ce sens, les probabilités constituent un outil qui permet d'aborder la prévision des phénomènes à risque de manière quantitative. Ce caractère quantitatif adopté par l'expert peut lui permettre d'être lucide et transparent dans la formulation puis la communication de ses incertitudes de prévision. Ce langage peut également éviter des incompréhensions liées au flou et à l'arbitraire du langage qualitatif et informel (Weber et Hilton, 1990).

## 9.3 Le prévisionniste face au pari : "Entre risque et incertitude, mon cœur balance".

### 9.3.1 Introduction

Nous avons vu en 9.1 que le protocole lui-même pouvait amener les prévisionnistes à des inconsistances. En particulier, ceux-ci ont commis des incohérences entre quantiles émis et paris choisis.

Dans cette section 9.3, nous analysons plus en détail quelles pourraient être les raisons de ces erreurs. S'agit-il seulement d'erreurs aléatoires dues à la difficulté de l'exercice, ou à son caractère déconnecté des conditions de travail opérationnel ? Ou serait-ce la marque que les prévisionnistes sont peu enclins à parier sur leurs propres prévisions ?

Ceux-ci sont en effet appelés à choisir entre un pari en univers risqué (où les probabilités sont précisément connues), et un pari en univers incertain (ou les probabilités sont inconnues, et ne sont que l'expression d'un degré de croyance aux yeux du prévisionniste, note n°103).

La distinction entre risque et incertitude a été faite pour la première fois par Knight (1921). Ellsberg (1961) a ensuite montré que les individus préféraient prendre des paris sur la base de probabilités connues plutôt que sur la base de probabilités inconnues[120]. De manière plus générale, les sujets préfèrent prendre des paris lorsqu'ils sont compétents sur les sources d'incertitude associées : les passionnés de football préfèreront prendre des paris sur le vainqueur de la coupe des clubs professionnels, alors que les passionnés de politique préfèreront parier sur le vainqueur de l'élection présidentielle (Health et Tversky, 1991).

Fox et Tversky (1995) ont affiné cette analyse en montrant que « l'aversion à l'ambiguité[121] est gouvernée par le sentiment de compétence ». Ces auteurs avancent l'hypothèse que le sentiment de compétence n'affecte les décisions du sujet que lorsque celui-ci est amené à prendre conscience de sa relative compétence (ou incompétence). Le signal qui produit chez le sujet le sentiment de compétence ou d'incompétence est la comparaison directe à des sources d'incertitudes qui lui sont plus familières, ou à des sujets plus cultivés ("comparative ignorance hypothesis", Fox et Tversky, 1995).

Fox et Weber (2002) ont procédé à diverses manipulations des sujets, en faisant varier des éléments du contexte dans lequel ceux-ci étaient amenés à prendre une décision. Ils ont par exemple obtenu le résultat suivant : les sujets trouvent les paris moins attractifs lorsque ceux-ci sont accompagnés d'informations qu'ils ne savent pas utiliser, comparativement aux mêmes paris qui leur sont proposés sans ces informations.

Ainsi, de manière générale, le contexte de la prise de décision peut comporter des signaux mettant en exergue des états de connaissance contrastés. Ces signaux induisent chez le sujet

---

[120] L'exemple (Ellsberg, 1961) des urnes contenant des boules rouges et noires est resté célèbre : l'urne A contient des boules rouges et noires en proportion 50-50, tandis que l'urne B contient des boules rouges et noires en proportion inconnue. Au jeu « Choisissez une couleur (rouge ou noire) ; on va tirer une boule dans une urne, et si c'est votre couleur qui sort, vous empochez $100 », les gens sont indifférent dans le choix d'une couleur pour chacune des urnes, mais préfèrent parier su l'urne A plutôt que sur l'urne B.

[121] La notion d'ambiguïté permet de dépasser la distinction entre le risque et l'incertitude, et de classer par source d'incertitude de plus en plus "vague" : les probabilités connues précisément ("gamble"), celles qui ne le sont pas (météo, course de chevaux, etc.), puis celles qui le sont de moins en moins (questions posées sur des domaines où la connaissance du sujet est de plus en plus vague).

un état d'esprit éphémère de compétence ou d'incompétence qui gouverne sa volonté d'agir en univers incertain (Fox et Weber, 2002).

Cette aversion à l'incertitude, à la nature de la source d'incertitude (ambiguïté), ou à des éléments de contexte est bien une irrationalité au sens d'une violation des axiomes de Savage (1954). Ces éléments de contexte peuvent être, pour le cas qui nous intéresse :

- univers risqué (pari sur un tirage dans une urne) ou univers incertain (pari sur des évolutions météorologiques)
- ordre suivant lequel on procède à l'élicitation de quantiles et la prise de pari. Suivant l'ordre dans lequel on réalise ces deux expériences, on peut soit créer l'atmosphère d'un examen, soit permettre un galop d'essai pour faciliter la formulation de la prévision.
- inter-comparaison entre prévisionnistes, entre situations météorologiques, ou entre bassins versants qui peuvent induire un sentiment relatif de compétence ou d'incompétence.

De tels éléments de contexte peuvent produire des irrationalités telles que des incohérences entre quantiles émis et paris pris. Nous analysons dans cette section en quoi le protocole adopté dans chaque journée d'expérience au CHA et au CHPMC a pu amener les prévisionnistes à une aversion à l'ambiguïté.

### 9.3.2 Protocole

L'expérience a été décrite en 7.3.1, les questionnaires sont fournis en Annexe 2. Le contexte plus précis décrit en 9.1.2 a montré que les prévisionnistes avaient été plus placés en situation d'examen dans l'expérience au CHA que dans celle au CHPMC.

Les termes utilisés au CHA, tels que « contrôle de cohérence », ainsi que le fait de faire parier les prévisionnistes sur leur propre prévision, pouvaient induire le doute dans l'esprit des prévisionnistes. Ils étaient invités lors du choix des paris à remettre en question la prévision qu'ils venaient d'émettre. Cette invitation peut constituer un signal produisant chez les prévisionnistes un sentiment de relative incompétence (selon les mots de Fox et Tversky, 1995).

Nous avons vu en 9.1.2 comment le protocole avait été modifié au CHPMC pour que les prévisionnistes soient mis en confiance :

- meilleure explication de ce nouveau concept des probabilités ;
- choix de paris sur des données "neutres", choisies par les expérimentateurs ;
- les paris, placés avant l'élicitation des quantiles, permettaient de faciliter la formulation de leur jugement.

Nous avions ainsi tenté de réduire au maximum tous les signaux qui pouvaient les amener à se sentir particulièrement incompétents.

### 9.3.3 Résultats obtenus au CHPMC

Voici deux exemples d'incohérence relevés lors du stage, pour une situation météorologique choisie :

1/ Exemple d'une prévision sur la Garonne à Palaminy :

Dans un premier temps, le prévisionniste confronté au choix ci-dessous choisit le jeu A.

***Choix n°-1***

**Jeu A** : Dans une urne, il y a 2 boules rouges, et 3 boules bleues. On tire une boule au hasard.

Si c'est une boule bleue qui sort, vous empochez 100 FF ; si c'est une rouge, vous êtes quitte.

**Jeu B** : Si la lame d'eau spatiale sur les 24h à venir (cumul 8h-8h) dépasse 5 mm sur la Garonne à Palaminy, vous empochez 100 FF ; sinon, vous êtes quitte.

***A quel jeu préférez-vous jouer ?***

Ce choix semble révéler que le prévisionniste estime qu'il y a moins de 3 chances sur 5 que la lame d'eau spatiale sur la Garonne à Palaminy dépasse 5 mm. Cela équivaut à considérer que le quantile qui a 3 chances sur 5 d'être dépassé, c'est à dire le quantile $L_{60}$, est inférieur à 5 mm.

Dans un second temps, le prévisionniste interrogé directement sur le quantile 50 donne l'estimation suivante : $L_{50} = 10$ mm

Finalement, le prévisionniste aura d'une part estimé que $L_{60} \leq 5$ mm en choisissant le pari risqué plutôt que le pari incertain. D'autre part, il a estimé que $L_{50} = 10$ mm. Dans cet exemple, le pari choisi est en contradiction avec le quantile estimé, puisque $L_{60} \leq L_{50}$.

Cette incohérence a dans cet exemple montré que le prévisionniste préférait prendre un pari risqué plutôt qu'un pari incertain. Ce résultat pourrait alors suggérer un attrait pour le risque au détriment de l'incertitude.

2/ Exemple d'une prévision sur l'Arn au Vintrou :

Dans un premier temps, le prévisionniste confronté au choix ci-dessous choisit le jeu B.

***Choix n°-3***

**Jeu A** : dans une urne, il y a 3 boules rouges, et 2 boules bleues. On tire une boule au hasard.

Si c'est une boule bleue qui sort, vous empochez 100 FF ; sinon vous êtes quitte.

**Jeu B** : Si la lame d'eau spatiale sur les 24h à venir dépasse 10 mm sur l'Arn au Vintrou, vous empochez 100 FF ; sinon vous êtes quitte.

***A quel jeu préférez-vous jouer ?***

Ce choix semble révéler que le prévisionniste estime qu'il y a plus de 2 chances sur 5 que la lame d'eau spatiale sur l'Arn au Vintrou dépasse 10mm. Cela équivaut à considérer que le quantile qui n'a 2 chance sur 5 d'être dépassé, c'est à dire le quantile $L_{40}$, est supérieur à 10 mm.

Dans un second temps, le prévisionniste interrogé directement sur le quantile 50 donne l'estimation suivante : $L_{50} = 15$ mm

Finalement, le prévisionniste aura d'une part estimé que $L_{40} \geq 10$ mm en choisissant le pari incertain plutôt que le pari risqué. D'autre part, il a estimé que $L_{50} = 15$ mm. Ici, on a donc $L_{40} \geq L_{50}$.

Cette incohérence a dans cet exemple montré un attrait pour l'incertitude au détriment du risque.

Globalement, sur les 7 possibilités d'erreur d'inversion des préférences entre quantiles émis et paris choisis (0), 4 erreurs marquent un attrait pour le risque, et 3 un attrait pour l'incertitude :

4 inversions des préférences marquant un attrait pour le pari risqué :

$L_{92} \leq b$ et $L_{90} \geq b$ $L_{86} \leq b$ et $L_{75} \geq b$ $L_{60} \leq a$ et $L_{50} \geq a$ $L_{40} \leq a$ et $L_{25} \geq a$

3 inversions des préférences marquant un attrait pour le pari incertain :

$L_{86} \geq b$ et $L_{90} \leq b$ $L_{60} \geq a$ et $L_{75} \leq a$ $L_{40} \geq a$ et $L_{50} \leq a$

Une première hypothèse est que ces incohérences sont la marque d'une réelle préférence des prévisionnistes pour le pari risqué ou le pari incertain.

L'hypothèse opposée est que les inversions ne témoignent que d'erreurs dues à des hésitations (en particulier au voisinage du point d'indifférence) ou à des étourderies, et seraient donc aléatoires. Dans ce cas, la probabilité $\boldsymbol{\rho}$ de choisir le pari risqué lors d'une inversion des préférences vaudrait 4/7, puisque des 7 possibilités d'erreur, 4 amènent à choisir un pari risqué (4/7 vaut environ 0,57).

Le Tableau 34 recensait le nombre d'inversions des préférences commises par les prévisionnistes avant et après feed-back : respectivement 29 et 25.

Le Tableau 35 estime à présent la probabilité $\boldsymbol{\rho}$ de choisir le pari risqué lors d'une inversion des préférences. On observe que 0,57 est dans l'intervalle de confiance à 95% sur $\boldsymbol{\rho}$, même avant feed-back.

| **Stage de formation au CHPMC** | **Avant feed-back** | **Après feed-back** |
|---|---|---|
| n, nombre d'erreurs commises d'inversions des préférences | 29 | 25 |
| k, nombre d'inversions des préférences marquant un attrait pour le pari risqué | 19 | 14 |
| Estimateur de la probabilité $\boldsymbol{\rho}$ de choisir le pari risqué lors d'une inversion des préférences | 0,66 | 0,56 |
| $IC_{95\%}$ [$\boldsymbol{\rho}$\|n,k] | [0,48 ; 0,81] | [0,37 ; 0,74] |

**Tableau 35 : Probabilité qu'un prévisionniste choisisse le pari risqué plutôt que le pari incertain lors d'une incohérence entre quantiles émis et paris pris. Stage de formation au CHPMC.**

Ces résultats montrent donc que même avant feed-back, les prévisionnistes n'ont montré aucune préférence pour le risque ou l'incertitude décelable. Nous concluons donc que les incohérences proviennent sans doute d'étourderie, et n'ont pas un caractère systématique.

Ni l'analyse des réponses pour chaque prévisionniste (Annexe 3), ni les discussions informelles qui ont suivi cette expérience n'ont suggéré d'aversion notable à l'incertitude.

### 9.3.4 Résultats obtenus au CHA

Pour un bassin donné, sur les quatre erreurs possibles, 2 marquent un attrait pour le risque, et 2 un attrait pour l'incertitude :

2 inversions des préférences marquant un attrait pour le pari risqué :

$L_{90} \geq L_{92}$ $L_{50} \geq L_{60}$

2 inversions des préférences marquant un attrait pour le pari incertain :

$L_{90} \leq L_{86}$ $L_{50} \leq L_{40}$

En notant **ρ** la probabilité de choisir le pari risqué lors d'une inversion des préférences, c'est cette fois l'hypothèse **ρ**=0,5 qui indique l'indifférence du prévisionniste entre risque incertitude.

Le Tableau 36 montre que cette hypothèse est acceptable (0,5 est proche du maximum de vraisemblance de la loi de **ρ** a posteriori).

| **Stage de formation au CHA** | **Avant feed-back** | **Après feed-back** |
|---|---|---|
| n, nombre d'erreurs commises d'inversions des préférences | 13 | 17 |
| k, nombre d'inversions des préférences marquant un attrait pour le pari risqué | 8 | 8 |
| Estimateur de la probabilité **ρ** de choisir le pari risqué lors d'une inversion des préférences | 0,62 | 0,47 |
| $IC_{95\%}$ [**ρ**\|n,k] | [0,35 ; 0,85] | [0,25 ; 0,70] |

**Tableau 36 : Probabilité qu'un prévisionniste choisisse le pari risqué plutôt que le pari incertain lors d'une incohérence entre quantiles émis et paris pris. Stage de formation au CHA.**

L'analyse des réponses pour chaque prévisionniste (Annexe 3), ne permet pas non plus de distinguer de prévisionniste montrant une aversion particulière à l'incertitude.

En revanche, une discussion informelle avec l'un des prévisionnistes a montré qu'il lui semblait bien plus délicat « vu la situation météorologique particulièrement complexe (marais barométrique) » de parier sur ses propres prévisions que sur le tirage de boules dans une urne. Ce prévisionniste a toujours choisi les paris risqués quels qu'en soient les enjeux, annotant dans la marge pour justifier son choix que « de toute façon [il] n'était pas sûr de ses prévisions ».

### 9.3.5 Interprétation et discussion

La "comparative ignorance hypothesis" de Fox et Tversky (1995), reprise par Fox et Weber (2002) indique que si les prévisionnistes reçoivent de leur environnement (i.e. leur contexte de décision) un signal qui les amène à se sentir incompétents à propos de la question qui leur était posée, ceux-ci montreront une aversion au pari associé. Ce type d'aversion est donc interprété par ces auteurs comme une aversion relative à une source d'incertitude, perçue comme "plus ambiguë", et pour laquelle le sujet se sent incompétent.

Au CHA le prévisionniste ayant montré une très forte aversion à l'incertitude pourrait avoir perçu ce type de signal. L'atmosphère d'examen que pouvait produire la phase de « vérification de cohérence » a pu l'amener à douter de ses compétences, et de ce fait produire une aversion à cette source d'incertitude.

Les résultats obtenus au CHPMC suggèrent en revanche une indifférence entre les paris sur les évolutions météorologiques et les paris sur le tirage d'une boule dans une urne dont la composition est connue.

Cela indique qu'aucun signal ne les a amenés à se sentir particulièrement compétents ou incompétents à propos de la prévision météorologique. Autrement dit, aucun effet comparatif ne s'est s'exercé :

- toutes les questions d'une même série étaient homogènes : même discipline (donc pas de possibilité d'intercomparaison avec ses compétences hippiques ou footbalistiques)

même situation météorologique (donc pas de possibilité de se sentir plus familier avec un type d'épisode météo plutôt qu'un autre) ;

- pas d'intercomparaison entre prévisionnistes : ceux-ci étaient situés dans des salles différentes, et on les a assurés que les questionnaires seraient traités anonymement ;
- explication plus approfondie des probabilités subjectives. Cela a ôté une part de mystère autour de ce concept (note n°103), diminuant ainsi l'éventuel sentiment d'incompétence par rapport à cette nouvelle notion ;
- les paris étaient placés avant l'élicitation des quantiles. Ceux-ci ne remettaient pas en cause les chiffres qu'auraient pu produire les prévisionnistes. A l'opposé d'un "contrôle de cohérence", placer d'abord les paris avant l'élicitation des quantiles constituait plutôt une aide au prévisionniste pour mieux formuler son jugement et les incertitudes associées.

Les deux derniers points peuvent expliquer que l'on ait observé de l'aversion à l'incertitude au CHA et non au CHPMC.

Les résultats confirment donc les conclusions de Fox et Tversky (1995) : l'aversion à l'incertitude n'est produite que lorsque les sujets perçoivent un signal qui les amène à se sentir incompétents.

Nous nous étions fixé pour objectif de « développer des méthodes interactives et intuitives d'élicitation des probabilités subjectives. » (7.6). Nous avons donc essayé d'être aussi interactifs que possible, comme le serait le prévisionniste au contact de l'usager, notamment en privilégiant les questions-réponses auxquelles ils sont habitués : « vraisemblance[122] que la pluie dépasse 10mm ? 30mm ?, etc. ». Dans notre cas, plutôt que d'avoir un usager qui leur fait part de ses enjeux opérationnels[123], on confrontait directement le prévisionniste à des paris dont les conséquences et les probabilités étaient connues, affichées, et internalisées par le prévisionniste (c'est lui, et non l'usager, qui gagnait ou qui perdait 100FF, même si ces jeux étaient factices).

Les résultats obtenus au CHPMC montrent que l'objectif a été partiellement atteint. Un tel protocole peut permettre à l'expert de se prononcer sans montrer d'aversion à l'incertitude, à condition de prendre garde à ce que des éléments de contexte n'induisent pas le sentiment d'incompétence chez la personne interrogée.

Malgré l'absence d'aversion à l'incertitude, notre protocole n'a malgré tout pas été suffisamment intuitif pour éviter les défauts d'inconsistance du prévisionniste (9.1). Confronter le prévisionniste à la fois à un choix de paris incertains ou risqués, et à une élicitation directe de plusieurs quantiles peut donc perturber son jugement. L'élicitation directe des quantiles est plus intuitive car plus proche des enjeux de l'usager (les quantiles élevés sont particulièrement appropriés aux enjeux de sûreté), et permet au prévisionniste un apprentissage plus facile (le prévisionniste peut lui-même réaliser plus aisément des feed-back, en regardant si ses quantiles émis sont dépassés avec la bonne fréquence). C'est

---

[122] Dans le langage courant de la prévision opérationnelle, classiquement qualitatif et informel, on emploie « vraisemblance », « risque », ou « probabilité » de la même façon pour exprimer le degré de certitude avec le quel on estime qu'un seuil ne sera pas dépassé.

[123] Les aller-retours d'information permettent à l'usager de communiquer sur ses enjeux opérationnels. D'une part, ses niveaux de seuil : « à 30$m^3$/s, il faut que je modifie mes plans de production, que je passe en consigne de crue, que je protège mon chantier dans le lit de la rivière, etc. » et d'autre part, ses préférences, qui révèlent les conséquences associées en terme de dommages (désutilité pour le décideur) : « si vous estimez que le risque est "suffisant", je préfère prendre les mesures nécessaires plutôt que de risquer de voir mes engins noyés »),

pourquoi, dans leur expérience « au fil de l'eau », les prévisionnistes du CHPMC fournissent des quantiles sur chacune de leurs prévisions, sans se prononcer explicitement sur des paris.

Des discussions informelles avec les prévisionnistes du CHPMC ont néanmoins montré que la notion de probabilité subjective avait fait son chemin : les prévisionnistes évaluent chaque quantile en y associant spontanément, sinon une idée de pari, des notions de conviction, de croyance ou de vraisemblance. Cela suggère donc que les quantiles ainsi émis au fil des jours reflètent bien le jugement du prévisionniste, et qu'en cela, la méthode semble constituer un bon outil de formalisation des incertitudes de prévision.

On pourra retenir la conclusion suivante : placé dans de bonnes conditions, un prévisionniste peut révéler de façon fiable son jugement sur les variables météorologiques à venir sans que sa pensée ne soit déformée par une forme d'aversion à l'incertitude, ou à l'ambiguïté.

## 9.4 Bibliographie

Atger, F., 2000. La prévision du temps à moyenne échéance en France, La Météorologie, vol. 8, 30, 61-86.

Atger, F., 1999. Tubing : an alternative to clustering for the classification of ensemble forecasts. Weather and forecasting, vol. 14.

Bachelard, G., 1934, Le nouvel esprit scientifique, Presses universitaires de France (6e Ed.), 183p.

Brunswik, E., 1956. Perception and the representative design of psychological experiments (2nd Ed.), Berkley: University of California Press.

Chomsky, N., 2000. The Architecture of Language, Oxford University Press.

Cosmides, L., Tooby, J., 1996. Are humans good intuitive after all? Rethinking some conclusions from the literature on judgment under uncertainty. Cognition, 58, 1-73.

Doswell, C. A., Flueck, J. A., 1989. Forecasting and verifying in a filed research project: DOPLIGHT '87. Weather and Forecasting, 4, 97-109.

Ellsberg, D., 1961. Risk, ambiguity and the Savage axioms. The Quarterly Journal of Economics, vol. 75, 3, 643-669.

Fox, C., Tversky, A., 1995. Ambiguity aversion and comparative ignorance. The Quarterly Journal of Economics, vol. 110, 3, 585-603.

Fox, C., Weber, M., 2002. Ambiguity aversion, comparative ignorance, and decision context. Organizational behavior and human decision processes, vol. 88, 1, 476-498.

Friedman, C., Gatti, G., Elstein, A., Franz, T., Murphy, G., Wolf, F., 2001. Are clinitians correct when they believe they are correct? Implications for medical decision support, Medinfo, Vol. 10, 1, 454-458.

Gelman, S., A., Byrnes, J., P., 1992. Perspectives on Language and Thought, Interrelations in Development, Cambridge University Press, 538 p.

Health, C., Tversky, A., 1991. Preference and belief: ambiguity and competence in choice under uncertainty. Journal of risk and uncertainty, vol. 4, 5-28.

Knight, F., H., 1921. Risk, uncertainty and profit. Boston, Houghton Mifflin.

Le Moigne, J.-L., 1990. La modélisation des systèmes complexes, Dunod, Paris, 178p.

Krzysztofowicz, R., Drzal, W., J., Drake, T., R., Weyman, J., C., Giordano, L., A., 1993. Probabilistic quantitative precipitation forecasts for river basins. Weather and Forecasting, vol. 8, 424-439.

Krzysztofowicz, R., Sigrest, A., 1999. Calibration of probabilistic quantitative precipitation forecasts. Weather and Forecasting, 14(3), 427-442.

Murphy, A. H., 1993. What is a good forecast ? An essay on the nature of goodness in weather forecasting. Weather and forecasting, vol. 8, no 2, 281-293.

Murphy, A. H., 1991. Probabilities, odds, and forecast of rare events, Weather and Forecasting, vol. 6, no 2, 302-307.

Murphy, A. H., 1985. Probabilistic weather forecasting, *in* Probability, statistics and decision making in the atmospheric science, Murphy, A. H and Katz, R., W., Eds, Westview Press, 337-377.

Murphy, A. H., Daan, H., 1984. Impacts of feedback and experience on the quality of subjective probability forecast : Comparison of results from the first and second years of the Zierikzee Experiment, Monthly Weather Review, vol. 112, 413-423.

Murphy, A., H., Winkler, R., L., 1974. Credible interval temperature forecasting: some experimental results. Monthly Weather Review, vol. 102, 784-794.

Palmarini, M., P., 1995, La réforme du jugement, ou comment ne plus se tromper, Odile Jacob (trad. De L'Illusione di sapere, Arnoldo Editore, 1993).

Piaget, J., 1947. Psychologie de l'intelligence, Armand Colin (Rééd. 2001), Pocket, 239 p.

Piaget, J., 1967. Biologie et connaissance. Gallimard, Paris, 1345 p.

Sanders, F.,1963. On subjective probability forecasting, Journal of Applied Meteorology, vol. 2, 2, 191-201.

Savage, L. J., 1954, The Foundations of statistics. Wiley.

Stanovich, K. E., 1999. Who is rational? Studies of individual differences in reasoning. Mahweh, NJ: Erlbaum.

Vygotsky, L., S., 1962. Thought and langage, Cambridge, MIT Press.

Weber, E., U., Hilton, D., J., 1990. Contextual effect in the interpretations of probability words: perceived base rate and severity of events, Journal of experimental psychology, Human perception and performance, vol. 16, 4, 781-789.

# 10 Conclusion générale

## 10.1 Bilan : finalement, que vaut le système de prévision ?

L'étude du cas classique des vecteurs de prévision déterministe des débits de J à J+4 (exemple de la vallée de l'Ain), nous a donné les ordres de grandeur suivants : la prévision permet d'augmenter de 1 à 3% les gains réalisés par le turbinage. En terme de valeur de l'information pour l'usager, la prévision déterministe opérationnelle se place à 30% voire 50%, sur une échelle allant de 0, pour la prévision naïve (distribution historique) à 1 (avenir connu).

Dans un cas où les incertitudes sont quantifiées (les prévisions hydrologiques saisonnières, données sous forme de 50 vecteurs de prévision équiprobables), l'exemple de la gestion du réservoir de Serre-Ponçon nous a permis d'estimer la valeur de l'information délivrée : la prévision permet d'augmenter de 1% à 8% les gains réalisés par turbinage (suivant le volume de neige de l'année, et le type de gestion : avec ou sans contrainte de cote touristique à respecter). La valeur relative d'une telle prévision est cette fois de l'ordre de 80% voire plus de 90%, entre prévision naïve et prévision parfaite.

*L'analyse économique développée sur ces deux exemples pourrait s'appliquer à l'ensemble de la production hydraulique. Compte-tenu du fait qu'une partie du parc de production n'est pas programmable (production au fil de l'eau, sans capacité de stockage), la prévision doit permettre de réaliser des gains de l'ordre du pour-cent pour l'ensemble de la production hydraulique.*

Quelle que soit la portée temporelle de la prévision, celle-ci offre à l'usager une meilleure anticipation des apports, lui permettant d'optimiser les placements de la production hydroélectrique, et de valoriser celle-ci aux moments les plus avantageux. La prévision à court terme lui permet également d'anticiper d'éventuels problèmes de saturation sur une vallée hydroélectrique (usine amont suréquipée par rapport aux capacités de production des usines avals, faible contenance des réservoirs avals, et bassins versants intermédiaires importants).

L'application de ces méthodes d'estimation de la valeur économique de la prévision à d'autres bassins permettra de réaliser un pilotage par la performance du système de prévision. (cf. voie d'amélioration n°1, en 10.2.1).

Cependant, l'estimation économique de ces deux types de prévisions opérationnelles (déterministe et probabiliste) pourrait nous laisser croire que les prévisions expertisées ne peuvent s'exprimer que sous ces deux formes :

- une prévision déterministe, issue de la critique de plusieurs modèles physiques de prévision et de données d'observations ; il s'agit de *la meilleure estimation possible du prévisionniste* compte tenu des données dont il dispose ;
- une prévision probabiliste, issue de *la critique d'un modèle de prévision d'ensemble*, sur le même mode que l'expertise physique habituelle : « le modèle hydrologique a ses réservoirs hydriques plus bas que ce que l'on observe », « il a fait démarrer l'onde de fusion trop tôt », etc.

Dans les deux cas, les prévisionnistes sont en quête de « la vraie valeur ». Une conséquence pratique est que les prévisionnistes réajustent les sorties des modèles de prévision d'ensemble en espérance bien plus souvent qu'en variance.

Pourtant, la prévision probabiliste n'est pas une simple critique de modèle.

Elle peut donner un rôle encore plus crucial à l'expert, en lui demandant de **s'impliquer personnellement dans la prévision par une prise de pari sur l'avenir**. Ainsi, lors de nos expérimentations, nous avons demandé à chaque prévisionniste d'EDF d'estimer la vraisemblance de tel scénario météorologique, ou la probabilité que tel seuil de débit soit dépassé ; à lui d'intégrer toutes les sorties de modèles déterministes ou probabilistes, ainsi que ses expérience et ses convictions pour répondre à de telles questions. Nous lui avons demandé d'exprimer ses prévisions à l'aide des probabilités subjectives, c'est à dire, en somme, d'estimer lui-même le risque qu'il se trompe.

Le protocole que nous avons mis en place pour recueillir de tels paris, avait deux objectifs :

1. **recueillir des informations qui étaient déjà existantes, mais qualitatives et non tracées.** Ces informations nous ont révélé
    - **la performance du système de prévision opérationnel dans la détection des crues.** *Vu l'état de l'art actuel en matière de détection de pluies intenses avec 24h d'anticipation, l'usager doit se résoudre à subir 60 à 80% de fausses alertes pour détecter plus qu'un épisode sur deux ;*
    - **le comportement des prévisionnistes lorsqu'ils sont confrontés aux limites de leur savoir**. *Ils sous-estiment leurs incertitudes (phénomène de surconfiance) et peuvent produire des prévisions biaisées (tous les prévisionnistes produisent les mêmes erreurs sur une situation donnée, et un prévisionniste donné commet des erreurs systématiques dans ses prévisions sur tel ou tel bassin versant).*
2. **améliorer le système de prévision :**
    - **améliorer la fiabilité des prévisions émises** (par la prise de conscience par les prévisionnistes de leurs biais et de leur surconfiance, réduire ces deux défauts) ;
    - **adapter la prévision aux enjeux spécifiques de chaque usager** (par exemple avec l'adoption de critères de performance mieux adaptés à l'enjeu de Sûreté Hydraulique).

Notre étude a permis de mettre en lumière un cercle vertueux : lorsque les prévisionnistes apprennent à mieux se connaître, et acquièrent une meilleure compréhension des enjeux des usagers (le pourquoi de la prévision), ils produisent une prévision à la fois plus fiable et plus performante pour certains usagers (meilleure détectabilité des phénomènes à risque).

Pour mieux appréhender ces phénomènes à risque (crues, vents violents ou neige collante) pouvant endommager le système de production, de transport et de distribution de l'électricité, nous préconisons de communiquer, en plus de la prévision centrale, un ou deux quantiles élevés. Le choix des quantiles 90 et 99 paraît opportun, à la stricte condition de les afficher comme tels : des niveaux qui n'ont qu'une probabilité donnée (une chance sur dix, une chance sur cent) d'être dépassés. Nous attirons ici l'attention du lecteur sur la dérive consistant à s'approprier les enjeux de l'usager à son insu, c'est à dire en lui communiquant des messages d'alerte basés sur un quantile, sans lui dire lequel. Le système de prévision serait dans ce cas opaque pour l'usager. Celui-ci constaterait que le prévisionniste "se couvre", mais il ne sait jamais à quel point. Prévisionniste et usager pourraient alors être précipités tous deux dans un abîme de perplexité, décrit par Savage (1971) : "You might discover with experience that your expert is optimistic or pessimistic in some respect and therefore temper his judgments.

Should he suspect you of this, however, you and he may well be on the escalator to perdition" (cf. note 61 p. 106).

L'usage explicite des probabilités permet d'éviter de tomber dans ce piège tout en adaptant les informations fournies aux enjeux de l'usager. Pourtant, ce mode de communication est accusé des deux maux suivants :

- **il ne serait pas limpide pour l'usager**
  il y a là une voie d'amélioration à travailler : le vecteur de communication. *(cf. voie d'amélioration n°2, en 10.2.2)*
- **il ne serait pas intuitif pour le prévisionniste**
  sur ce point, nous nous accordons un satisfecit : nous avons montré que les prévisionnistes n'étaient pas complètement déboussolés par cette nouvelle manière d'exprimer leurs incertitudes. Une fois compris le pourquoi de ce nouveau langage, et avec un peu de pratique, ils réalisent assez peu d'erreurs, et ne montrent pas d'aversion caractérisée à l'incertitude. Mieux, ce langage leur permet de corriger certaines mauvaises habitudes, comme celle de sous-estimer systématiquement leurs incertitudes.

Progressivement, la prévision probabiliste entre dans les mœurs des prévisionnistes. Du long terme (prévisions d'apport de remplissage des grands réservoirs), celle-ci gagne peu à peu le moyen terme sur des bassins associés à de forts enjeux énergétiques (prévisions probabilistes de J à J+15 sur le lac de Serre-Ponçon), voire le court terme, pour la prévision de crue sur des bassins sensibles et très réactifs (proposition à l'usager d'un quantile 90 pour la prévision des débits du Tarn).

Après avoir mis en place de manière semi-opérationnelle un système de prévision probabiliste, nous nous devions d'estimer la valeur de ce nouveau type d'information pour l'usager sensible à la Sûreté Hydraulique. Mais la complexité du système de gestion de crue est telle qu'on ne peut pas traduire la contribution du système de prévision en terme monétaires, ni même en termes quantitatifs (« la prévision améliore de X% la Sûreté Hydraulique »). En revanche, nous avons montré comment la prévision contribuait aux mécanismes garants de la Sûreté Hydraulique : augmentation de la résilience de l'organisation, renforcement de la défense en profondeur, et augmentation des marges de manœuvre de l'exploitant hydroélectrique. Nous avons montré de plus que la prévision probabiliste permettait de produire des informations mieux adaptées (meilleure détectabilité des risques météorologiques grâce au quantile 90).

*Si le quantile 90 a été un premier progrès vis-à-vis de l'enjeu de sûreté par rapport à la simple prévision déterministe, nous pourrions imaginer généraliser la démarche à la formulation de quantiles plus élevés, pour anticiper sur des événements plus sévères pour l'usager. C'est pourquoi nous préconisons de demander au prévisionniste de fournir quotidiennement un quantile 99. Cependant, nous avons analysé les difficultés qu'éprouvaient les prévisionnistes à se prononcer sur la vraisemblance d'événements s'éloignant notablement de leur prévision centrale : suivant la vulnérabilité des bassins, ils placent le quantile 90 plus ou moins haut, produisant ainsi une prévision biaisée. De telles difficultés risquent fort de se retrouver d'autant plus importantes que les probabilités sont d'autant moins « palpables » (peu d'occurrence de phénomènes extrêmes, donc peu de retours d'expérience ni d'occasions d'apprentissage). Nous verrons en 10.2.3 comment l'utilisation des sorties du modèle ANALOG pourrait aider le prévisionniste à quantifier les événements associés à des probabilités très faibles.*

En effet, grande a été notre déception lorsque nous avons constaté (en utilisant le quantile 90) que le modèle ANALOG donnait bien souvent des résultats meilleurs en terme de détectabilité que les prévisionnistes, alors même que ceux ci disposaient chaque jour des sorties de ce modèle. Cela nous invite à une dernière *voie d'amélioration (en 10.2.3)*.

En conclusion, nous ne voudrions pas réduire le problème de la prévision des événements à risque à une guerre de chiffres de performance entre les prévisionnistes opérationnels et le modèle ANALOG. Nous avons vu que la prévision probabiliste faite par les experts permettait en théorie une amélioration de la Sûreté Hydraulique par une meilleure détectabilité des phénomènes à risque.

Veillons à ce que la mise en pratique de cette méthode ne fragilise pas le système en déshumanisant le canal d'information. Il ne s'agit pas faire passer des chiffres (voire une pdf entière) par mail, fax, serveur ou intranet, et de supprimer toute confrontation par téléphone des enjeux courants de l'usager et de la situation hydrométéorologique vue (et prévue) par le prévisionniste. Celle-ci semble encore utile pour que l'usager puisse aboutir à une décision satisfaisante.

N'oublions pas que la capacité des sujets (prévisionnistes ou usagers) à faire face à des imprévus, défaillances de systèmes techniques, ou de communications, apporte une grande robustesse à l'ensemble du système de gestion de crue (système de prévision et système d'exploitation hydroélectrique). Particulièrement avant l'occurrence d'une crise, ou au cœur de celle-ci, il est primordial de disposer de boucles de rattrapage : le système ne doit pas s'écrouler sur une simple faute de frappe (défaillance humaine) ou sur un bug informatique lors de la transmission des informations (défaillance technique).

La finalité de notre démarche est donc, en bonne intelligence avec l'usager, de proposer une nouvelle base de dialogue à partir d'informations plus riches, plutôt que de remplacer le système actuel par un système qui acheminerait automatiquement les prévisions probabilistes expertisées à l'usager, en coupant tout autre lien entre les protagonistes.

## 10.2 Voies d'amélioration du système de prévision

### 10.2.1 Piloter les améliorations du système de prévision par leur performance pour l'usager

Comme nous l'avions préconisé en conclusion du Chapitre 3, l'estimation de la valeur économique de la prévision permettrait d'arbitrer entre les différentes actions d'amélioration du système de prévision en fonction de leur efficacité pour l'usager.

Cette démarche de modélisation de la valeur économique de l'information de prévision est aujourd'hui reprise par un agent EDF, anciennement prévisionniste, en relation avec divers usagers de la prévision.

### 10.2.2 Améliorer l'efficacité du vecteur de communication de la prévision

Comment le prévisionniste peut-il se faire comprendre alors même qu'il n'est pas sûr de ce qu'il avance ?

Pour répondre à une telle question, on ne peut se limiter à la question du formalisme :

- le choix de la présentation des prévisions sous forme de tableaux de chiffres, ou de manière graphique ;

- l'accompagnement des prévisions chiffrées par un texte donnant la situation générale et ses évolutions ;
- la représentation spatiale des processus hydrologiques ou météorologiques en cours ;
- le choix du canal d'information : écrit (fax, courrier électronique, ou forum intranet) ou oral (contact téléphonique), etc.

Il nous faut indubitablement éduquer l'usager afin que celui-ci puisse utiliser un nouveau type de prévision qui lui soit mieux adapté : la prévision probabiliste, quelle qu'en soit la forme (ensemble de vecteurs de prévision équiprobable, pdf sur un volume d'apports attendu, quantile élevé sur un débit prévu, etc.). Heureusement, certains usagers sont déjà sensibilisés à ce nouveau mode de communication des incertitudes :

- soit qu'ils utilisaient déjà une information probabiliste pour des échéances lointaines,
- soit qu'ils ont dû se résoudre à entendre parler de cette problématique étant donnée leur sensibilité à certains phénomènes à risque, et à la performance actuelle du système de prévision, pour les détecter avec une anticipation suffisante.

Pour améliorer l'efficacité du vecteur de communication de la prévision, il nous faut également, à force d'aller-retour entre prévisionniste et usager, trouver le domaine de recouvrement maximum de leurs savoirs. Nous pourrions ainsi **ajuster** le **niveau de technicité** de l'information formulée par le prévisionniste au **niveau de compréhension** de l'usager. L'usager deviendrait alors un complice de la prévision, en partageant une part de savoir voire d'expertise avec le prévisionniste.

Le prévisionniste pourrait par exemple partager une part de son expertise en conditionnant la pdf sur la prévision d'une variable qui intéresse l'usager à l'observation d'une autre variable que peut percevoir l'usager. C'est ainsi que l'usager pourrait lui-même réévaluer le risque de crue durant un épisode pluvio-neigeux en fonction de la température qu'il observe (puisque celle-ci est indicative de l'altitude de la limite pluie-neige, et donc l'extension spatiale du bassin soumis aux précipitations liquides).

### 10.2.3 Choisir des priors informatifs et utiles au prévisionniste

Nous avons constaté que les prévisionnistes ne regardaient pas assez le modèle ANALOG… et que les sorties d'ANALOG étaient meilleures que les prévisionnistes lorsqu'il fallait détecter des phénomènes à risque.

Nous pourrions donc imaginer que les prévisionnistes disposent des sorties du modèle ANALOG comme matière première pour produire leur prévision[124]. Ils pourraient s'aider de la fonction de répartition empirique ainsi produite (Figure 30), ou, si cela leur est plus intuitif, de l'histogramme de cette fonction de densité (Figure 31).

En calant une loi paramétrique telle que celle proposée par Krzysztofowicz (2002) : une fonction Dirac en 0, associé à une loi de Weibull sur ]0 ; +∞[, les prévisionnistes pourraient

[124] R. Krzysztofowicz (comm.pers.) insiste sur la compatibilité des différentes informations que le prévisionniste doit manipuler. En particulier, il faut que les informations a priori (distribution climatologique, sortie du modèle ANALOG, ou autre adaptation statistique d'un modèle physique) soient présentées dans la même forme que celle suivant laquelle le prévisionniste doit fournir sa prévision : par exemple les 3 quantiles 10, 50 et 90, avec les mêmes agrégations spatiales et pas de temps.

alors en modifier les caractéristiques : la probabilité de non-pluie, l'espérance des pluies, la probabilité de pluies intenses, etc.

Ce type de d'aide au prévisionniste a été testé avec un certain succès depuis bon nombre d'années (Murphy et Winkler, 1974)

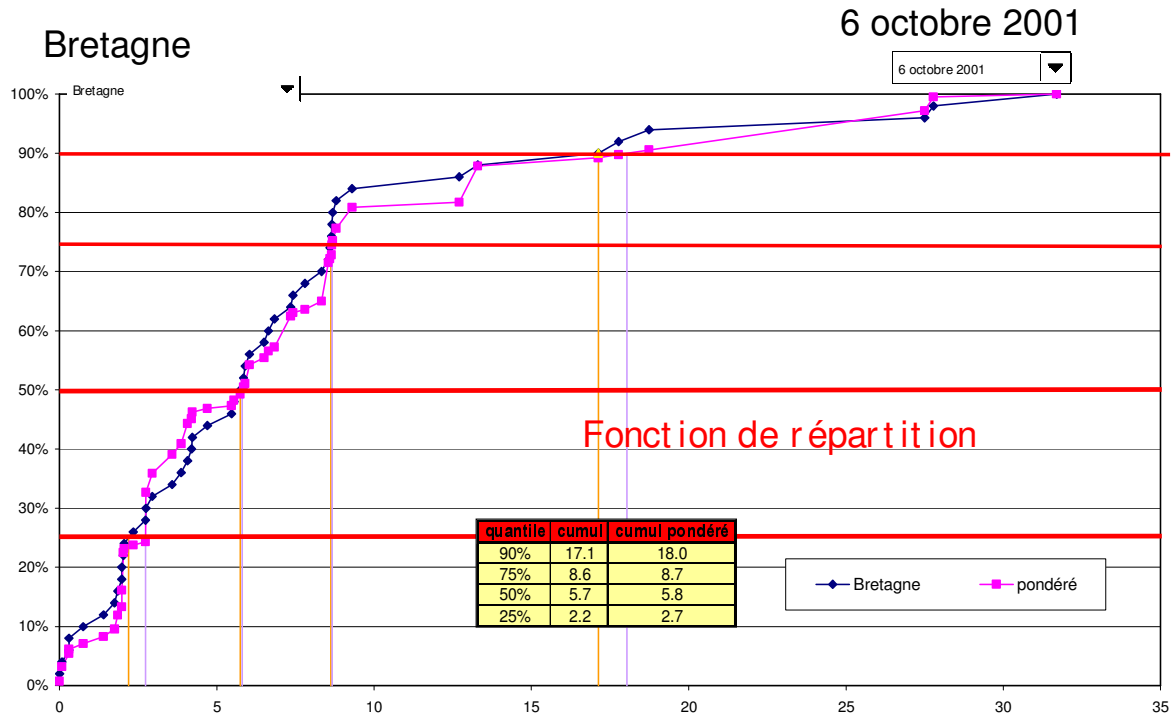



**Figure 30 : Fonction de répartition d'une prévision ANALOG**

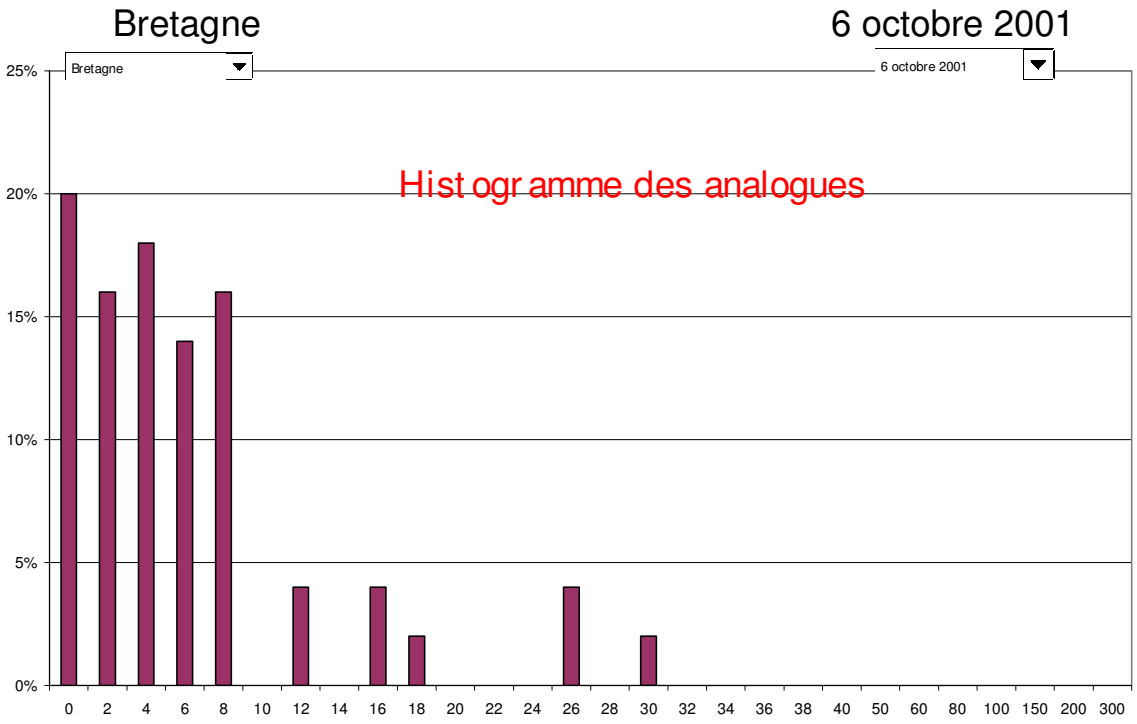



**Figure 31 : Histogramme de la fonction de densité d'une prévision ANALOG**

Inviter le prévisionniste à produire des prévisions en utilisant tel ou tel type de loi paramétrique aurait pour avantage de mieux cerner les queues de distributions, et permettrait de travailler sur la prévision d'événements extrêmes, cette fois-ci d'une ampleur bien plus grande que celles susceptibles d'être décelées avec le quantile 90%. Pour autant, une telle paramétrisation, est loin d'être évidente :

- soit on demande des informations au prévisionniste sur les queues de distribution… et il risque fort de ne pas avoir les éléments pour garnir cette queue de distribution : Comment en effet faire distinguer à l'expert une probabilité de $10^{-3}$ d'une probabilité de $10^{-5}$ ?
- soit, à partir seulement des quantiles fournis on cale une loi donnée (Weibull, Bêta, exponentielle, etc.). Dans ce cas, la queue de distribution serait très liée au choix de la loi retenue par le modélisateur, ce qui prendrait un caractère arbitraire dès que l'on s'éloignerait des paris effectivement pris par le prévisionniste (extrapoler un quantile 99 à partir d'un quantile 90 par exemple).

R. Garçon (comm. pers.) propose un compromis entre ces deux extrêmes en suivant la méthode :

- réaliser une partition des types de sorties du modèle ANALOG (front océanique, retour d'Est, flux de Sud, etc.) ;
- associer à chacune des classes une loi paramétrique dont la forme est la mieux ajustée aux éléments de cette classe ;
- fournir au prévisionniste une telle loi de probabilité, dont le caractère arbitraire serait limité par le caractère judicieux du choix de la loi étant donné les sorties en cours du modèle ANALOG ;
- le prévisionniste serait alors libre de modifier les paramètres de cette loi, voire de changer de loi en particulier s'il veut mieux préciser les queues de distribution.

Une telle méthode méritera d'être testée sitôt que les prévisionnistes manipuleront avec aisance, et dans leur pratique opérationnelle, les probabilités pour formuler leurs prévisions. Une application prématurée risquerait au contraire de perdre définitivement les prévisionnistes pour la prévision probabiliste, celle-ci étant jugée peu intuitive et loin de leur pratique quotidienne de l'hydrométéorologie[125].

[125] R. Krzysztofowicz (comm. pers.) suggère de ne travailler avec les prévisionnistes qu'avec la notion de quantiles. Cette notion peut leur être intuitive (cf. Chapitre 9), grâce aux paris que l'on peut associer. En revanche, des notions telles que la variance, ainsi que les grandes familles de lois (Weibull, Bêta, Exponentielle, etc.) sont des notions abstraites pour les prévisionnistes opérationnels

# Table des illustrations

**Figure 1 : Elaboration des prévisions opérationnelles à EDF ..... 18**
**Figure 2 : Localisation du bassin versant de l'Ain à Vouglans et du BVI à Bolozon. ..... 31**
**Figure 3 : Schéma hydraulique de la vallée de l'Ain. ..... 32**
**Figure 4 : Chronique des débits moyens journaliers sur l'Ain à Vouglans, durant les années 2000 et 2001 ..... 33**
**Figure 5 : Chronique des débits moyens journaliers du bassin versant intermédiaire de l'Ain à Bolozon, durant les années 2000 et 2001 ..... 33**
**Figure 6 : Exemple de vecteurs de prévision du jour J au jour J+4. Débit moyen journalier de l'Ain à Vouglans. ..... 35**
**Figure 7 : Exemple de vecteurs de prévision du jour J au jour J+4. Débit moyen journalier du bassin versant intermédiaire (BVI) de l'Ain à Bolozon. ..... 35**
**Figure 8 : Trajectoires des stocks du lac de Vouglans du 1er janvier au 15 août 2000. ... 38**
**Figure 9 : Trajectoires des stocks du lac de Vouglans du 1er janvier au 15 août 2001. ... 38**
**Figure 10 : Courbes de remplissage établies du 1er janvier 2001 au 30 juin 2001 (semaine 1 à 26). ..... 46**
**Figure 11 : Courbes de remplissage établies du 1er janvier 2002 au 30 juin 2002 (semaine 1 à 26). ..... 47**
**Figure 12 : Trajectoires des stocks du 1er avril au 31 décembre 2001, sans l'intégration de la contrainte touristique dans la règle de gestion. Les 4 trajectoires résultent des différentes gestions réalisées : sur la base des connaissances historiques, grâce à une prévision brute ou expertisée, ou en avenir certain. Le stock initial est pris égal à 122 hm3. ..... 48**
**Figure 13 : Trajectoires des stocks du 1er avril au 31 décembre 2001, avec intégration de la contrainte touristique dans la règle de gestion. ..... 49**
**Figure 14 : Trajectoires des stocks du 1er avril au 31 décembre 2001, avec contrainte touristique. Visualisation des courbes de remplissage. ..... 49**
**Figure 15 : Trajectoires des stocks du 1er avril au 31 décembre 2002, sans l'intégration de la contrainte touristique dans la règle de gestion. Les 4 trajectoires résultent des différentes gestions réalisées : sur la base des connaissances historiques, grâce à une prévision brute ou expertisée, ou en avenir certain. Le stock initial est pris égal à 122 hm3. ..... 52**
**Figure 16 : Trajectoires des stocks du 1er avril au 31 décembre 2002, avec intégration de la contrainte touristique dans la règle de gestion. Le stock initial est pris égal à 122 hm3. ..... 52**
**Figure 17 : Trajectoires des stocks du 1er avril au 31 décembre 2002, avec intégration de la contrainte touristique dans la règle de gestion. Le stock initial est pris égal à 500 hm3. ..... 53**
**Figure 18 : Trajectoires des stocks du 1er avril au 31 décembre 2002, avec contrainte touristique (stock initial 500 hm3). Superposition des courbes de remplissage. ..... 54**
**Figure 19 : Trajectoire de la cote de la retenue de Serre-Ponçon au cours e l'année 2001. Courbe haute et courbe basse sont les maxima et minima historiques depuis la mise en eau (1961). La courbe moyenne est la moyenne des cotes historiques relevées. .. 55**
**Figure 20 : Trajectoire de la cote de la retenue de Serre-Ponçon au cours e l'année 2002. ..... 56**

**Figure 21 : Localisation des 6 bassins versants du CHPMC (en jaune) et des 6 du CHA (en bleu). Prévision probabiliste de précipitation. Stage de sensibilisation-formation à EDF (les points localisent les pluviomètres). ..... 67**
**Figure 22 : Localisation des bassins versants pour la prévision probabiliste de précipitation. Expérience semi-opérationnelle à EDF..... 70**
**Figure 23 : Détection par les experts des pluies intenses (>20mm) avec 24h d'anticipation. Règle de décision basée soit sur le quantile 50 soit sur le quantile 90 En abscisses, le taux de fausses alertes (TFA), en ordonnée, le taux d'événements manqués (TEM) ..... 99**
**Figure 24 : Détection par le modèle des analogues des pluies intenses (>20mm) avec 24h d'anticipation. Règle de décision basée soit sur le quantile 50 soit sur le quantile 90 En abscisses, le taux de fausses alertes (TFA), en ordonnée, le taux d'événements manqués (TEM) ..... 101**
**Figure 25 : Diagramme de fiabilité pour les quantiles 50% et 90% fournis par les prévisionnistes d'EDF. Lames d'eau spatiales aux échéances J et J+1. ..... 118**
**Figure 26 : Diagramme de fiabilité pour les quantiles 10%, 50% et 90% fournis par les prévisionnistes d'EDF. Débits prévus aux échéances J et J+1. ..... 121**
**Figure 27 : Résultat du test de culture générale pour chacune des quatre populations interrogées. ..... 133**
**Figure 28 : Biais dans les réponses au questionnaire de culture générale, pour les quatre populations. Estimation Bayésienne et des incertitudes associées (intervalle de confiance à 95% sur b+)..... 136**
**Figure 29 : Biais dans les réponses au questionnaire de culture générale, pour la population française agrégée. Estimation Bayésienne et des incertitudes associées (intervalle de confiance à 95% sur b+). ..... 137**
**Figure 30 : Fonction de répartition d'une prévision ANALOG..... 176**
**Figure 31 : Histogramme de la fonction de densité d'une prévision ANALOG ..... 176**
**Figure 32 : Evolution de la largeur des intervalles de J à J+1. Production semi-opérationnelle de prévisions probabilistes à EDF. ..... 191**
**Figure 33 : Probabilité θ que l'observé tombe dans l'IC sachant que celui-ci est large ou étroit. Comparaison des 10% des IC les plus étroits avec le reste de l'échantillon. ..... 192**
**Figure 34 : Détection par les experts des pluies intenses (>20mm) avec 24h d'anticipation. Règle de décision basée soit sur le quantile 50 soit sur le quantile 90 En abscisses, le taux de fausses alertes (TFA), en ordonnée, le taux d'événements manqués (TEM) ..... 195**

# Annexes

Annexe 1 Estimation Bayésienne de l'incertitude sur un paramètre 182

Annexe 2 Présentation du stage de sensibilisation aux prévisionnistes d'EDF : Formaliser et communiquer l'incertitude liée aux prévisions hydrométéorologiques 184

Annexe 3 Attrait pour le risque ou pour l'incertitude lors d'une inversion des préférences ? Résultats par prévisionniste. 190

Annexe 4 Incertitude et échéance de la prévision : 191

Annexe 5 Surconfiance et précision affichée 192

Annexe 6 Détectabilité des pluies intenses, et biais dans les prévisions. Prévision semi-opérationnelle au CHPMC 193

Annexe 7 Graphes prévus-observés pour chacune des variables à prévoir. Prévision semi-opérationnelle au CHPMC 196

Annexe 8 Métrologie du risque Comment évaluer le risque juridique ? 205

# Annexe 1
# Estimation Bayésienne de l'incertitude sur un paramètre

Au cours de cette thèse, nous avons plusieurs fois eu à donner une évaluation de la probabilité d'un événement. Cette probabilité, qui nous était a priori inconnue, pouvait être la probabilité des événements suivants :

- E1 : « L'observé tombe dans l'intervalle de crédibilité fourni dans la prévision » (Chapitre 7)
- E2 : « L'observé tombe au dessous de l'IC » (Chapitre 8)
- E3 : « L'observé tombe au dessous du quantile 50% » (section 6.2.4 et Annexe 6)

A chaque fois, on ne considérait que deux événements complémentaires E et $\underline{E}$ dans un univers $\Omega$ donné (tirage de Bernoulli à deux modalités ) :

| | E | $\underline{E}$ | $\Omega=E \cup \underline{E}$ | Prob(E) | Prob(E) si la prévision est calibrée |
|---|---|---|---|---|---|
| 1 | « L'observé tombe dans l'IC » | « L'observé tombe à côté de l'IC » | Toutes les prévis | $\theta$ | 0,8 ; 0,7 ou 0,5 |
| 2 | « L'observé tombe au dessous de l'IC » | « L'observé tombe au dessus de l'IC » | Les prévis pour lesquelles l'obs tombe à côté de l'IC | b+ | 0,5 |
| 3 | « L'observé tombe au dessous du q50% » | « L'observé tombe au dessus du q90% » | Les prévis pour lesquelles q50%>0 et (obs<q50% ou obs>q90%) | b+ | 5/6 ≈ 0.83 |

Pour déterminer la probabilité de E (que l'on notera $\theta$), on dispose de n expériences que l'on suppose identiques et indépendantes, qui ont conduit à k occurrences de E.

On montre ici les différentes étapes de la méthode (inférence Bayésienne) qui nous a permis d'estimer la probabilité de E, grâce aux données recueillies par l'expérience de prévision probabiliste :

$\underline{1^{ère} \text{ étape}}$

Si la probabilité d'occurrence de l'événement E vaut bien $\theta$, la probabilité d'obtenir k occurrences de E en n prévisions vaut alors :

**(1)** **Prob (K=k) = [n,k| θ] = $C_{n,k}$ . $\theta^k$ .(1- θ)$^{n-k}$** (après un rapide calcul combinatoire)

2ème étape

Ne connaissant pas la valeur de θ, la probabilité de E, on lui attribue une loi de densité de probabilité non informative sur [0,1][126].

On choisit par exemple une loi Bêta de paramètres (1/2 ;1/2)[127] :

**(2)** **[θ] ~ $\theta^{-1/2}$ .(1- θ)$^{-1/2}$** couramment appelé prior de Jeffrey

3ème étape

On applique le théorème de Bayes :

**(3)** **[θ | n,k] ~ [n,k| θ] . [θ]**

de (1), (2), et (3), on obtient donc : **[θ | n,k] ~ $\theta^k$ . (1- θ)$^{n-k}$ . $\theta^{-1/2}$ .(1- θ)$^{-1/2}$**

**[θ | n,k] ~ $\theta^{-1/2+k}$ . (1- θ)$^{-1/2+n-k}$**

**[θ | n,k] ~ β(1/2+k ; ½+n-k)**

Après application de la formule de Bayes (qui nous a permis d'injecter les données prévues et observées issues de l'expérience), on montre donc que θ suit à nouveau une loi Bêta avec une simple mise à jour des paramètres.

Par cette méthode, nous avons tiré des données plus qu'une simple estimation de θ ($\underline{\theta}$=k/n) : on dispose d'une distribution de probabilité sur la valeur de θ, dont l'on peut tirer par exemple un intervalle de confiance à 95%.

---

[126] On appelle « prior » cette loi de densité de probabilité, car c'est une loi qu'on se fixe a priori, avant de se confronter aux données. On choisit un « prior non informatif » lorsqu'on refuse de donner une idée a priori sur la valeur de la probabilité en question. On peut alors prendre (entre autres) soit une loi uniforme sur [0,1], soit une loi qui a même espérance que la loi uniforme (cad. 0,5), mais une variance plus élevée, comme la loi β(1/2 ;1/2) (la variance de cette loi vaut 1/8 contre 1/12 pour la loi uniforme). Cette dernière loi, classiquement utilisée comme prior non-informatif, est connue sous le nom de prior de Jeffrey.

[127] Une variable X suit une loi Bêta de paramètres (p,q) si sa loi de densité de probabilité s'écrit :

$$f(x) = \frac{1}{B(p,q)} x^{p-1}(1-x)^{q-1} \qquad \text{avec} \qquad B(p,q) = \int_0^1 t^{p-1}(1-t)^{q-1}\,dt$$

# Annexe 2

## Présentation du stage de sensibilisation aux prévisionnistes d'EDF : Formaliser et communiquer l'incertitude liée aux prévisions hydrométéorologiques

Dans l'expérience décrite en 7.3, les prévisionnistes d'EDF avaient à leur disposition :

- ◊ l'image satellitaire IR de 6h (TU) sur l'Europe
- ◊ l'image Radar 6h sur la France
- ◊ les sorties du modèle ANALOG (d'EDF)
- ◊ les sorties papier des modèles Arpège (de Météo-France) et CEP (du centre européen de prévision à moyen-terme), donnant les champs des différentes grandeurs suivantes, à échéance 12h et 24h :
    - PMER, la pression au niveau de la mer
    - E107, l'épaisseur de la couche 1000-700 hPa
    - Z500, l'altitude de la couche 500 hPa
    - T500, la température de la couche 500 hPa
    - Z700, l'altitude de la couche 700 hPa
    - HU700, l'humidité de la couche 700 hPa
    - Z850, l'altitude de la couche 850 hPa
    - T850, la température de la couche 850 hPa
    - TA500, le tourbillon absolu à 500hPa
    - VV800, la vitesse verticale à 800 hPa
- ◊ certaines images visualisant des champs de prévision fournis par le modèle Arpège sur la France à diverses échéances :
    - Précipitations
    - θ'w à 700hPa, la température pseudo adiabatique du thermomètre mouillé
    - IA 1000 à 300 hPa, l'épaisseur de l'instabilité absolue de la couche 1000 - 300 hPa
    - une carte présentant plusieurs tampons masse d'air

Le protocole a été différent entre l'expérience faite au CHA et celle réalisée au CHPMC. Dans chaque cas, on proposait d'une part d'estimer directement divers quantiles de probabilité, et d'autre part de révéler ses préférences entre un pari risqué (probabilités connues) et un pari sur des évolutions météorologiques. Cependant, l'ordre différait entre les deux expériences : au CHA, l'élicitation des quantiles a précédé les paris, et au CHPMC, les prévisionnistes ont commencé par parier. On décrit ci-dessous plus en détail chaque protocole.

Expérience réalisée au CHA

Au CHA, les prévisionnistes ont directement estimé les 4 quantiles L25, L50, L75 et L90 sur leur prévision de la lame d'eau spatiale sur 6 bassins versants[128].

Ils ont ensuite été amenés à parier sur leurs quantiles émis.

On donne ci-dessous les questions auxquelles les prévisionnistes étaient confrontés pour un bassin versant donné (ici, le bassin versant de la Bienne à Saint Claude).

[128] La Bienne à Saint-Claude, l'Arc à Modane, la Durance à Serre-Ponçon, la Loire à Pont la Borie, le Chassezac à Pied de Borne, BVI Durance de Serre-Ponçon à Cadarache (sauf Verdon).

## *1/ Formalisation des probabilités subjectives sur la Bienne à Saint Claude*

***Question 1 :***

**Sur le bassin BV1 *(Bienne à Saint Claude)*, quelle est la lame d'eau L50 (cumul 8h-8h) médiane que vous prévoyez ?**

**L50 = .....mm**

*(L50 est telle que la lame d'eau qui tombera effectivement sur ce bassin a autant de chance de lui être supérieure qu'inférieure ou égale)*

*Par exemple, répondre L50=0 signifie que vous pensez qu'il y a au moins une chance sur deux qu'il ne pleuve pas sur le bassin durant les 24h à venir.*

*En répondant par un L50 strictement positif, vous pensez que les chacun des deux événements suivants ont autant de chance de se produire :*

◊ *Evénement 1 : Il va y avoir ces prochaines 24h une lame d'eau spatiale supérieure à L50*

◊ *Evénement 2 : Soit il ne pleuvra pas, soit il pleuvra moins de L50 durant ces prochaines 24h.*

***Question 2 :***

**Si vous avez répondu à la première question un L50>0 , répondez aux questions a1) et a2) ; sinon, répondez à la question b)**

**a1) Je vous informe que la lame d'eau effective va être supérieure au L50 que vous venez d'écrire.**
**Conditionnellement à cette nouvelle information, je vous demande la nouvelle médiane L75 que vous prévoyez.**

**L75 = .....mm**

*En d'autres termes, le chiffre L75 est celui qui partage en deux parties d'égale probabilité la lame d'eau spatiale conditionnellement a cette nouvelle hypothèse.*

**a2) Je vous informe que la lame d'eau effective va être inférieure à L50.**
**Conditionnellement à cette nouvelle information, je vous demande la nouvelle médiane L25 que vous prévoyez.**

**L25 = .....mm**

*En d'autres termes, le chiffre L25 est celui qui partage en deux parties d'égale probabilité la lame d'eau spatiale conditionnellement a cette nouvelle hypothèse.*

**b) Vous pensez qu'il y a au moins une chance sur deux pour qu'il ne pleuve pas sur le bassin BV1 *(Bienne à Saint Claude)*. Pensez vous qu'il y a trois chances sur quatre (ou plus) qu'il ne pleuve pas sur ce bassin ?**

**i) Plus de trois chances sur quatre qu'il ne pleuve pas** □

**ii) Plus d'une chance sur quatre qu'il pleuve** □

**Dans ce cas, quelle serait la lame d'eau L75 qui n'aurait qu'une chance sur quatre d'être dépassée ?**

**L75 = .....mm**

**Question 3 :**

**Quelle sera la valeur L90 qui n'aura qu'une chance sur dix d'être dépassée ?**

**L90 = .....mm**

## ***2/ Contrôle de cohérence de la lame d'eau spatiale sur la Bienne à Saint Claude***

*Par la suite, on va vous demander de nous donner vos préférences entre plusieurs jeux. Pesez bien ce que chacun des jeux signifie, et donnez vos préférences en vous mettant vraiment dans la peau du parieur que vous seriez en telle situation.*

***Test n°-1***

**Jeu A** : dans une urne, il y a 3 boules rouges, et 2 boules vertes. On tire une boule au hasard.

Si c'est une boule verte qui sort, vous empochez 100 FF ; si c'est une rouge, vous êtes quitte.

**Jeu B** : Si la lame d'eau spatiale sur les 24h à venir dépasse L50 (tel que vous l'avez vous-même défini), vous empochez 100 FF ; sinon, vous êtes quitte.

***A quel jeu préférez-vous jouer ?***

***Test n°-2***

**Jeu A** : dans une urne, il y a 2 boules rouges, et 3 boules vertes. On tire une boule au hasard.

Si c'est une boule verte qui sort, vous empochez 100 FF ; si c'est une rouge, vous êtes quitte.

**Jeu B** : Si la lame d'eau spatiale sur les 24h à venir dépasse L50 (tel que vous l'avez vous-même défini), vous empochez 100 FF ; sinon, vous êtes quitte.

***A quel jeu préférez-vous jouer ?***

***Test n°-3***

**Jeu A** : Dans une urne, il y a 14 boules rouges, et une boule verte. On tire une boule au hasard.

Si c'est la boule verte qui sort, vous perdez 100 FF ; si c'est une rouge, vous gagnez 30 FF.

**Jeu B** : Si la lame d'eau spatiale sur les 24h à venir dépasse L90 (tel que vous l'avez vous-même défini), vous perdez 100 FF; sinon, vous gagnez 30 FF.

***A quel jeu préférez-vous jouer ?***

***Test n°-4***

**Jeu A** : dans une urne, il y a 6 boules rouges, et une boule verte. On tire une boule au hasard.

Si c'est la boule verte qui sort, vous perdez 100 FF ; si c'est une rouge, vous gagnez 30 FF.

**Jeu B** : Si la lame d'eau spatiale sur les 24h à venir dépasse L90 (tel que vous l'avez vous-même défini), vous perdez 100 FF ; sinon, vous gagnez 30 FF.

***A quel jeu préférez-vous jouer ?***

Expérience réalisée au CHPMC

Au CHPMC, les prévisionnistes ont tout d'abord été amenés à prendre des paris sur la lame d'eau spatiale prévue sur chacun des 6 bassins. Ils avaient ainsi le choix de prendre soit un pari en univers incertain (la prévision météorologique), soit un pari en univers risqué (tirage d'une boule dans une urne).

On leur a ensuite demandé de fournir plusieurs quantiles de probabilité sur la lame d'eau attendue sur chacun des 6 bassins

On donne ci-dessous les questions auxquelles les prévisionnistes étaient confrontés pour un bassin versant donné (ici, le bassin versant du Blavet à Guerledan).

***Des paris sur la lame d'eau sur le bassin versant du Blavet à Guerledan***

*Par la suite, on va vous demander de nous donner vos préférences entre plusieurs jeux. Pesez bien ce que chacun des jeux signifie, et donnez vos préférences en vous mettant vraiment dans la peau du parieur que vous seriez en telle situation.*

***Choix n°-1***

**Jeu A** : dans une urne, il y a 2 boules rouges, et 3 boules bleues. On tire une boule au hasard.

Si c'est une boule bleue qui sort, vous empochez 100 FF ; si c'est une rouge, vous êtes quitte.

**Jeu B** : Si la lame d'eau spatiale sur les 24h à venir (cumul 8h-8h) dépasse 5 mm sur le Blavet à Guerledan, vous empochez 100 FF ; sinon, vous êtes quitte.

***A quel jeu préférez-vous jouer ?***

***Choix n°-2***

**Jeu A** : dans une urne, il y a 2 boules rouges, et 3 boules bleues. On tire une boule au hasard.

Si c'est une boule rouge qui sort, vous empochez 100 FF ; si c'est une bleue, vous êtes quitte.

**Jeu B** : Si la lame d'eau spatiale sur les 24h à venir (cumul 8h-8h) dépasse 5 mm sur le Blavet à Guerledan, vous empochez 100 FF ; sinon, vous êtes quitte.

***A quel jeu préférez-vous jouer ?***

***Choix n°-3***

**Jeu A** : dans une urne, il y a 6 boules rouges, et 1 boule bleue. On tire une boule au hasard.

Si c'est la boule bleue qui sort, vous perdez 100 FF ; si c'est une rouge, vous gagnez 50 FF.

**Jeu B** : Si la lame d'eau spatiale sur les 24h à venir dépasse 30 mm sur le Blavet à Guerledan, vous perdez 100 FF ; sinon, gagnez 50 FF.

***A quel jeu préférez-vous jouer ?***

***Choix n°-4***

**Jeu A** : dans une urne, il y a 12 boules rouges, et 1 boule bleue. On tire une boule au hasard.

Si c'est la boule bleue qui sort, vous perdez 100 FF ; si c'est une rouge, vous gagnez 50 FF.

**Jeu B** : Si la lame d'eau spatiale sur les 24h à venir dépasse 30 mm sur le Blavet à Guerledan, vous perdez 100 FF ; sinon, vous gagnez 50 FF.

***A quel jeu préférez-vous jouer ?***

## ***Elicitation des quantiles de probabilité de la lame d'eau sur le bassin versant du Blavet à Guerledan***

***Question 1 :***

**Sur le bassin BV1 *(Blavet à Guerledan)*, quelle est la lame d'eau L50 (cumul 8h-8h) médiane que vous prévoyez ?**

**L50 = .....mm**

*(L50 est telle que la lame d'eau qui tombera effectivement sur ce bassin a autant de chance de lui être supérieure qu'inférieure ou égale)*

*Par exemple, répondre L50=0 signifie que vous pensez qu'il y a au moins une chance sur deux qu'il ne pleuve pas sur le bassin durant les 24h à venir.*

*En répondant par un L50 strictement positif, vous pensez que les chacun des deux événements suivants ont autant de chance de se produire :*

◊ *Evénement 1 : Il va y avoir ces prochaines 24h une lame d'eau spatiale supérieure à L50*

◊ *Evénement 2 : Soit il ne pleuvra pas, soit il pleuvra moins de L50 durant ces prochaines 24h.*

***Question 2 :*** **Si vous avez répondu à la première question un L50>0 , répondez aux questions a1) et a2) ; sinon, répondez à la question b)**

**a1) Je vous informe que la lame d'eau effective va être supérieure au L50 que vous venez d'écrire.**
**Conditionnellement à cette nouvelle information, je vous demande la nouvelle médiane L75 que vous prévoyez.**

**L75 = .....mm**

*En d'autres termes, le chiffre L75 est celui qui partage en deux parties d'égale probabilité la lame d'eau spatiale conditionnellement a cette nouvelle hypothèse.*

**a2) Je vous informe que la lame d'eau effective va être inférieure à L50.**
**Conditionnellement à cette nouvelle information, je vous demande la nouvelle médiane L25 que vous prévoyez.**

**L25 = .....mm**

*En d'autres termes, le chiffre L25 est celui qui partage en deux parties d'égale probabilité la lame d'eau spatiale conditionnellement a cette nouvelle hypothèse.*

**b) Vous pensez qu'il y a au moins une chance sur deux pour qu'il ne pleuve pas sur le bassin BV1 *(Blavet à Guerledan)*. Pensez vous qu'il y a trois chances sur quatre (ou plus) qu'il ne pleuve pas sur ce bassin ?**

**i) Plus de trois chances sur quatre qu'il ne pleuve pas** □

**ii) Plus d'une chance sur quatre qu'il pleuve** □

**Dans ce cas, quelle serait la lame d'eau L75 qui n'aurait qu'une chance sur quatre d'être dépassée ?**

**L75 = .....mm**

**Question 3 :**

**Quelle sera la valeur L90 qui n'aura qu'une chance sur dix d'être dépassée ?**

**L90 = .....mm**

# Annexe 3

## Attrait pour le risque ou pour l'incertitude lors d'une inversion des préférences ? Résultats par prévisionniste.

| Expérience au CHA | Avant feed-back | | Après feed-back | |
|---|---|---|---|---|
| | **Attrait pour l'incertitude** | **Attrait pour le risque** | **Attrait pour l'incertitude** | **Attrait pour le risque** |
| Prévisionniste A | Absent | | Pas de réponse | |
| Prévisionniste B | 0 | 0 | Pas de réponse | |
| Prévisionniste C | 1 | 1 | 2 | 2 |
| Prévisionniste D | 0 | 2 | 2 | 3 |
| Prévisionniste E | 2 | 3 | 3 | 1 |
| Prévisionniste F | 2 | 2 | 2 | 2 |
| **TOTAL** | **5** | **8** | **9** | **8** |

**Tableau 37 : Nombre d'erreurs marquant soit un attrait pour l'incertitude, soit un attrait pour le risque. Décomposition par prévisionniste. Expérience au CHA.**

| Expérience au CHPMC | Avant feed-back | | Après feed-back | |
|---|---|---|---|---|
| | **Attrait pour l'incertitude** | **Attrait pour le risque** | **Attrait pour l'incertitude** | **Attrait pour le risque** |
| Prévisionniste 1 | 0 | 0 | 2 | 0 |
| Prévisionniste 2 | 6 | 0 | 3 | 6 |
| Prévisionniste 3 | 0 | 7 | 1 | 2 |
| Prévisionniste 4 | 2 | 3 | 1 | 1 |
| Prévisionniste 5 | 2 | 2 | 0 | 0 |
| Prévisionniste 6 | 0 | 1 | 2 | 3 |
| Prévisionniste 7 | 0 | 2 | 2 | 2 |
| Prévisionniste 8 | 0 | 4 | 0 | 0 |
| **TOTAL** | **10** | **19** | **11** | **14** |

**Tableau 38 : Nombre d'erreurs marquant soit un attrait pour l'incertitude, soit un attrait pour le risque. Décomposition par prévisionniste. Expérience au CHPMC.**

# Annexe 4

## Incertitude et échéance de la prévision : Comment le biais d'ancrage peut contribuer au Hard-easy effect

Nous avons vu en 7.4.3 que durant l'expérience semi-opérationnelle à EDF (au CHPMC), les prévisionnistes avaient tendance à sous-estimer leurs incertitudes d'autant plus que l'échéance était lointaine.

L'histogramme présenté en Figure 32 donne la proportion de prévisions probabilistes pour lesquelles la largeur de l'intervalle augmente de J à J+1, reste constante ou diminue :

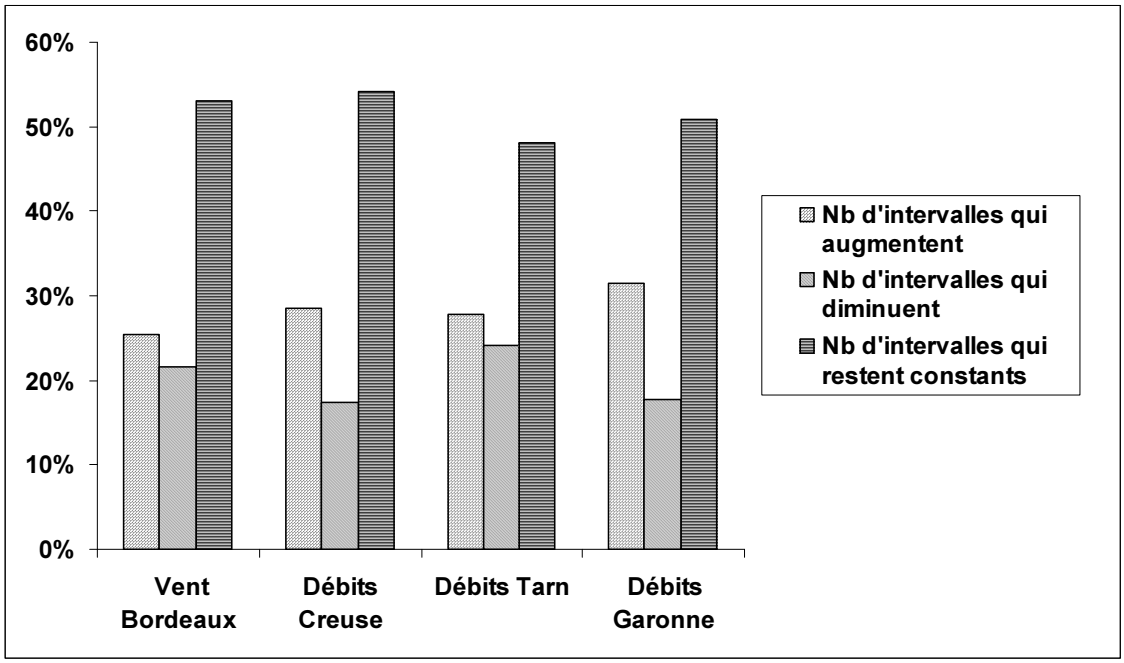



**Figure 32 : Evolution de la largeur des intervalles de J à J+1. Production semi-opérationnelle de prévisions probabilistes à EDF.**

Le comportement est très similaire pour chacun des paramètres à prévoir :

- présence d'un biais d'ancrage : un jour sur deux les prévisionnistes conservent la même largeur pour fournir un intervalle de confiance à J et J+1.
- dans 20% des cas, ils diminuent même cette largeur de J à J+1
- les prévisionnistes n'augmentent pas la largeur de ses intervalles que dans 20 à 30% des cas

Il y a donc un « effet ciseau » : alors que dans la grande majorité des cas, l'incertitude est plus grande pour l'échéance J+1 qu'à J, le prévisionniste choisit la plupart du temps de conserver la même largeur de ses IC de prévision, voire de la diminuer.

Faute d'élargir suffisamment leurs intervalles sur les prévisions de J à J+1, les prévisionnistes sous-estiment plus leurs incertitudes à J+1 qu'à J.

# Annexe 5
## Surconfiance et précision affichée

Lorsque le prévisionniste fournit un intervalle très étroit, son estimation des incertitudes de prévision est-elle aussi fiable qu'à l'accoutumée ?

Pour y répondre, on utilise la méthode suivante :

Pour chaque paramètre, on isole les 10% des prévisions pour lesquelles l'intervalle de crédibilité fourni était le plus étroit.

Cela revient à considérer les prévisions pour lesquelles la largeur des intervalles fournis est

- 10km/h pour le vent à Bordeaux
- $5m^3/s$ pour le débit de la Creuse
- $5m^3/s$ pour le débit du Tarn
- $10m^3/s$ pour le débit de la Garonne

On compare ensuite le nombre de réalisations dans l'intervalle de prévision pour ces prévisions fournissant un IC très étroit, et pour le reste de l'échantillon.

On visualise dans la Figure 33 l'estimateur de θ, et l'intervalle de confiance à 80% sur θ pour les deux échantillons de prévision : celles associées aux IC les plus étroits, et les autres.

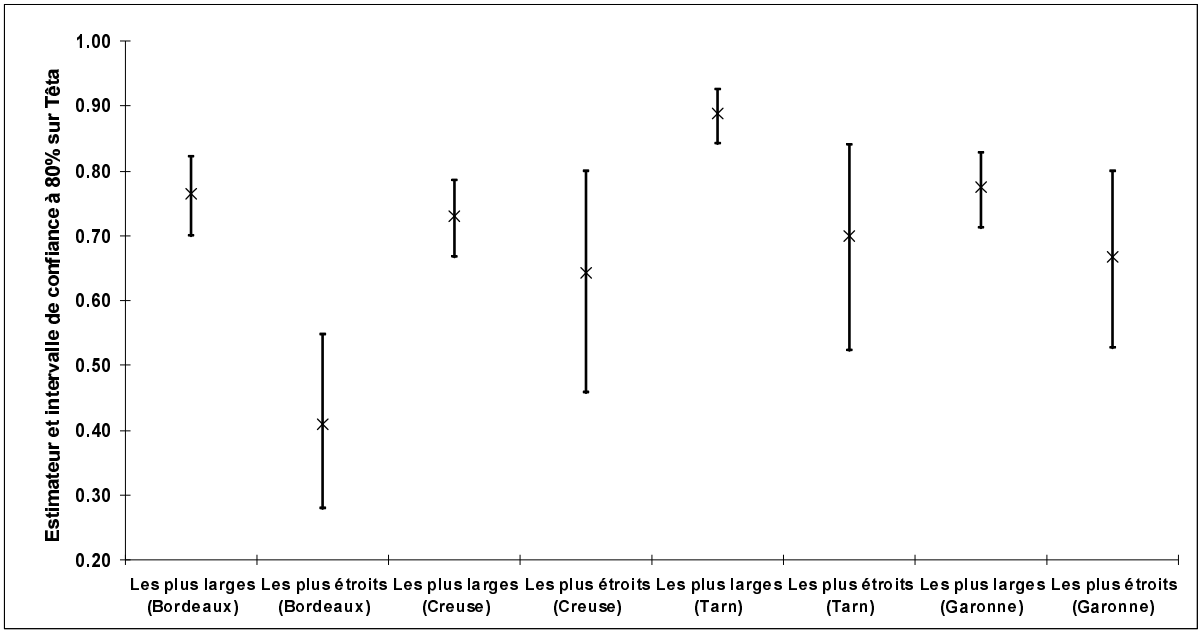



**Figure 33 : Probabilité θ que l'observé tombe dans l'IC sachant que celui-ci est large ou étroit. Comparaison des 10% des IC les plus étroits avec le reste de l'échantillon.**

Pour tous les paramètres, les prévisionnistes sous-estiment bien plus leurs incertitudes lorsqu'ils fournissent un IC très étroit.

Pour le vent à Bordeaux, ou la Garonne à Palaminy, la sous-estimation globale des incertitudes est portée uniquement par les prévisions fournissant un IC très étroit : si l'on retire celles-ci, l'estimation des incertitudes est parfaitement fiable.

# Annexe 6

## Détectabilité des pluies intenses, et biais dans les prévisions. Prévision semi-opérationnelle au CHPMC

### 1. Définition d'un biais adapté aux précipitations

Pour étudier le biais des prévisions de précipitations émises par les prévisionnistes, nous sommes confrontés au problème des prévisions égales à 0. Si par exemple le prévisionniste place le quantile 50% à 0, c'est qu'il estime qu'il y a **plus** d'une chance sur deux pour qu'il ne pleuve pas.

- le quantile 50% doit donc être dépassé moins d'une fois sur deux s'il est nul (et une fois sur deux sinon)
  *NB : dans la moitié des prévisions le quantile 50% vaut zéro.*
- idem pour le quantile 90% : il doit être dépassé moins d'une fois sur 10 s'il est nul (et une fois sur dix sinon)
  *NB : dans 20% des prévisions, le quantile 90% vaut zéro.*

Globalement, et dans le bénéfice du doute, on pourra penser que l'estimation de l'incertitude est correcte si

- q50% est dépassé plus d'une fois sur deux,
- q90% est dépassé plus d'une fois sur 10,

Et si on ôte les prévisions à 0, on retrouve la même analyse que pour le vent ou les débits : théoriquement,

- q50% doit être dépassé « exactement » une fois sur deux,
- q90% « exactement » une fois sur 10

Si nous ôtons les prévisions pour lesquelles le quantile 50 est nul, q50% doit être dépassé exactement une fois sur deux, et q90% exactement une fois sur dix.

Définissons à présent ce que nous entendons le biais dans les prévisions, en ne considérant que les prévisions non nulles. De telles prévisions sont calibrées si :

- 50% des réalisations tombent sous le quantile 50
- 40% des réalisations tombent dans la fourchette donnée
- 10% des réalisations tombent au dessus du quantile 90 (k-)

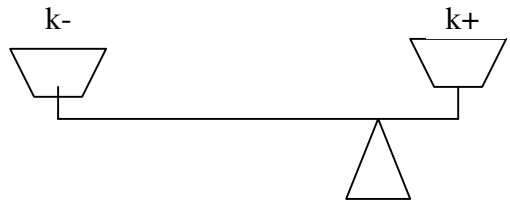
Si tel n'est pas le cas, on distingue la **sous-estimation des incertitudes** du **biais** de la prévision probabiliste :

- les prévisions présentent une sous-estimation des incertitudes lorsque moins de 40% des réalisations tombent dans les fourchettes fournies

- les prévisions présentent un biais lorsque l'intervalle fourni est décalé systématiquement dans le même sens par rapport aux observés : il devrait y avoir 5 fois plus souvent des réalisations qui tombent sous la fourchette fournie que de réalisations qui tombent au dessus. Lorsque la proportion est très déplacée par rapport à ce ratio d'1/5, on dit que la prévision est biaisée

Soit k+ le nombre des réalisations qui tombent effectivement sous le quantile 50 (surestimations), et k- le nombre des réalisations qui tombent effectivement au dessus du quantile 90 (sous-estimations)

Pour fixer les idées, on peut prendre l'illustration d'une balance ayant un bras de levier cinq fois plus grand à gauche qu'à droite. La prévision n'est pas biaisée lorsque la balance est équilibrée




*A l'équilibre, le pointeur devrait se situer aux 5/6 de la balance ; s'il est très décalé par rapport à cette référence, la prévision est biaisée.*

## 2. Estimation du biais dans les prévisions expertisées

| | Pluie Creuse | **Pluie Tarn** | Pluie Arn | Pluie Garonne |
|---|---|---|---|---|
| Nombre de réalisations qui sortent de l'IC (**n***) | 79 | **67** | 58 | 72 |
| Nombre de surestimations : **k+** | 71 | **55** | 53 | 65 |
| Estimation de **b+** | 0,90 | **0,82** | 0,91 | 0,90 |
| $IC_{80\%}$ [**b+**\|**n***,**k+**] | [0,86 ; 0,94] | **[0,79 ; 0,87]** | [0,87 ; 0,95] | [0,86 ; 0,94] |

**Tableau 39 : Biais dans les prévisions de précipitation expertisées (l'absence de biais correspond à b+=0,83).**

En ne considérant que les prévisions donc le quantile 50 est strictement positif, **seule la prévision de la lame d'eau sur le Tarn n'est pas biaisée**. **Sur tous les autres bassins, le prévisionniste surestime trop souvent la pluie prévue** dès qu'il prévoit l'occurrence de pluie significative sur le bassin (quantile 50 non nul).

Remarque :

Il y a également un "biais de sélection" à choisir les prévisions non nulles : on sélectionne les jours où le prévisionniste a pressenti, à tort ou a raison, l'occurrence de pluie. Dans de telles situations, on comprend bien qu'il a plus de chances de surestimer la pluie qu'à l'accoutumée.

On attend cependant du prévisionniste qu'il ait une estimation fiable de son incertitude, quel que soit le type de situation météorologique en cours, c'est à dire pour toutes les classes de prévision (pas de pluie prévue, peu de pluie prévue, beaucoup, etc.)

### 3. Rappel de la performance des prévisions sur chaque bassin

On présente à nouveau la Figure 23, tirée de 6.2.4, présentant les couples (TFA ;TEM) des prévisions sur chaque bassin, suivant le quantile choisi (quantile 50 et quantile 90).

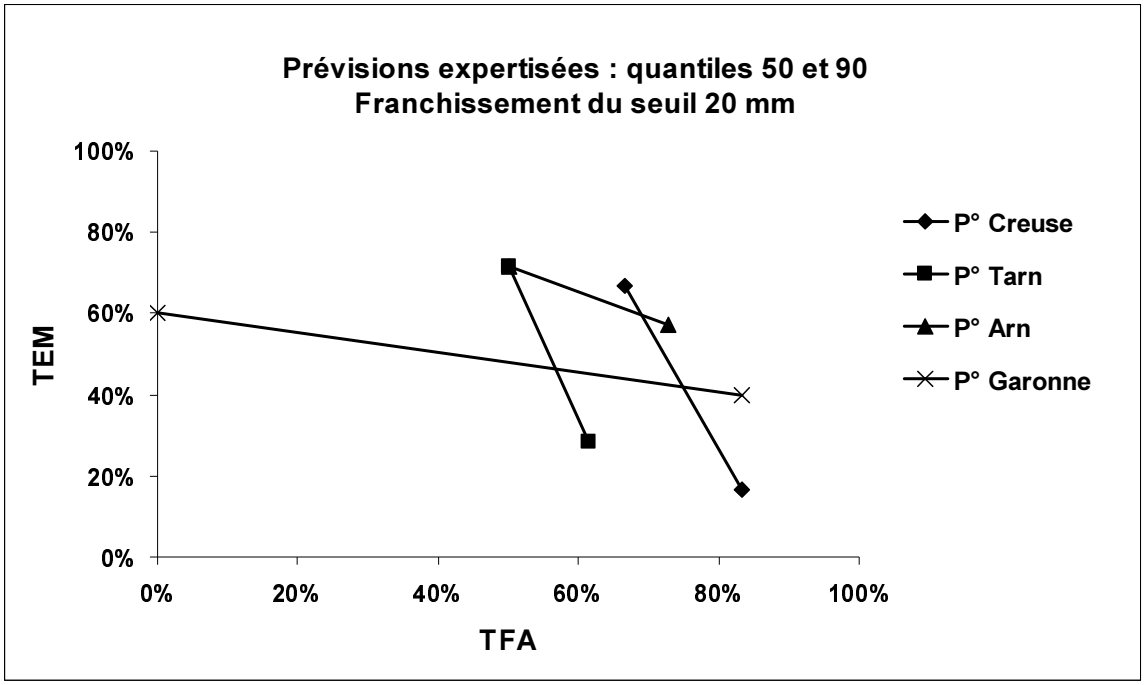



**Figure 34 : Détection par les experts des pluies intenses (>20mm) avec 24h d'anticipation.**
**Règle de décision basée soit sur le quantile 50 soit sur le quantile 90**
**En abscisses, le taux de fausses alertes (TFA), en ordonnée, le taux d'événements manqués (TEM)**

### 4. Conclusion : les prévisionnistes n'ont pas abandonné leur tendance à « piloter par le biais » leurs prévisions pour les adapter aux enjeux opérationnels.

Les prévisions sur le Tarn et sur la Creuse, toutes deux très performantes, répondent à des enjeux différents qui sont intégrés dans le jugement du prévisionniste lorsqu'il livre sa prévision probabiliste ; en effet :

- la prévision de la lame d'eau sur le Tarn est calibrée alors que la prévision sur la Creuse présente un biais de surestimation
- la prévision sur la Creuse produit un TFA bien plus élevé que celle sur le Tarn, et un TEM bien plus faible.

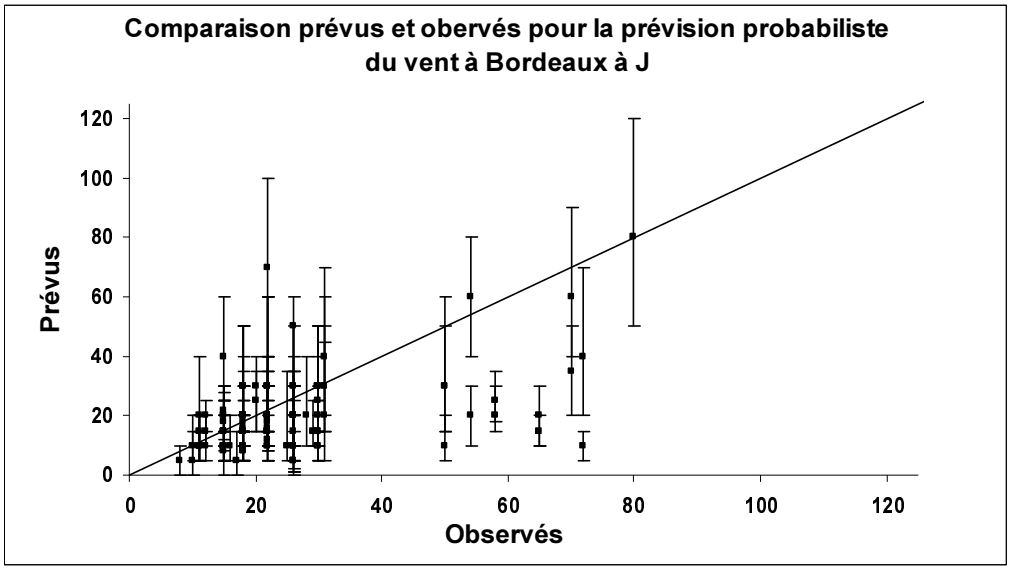
C'est le signe que le bassin de la Creuse à Eguzon est un bassin plus « sensible » que celui du Tarn à Millau, au sens où les événements manqués sont bien plus lourds de conséquences. Le prévisionniste décale alors sa prévision de lame d'eau vers le haut, ce qui confère à celle-ci une très bonne détectabilité pour un TFA qui reste cependant acceptable.

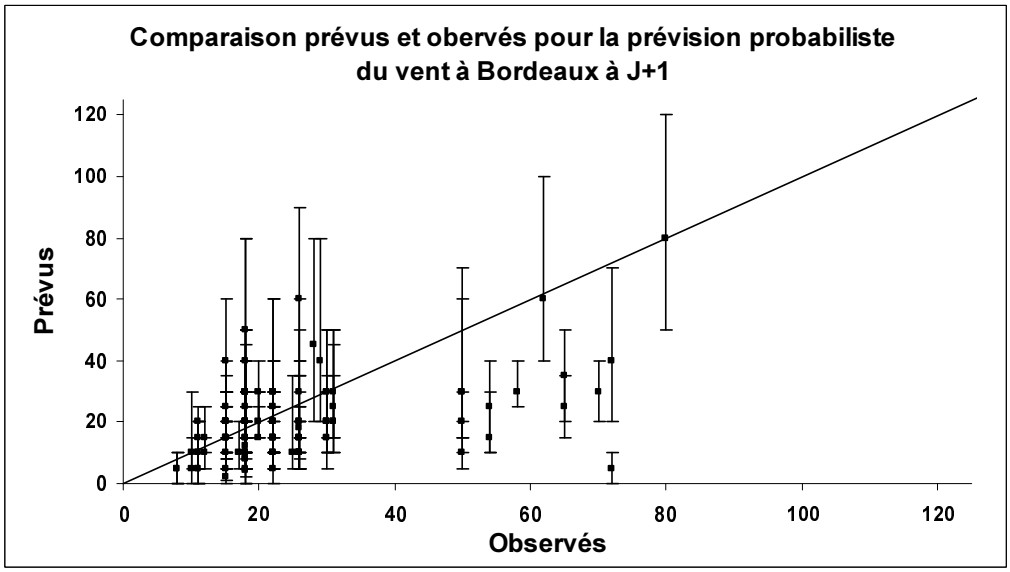
# Annexe 7

## Graphes prévus-observés pour chacune des variables à prévoir. Prévision semi-opérationnelle au CHPMC

Pour « visualiser » la performance d'une prévision, on a coutume de représenter le nuage de points (prévus ; observés). Ici, les prévus sont des intervalles accompagnés d'une valeur centrale.

### *Prévision du vent à Bordeaux*



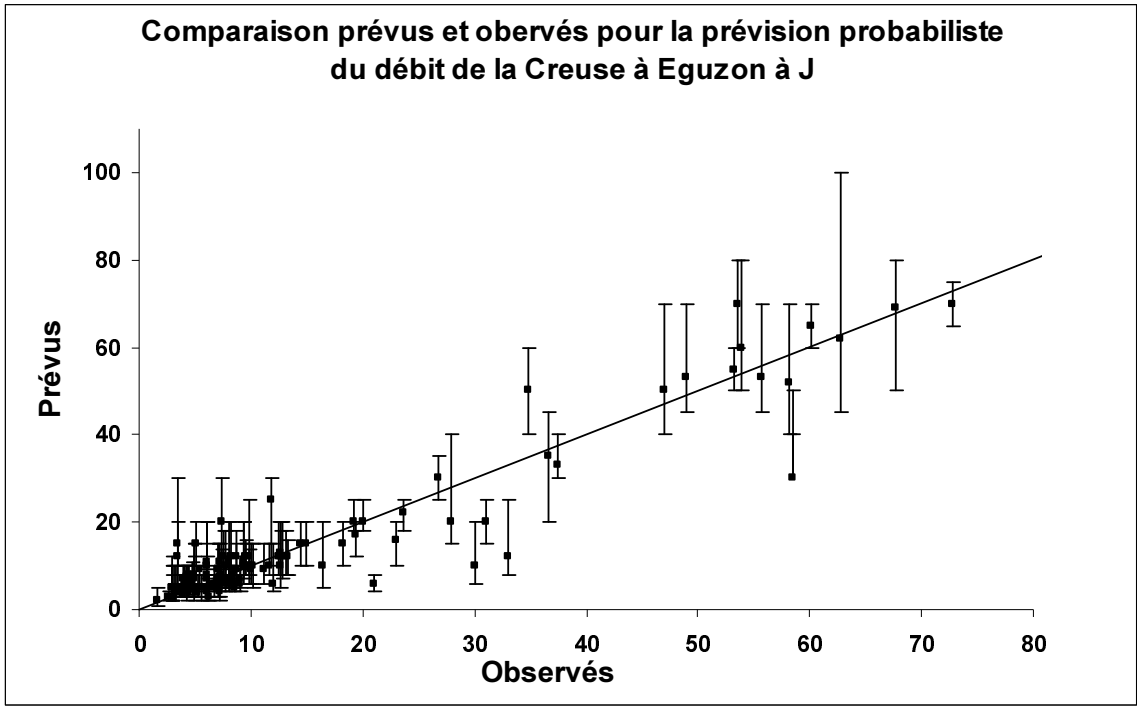

Le biais de sous-estimation est assez « visible », sur toute la plage des observés.

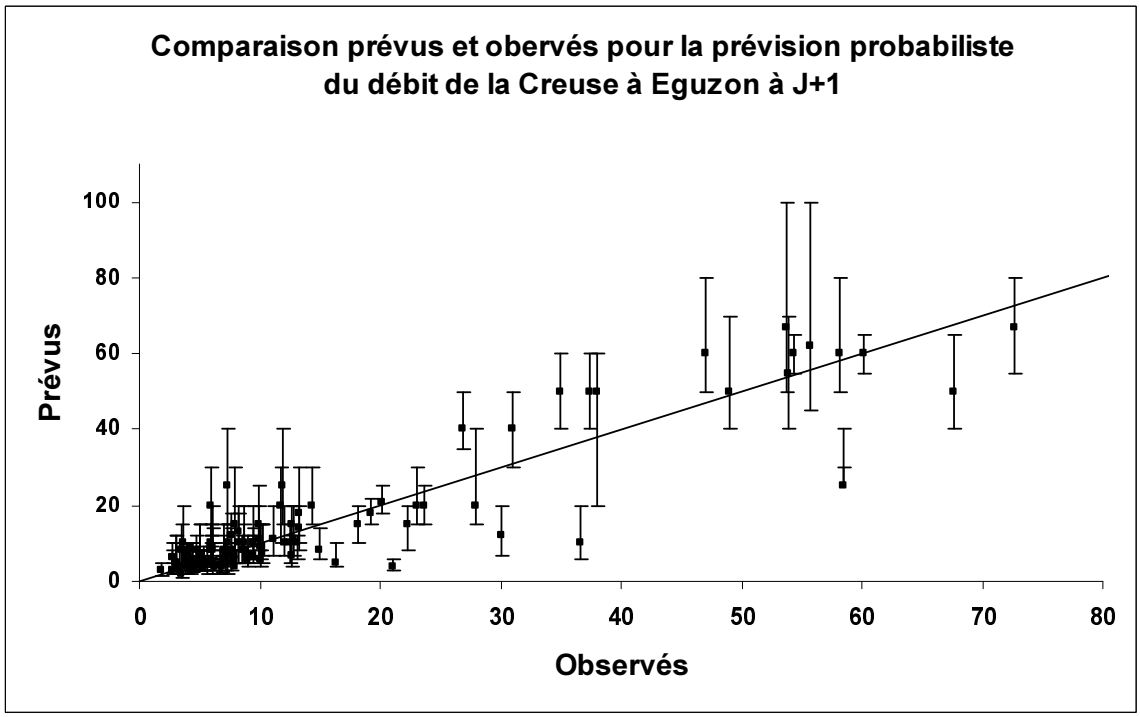
Les fausses alertes ont toujours été assorties d'une très large incertitude (ce qui est bon signe !), mais lors de plusieurs événements manqués, le prévisionniste était très sûr de lui.

Les prévisions sont « ancrées » à des valeurs rondes.

Pas d'observé entre 30 et 50km/h, ce qui est dû à l'hétérogénéité des sorties de météoplus (vent moyen/ vent en rafales).

## *Prévision du débit de la Creuse à Eguzon*



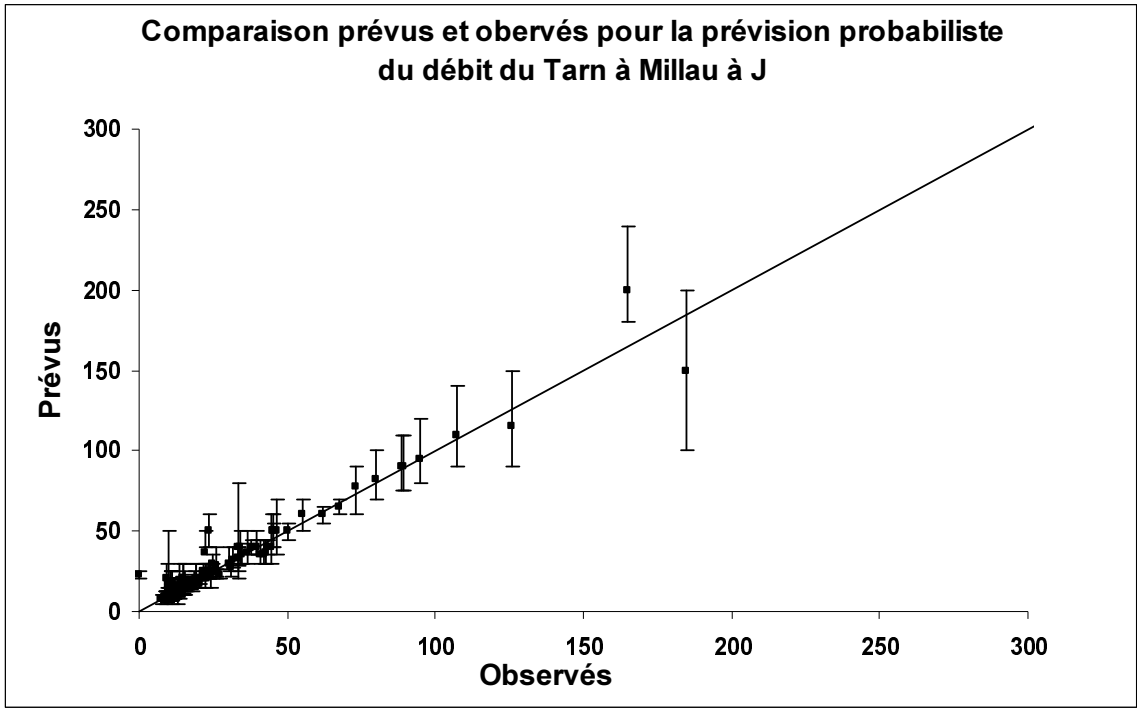



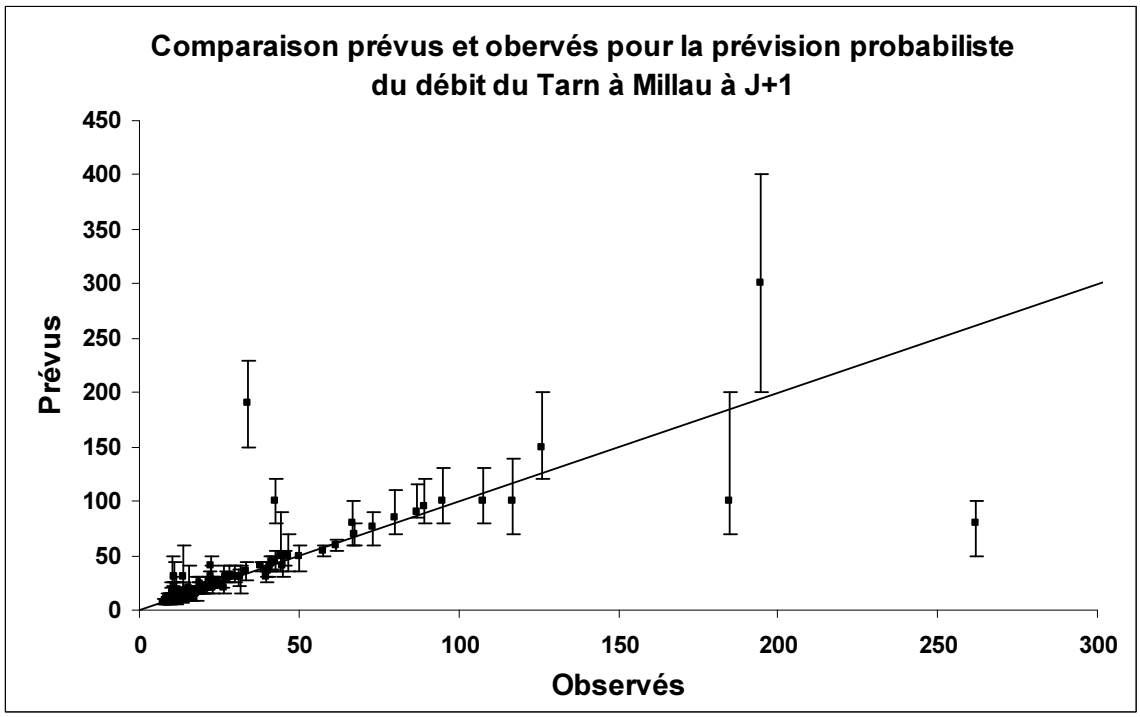

Le biais de surestimation est assez visible : les prévus sont bien plus souvent supérieurs aux observés (au dessus de la droite) que le contraire.

Jusque dans les valeurs élevées, l'estimation des incertitudes est fiable : les erreurs commises sont toujours du même ordre de grandeur que la largeur des intervalles.

Pour l'observé égal à 58 m3/s, le prévisionniste a interverti les quantiles 10% et 50%. C'est la seule erreur de ce type durant ces 6 premiers mois d'expérimentation.

## *Prévision du débit du Tarn à Millau*



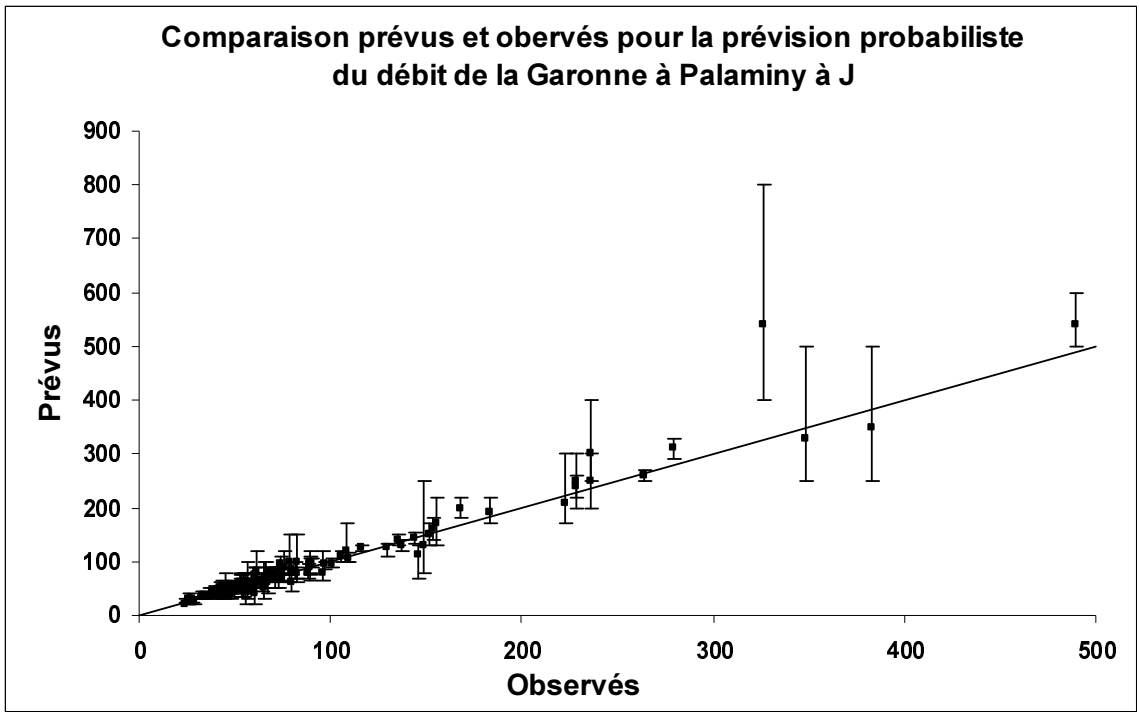



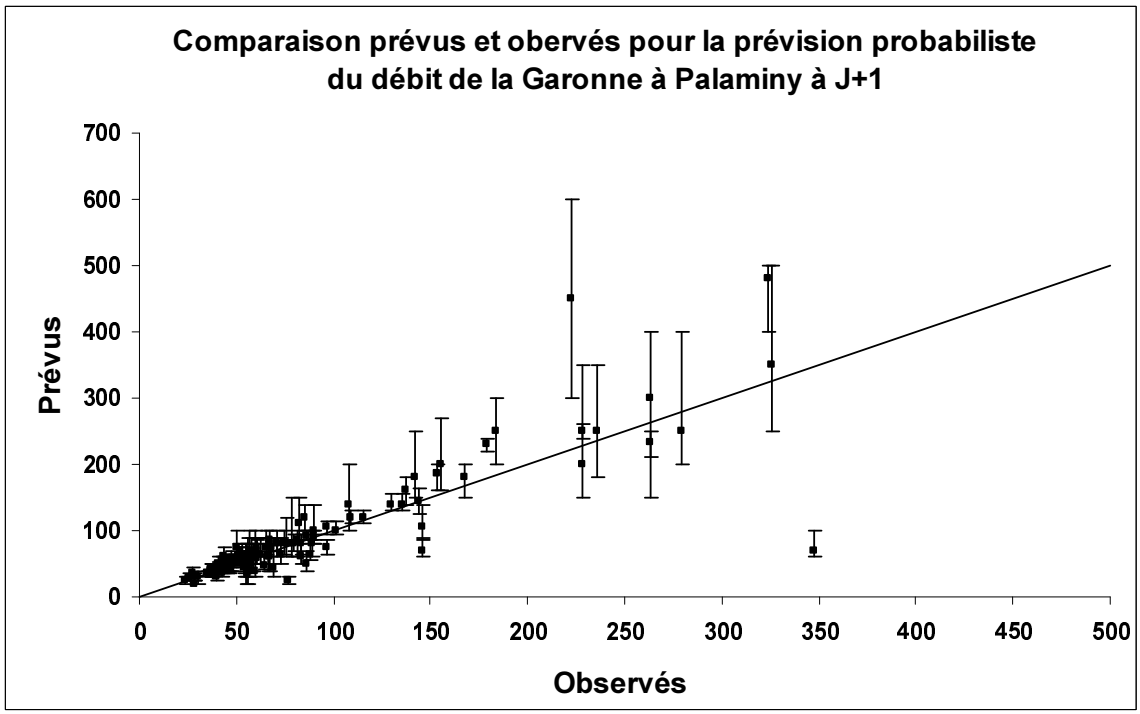

Quel que soit le débit observé, le prévisionniste a une très bonne appréhension des débits à venir (quantile 50 très proche de l'observé pour une large gamme de débits).

Particulièrement pour les débits inférieurs à 100m3/s, le prévisionniste associe une marge d'incertitude bien trop grande comparativement à la fiabilité de ses prévisions. Pour la prévision du Tarn à Millau, le prévisionniste surestime donc ses incertitudes.

On note cependant un événement manqué à J+1, pour lequel le prévisionniste avait pourtant affiché une grande confiance.

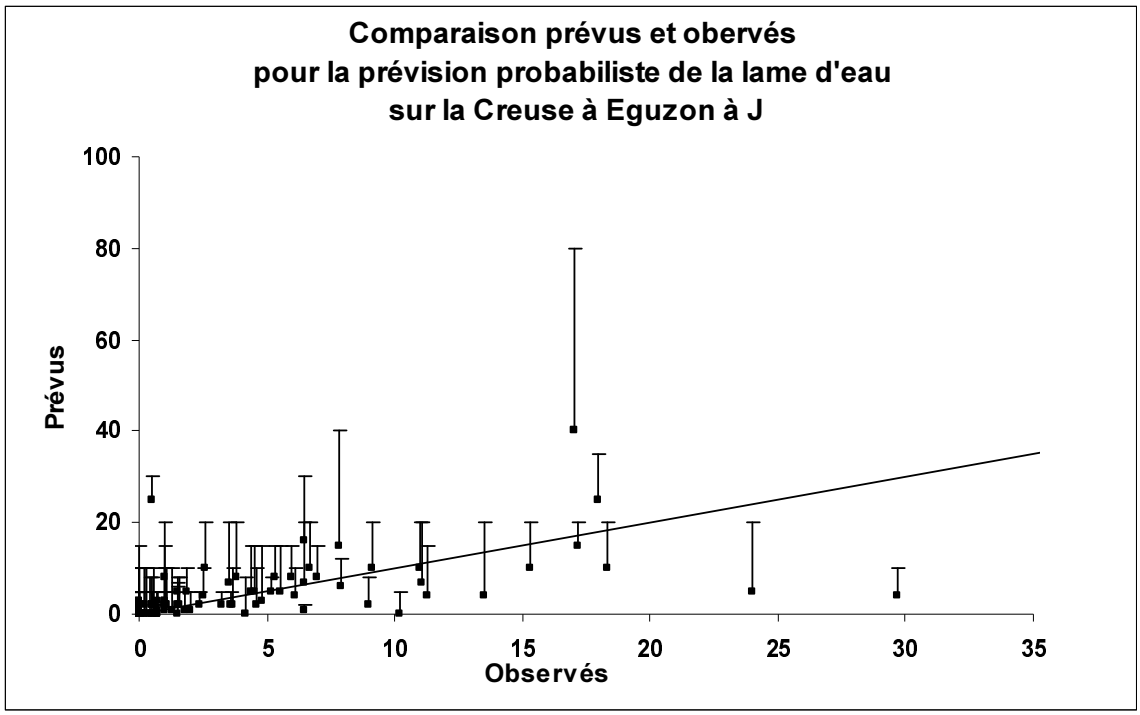
## *Prévision du débit de la Garonne à Palaminy*

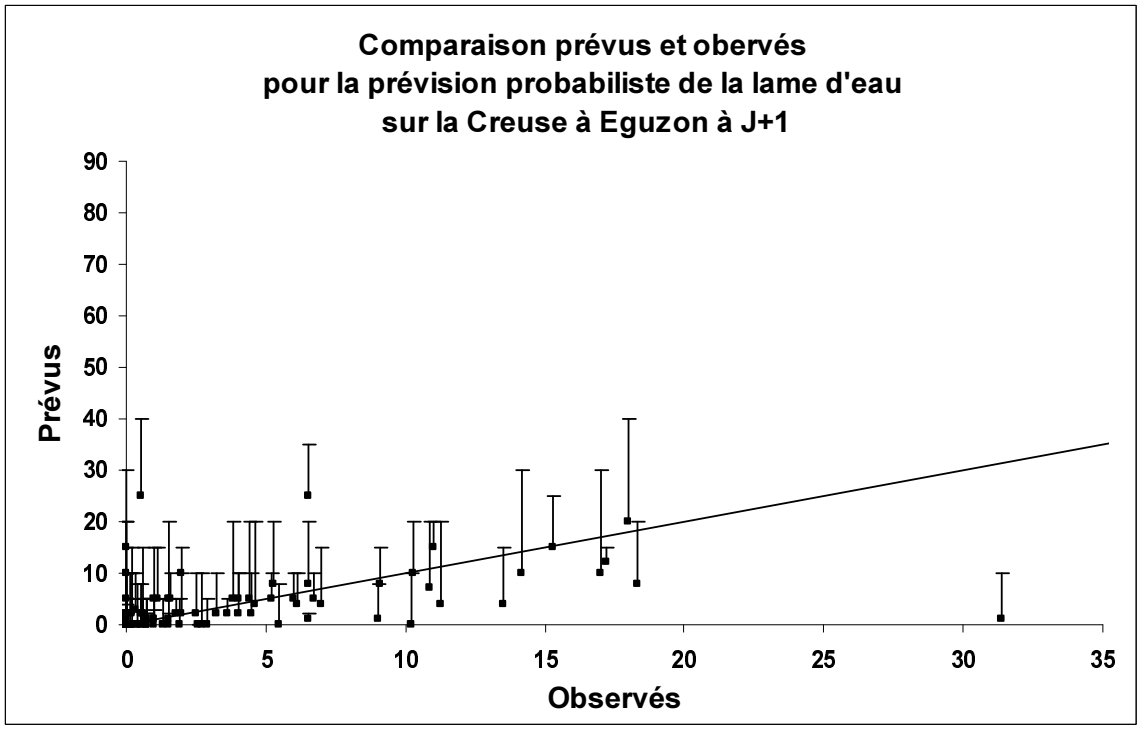



L'épisode du 16 juillet (348m3/s en débit moyen journalier) fut un événement manqué à J-1 ; ce jour là le prévisionniste avait pourtant affiché une grande confiance.

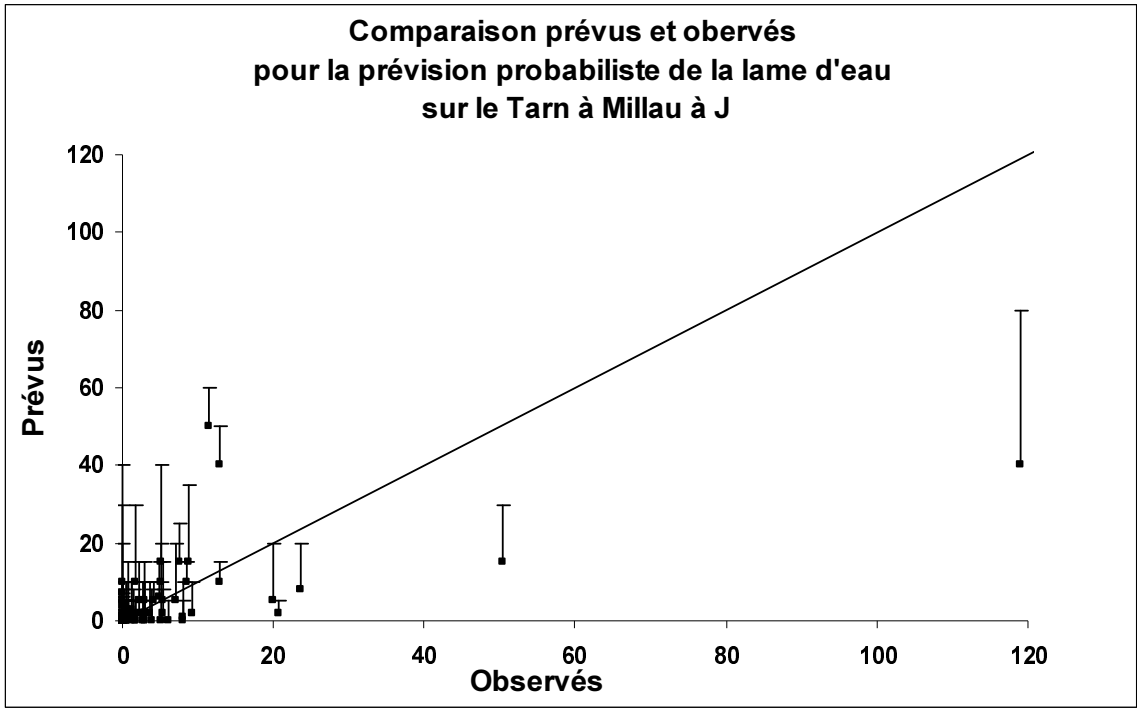
## *Lame d'eau spatiale sur la Creuse à Eguzon*

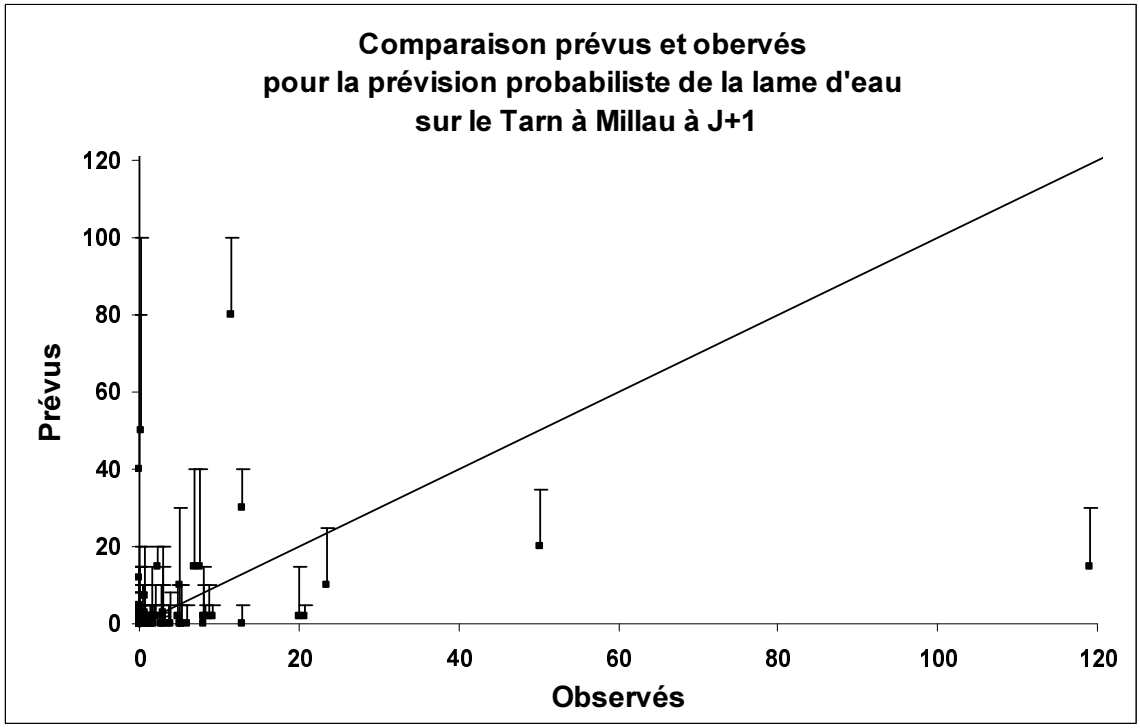



On observe un biais de surestimation, surtout pour les faibles valeurs d'observés.

## *Prévision de la lame d'eau spatiale sur le Tarn à Millau*



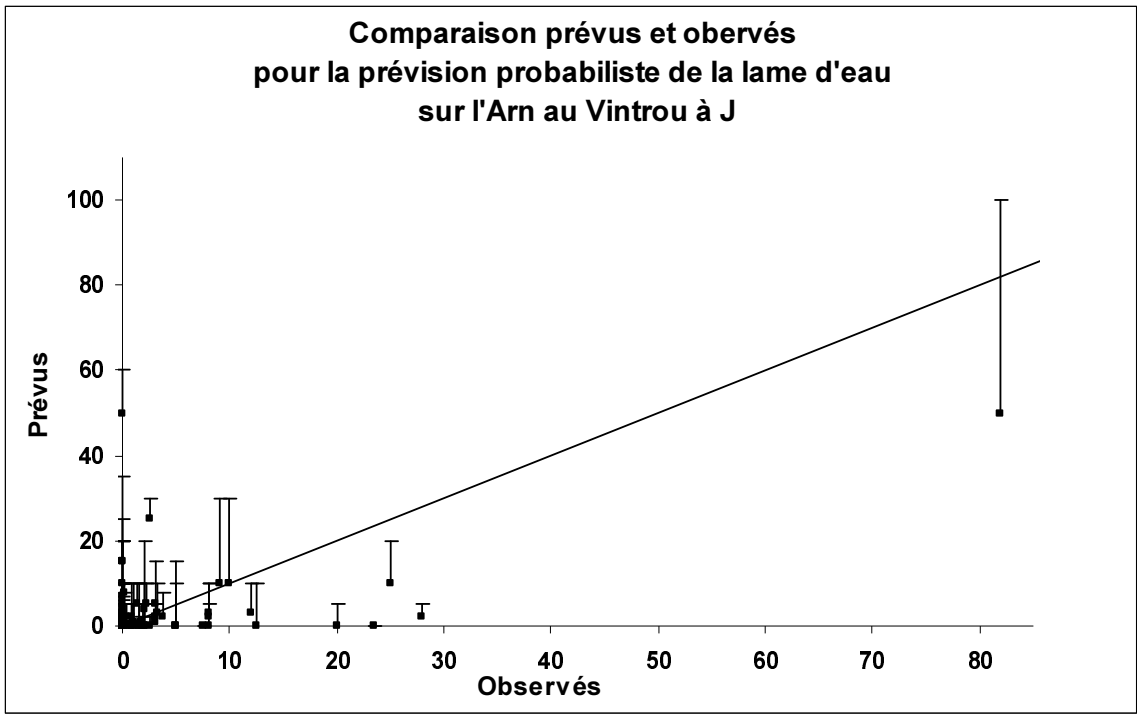



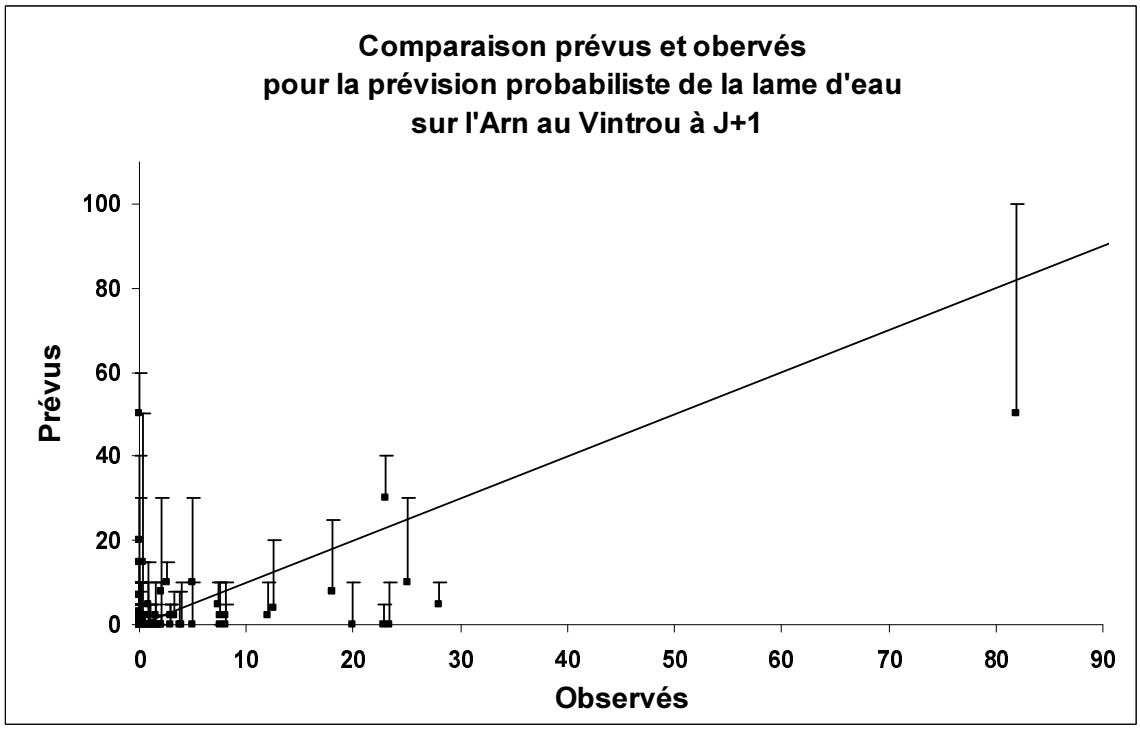

Sur la période, peu d'observés >10mm, sinon quelques très fortes valeurs.

L'épisode à 120 mm du 9 octobre 2002 a été très mal anticipé à J-1 ; bien mieux à J.

Quelques fausses alertes « record » : 40 à 50 mm en prévision centrale, et jusqu'à 100mm en quantile 90%, et une lame d'eau observée nulle.

## *Prévision de la lame d'eau spatiale sur l'Arn au Vintrou*

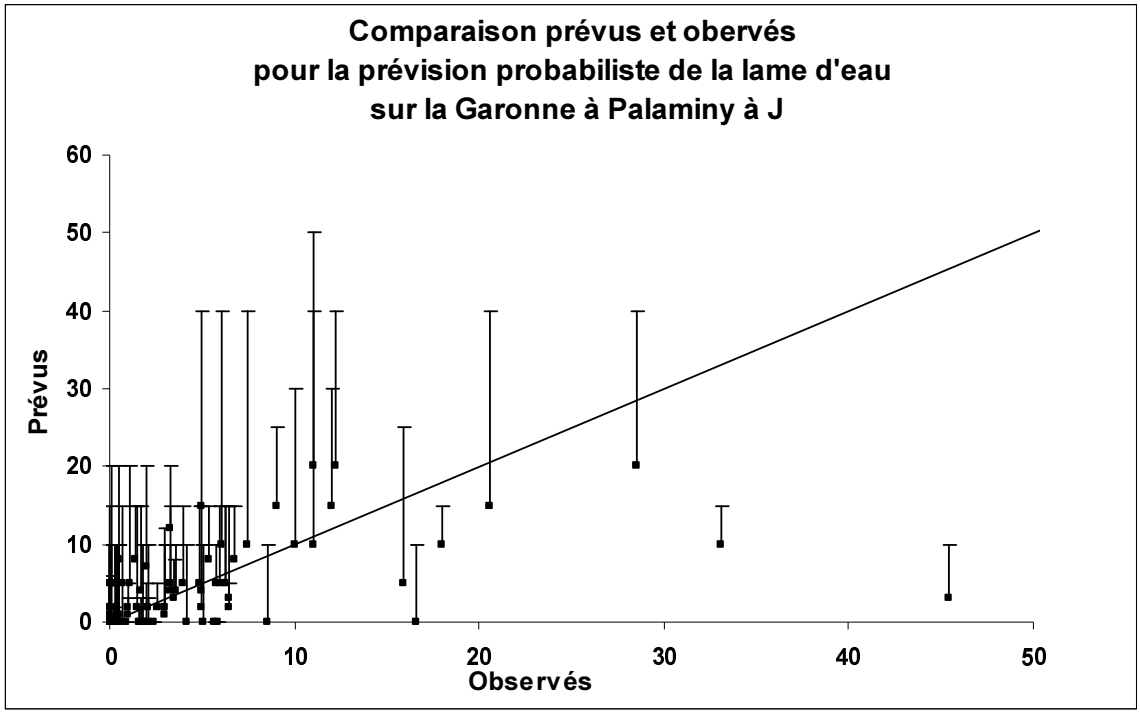


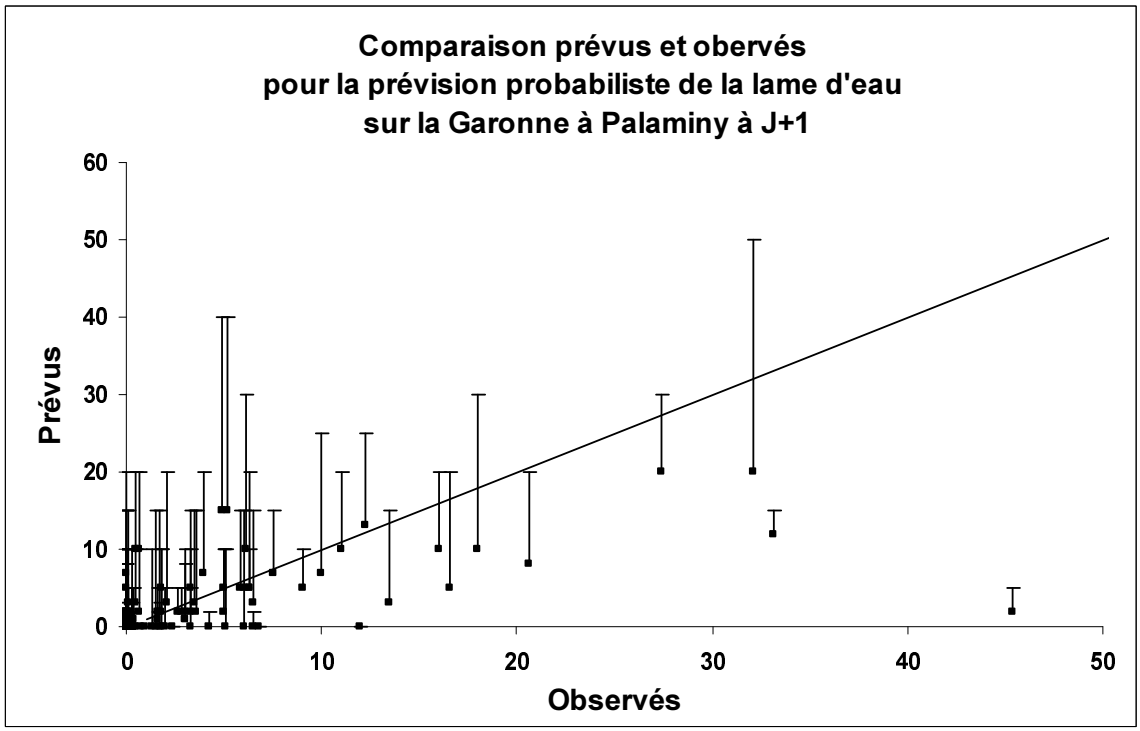




La largeur q50-q90 diminue beaucoup de J à J+1… à tort, surtout pour les valeurs de pluie respectables.

Même à J-1, l'épisode du 9 octobre 2002 sur l'Arn (80 mm) avait été très bien prévu.

## *Prévision de la lame d'eau spatiale sur la Garonne à Palaminy*




On observe un biais de surestimation, particulièrement visible pour les faibles valeurs de l'observé.

Les deux épisodes à 35 mm et 45 mm ont mal été anticipés, avec pourtant une très bonne confiance affichée. Pour les autres fortes valeurs d'observé, l'évaluation des incertitudes est correcte.

# Annexe 8

## Métrologie du risque
## Comment évaluer le risque juridique ?

Ces réflexions reprennent largement des points abordés avec Valérie Godfrin, du Pôle Cindyniques de l'Ecole des Mines de Paris, lors d'une réunion datée du lundi 23 avril 2001.

Dans le domaine du risque, les visions de l'ingénieur et du juriste se télescopent.

Pour l'ingénieur, le risque s'exprime comme la conjonction entre l'intensité et la probabilité d'un événement. Il cherchera par exemple à prémunir un ouvrage hydraulique contre le risque " crue décamillénale ", c'est à dire contre une crue dont l'intensité (ici, le débit de pointe) n'a qu'une chance sur dix mille de se produire chaque année. L'ingénieur quantifie également le risque lié à une décision (le mot décision étant entendu dans l'acception qu'en font les mathématiciens, c'est à dire une application d'un ensemble d'informations dans un ensemble d'actions) : c'est l'espérance des pertes (ou des gains) engendrés par cette action suivant les états de la nature.

Le juriste quant à lui dit qu'une décision comporte un risque juridique dès que la responsabilité civile, pénale ou administrative d'une personne peut être engagée suite à cette décision, c'est à dire dès lors qu'il y a fait générateur, dommage, et lien de causalité entre les deux.

Il existe des responsabilités pour faute, en vertu desquelles l'auteur sera sanctionné pour les faits répréhensibles commis (faute qui découle d'une action ou d'une abstention). C'est la gravité de la faute - et non du dommage - qui conditionne la sanction .

Il existe aussi des responsabilités sans faute, en vertu desquelles la responsabilité est engagée en raison du risque créé par une personne, une activité ou une chose ou du profit retiré par une personne d'une activité ou d'une chose .

Si le juriste est peu enclin à se prononcer sur une probabilité de sanction a priori, c'est qu'il est difficile de connaître a priori la qualification juridique des faits retenue qui va déterminer le régime juridique applicable et les éventuelles poursuites judiciaires ou leur issue. Cela pour de multiples raisons :

### Fait et acte juridique

Il faut faire la distinction entre un fait et un acte juridique : l'acte juridique (par exemple, un contrat) se définit comme une manifestation de volonté destinée à produire des effets de droit. Quant au fait, il désigne un événement pas forcément prévisible, mais qui génère tout de même des effets juridiques sans que ces effets aient été voulus (un accident de voiture par exemple). Prévoir une sanction probable, et sa probabilité associée, nécessiterait donc, outre de connaître les différents actes juridiques, de prévoir à l'avance les faits qui vont survenir. Cela est contradictoire avec la définition même des faits, en Droit.

## La procédure d'instruction

Il peut être difficile de présager de ce qui émergera de la procédure d'instruction durant laquelle le juge cherchera à retracer les faits et leur déroulement en matière pénale ou administrative. Savoir comment la responsabilité va être appréciée demande de pouvoir raisonner par analogie. Même si l'on se place dans un domaine où il existe une jurisprudence bien établie, il semble difficile au juriste de savoir quels seront les faits retenus contre une personne, la qualification juridique appliquée et quelles seront en matière pénale les circonstances atténuantes accordées. En fait, il est difficile de pouvoir établir a priori que les faits seront exactement les mêmes que ceux qui auront été jugés dans une affaire antérieure, qu'ils recevront de ce fait la même qualification et que les sanctions seront les mêmes.

## La qualification

La phase de qualification des faits est essentielle : elle a pour effet de transcrire les événements qui se sont produits en " langage juridique " et de leur attacher un régime juridique. Si, en matière pénale, le juge d'instruction qualifie les faits (coups et blessures involontaires par exemple), le juge ne sera saisi, en matière civile ou administrative, que sur qualification proposée par les requérants. Il convient donc, si l'on veut prévoir les issues probables du jugement, de commencer par s'interroger sur la qualification des faits qui seront reprochés ; ce n'est qu'en fonction de cette qualification que le juge pourra apprécier la mise en cause de la responsabilité et le régime juridique qui en découle. En matière civile, les plaignants proposent une qualification des faits qui peut être erronée par rapport à celle qui pourrait être retenue. Or, si le juge a la possibilité de requalifier les faits par rapport à ce que propose le plaignant, il est lié par rapport à la demande des parties et ne peut se saisir de lui-même d'un problème qui n'aurait pas été soulevé.

De plus, en vertu, du principe de l'autorité de la chose jugée, il n'est pas possible de rejuger le même événement, même sous une autre qualification, à moins de disposer d'éléments nouveaux à l'issue du litige.

## Les vices de procédure

Une fois le procès engagé, l'incertitude sur son issue provient en partie des vices de procédure que pourront dénoncer l'une ou l'autre des parties. Si, en matière civile, une règle déontologique tacitement acceptée par les avocats est d'éviter d'avoir recours à ce type d'argument, en revanche, en matière pénale, il est extrêmement courant de se servir des vices de procédure pour tenter d'annuler cette procédure.

En matière administrative, le juge sanctionne par l'intermédiaire du recours pour excès de pouvoir, le non respect des procédures et du formalisme requis pour les actes administratifs.

## Lien de causalité

Le lien de causalité est la relation de cause à effet entre le fait générateur et le préjudice invoqué. Ce lien de causalité peut être apprécié de différentes façons.

En vertu de la théorie de l'équivalence des conditions, tout événement qui a été une condition indispensable du dommage en est une cause, quelle que soit son importance ou sa proximité avec ce dommage. Le juge pénal apprécie le lien de causalité en se référant à cette théorie : il étudie non seulement les causes très proches d'un dommage (directes), mais aussi celles

indirectes qui ont pu en être à l'origine. Il est dans ces conditions difficile de prévoir par quelles relations de causalité la responsabilité d'un individu ou d'une entreprise pourra être mise en cause suite à un dommage. Cependant, en matière d'infraction non intentionnelle (c'est-à-dire lorsque l'auteur d'un fait répréhensible ne voulait pas causer de préjudice), depuis la loi Fauchon du 10 juillet 2000, la responsabilité pénale de l'auteur d'une faute indirecte non intentionnelle ne peut être engagée que si cette faute est d'une exceptionnelle gravité (le procès du Drac, la faute commise par l'institutrice n'a pas été considérée comme la cause directe du décès des enfants et de surcroît cette faute indirecte n'a pas été jugée d'une exceptionnelle gravité donc relaxe).

Le juge civil quant à lui, aurait tendance à s'appuyer sur la théorie de la causalité adéquate : il opère une sélection des causes et ne retient comme cause du dommage que les événements qui devaient normalement le produire, c'est-à-dire les facteurs patents.

## L'évaluation des dommages

En théorie, la responsabilité se mesure à l'aune de la gravité de la faute commise ou en vertu de l'intervention d'un fait dans la réalisation d'un dommage (pour les responsabilités sans faute) et non en fonction de la gravité des dommages. Cela dit, la jurisprudence est de plus en plus favorable à la victime et de manière déviante, les juges sont enclins à apprécier la responsabilité en fonction de la gravité des dommages. Logiquement, l'évaluation du dommage doit seulement servir à déterminer les réparations ou l'indemnisation que devra supporter l'auteur du dommage.

Pour évaluer l'ampleur d'un dommage peut se poser le problème de l'incertitude liée aux outils, méthodes ou techniques de mesure du dommage. Par exemple, suite à un accident ayant provoqué la pollution d'une rivière, on prélève dans cette rivière plusieurs échantillons dont on mesure la concentration en un polluant. Ces échantillons peuvent présenter différents taux de concentration, la dispersion des diverses mesures étant due aux incertitudes de mesure, et à la variabilité naturelle. Dans ces conditions, sur quel chiffre pourront s'appuyer chacune des parties ? Est-il raisonnable de ne se contenter que d'une mesure d'un échantillon, ou de leur moyenne ? Ces questions ne semblent pas faire l'objet à l'heure actuelle d'une réponse consensuelle, ce qui autorise certains juges à ne pas considérer du tout les incertitudes liées aux mesures dans l'évaluation des dommages.

Hors du problème de l'échantillonnage, la seule évaluation des dommages peut faire l'objet de batailles d'experts, dans un ou plusieurs domaines (impact environnemental, socio-économique, etc.) surtout si les dommages ne sont pas uniquement matériels (cas du préjudice moral).

## Le partage de responsabilité

Si la mise en cause de la responsabilité civile ou administrative est avérée, restera à prévoir comment se décidera le partage de responsabilité, en cas de pluralité de causes du dommage ou de plusieurs auteurs d'un même dommage (ex. si le préjudice est estimé à 100 000 F, le juge procède à un partage de l'indemnisation à verser entre les co-auteurs du dommage, en fonction de leur intervention dans la réalisation du dommage) .

## L'interprétation de la loi

Tout l'art du juge consiste à interpréter la loi, générale et impersonnelle, à la lumière des faits. Cette interprétation, suffisamment délicate pour qu'on parle d'art, ne permet pas de prévoir a priori les sanctions appliquées. Il est vrai que la jurisprudence antérieure sert notamment de référence et permet d'avoir une idée de l'issue d'un litige. Cependant, le problème de la qualification des faits (supra) montre qu'il est difficile a priori de déterminer une sanction si les faits survenus ne sont pas tout à fait les mêmes.

De plus, une décision juridique s'inscrit dans un contexte socio-économique qui peut modifier l'issue d'un litige (la catastrophe de Val d'Isère en 1970 n'a fait l'objet d'aucune sanction pénale, alors que les mêmes faits, jugés aujourd'hui, déboucheraient certainement sur des condamnations, compte tenu de la tendance à rechercher des responsabilités pénales dès lors qu'il y a atteinte à l'intégrité physique).

Au tribunal civil, la responsabilité est appréciée in abstracto, c'est à dire qu'on raisonnera par rapport à ce qu'aurait fait un homme raisonnement intelligent, muni d'une diligence normale. En revanche, en matière pénale, pour les infractions non intentionnelles, le juge apprécie la responsabilité d'une personne in concreto, en ne s'intéressant qu'à la personne seule : quelles étaient ses connaissances, ses moyens d'action, etc.

## L'Appel et la Cassation

Si, malgré toutes les difficultés énoncées plus haut, un juriste avisé arrivait, à prévoir des issues probables d'un jugement en première instance, reste à savoir si l'une des parties se pourvoira en Cour d'Appel, puis en Cour de Cassation, sachant que la jurisprudence de cette Cour sert de référence et que celle-ci ne revient plus sur les faits, mais juge uniquement en droit (elle examine les vices de procédures, la violation de la loi, la mauvaise interprétation des textes…).

## Conclusions et conséquences dans le domaine de la Sûreté Hydraulique à EDF

En bon Bayésien, on pourrait dire que puisqu'on peut difficilement déterminer à l'avance l'issue d'un procès, on en est réduit à utiliser un prior non informatif sur la sanction encourue suite à une décision.

Ce prior non informatif pourrait prendre la forme d'une loi uniforme sur [0, sanction maximale prévue par la loi] ; ce serait faire l'impasse notamment sur le premier point évoqué : comment prévoir les faits ? Lorsque EDF décide de déverser de l'eau dans le lit naturel d'une rivière, dans le respect des consignes d'exploitation et du cahier des charges, cette décision n'entraîne a priori aucune sanction judiciaire... sauf si le lit est fréquenté par quelque touriste inconscient du danger. Dans ce cas, la machine judiciaire peut se mettre en route et les sanctions potentielles précises qu'encourent les auteurs du dommage peuvent être difficiles à prévoir. Le procès qui a suivi la catastrophe du Drac de décembre 1995 est suffisamment révélateur à ce sujet.

Il faudrait donc, suite à chaque manœuvre d'exploitation présentant un impact à l'extérieur de l'usine, lui associer une loi de probabilité uniforme sur [0, $+\infty$] pour la sanction judiciaire encourue... ce qui ne serait pas d'un grand secours pour l'exploitant.

En matière de Sûreté Hydraulique, le seul déclenchement d'un procès étant lui-même extrêmement difficile à prévoir, on ne cherchera pas à quantifier le risque lié à une situation ou à une décision au sens du risque juridique tel qu'il a été défini en introduction. En cherchera en revanche à exploiter au mieux les recherches récentes en métrologie du risque dans le domaine des sciences du danger, afin de caractériser le risque lié à une situation (ie. le niveau de dégradation de la situation), et celui lié à une décision (ie. la variation de risque lié à la situation avant et après la décision).