

# La prévision numérique à l'échelle saisonnière : que sait-on faire et que peut-on espérer ?

**Michel Déqué**

Météo-France

Centre national de recherches météorologiques (CNRM)

42, avenue Gaspard-Coriolis – 31057 Toulouse Cedex

michel.deque@meteo.fr

## Résumé

Au-delà de la limite théorique de prévisibilité déterministe des phénomènes météorologiques, la prévision numérique reste possible parce que l'évolution lente des conditions aux limites de l'atmosphère – essentiellement la température de surface des océans tropicaux – induit une réponse atmosphérique spécifique. À l'échelle quotidienne, cette réponse est masquée par les fluctuations chaotiques, mais elle peut être mise en évidence par des moyennes temporelles et des moyennes d'ensemble. Les résultats d'expériences conduites avec plusieurs modèles de prévision numérique montrent que la prévisibilité à l'échelle saisonnière est meilleure dans les régions tropicales et, pour les régions tempérées, pendant la saison hivernale.

## Abstract

**Numerical weather prediction in the seasonal time scale: state of the art and perspectives**

Numerical weather prediction beyond the theoretical limit of deterministic predictability is possible because the slow evolution of atmospheric boundary conditions (mainly tropical sea-surface temperature) induces a specific atmospheric response. This response is blurred at daily time scale by chaotic fluctuations, but appears if time averaging and ensemble averaging are performed. Results of experiments based on several models show a higher predictability in the tropical regions and, for the midlatitudes, during winter.

## Les premiers pas

Le besoin en prévisions météorologiques à l'échéance de quelques mois est aussi ancien que le besoin en prévisions à court terme. Ainsi, le nombre de dictons consacrés aux mois futurs est au moins aussi élevé que le nombre de dictons fondés sur l'observation de l'évolution du temps sur quelques heures. L'activité agricole ayant été dominante dans notre pays pendant des siècles, la planification de travaux pénibles dont le résultat reste aléatoire a été porteuse d'une demande devant laquelle la science est restée muette. C'est pour répondre aux marins, puis aux aviateurs, que la science météorologique s'est faite science appliquée.

Avec la naissance de la météorologie opérationnelle moderne dans les années 1950, certains pays, dont les États-Unis mais surtout les pays à économie dirigée fondée sur la planification, ont développé une pratique opérationnelle de la prévision à longue échéance. D'autres, comme le Royaume-Uni ou la France, se sont limités à des travaux de recherche (voir Gilchrist, 1986, pour une synthèse). La technique employée pour produire ces prévisions était la méthode de **prévision statistique**. Bien que le présent article traite de **prévision numérique**, l'application des méthodes statistiques en prévision saisonnière sera évoquée page 27. Dès les années 1960, les prévisions à courte échéance ont été produites à partir de modèles



Pour l'agriculture, la demande de prévisions météorologiques saisonnières est très ancienne.  
(Photo Météo-France, P. Taburet)

numériques fondés sur des équations plus ou moins simplifiées des mouvements de l'atmosphère. Dans les années 1970, ont été développés des **modèles de circulation générale** (MCG) couvrant le globe et pouvant être intégrés sur plusieurs mois (Smagorinsky, 1969). Mais il faut attendre dix ans de plus pour que ces modèles puissent être employés dans le but d'élaborer des prévisions à quelques mois d'échéance. En 1979, la première expérience mondiale du Global Atmospheric Research Project (Garp) fournit des observations quotidiennes sur le globe. Le Centre européen pour les prévisions météorologiques à moyen terme (CEPMMT) complète ensuite cette couverture mondiale, en diffusant et en archivant des analyses planétaires opérationnelles. Il avait été précédé de quelques années par le National Meteorological Center (NMC) américain, devenu depuis National Center for Environmental Prediction (NCEP), dont les analyses sont moins utilisées en Europe. Tous les ingrédients sont alors en place pour produire de façon systématique des prévisions numériques saisonnières. Un des premiers articles à rendre compte de cette activité est celui de Shukla (1981). Le phénomène El Niño de 1982-1983 sert de révélateur de la capacité des modèles à produire une anomalie atmosphérique de grande échelle en réponse à une anomalie de la température de surface de l'océan (Owen et Palmer, 1987). À cette époque, on employait l'expression **prévision dynamique à échéance étendue** (acronyme anglais DERF) pour désigner la prévision à échéance 30-40 jours. La température de surface de l'océan, qui est la source de la prévisibilité, était considérée comme constante au cours de l'intégration du modèle.

Au début des années 1990, apparaissent les premiers **modèles couplés océan-atmosphère** capables de faire évoluer de façon réaliste un phénomène comme El Niño. C'est à cette époque que l'on comprend que la prévision d'une moyenne mensuelle ne présente que peu d'intérêt, car le résidu constitué par les fluctuations quotidiennes imprévisibles au-delà de 10-15 jours est plus important que la réponse atmosphérique au forçage de surface. La **prévision saisonnière** devient alors le point de passage obligé des études de prévisibilité à longue échéance (Palmer et Anderson, 1994). Toutefois, pendant la première moitié des années 1990, on ne pouvait pas utiliser les modèles couplés océan-atmosphère car on ne connaissait pas les conditions initiales pour l'océan ; on se trouvait alors dans une situation

La grande puissance de calcul nécessaire pour élaborer des prévisions à longue échéance avec un modèle numérique n'était pas disponible avant l'apparition des superordinateurs vectoriels, de type Cray 1, à la fin des années 1970. Sur la photo, le Cray 2 de Météo-France installé à Toulouse en 1992. (Photo Météo-France)



analogue à celle des MCG atmosphériques des années 1970 qui ne disposaient pas de mesures tridimensionnelles sur le globe. La question même de l'existence d'une prévisibilité à l'échelle de la saison, alors que la prévision numérique ne donnait rien d'intéressant au-delà de cinq jours, restait posée car les expériences numériques portaient sur très peu de cas ; les équipes qui présentaient de bons scores pouvaient être soupçonnées d'avoir opéré une sélection des cas les plus favorables.

C'est dans ce contexte qu'en 1993 Jean-Claude André (Météo-France) et David Burridge (CEPMMT) proposent de conduire une expérience de grande ampleur que l'on ne pourra mettre en doute et qui permettra d'évaluer l'intérêt pour un service météorologique de

mettre en place un système opérationnel de prévision numérique saisonnière. Ce projet, appelé **Seasonal Forecasting**, devient quelques années plus tard un projet soutenu par la Commission européenne sous le nom de **Provost**. Nous reviendrons un peu plus loin sur le principe de cette expérience et sur ses résultats. Après les progrès de l'observation océanique et le succès de la prévision du phénomène El Niño de 1997-1998 par les modèles couplés océan-atmosphère, un nouveau projet baptisé **Demeter** a vu le jour en 2000, pour prendre la suite de Provost en incluant la prévision de l'état de l'océan. Avant d'aller plus loin, tentons de comprendre pourquoi la prévision saisonnière est possible et quels sont les mécanismes mis en jeu.

## Mesurer le score d'une prévision saisonnière

Les champs météorologiques bruts produits par une prévision saisonnière sont assez éloignés de la réalisation qu'ils sont censés prévoir. Tout d'abord, les modèles présentent une erreur systématique non négligeable, par rapport à une prévision à courte échéance. Ensuite, la méthode de prévision d'ensemble, destinée à atténuer l'effet papillon et les divergences de conception des modèles (dans le cas d'un multi-modèle), se traduit par des champs plus lisses et des extrema émoussés. Ces champs prévus sont en fait une matière première à partir de laquelle on élabore des produits plus raffinés (indices, probabilités). On n'utilise pas l'erreur quadratique comme instrument de mesure du succès d'une prévision, mais le coefficient de corrélation spatio-temporelle entre la prévision et l'observation. Ce coefficient synthétise l'erreur quadratique et les variances des champs prévus et observés, tout en faisant abstraction du biais des prévisions. Si  $P$  et  $O$  désignent la prévision et l'observation, si  $\bar{P}$  et  $\bar{O}$  sont des moyennes pluriannuelles de ces deux quantités (on dit aussi des climatologies) et si  $d(X,Y)$  est un critère de distance moyenne entre  $X$  et  $Y$  (en pratique, la moyenne quadratique spatiale, temporelle ou spatio-temporelle de leur différence), le coefficient de corrélation peut se définir par :

$$\text{Corr}(P, O) = \frac{d^2(P, \bar{P}) + d^2(O, \bar{O}) - d^2(P - \bar{P}, O - \bar{O})}{2d(P, \bar{P})d(O, \bar{O})}$$

Il existe une formule plus brève, mais qui n'utilise pas le concept de distance, utile pour réaliser l'agrégation spatio-temporelle. Ce coefficient est une mesure de la qualité potentielle d'une prévision (Déqué et Royer, 1992) ; il ne faut donc pas se laisser abuser par sa valeur. Le coefficient de corrélation prend la valeur maximale de 1 quand la prévision est liée à l'observation par une formule linéaire. Deux séries qui se ressemblent à l'œil ont en général une corrélation supérieure à 0,80. À l'échelle synoptique, les prévisionnistes considèrent qu'en dessous de 0,60, une carte prévue n'est pas exploitable. En prévision saisonnière, on est moins difficile. Les tests utilisant la méthode de chaotisation montrent qu'au-dessus de 0,20, la prévision apporte une certaine information. Il dépend ensuite de son utilisateur de déterminer si cette information est utile ou s'il vaut mieux se fier au hasard ou à la climatologie.

## Le paradoxe de la prévision saisonnière

Les météorologistes tentent depuis longtemps d'expliquer au public que si leurs prévisions se limitent à quelques jours, ce n'est pas du fait de leur paresse ni de leur incompetence, mais à cause du caractère fondamentalement imprévisible des mouvements atmosphériques. On cite la non-linéarité, le chaos, l'attracteur de Lorenz (1963) et le célèbre effet papillon. Il est possible de prévoir une éclipse des années à l'avance avec une précision horaire, mais on ne peut affirmer à quelle heure il pleuvra demain. Ce discours a été relativement facile à tenir tant que la prévision saisonnière en France était laissée aux almanachs (dont celui du service public de La Poste).

On peut illustrer la chute brutale de prévisibilité en fonction de l'échéance en utilisant un score très répandu en prévision météorologique du fait de sa stabilité : le coefficient de corrélation du géopotential à 500 hPa sur l'hémisphère nord. Les lecteurs qui ne sont pas familiers avec le géopotential à 500 hPa peuvent imaginer ce champ, aux latitudes tempérées, comme une forme lissée de la pression réduite au niveau de la mer. L'expérience numérique décrite ici n'est pas Provost, mais une expérience que nous appellerons Persist et que nous détaillerons page 25. Pour l'instant, considérons seulement qu'il s'agit d'un jeu de 15 prévisions d'hiver à 120 jours d'échéance, dans lequel la température de la mer est prévue par persistance<sup>(1)</sup>. Comme le montre la figure 1, le coefficient de corrélation décroît fortement au cours des vingt premiers jours, puis oscille autour de 0 avec une faible amplitude. Un coefficient de corrélation empirique n'étant jamais rigoureusement nul, une zone de non-corrélation peut être estimée par une méthode de **chaotisation** : on effectue des permutations aléatoires des années pour les prévisions et on calcule les coefficients de corrélation correspondants. Parmi les 15! permutations possibles, on n'en prend que 200 au hasard ; on dispose alors d'une série de 200 coefficients pour lesquels la ressemblance entre la prévision et l'observation ne peut être que fortuite. En considérant les quantiles à 97,5 % et 2,5 %, on obtient la zone hachurée qui correspond

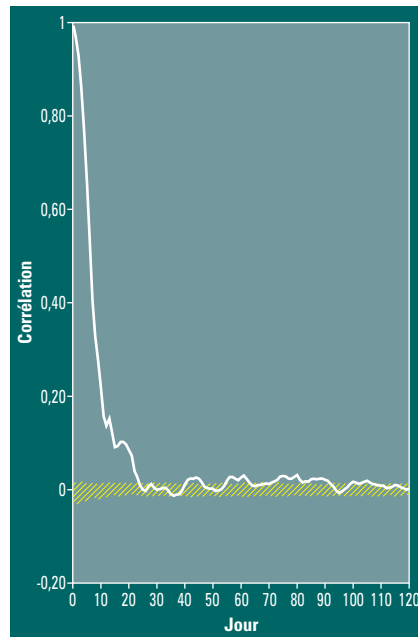


Figure 1 - Coefficient de corrélation du géopotential à 500 hPa sur l'hémisphère nord en fonction de l'échéance de prévision (jours). La zone hachurée correspond à l'intervalle de confiance à 95 % d'une prévision aléatoire.

à un intervalle de confiance à 95 % pour un coefficient de corrélation empirique dont la véritable valeur est zéro. On constate que la courbe du coefficient de corrélation se situe assez souvent au-dessus de la zone hachurée et ne passe jamais en dessous. À ce stade, nous avons une présomption que toute la prévisibilité n'a pas disparu, mais pas encore une preuve.

Si les prévisions quotidiennes étaient complètement indépendantes de l'observation, leur moyenne temporelle ne serait pas corrélée avec la moyenne des observations correspondantes. Considérons la moyenne des jours 31 à 120, pour éliminer les cas où la prévisibilité des données quotidiennes reste visible. Le coefficient de corrélation atteint alors 0,10. Ce n'est pas une corrélation élevée, mais, comme on dispose de 60 prévisions du modèle pour chaque cas, on peut calculer un intervalle de confiance à 95 % et on trouve [0,08 ; 0,12]. Une autre façon de se convaincre qu'un tel score ne peut être obtenu par hasard est d'appliquer la même technique de chaotisation pour les moyennes de 90 jours. On trouve un intervalle similaire à la zone hachurée de la figure 1, à savoir [-0,02 ; 0,02]. Cela montre que, compte tenu de la taille des échantillons et de la stabilité statistique de l'estimateur, une corrélation de 0,10 ne peut pas être obtenue par hasard.

D'où vient cette infime prévisibilité quotidienne qu'un opérateur de moyenne temporelle met en évidence en réduisant

la partie imprévisible des fluctuations ? Des expériences de type Provost ont été menées en prenant, soit la situation initiale de l'année précédente et la température de la mer de l'année en cours, soit la situation initiale de l'année en cours et la température de la mer de l'année précédente. Les résultats montrent que si les températures de la mer ne sont pas les bonnes, il n'existe aucune prévisibilité à l'échelle saisonnière. Mais, si les températures de la mer sont les bonnes, il y a de la prévisibilité. Sous les tropiques, cette prévisibilité ne dépend pas du fait que les conditions initiales atmosphériques soient ou non les bonnes. Aux latitudes tempérées, le fait de dégrader les conditions initiales dégrade un peu les scores. Il est donc clair que les températures de la mer, par leur évolution lente, impriment à l'atmosphère un signal de grande échelle auquel se superposent des fluctuations prévisibles seulement à une échéance inférieure à la vingtaine de jours. Le rôle, de second plan, joué par les conditions initiales dans la prévisibilité des latitudes tempérées s'explique par la présence de grandes ondes atmosphériques quasi stationnaires ainsi que par le caractère persistant du contenu en eau du sol et de la couverture de neige qui sont des variables évolutives dans un MCG.

Puisqu'un opérateur de moyenne temporelle augmente la prévisibilité par effet de lissage, les chercheurs ont tout de suite pensé à produire un ensemble de prévisions au lieu d'une seule. Cette technique est appelée **méthode de Monte-Carlo** ou **prévision d'ensemble**. Pour produire un ensemble de petite taille, on peut effectuer des prévisions à partir de conditions initiales décalées dans le temps (dans Provost, le décalage est de 24 heures). Pour produire des ensembles dont la taille dépasse la dizaine, on introduit une petite perturbation aléatoire dans les conditions initiales ou dans les conditions aux limites. Dans tous les cas, la distance entre deux prévisions individuelles appartenant à un même ensemble ne croît plus au-delà d'une vingtaine de jours d'intégration. Dans l'expérience Persist que nous analysons ici, 60 prévisions ont été effectuées pour chacun des 15 hivers, en perturbant la température de surface de la mer. Le score de la prévision saisonnière obtenu en moyennant les 60 prévisions individuelles passe de 0,10 à 0,24. On ne peut pas espérer aller beaucoup plus loin en augmentant la taille de l'ensemble car l'observation ne contient qu'un seul membre. Pour améliorer le score, il faut améliorer le MCG et affiner la prévision de température de la

(1) La **prévision par persistance** consiste à prévoir, dans le futur, des états identiques à l'état initial. Dans ce cas précis, la température de la mer est considérée comme constante au cours de l'évolution.

Tableau 1 - Caractéristiques des différentes expériences numériques mentionnées dans le texte.

	Nombre de modèles	Taille des ensembles	Température de la mer	Nombre d'années	Observation de référence
<b>Provost</b>	4 (3 en été)	9	Observée	15	ERA15
<b>Elmasifa</b>	1	3	Prévision statistique	15 (+ temps réel)	Archives des SMN
<b>Provost 2</b>	1	60	Observée	15	ERA15
<b>Persist</b>	1	60	Prévision statistique	15	ERA15
<b>Couple</b>	1	15	Couplage	15	ERA15
<b>Demeter</b>	7	9	Couplage	30	ERA40

mer. On peut avoir une idée du second facteur d'amélioration en utilisant la température de la mer réellement observée. Ce n'est plus vraiment une prévision, mais c'est ainsi que l'expérience Provost a fonctionné. Le coefficient de corrélation grimpe alors à 0,41. Les caractéristiques techniques des différentes expériences mentionnées ici (Provost, Persist, etc.) sont résumées dans le tableau 1.

## L'expérience Provost et ses résultats

### Les conditions de l'expérience

L'absence de données homogènes sur le globe pendant une longue période constituait la plus grande difficulté pour produire des prévisions rétrospectives. En 1994, le CEPMMT a lancé une vaste expérience pour retraiter les observations de la période 1979-1993. Cet exercice est connu sous le nom de **réanalyse** du CEPMMT ou ERA15. Dans le projet Provost, les données ERA15 ont servi de conditions initiales atmosphériques, de conditions aux limites océaniques et de données de vérification. L'expérience de prévisibilité a reposé sur quatre modèles : Arpège (Météo-France), IFS (CEPMMT), Unified Model (Met Office britannique) et

Arpège à plus haute résolution horizontale (appelé ici EDF). Pour chacune des quinze années disponibles, quatre prévisions ont été effectuées (une pour chaque saison, EDF ne traitant que les cas d'hiver). Enfin, chaque prévision consistait en neuf intégrations sur quatre mois du modèle, à partir de conditions initiales décalées de 24 heures. Ainsi, la prévision de l'hiver 1979-1980 utilise neuf simulations numériques démarrant les 22, 23, ..., 30 novembre 1979. Quand on agrège les quatre modèles, on dispose de 36 simulations d'un même hiver. Cette approche, que l'on nomme **multimodèle**, offre l'avantage de prendre à la fois en compte l'incertitude sur les conditions initiales et celle sur la formulation des modèles. En revanche, dans cet exercice, l'incertitude sur les conditions aux limites n'est pas prise en compte puisque les 36 simulations font appel aux mêmes températures de surface de la mer.

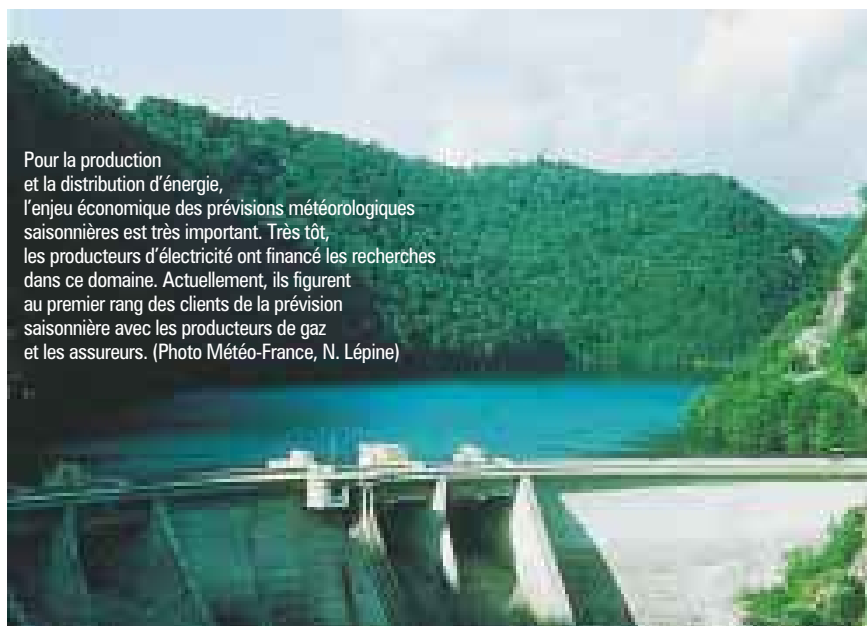
L'expérience Provost, qui représente des centaines d'années simulées, s'est déroulée pendant deux ans sur le calculateur du CEPMMT. Une expérience

du même type a été menée aux États-Unis, sous le nom de Dynamical Seasonal Prediction (DSP), uniquement pour les cas d'hiver et à partir d'une situation initiale fixée à la mi-décembre. Elle regroupait le National Center for Atmospheric Research (Ncar), le Center for Ocean-Land-Atmosphere Studies (Cola), le Goddard Space Flight Center (GSFC) et le NCEP. Les résultats des expériences Provost et DSP ont fait l'objet d'un ouvrage collectif (Palmer et Shukla, 2000).

### Les résultats

Pour illustrer les conclusions de l'expérience Provost, nous nous limiterons à deux paramètres, la température de la basse troposphère, au niveau 850 hPa, et les précipitations, ainsi qu'à deux saisons, l'hiver et l'été. Dans cette expérience, l'hiver désigne les mois de janvier, février et mars et l'été les mois de juillet, août et septembre. Ce décalage d'un mois par rapport aux saisons météorologiques traditionnelles est destiné à éviter des scores artificiellement élevés pour une prévision saisonnière. En effet, les simulations débutant respectivement fin novembre et fin mai, les situations prévues pour les quinze premiers jours des mois de décembre et de juin sont assez proches des observations correspondantes. Par ailleurs, nous considérons ici uniquement le score de la moyenne du multimodèle (36 cas d'hiver ou 27 cas d'été).

Les figures 2 et 3 concernent les prévisions d'hiver. On utilise comme score le coefficient de corrélation simple entre les 15 valeurs observées et prévues pour chaque position géographique. Le coefficient de corrélation pour la température est positif pratiquement sur tout le globe. Des valeurs supérieures à 0,6 sont obtenues sur les océans Indien et Pacifique. Sur l'Europe, à l'exception de l'Angleterre, le coefficient dépasse 0,2. La qualité de la prévision est



Pour la production et la distribution d'énergie, l'enjeu économique des prévisions météorologiques saisonnières est très important. Très tôt, les producteurs d'électricité ont financé les recherches dans ce domaine. Actuellement, ils figurent au premier rang des clients de la prévision saisonnière avec les producteurs de gaz et les assureurs. (Photo Météo-France, N. Lépine)

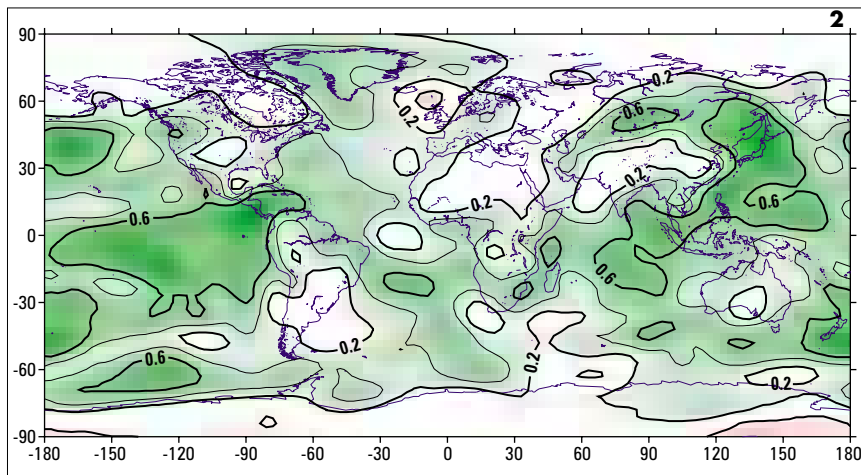
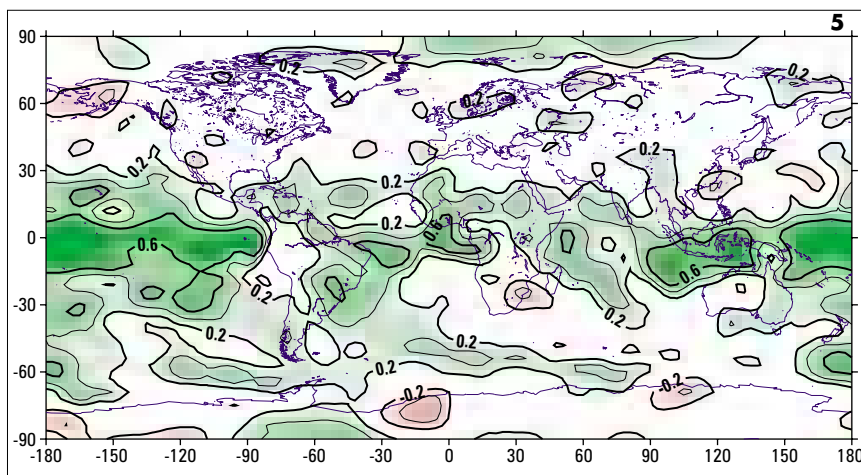
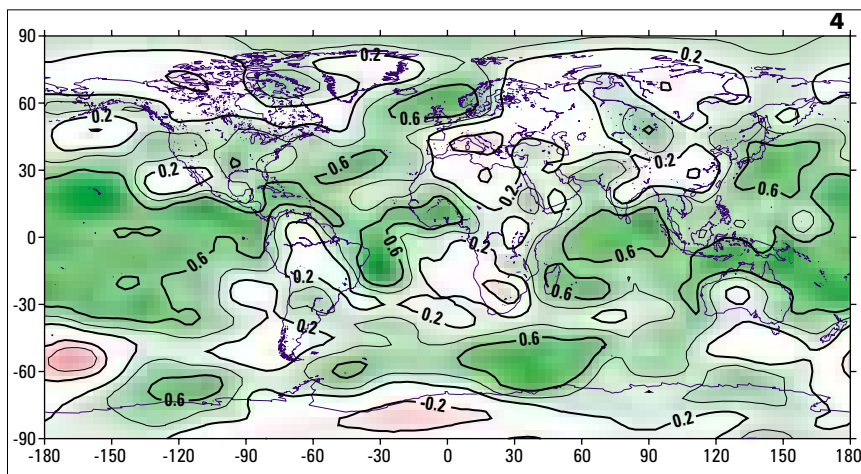
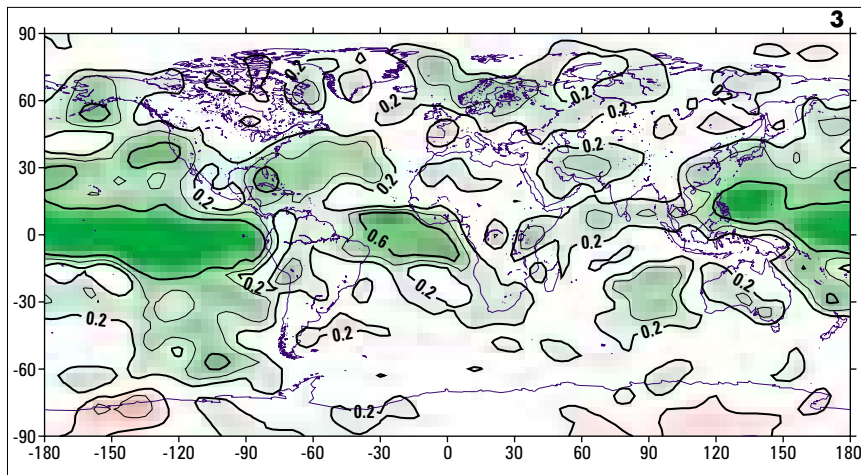


Figure 2 - Coefficients de corrélation pour une prévision saisonnière de température à 850 hPa en hiver. Isolignes  $\pm 0,2$ ,  $\pm 0,4$  et  $\pm 0,6$  avec coloration progressive en vert au-dessus de 0,2 et en rouge en dessous de  $-0,2$ .

Figure 3 - Idem figure 2 pour les précipitations en hiver.

Figure 4 - Idem figure 2 pour la température en été.

Figure 5 - Idem figure 2 pour les précipitations en été.



moins en ce qui concerne les précipitations. Des noyaux négatifs apparaissent çà et là, en particulier sur la France. C'est sous les tropiques que l'on trouve les corrélations les plus élevées.

En été (figures 4 et 5), les corrélations sont un peu plus faibles qu'en hiver, notamment aux latitudes tempérées : en particulier, la prévisibilité de la température sur l'Europe se limite au nord du continent. En ce qui concerne les précipitations tropicales, le Brésil et le Sahel ont des scores honorables, mais on ne peut pas en dire autant de la mousson indienne.

Le caractère prévisible d'un paramètre est dépendant de la région et de la saison. On va voir que la prévisibilité est aussi variable d'une année sur l'autre. Revenons au géopotential en hiver sur l'hémisphère nord, afin d'avoir un score suffisamment stable pour être statistiquement significatif. La figure 6 montre, pour chaque hiver, la valeur du coefficient de corrélation d'anomalie de chaque modèle et du multimodèle. On constate que, pour quatre hivers, le coefficient du multimodèle dépasse 0,5 ; trois de ces hivers correspondent à un phénomène El Niño, ce qui n'est pas un hasard. On constate aussi que, pour deux hivers, le score est négatif, ce qui caractérise une mauvaise prévision. Il faut quand même noter que le modèle 1 a quatre scores négatifs, le modèle 2 en a trois, le modèle 3 en a cinq et le modèle 4 en a deux. On peut se poser la question de l'intérêt de prendre plusieurs modèles, plutôt que de prendre le meilleur. Le tableau 2 montre qu'en moyenne, le multimodèle ne fait pas mieux que le modèle 4 pour le géopotential. Mais il indique aussi que le modèle 4 n'est pas le meilleur pour la prévision de température et que le multimodèle reste toujours au moins aussi bon que le meilleur des quatre. Ce résultat se retrouve pour les autres saisons et pour les régions tropicales.

## Les expériences similaires

Les scores du projet DSP ne concernent pas les mêmes années ni la même période de l'année. Ils ne sont donc pas directement comparables au tableau 2.

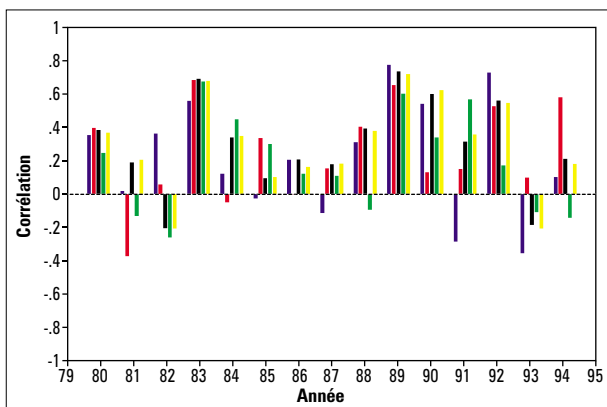


Figure 6 - Coefficient de corrélation du géopotential à 500 hPa sur l'hémisphère nord en hiver, pour chaque année de l'expérience Provost. Le score de chacun des quatre modèles est représenté par une barre de couleur et celui du multimodèle par une barre noire au centre de l'ensemble.

Cependant, pour l'hémisphère nord, on retrouve les meilleurs scores pour les hivers 1982-1983, 1988-1989 et 1991-1992. En dehors de Provost et de DSP, des expériences ont été menées au Japon (Kobayashi et al., 2000) et au Canada (Derome et al., 2001). La communauté scientifique s'accorde pour constater qu'avec les modèles actuels et une prévision parfaite de température de la mer, la corrélation de géopotential à 500 hPa sur l'hémisphère nord en hiver ne dépasse pas en moyenne 0,4, mais qu'elle présente de fortes variations d'une année sur l'autre et peut atteindre 0,7 certaines années. En ce qui concerne des scores plus localisés, la grande diversité traduit sans doute le fait qu'il n'existe pas encore de série assez longue de couples prévision-observation pour obtenir des scores stables.

## Après Provost

### Une nouvelle version du modèle atmosphérique

L'expérience Provost a fait usage de modèles atmosphériques développés au début des années 1990. En cinq ans, de gros progrès ont été accomplis en modélisation. En ce qui concerne Météo-France, le modèle Arpège-Climat a été profondément remanié

dans la partie dynamique, la discrétisation horizontale, verticale et tempo-

relle, ainsi que dans la paramétrisation des processus physiques. On trouvera dans Déqué (2001) une description sommaire des changements et plus de détails sur les expériences mentionnées ci-dessous. Avec la version 3 d'Arpège-Climat, la simulation du climat présent est plus réaliste qu'avec la version 1, aussi bien sur l'Europe que sous les tropiques. Nous n'avons pas résisté à la tentation de chercher si cette nouvelle version améliorerait la prévisibilité saisonnière par rapport à son aînée de cinq ans. Une expérience du même type que Provost, limitée à l'hiver et à l'été, mais comprenant 60 prévisions individuelles au lieu de 9, a été réalisée en 2000. Quand on choisit au hasard 9 prévisions parmi les 60, on obtient en moyenne un coefficient de corrélation de géopotential à 500 hPa sur l'hémisphère nord en hiver de 0,37, ce qui est un progrès par rapport à la valeur de 0,31 obtenue dans Provost avec Arpège-Climat version 1. Il existe donc un espoir de voir les progrès de la modélisation se traduire par des progrès de la prévisibilité. Deux remarques viennent tempérer cet optimisme. Si l'on calcule un intervalle de confiance à 95 % lié au tirage des 9 prévisions parmi les 60, on trouve [0,30 ; 0,43]. La valeur plus faible obtenue dans Provost peut donc être la conséquence d'un manque de chance. Si l'on regarde le score de température sur l'Europe, il est plus faible dans la nouvelle version : l'amélioration n'est pas uniforme ou, plus vraisemblablement, les estimateurs

locaux de score ne sont pas assez stables. Avec les 60 prévisions, le score passe de 0,37 à 0,41, mais ce n'est pas en augmentant encore la taille de l'ensemble de prévisions que l'on atteindra 0,9 (voir Déqué, 1997 pour une étude de l'asymptote).

## La prévision statistique de la température de la mer

Une autre expérience menée avec la version 3 d'Arpège-Climat a consisté à se rapprocher d'une prévision opérationnelle en ne faisant plus l'hypothèse d'un océan parfaitement connu. Depuis 1999, le service de prévision opérationnelle de Météo-France produit chaque mois une prévision à quatre mois d'échéance, à des fins d'analyse interne pour l'instant, mais bientôt pour des partenaires extérieurs. Contrairement à Provost, il n'est évidemment plus possible d'utiliser la température de la mer observée. Un algorithme de prévision statistique a été mis au point au CNRM dans le cadre du projet Elmasifa de la Commission européenne, ciblé sur la prévision de précipitations hivernales sur le Maghreb. Il a été récemment amélioré pour créer différents ensembles de température de la mer : il est naturel de faire porter l'incertitude sur cette variable qui donne une information essentielle à l'échelle saisonnière. Dans les deux cas, c'est une simple régression linéaire à partir de la valeur du mois précédent qui est mise en œuvre. L'expérience Persist, mentionnée au début de cet article, consiste à refaire l'expérience Provost avec des températures de la mer prévues et avec la version 3 d'Arpège-Climat. Les ensembles contiennent chacun 60 prévisions individuelles, mais, ici, chaque prévision individuelle a sa propre température de la mer : au cours du premier mois, les différences sont minimes, mais, au bout du quatrième mois, les anomalies par rapport au climat moyen peuvent être très différentes. Le score du géopotential à 500 hPa sur l'hémisphère nord en hiver n'est plus que 0,24. On mesure ainsi le gain potentiel apporté par une bonne prévision de la température de la mer.

### Prévisions couplées océan-atmosphère

Pour espérer faire une bonne prévision de la température de la mer, il faut disposer d'un modèle couplé océan-atmosphère, surtout si l'on vise une échéance supérieure à quatre mois. Le projet Demeter est un projet européen, coordonné par le CEPMMT, qui se propose de poursuivre l'expérience Provost dans deux directions : s'appuyer sur la nouvelle réanalyse ERA40 du CEPMMT, qui s'étendra sur quarante ans (1960-1999), et introduire des modèles couplés océan-atmosphère. Les

Tableau 2 - Coefficients de corrélation moyens de Provost en hiver sur l'hémisphère nord pour chaque modèle individuel et pour le multimodèle.

	Géopotential	Température	Précipitations
Modèle 1	0,31	0,24	0,26
Modèle 2	0,31	0,28	0,29
Modèle 3	0,28	0,14	0,28
Modèle 4	0,38	0,24	0,36
Multi-modèle	0,38	0,29	0,37

partenaires sont plus nombreux que dans Provost, mais certains se limiteront à la période 1990-1999, beaucoup plus riche en observations océaniques. Le projet devrait aboutir fin 2003.

En attendant la fin du projet, nous avons mené au CNRM une expérience pré-Demeter. Elle couvre les 15 hivers et les 15 étés de Provost, avec des ensembles de 15 prévisions individuelles. Le score du géopotential à 500 hPa sur l'hémisphère nord en hiver est de 0,23, ce qui est comparable à ce que l'on obtient avec une prévision statistique de température de la mer. C'est sur la bande intertropicale (30° N-30° S) que l'on peut mesurer l'apport de la méthode couplée, seule capable de prévoir une évolution correcte du phénomène El Niño comme l'a montré, en temps réel, le cas des prévisions de 1997. Le tableau 3 contient les scores de précipitations pour janvier-février-mars et juillet-août-septembre. Toutes ces corrélations ont été calculées à partir d'ensembles de 15 prévisions individuelles, pour comparer des choses comparables.

	Janvier Février Mars	Juillet Août Septembre
Provost	0,53	0,42
Provost 2	0,47	0,40
Persist	0,39	0,34
Couple	0,47	0,34

Tableau 3 - Coefficients de corrélation moyens pour les précipitations tropicales sur deux saisons (janvier-février-mars et juillet-août-septembre) et diverses expériences : Provost en mode multimodèle, Provost refait avec la version 3 d'Arpège-Climat (Provost2), prévision statistique des températures de la mer (Persist) et modèle couplé océan-atmosphère (Couple).

Contrairement au géopotential sur l'hémisphère nord, les précipitations de la bande intertropicale sont très stables vis-à-vis du choix de l'ensemble, car chaque prévision individuelle d'un ensemble réagit à peu près de la même façon au forçage océanique. Les comparaisons du tableau 3 sont donc pertinentes. On y voit la supériorité du multimodèle sur chacun des modèles, de la température de la mer observée sur la température prévue et de la méthode couplée sur la méthode statistique. Dans le dernier cas, cette supériorité ne se manifeste qu'en hiver, car c'est la période de prédilection des événements El Niño.

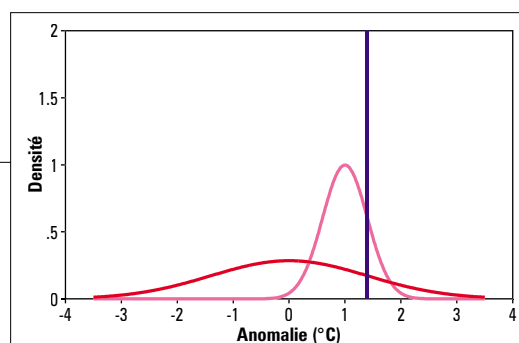
Figure 7 - Illustration d'une prévision probabiliste : le paramètre est l'anomalie hivernale de température sur la France. La courbe rouge est la distribution climatologique, la courbe magenta est la distribution probabiliste prévue, la barre bleue est l'anomalie de température observée.

## Les limites de la prévision saisonnière

### La prévision probabiliste

Les scores mentionnés jusqu'ici sont encourageants sur le plan scientifique, car ils démontrent que le système, d'équations employé pour modéliser l'atmosphère n'a pas un comportement aberrant lorsqu'il porte sur plusieurs mois de simulation. Sur le plan pratique, c'est une autre histoire. Un utilisateur de ce type de prévision va connaître à peu près autant de satisfactions que de déconvenues et il lui faudra de nombreuses années avant de réaliser qu'une prévision dotée d'une corrélation de 0,2 est meilleure qu'une prévision fantaisiste. On voit sur la figure 6 que certaines années sont intrinsèquement plus prévisibles que d'autres.

La méthode scientifique pour gérer l'incertitude liée à une information consiste à faire appel à la théorie des probabilités. La figure 7 illustre ce que peut être une prévision probabiliste. Pour un hiver donné, la température moyenne sur la France est 1,4 °C au-dessus de la normale saisonnière (ligne bleue). Il s'agit donc d'un hiver modérément doux. Au lieu de chercher à prévoir exactement cette valeur avec un modèle, on va produire une courbe de densité de probabilité (courbe magenta). Dans cet exemple, on est pratiquement certain que l'anomalie de température sera positive et qu'elle ne dépassera pas 2 °C au-dessus de la normale. Même si le maximum de densité ne coïncide pas avec la valeur observée, on peut dire que l'on a affaire ici à une bonne prévision. Pour lui attribuer un score, il faut la comparer à une prévision de référence disponible à un coût quasi nul. Dans le cas d'une prévision déterministe, la référence est implicitement la prévision d'une anomalie nulle. Dans le cas d'une prévision probabiliste, la prévision à battre est matérialisée par



la courbe rouge qui correspond à la distribution des anomalies observées sur une période de référence (ici 15 ans avec lissage par une courbe de Gauss). Il s'agit de la prévision climatologique au sens probabiliste.

Pour Provost, une présentation plus complète des méthodes et des résultats serait assez austère et plutôt décevante. Nous nous limiterons à quelques généralités. L'évaluation par un score est de la plus haute importance en approche probabiliste, car une prévision probabiliste peut toujours être considérée comme juste : si l'on prédit avec une probabilité de 0,9 un hiver froid, on n'exclut pas pour autant la possibilité d'un hiver doux. Le critère le plus naturel pour mesurer un écart entre une distribution prévue (courbe magenta) et observée (ligne bleue) est la distance quadratique. Cette distance est ensuite comparée à la distance entre la distribution de référence (courbe rouge) et l'observation. On obtient alors le **score de Brier** (1950), qui est en quelque sorte le pendant de l'erreur quadratique moyenne pour la prévision déterministe. Le problème est que, dans Provost et ses successeurs, le score probabiliste des modèles est souvent inférieur à celui de la prévision climatologique. Il y a deux raisons pour cela. La première est que les faibles niveaux de corrélation traduisent une prévisibilité assez faible de la position du pic de densité. Quand on utilise l'erreur quadratique moyenne pour évaluer une prévision saisonnière déterministe, on se retrouve souvent en dessous de la prévision climatologique alors que le coefficient de corrélation est positif. La deuxième raison est plus fondamentale. Quand on considère que les diverses prévisions individuelles d'un ensemble sont des réalisations possibles du prédicand<sup>(1)</sup>, on commet l'erreur de croire que la seule source d'incertitude provient de l'état initial. Autrement dit, on fait comme si le modèle représentait parfaitement les lois de la nature et on sous-estime la dispersion autour du pic de densité qui devrait aussi tenir compte du fait que le modèle peut avoir une réponse à son forçage externe différente de celle de la réalité. Par exemple, un modèle peut produire en moyenne des hivers froids sur l'Europe lors des phénomènes El Niño tandis que, dans la réalité, les hivers ne sont ni plus chauds ni plus froids en moyenne sur cette région.

Pour corriger le premier travers, il suffit d'utiliser un critère d'évaluation à la fois moins sévère et plus proche du besoin

(1) On appelle **prédicand** la variable que l'on cherche à prévoir.

des utilisateurs. Pour la prévision déterministe, le coefficient de corrélation joue bien ce rôle. Pour la prévision probabiliste, le candidat le plus sérieux est le critère de valeur économique. Il s'appuie sur la théorie de la décision (Murphy, 1977). Ses limitations sont le choix arbitraire d'un seuil, puisqu'il ne s'applique qu'aux prévisions à deux modalités, et le choix du rapport coût sur perte, qui dépend de l'utilisateur. On peut quand même produire des résultats robustes en prenant la partition qui distingue anomalie positive et anomalie négative et en choisissant le rapport coût sur perte égal à 1/2 ; en effet, c'est pour cette valeur du rapport coût sur perte que l'utilisateur a le plus besoin d'une prévision probabiliste pour optimiser sa stratégie de décision. Avec ce critère, on montre dans Provost que, là où la corrélation est significativement positive, par exemple pour les pluies tropicales, on obtient des gains de 20 % : si une stratégie de décision fondée sur la seule climatologie apporte en moyenne un bénéfice B, et si la stratégie idéale (supposant le futur parfaitement connu) apporte un bénéfice B+100, alors la stratégie utilisant les prévisions numériques apporte B+20 en moyenne.

Le deuxième défaut de l'approche classique est plus difficile à corriger : il ne suffit pas de changer de thermomètre, il faut trouver le bon médicament. L'utilisation de plusieurs modèles est la première méthode qui vient à l'esprit. Avec des modèles différents, on va implicitement prendre en compte l'incertitude sur les équations car chaque modèle utilise une formulation différente pour représenter les phénomènes. Hélas, cela ne suffit pas, car, en augmentant artificiellement l'écart type des distributions prévues, on parvient encore à améliorer un peu les scores, ce qui montre que la dispersion du multimodèle est toujours sous-estimée.

La figure 6 semble montrer que, lors des hivers pour lesquels les modèles sont d'accord, il faut attribuer une plus grande confiance à la prévision. Cependant, l'approche probabiliste classique, qui considère que les prévisions individuelles de l'ensemble forment un échantillon de la loi probabiliste qui produit la réalité observée, n'exploite pas bien cette propriété. En effet, quand on produit à partir de Provost des prévisions probabilistes gaussiennes dans lesquelles l'écart type est remplacé par la moyenne des 15 écarts types, les scores ne sont pas dégradés. Cela veut dire que l'information contenue dans la variabilité, d'un hiver sur l'autre, de la dispersion n'est pas source de prévisibilité. Une consé-

quence de ce résultat est qu'à l'heure actuelle, la meilleure prévision probabiliste que l'on puisse obtenir à partir de Provost consiste à prendre une loi de Gauss ayant pour moyenne la prévision déterministe de l'hiver considéré et pour écart type un écart type constant issu de la prévision climatologique.

Pour essayer de faire de vraies prévisions probabilistes, deux approches sont explorées. Elles combinent les résultats du ou des modèles numériques et une connaissance statistique des prévisions passées. La première consiste à chercher à prévoir l'écart type de la distribution prévue par une méthode statistique. Les prédicteurs<sup>(1)</sup> évidents sont la dispersion du modèle et l'état de l'indice d'oscillation australe. Mais il en existe sûrement d'autres et l'on tombe sur un problème de stabilité statistique avec les 15 cas de Provost. La deuxième approche reprend le concept d'analogues utilisé en prévision statistique non linéaire (Clark et Déqué, 2003). Si l'on dispose d'une série d'hivers assez longue, on va chercher dans cette série tous les cas où le modèle s'est comporté de manière similaire et l'on prend comme ensemble de prévisions les observations de ces années-là. La méthode offre deux avantages. Si un paramètre n'a aucune prévisibilité, les années analogues observées seront obtenues au hasard et la prévision probabiliste ne sera pas plus mauvaise que la prévision climatologique (avec des ensembles de taille supérieure à 10), contrairement à l'approche classique qui peut être pire que la prévision climatologique, par excès de confiance. Le deuxième avantage vient de l'utilisation d'ensembles de prévisions. La difficulté avec les méthodes d'analogues vient de ce que l'atmosphère a tellement de degrés de liberté que les meilleurs analogues ne se ressemblent pas vraiment quand ils sont issus d'un échantillon de quelques dizaines. Avec des ensembles de 10 membres, la recherche des analogues se fait dans un échantillon 10 fois plus grand. Cependant, Provost est trop court pour offrir une variété suffisante en termes de forçage par les températures de la mer (trois épisodes El Niño et deux épisodes La Niña en quinze ans). Là encore, les résultats de Demeter sont attendus avec impatience.

## La prévision statistique

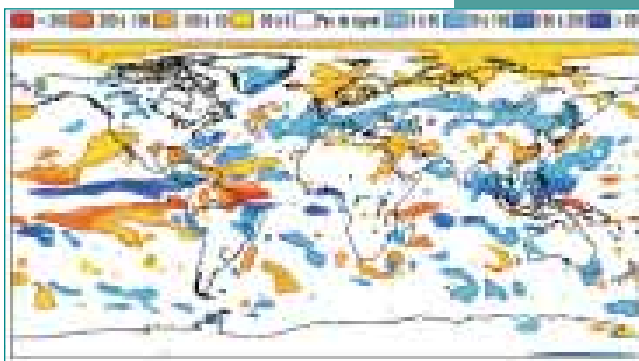
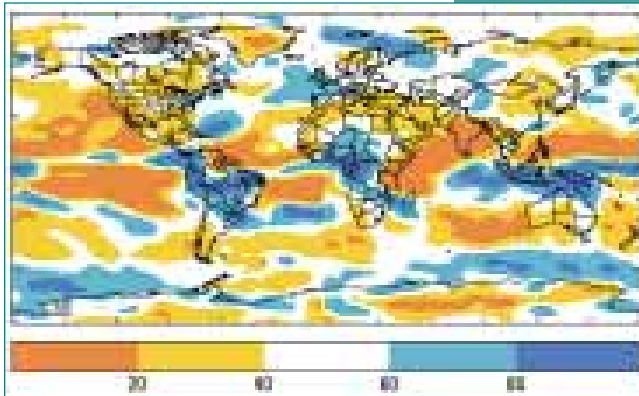
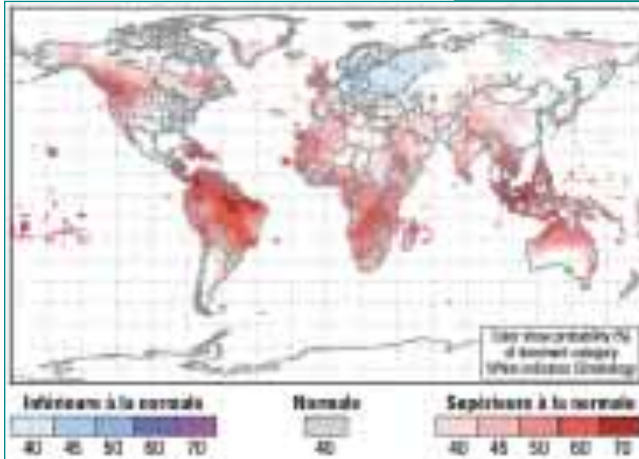
Les prévisionnistes à longue échéance ont d'abord cherché des lois empiriques pour obtenir les prévisions que le public ou l'État leur demandait. Le découvreur de l'oscillation australe,

G. Walker, cherchait à prévoir la mousson indienne. Comme pour les proverbes, on a d'abord fait appel à la mémoire humaine. Puis la représentation graphique de certains indices, comme l'indice d'oscillation australe ou l'indice de l'oscillation nord-atlantique, a permis de faire des synthèses plus efficacement. Enfin, les méthodes d'analyse de données, comme la régression linéaire multiple avec sélection des prédicteurs ou l'analyse factorielle discriminante, ont permis d'optimiser la solution par rapport à un critère donné. Tant que les calculs étaient faits à la main ou avec des machines à calculer rudimentaires, le danger de cette approche empirique ne s'est pas manifesté. De nos jours, n'importe quel micro-ordinateur est capable, en quelques minutes, de tester toutes les combinaisons de prédicteurs et de coefficients jusqu'à obtenir une série prévue qui « colle » pratiquement à la série observée. Il existe un garde-fou qui consiste à séparer une série d'apprentissage (pour l'optimisation) et une série test (pour la vérification). Mais, si l'on n'est pas content du score sur la série test, rien n'empêche de recommencer avec d'autres prédicteurs sur la série d'apprentissage. Cela n'implique pas que la méthode de prévision ainsi mise au point n'ait aucune valeur, mais le score promis est certainement au-dessus du score que l'on obtiendra en exploitation réelle. C'est pourquoi les méthodes numériques annoncent des corrélations d'au mieux 0,4, tandis que les méthodes statistiques promettent des corrélations de 0,6, voire de 0,8. Le jour où l'on disposera d'une puissance de calcul suffisante pour mener des expériences de type Provost en 48 heures, la crédibilité des scores correspondants sera affectée car on pourra tenter toutes sortes de modifications dans le modèle et garder la version qui donne le meilleur score sur la période de référence.

Dans le cadre du projet européen Elmasifa, une expérience intéressante a été menée pour prévoir les précipitations sur le Maroc en hiver. Une équipe a reçu les données de la période 1950-1978 et a mis en place deux algorithmes : l'un pour prévoir les précipitations sur le Maroc, l'autre pour prévoir les températures de la mer sur le globe. Les algorithmes ont été fournis à une deuxième équipe qui disposait de données pour la période 1979-1993. Des prévisions statistiques de précipitations et de température de la mer ont alors été effectuées.

(1) Les **prédicteurs** sont les variables explicatives employées pour prévoir la variable considérée (le *prédicteand*).

## Quelques exemples de prévisions saisonnières



À l'heure actuelle, seuls quelques organismes produisent des prévisions saisonnières à l'aide de modèles climatiques :

- Le CEPMMT, le NCEP et le Met Office disposent de modèles océaniques et atmosphériques couplés ; le modèle du CEPMMT s'est montré particulièrement apte à prédire les derniers épisodes El Niño et La Niña.
- Les autres, dont Météo-France, utilisent des modèles où la température de la surface de la mer est prescrite en début de simulation ; elle peut cependant évoluer, avec en général un rappel vers la climatologie.

Les prévisions saisonnières, qui sont obtenues par la technique de prévision d'ensemble, peuvent être présentées sous forme déterministe (la prévision moyenne par exemple) ou probabiliste. Les modèles employés ont une résolution spatiale de 200 à 300 km. Les prévisions sont en général de meilleure qualité pour la température que pour les précipitations et pour l'hiver (de l'hémisphère considéré) que pour l'été.

Les figures représentent quatre exemples de prévision saisonnière, calculés à partir d'états initiaux de janvier ou février 2003 pour la période mars-mai 2003.

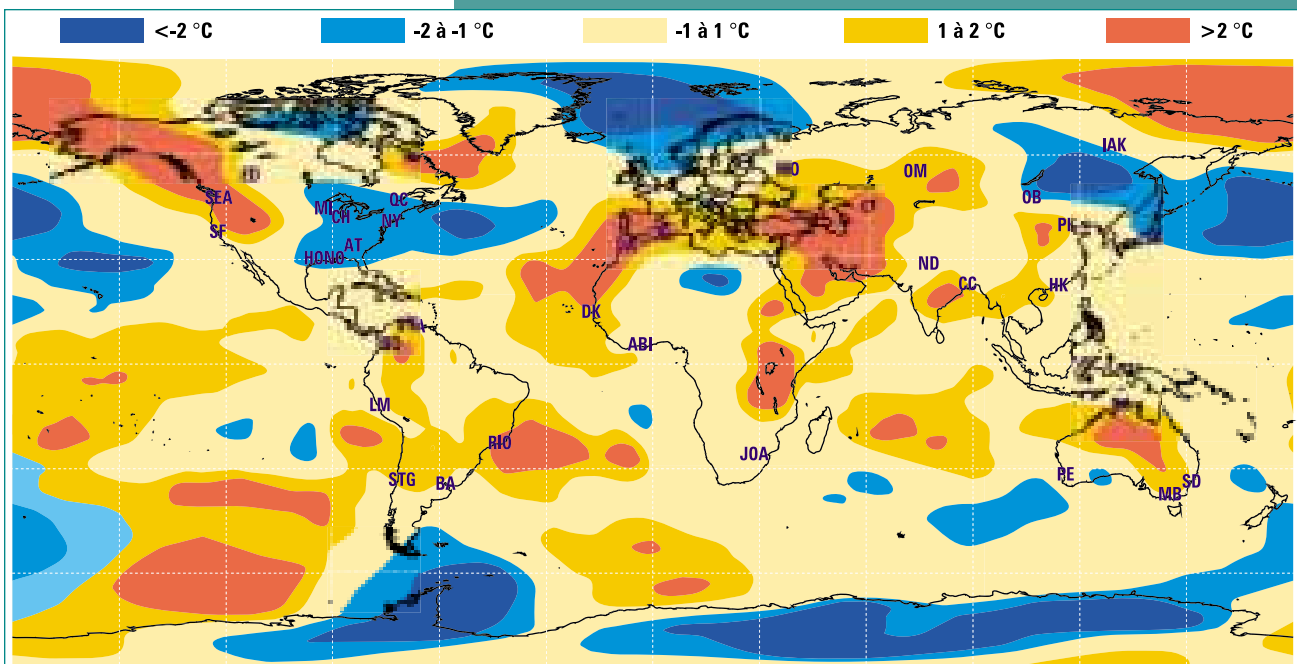
De haut en bas :

International Research Institute (IRI). Prévision saisonnière multimodèle calculée à partir d'états initiaux de février 2003 pour la période mars-mai 2003. Probabilité (en %) d'anomalie de température par rapport à la normale climatologique (probabilité du tercile le plus probable). Les anomalies positives sont représentées en rouge, les anomalies négatives en bleu.

Met Office. Prévision saisonnière de quantité de précipitations, calculée le 28 février 2003 pour la période mars-mai 2003. Probabilité (en %) que le cumul de précipitations durant cette période soit supérieur à la normale climatologique. Les couleurs orange et rouge indiquent des conditions plus sèches que la normale, tandis que les couleurs bleu clair et bleu foncé sont représentatives de conditions plus pluvieuses que la normale. (© Crown copyright, UK Met Office)

CEPMMT. Prévision saisonnière d'anomalie du cumul trimestriel de précipitations (en mm), calculée à partir du 14 février 2003 pour la période mars-mai 2003. (© CEPMMT)

Météo-France. Prévision saisonnière d'anomalie de température à 850 hPa (en °C), calculée à partir de dates de janvier 2003 pour la période mars-mai 2003. (© Météo-France)



Puis un ensemble de prévisions obtenues avec un modèle numérique forcé par ces températures de la mer a été constitué et l'on en a extrait les prévisions de précipitations. Avec la méthode purement statistique, la corrélation est  $-0,26$ . Avec la prévision numérique forcée par les températures de la mer prévues, la corrélation est  $0,22$ . En extrayant les précipitations prévues par Provost, on obtient une corrélation de  $0,48$  (Provost utilise les températures de la mer observées). Le protocole expérimental qui consiste à isoler, non seulement les fichiers, mais aussi leur traitement est un bon garant de la crédibilité des scores.

S'il est clair que les méthodes statistiques pures ont peu d'avenir, ne serait-ce que parce que les fichiers actuels de données couvrent à peu près cinquante ans et ne sont extensibles que vers le futur, les méthodes hybrides vont jouer un rôle croissant. Tout d'abord, les modèles ne produisent une information crédible qu'à grande échelle. Pour obtenir des paramètres utiles comme la température à 2 mètres à Orly, la solution passe par l'adaptation statistique. Ensuite, la technique de prévision d'ensemble est nécessaire mais pas suffisante pour produire une prévision probabiliste plus efficace que la prévision climatologique. Comme on l'a vu p. 27, une partie de l'information sur la dispersion ne peut venir que sous forme statistique.

## Conclusion

Ce qui précède montre que le choix du mot prévision pour désigner la prévision saisonnière est mal approprié et source

de malentendus car il évoque une annonce de ce qui va se passer dans le futur. Le futur de l'atmosphère est gouverné essentiellement par le chaos et l'objet de la prévision saisonnière est seulement de dire en quoi le climat de la saison à venir se distingue du climat des dernières années. Ce futur n'est qu'une réalisation possible d'un ensemble que l'on peut appeler le climat d'une année particulière. Sur le plan pratique, ce climat peut s'estimer par le calcul d'une moyenne saisonnière, mais le climat ne se limite pas à la seule moyenne ; il contient aussi de l'information sur la variabilité à l'intérieur de la saison et sur la fréquence de certains événements. À partir d'un ensemble de prévisions, on peut produire une information assez riche et statistiquement fiable, à défaut d'être physiquement sûre. La difficulté vient de ce que les caractéristiques réelles de la saison à prévoir ne sont estimées que sur une seule réalisation et n'ont pas cette stabilité statistique.

Une comparaison un peu osée mais instructive peut être faite avec les pronostics hippiques. Aucun pronostiqueur sérieux ne prétendra donner à l'avance le tiercé gagnant. Mais le pronostiqueur possède des informations sur le terrain, la santé et les performances passées qui font que les résultats de tous les chevaux ne sont pas équiprobables. La connaissance, même approximative car les modèles numériques sont loin d'être parfaits, du climat, c'est-à-dire de l'espérance mathématique des paramètres météorologiques, permet de choisir une stratégie de décision qui peut être fructueuse sur le long terme. Le grand public, c'est-à-dire les personnes qui ne sont pas placées devant un choix dont les conséquences sont chiffrables en

fonction du temps qu'il va faire au cours de la saison à venir, n'est pas concerné par la prévision saisonnière.

La grande question scientifique liée à la prévision saisonnière n'est pas la production de prévisions, mais l'évaluation de ce qui peut être utilisé dans une prévision. Il faut trier entre le bon grain de la réponse correcte aux forçages et l'ivraie des fantaisies du modèle. Un projet comme Demeter couvrant au moins trente ans, mettant en jeu la plupart des modèles européens et utilisant le couplage avec l'océan va permettre d'aller plus loin dans cette voie. Les progrès majeurs de la prévision au cours des dix prochaines années viendront alors d'une bonne initialisation de l'état de l'océan en amont (on peut citer le projet Mercator) et d'un bon traitement statistique en aval pour obtenir des probabilités crédibles et une adaptation aux paramètres locaux de l'utilisateur. Pendant ce temps, les modèles atmosphériques et océaniques continueront de se perfectionner. On peut raisonnablement espérer que ces progrès se traduiront par une amélioration des scores.

## Remerciements

Je remercie mes collègues du CNRM et du CEPMMT pour leur contribution dans les projets Provost et Demeter et pour l'aide qu'ils m'ont apportée à diverses étapes. La lisibilité de ce travail a été bien améliorée par les remarques pertinentes d'un relecteur anonyme. Le projet Demeter est cofinancé par la Commission européenne sous le numéro EVK2-1999-00197.

## Bibliographie

- **Brier G. W.**, 1950 : Verification of forecasts expressed in terms of probabilities. *Mon. Wea. Rev.*, 78, 1-3.
- **Clark R. T. et M. Déqué**, 2003 : Conditional probability seasonal predictions of precipitation. *Quart. J. Roy. Meteor. Soc.*, 129, 1-15.
- **Déqué M.**, 1997 : Ensemble size for numerical seasonal forecasts. *Tellus*, 49A, 74-86.
- **Déqué M.**, 2001 : Seasonal predictability of tropical rainfall : probabilistic formulation and validation. *Tellus*, 53A, 500-512.
- **Déqué M. et J.-F. Royer**, 1992 : The skill of extended-range extratropical winter dynamical forecasts. *J. Climate*, 5, 1346-1356.
- **Derome J., G. Brunet, A. Plante, N. Gagnon, G. J. Boer, F. W. Zwiers, S. J. Lambert, J. Sheng et H. Ritchie**, 2001 : Seasonal predictions based on two dynamical models. *Atmos.-Ocean*, 39, 4, 485-501.
- **Gilchrist A.**, 1986 : Long-range forecasting. *Quart. J. Roy. Meteor. Soc.*, 112, 567-592.
- **Kobayashi C., K. Takano, S. Kusunoki, M. Sugi et A. Kitoh**, 2000 : Seasonal predictability in winter over eastern Asia using the JMA global model. *Quart. J. Roy. Meteor. Soc.*, 126, 2111-2124.
- **Lorenz E. N.**, 1963 : Deterministic non-periodic flow. *J. Atmos. Sci.*, 20, 130-141.
- **Murphy A. H.**, 1977 : The value of climatological, categorical and probabilistic forecasts in the cost-loss ratio situation. *Mon. Wea. Rev.*, 105, 803-816.
- **Owen J. A. et T. N. Palmer**, 1987 : The impact of El Niño on an ensemble of extended-range forecasts. *Mon. Wea. Rev.*, 115, 2103-2117.
- **Palmer T. N. et D. L. T. Anderson**, 1994 : The prospects for seasonal forecasting - a review paper. *Quart. J. Roy. Meteor. Soc.*, 120, 755-793.
- **Palmer T. N. et J. Shukla**, 2000 : DSP/PROVOST special volume. *Quart. J. Roy. Meteor. Soc.*, 126, 1989-2351.
- **Shukla J.**, 1981 : Dynamical predictability of monthly means. *J. Atmos. Sci.*, 1981, 2547-2572.
- **Smagorinsky J.**, 1969 : Problems and promises of deterministic extended range forecasting. *Bull. Amer. Meteorol. Soc.*, 50, 286-311.