

# Revue bibliographique des méthodes de prévision des débits A review of streamflow forecasting methods

V. Fortin, T. B.M.J. Ouarda, P. F. Rasmussen and B. Bobée

Volume 10, Number 4, 1997

URI: <https://id.erudit.org/iderudit/705289ar>

DOI: <https://doi.org/10.7202/705289ar>

[See table of contents](#)

## Publisher(s)

Université du Québec - INRS-Eau, Terre et Environnement (INRS-ETE)

## ISSN

0992-7158 (print)

1718-8598 (digital)

[Explore this journal](#)

## Cite this article

Fortin, V., Ouarda, T. B., Rasmussen, P. F. & Bobée, B. (1997). Revue bibliographique des méthodes de prévision des débits. *Revue des sciences de l'eau / Journal of Water Science*, 10(4), 461–487. <https://doi.org/10.7202/705289ar>

## Article abstract

A large number of models are available for streamflow forecasting. In this paper we classify and compare nine types of models for short, medium and long-term flow forecasting, according to six criteria:

1. validity of underlying hypotheses,
2. difficulties encountered when building and calibrating the model,
3. difficulties in computing the forecasts,
4. uncertainty modeling,
5. information required by each type of model, and
6. parameter updating.

We first distinguish between empirical and conceptual models, the difference being that conceptual models correspond to simplified representations of the watershed, while empirical model only try to capture the structural relationships between inputs to the watershed and outputs, such as streamflow.

Amongst empirical models, we distinguish between stochastic models, i.e. models based on the theory of probability, and non-stochastic models. Three types of stochastic models are presented: statistical regression models, Box-Jenkins models, and the nonparametric k-nearest neighbor method. Statistical linear regression is only applicable for long term forecasting (monthly flows, for example), since it requires independent and identically distributed observations. It is a simple method of forecasting, and its hypotheses can be validated a posteriori if sufficient data are available. Box-Jenkins models include linear autoregressive models (AR), linear moving average models (MA), linear autoregressive - moving average models (ARMA), periodic ARMA models (PARMA) and ARMA models with auxiliary inputs (ARMAX). They are more adapted for weekly or daily flow forecasting, since the yellow for the explicit modeling of time dependence. Efficient methods are available for designing the model and updating the parameters as more data become available. For both statistical linear regression and Box-Jenkins models, the inputs must be uncorrelated and linearly related to the output. Furthermore, the process must be stationary. When it is suspected that the inputs are correlated or have a nonlinear effect on the output, the k-nearest neighbor method may be considered. This data-based nonparametric approach simply consists in looking, among past observations of the process, for the k events which are most similar to the present situation. A forecast is then built from the flows which were observed for these k events. Obviously, this approach requires a large database and a stationary process. Furthermore, the time required to calibrate the model and compute the forecasts increases rapidly with the size of the database. A clear advantage of stochastic models is that forecast uncertainty may be quantified by constructing a confidence interval.

Three types of non-stochastic empirical models are also discussed: artificial neural networks (ANN), fuzzy linear regression and multivariate adaptive regression splines (MARS). ANNs were originally designed as simple conceptual models of the brain.

However, for forecasting purposes, these models can be thought of simply as a subset of non linear empirical models. In fact, the ANN model most commonly used in forecasting, a multi-layer feed-forward network, corresponds to a non linear autoregressive model (NAR). To capture the moving average components of a time series, it is necessary to use recurrent architectures. ANNs are difficult to design and calibrate, and the computation of forecasts is also complex. Fuzzy linear regression makes it possible to extract linear relationships from small data sets, with fewer hypotheses than statistical linear regression. It does not require the observations to be uncorrelated, nor does it ask for the error variance to be homogeneous. However, the model is very sensitive to outliers. Furthermore, a posteriori validation of the hypothesis of linearity is not possible for small data sets. MARS models are based on the hypothesis that time series are chaotic instead of stochastic. The main advantage of the method is its ability to model non-stationary processes. The approach is non-parametric, and therefore requires a large data set.

Amongst conceptual models, we distinguish between physical models, hydraulic machines, and fuzzy rule-based systems. Most conceptual hydrologic models are hydraulic machines, in which the watershed is considered to behave like a network of reservoirs. Physical modeling of a watershed would imply using fundamental physical equations at a small scale, such as the law of conservation of mass. Given the complexity of a watershed, this can be done in practice only for water routing. Consequently, only short term flow forecasts can be obtained from a physical model, since the effects of precipitation, infiltration and evaporation must be negligible. Fuzzy rule-based systems make it possible to model the water cycle using fuzzy IF-THEN rules, such as IF it rains a lot in a short period of time, THEN there will be a large flow increase following the concentration time. Each fuzzy quantifier is modeled using a fuzzy number to take into account the uncertainty surrounding it. When sufficient data are available, the fuzzy quantifiers can be constructed from the data. In general, conceptual models require more effort to develop than empirical models. However, for exceptional events, conceptual models can often provide more realistic forecasts, since empirical models are not well suited for extrapolation.

A fruitful approach is to combine conceptual and empirical models. One way of doing this, called extended streamflow prediction or ESP, is to combine a stochastic model for generating meteorological scenarios with a conceptual model of the watershed. Based on this review of flow forecasting models, we recommend for short term forecasting (hourly and daily flows) the use of the k-nearest neighbor method, Box-Jenkins models, water routing models or hydraulic machines. For medium term forecasting (weekly flows, for example), we recommend the k-nearest neighbor method and Box-Jenkins models, as well as fuzzy-rule based and ESP models. For long term forecasting (monthly flows), we recommend statistical and fuzzy regression, Box-Jenkins, MARS and ESP models. It is important to choose a type of model which is appropriate for the problem at hand and for which the information available is sufficient. Each type of model having its advantages, it can be more efficient to combine different approaches when forecasting streamflow.

# Revue bibliographique des méthodes de prévision des débits

## A review of streamflow forecasting methods

V. FORTIN<sup>1</sup>, T. B.M.J. OUARDA<sup>1</sup>, P. F. RASMUSSEN<sup>1</sup> et B. BOBÉE<sup>1\*</sup>

Reçu le 30 novembre 1996, accepté le 27 mai 1997\*.

### SUMMARY

A large number of models are available for streamflow forecasting. In this paper we classify and compare nine types of models for short, medium and long-term flow forecasting, according to six criteria: 1) validity of underlying hypotheses, 2) difficulties encountered when building and calibrating the model, 3) difficulties in computing the forecasts, 4) uncertainty modeling, 5) information required by each type of model, and 6) parameter updating.

We first distinguish between empirical and conceptual models, the difference being that conceptual models correspond to simplified representations of the watershed, while empirical model only try to capture the structural relationships between inputs to the watershed and outputs, such as streamflow.

Amongst empirical models, we distinguish between stochastic models, *i.e.* models based on the theory of probability, and non-stochastic models. Three types of stochastic models are presented: statistical regression models, Box-Jenkins models, and the nonparametric *k*-nearest neighbor method. Statistical linear regression is only applicable for long term forecasting (monthly flows, for example), since it requires independent and identically distributed observations. It is a simple method of forecasting, and its hypotheses can be validated *a posteriori* if sufficient data are available. Box-Jenkins models include linear autoregressive models (AR), linear moving-average models (MA), linear autoregressive - moving average models (ARMA), periodic ARMA models (PARMA) and ARMA models with auxiliary inputs (ARMAX). They are more adapted for weekly or daily flow forecasting, since they allow for the explicit modeling of time dependence. Efficient methods are available for designing the model and updating the parameters as more data become available. For both statistical linear regression and Box-Jenkins models, the inputs must be uncorrelated and linearly related to the output. Furthermore, the process must be stationary. When it is suspected that the inputs are correlated or have a nonlinear effect on the output, the *k*-nearest neighbor method may be considered. This data-based nonparametric approach simply consists in looking, among

1. Chaire en Hydrologie Statistique, Institut National de la Recherche Scientifique, INRS-Eau, 2800 rue Einstein, C.P. 7500, Sainte-Foy, Québec, G1V 4C7, Canada.

\* Les commentaires seront reçus jusqu'au 20 juillet 1998.

past observations of the process, for the  $k$  events which are most similar to the present situation. A forecast is then built from the flows which were observed for these  $k$  events. Obviously, this approach requires a large database and a stationary process. Furthermore, the time required to calibrate the model and compute the forecasts increases rapidly with the size of the database. A clear advantage of stochastic models is that forecast uncertainty may be quantified by constructing a confidence interval.

Three types of non-stochastic empirical models are also discussed; artificial neural networks (ANN), fuzzy linear regression and multivariate adaptive regression splines (MARS). ANNs were originally designed as simple conceptual models of the brain. However, for forecasting purposes, these models can be thought of simply as a subset of non linear empirical models. In fact, the ANN model most commonly used in forecasting, a multi-layer feed-forward network, corresponds to a non linear autoregressive model (NAR). To capture the moving average components of a time series, it is necessary to use recurrent architectures. ANNs are difficult to design and calibrate, and the computation of forecasts is also complex. Fuzzy linear regression makes it possible to extract linear relationships from small data sets, with fewer hypotheses than statistical linear regression. It does not require the observations to be uncorrelated, nor does it ask for the error variance to be homogeneous. However, the model is very sensitive to outliers. Furthermore, a posteriori validation of the hypothesis of linearity is not possible for small data sets. MARS models are based on the hypothesis that time series are chaotic instead of stochastic. The main advantage of the method is its ability to model non-stationary processes. The approach is non-parametric, and therefore requires a large data set.

Amongst conceptual models, we distinguish between physical models, hydraulic machines, and fuzzy rule-based systems. Most conceptual hydrologic models are hydraulic machines, in which the watershed is considered to behave like a network of reservoirs. Physical modeling of a watershed would imply using fundamental physical equations at a small scale, such as the law of conservation of mass. Given the complexity of a watershed, this can be done in practice only for water routing. Consequently, only short term flow forecasts can be obtained from a physical model, since the effects of precipitation, infiltration and evaporation must be negligible. Fuzzy rule-based systems make it possible to model the water cycle using fuzzy IF-THEN rules, such as IF it rains a lot in a short period of time, THEN there will be a large flow increase following the concentration time. Each fuzzy quantifier, here in italics, is modeled using a fuzzy number to take into account the uncertainty surrounding it. When sufficient data are available, the fuzzy quantifiers can be constructed from the data. In general, conceptual models require more effort to develop than empirical models. However, for exceptional events, conceptual models can often provide more realistic forecasts, since empirical models are not well suited for extrapolation.

A fruitful approach is to combine conceptual and empirical models. One way of doing this, called extended streamflow prediction or ESP, is to combine a stochastic model for generating meteorological scenarios with a conceptual model of the watershed.

Based on this review of flow forecasting models, we recommend for short term forecasting (hourly and daily flows) the use of the  $k$ -nearest neighbor method, Box-Jenkins models, water routing models or hydraulic machines. For medium term forecasting (weekly flows, for example), we recommend the  $k$ -nearest neighbor method and Box-Jenkins models, as well as fuzzy-rule based and ESP models. For long term forecasting (monthly flows), we recommend statistical and fuzzy regression, Box-Jenkins, MARS and ESP models. It is important to choose a type of model which is appropriate for the problem at hand and for which the information available is sufficient. Each type of model

having its advantages, it can be more efficient to combine different approaches when forecasting streamflow.

**Key words :** forecasting, flow, stochastic model, conceptual model, artificial neural network, fuzzy set theory.

## RÉSUMÉ

Dans le domaine de la prévision des débits, une grande variété de méthodes sont disponibles : des modèles stochastiques et conceptuels mais aussi des approches plus novatrices telles que les réseaux de neurones artificiels, les modèles à base de règles floues, la méthode des  $k$  plus proches voisins, la régression floue et les splines de régression. Après avoir effectué une revue détaillée de ces méthodes et de leurs applications récentes, nous proposons une classification qui permet de mettre en lumière les différences mais aussi les ressemblances entre ces approches. Elles sont ensuite comparées pour les problèmes différents de la prévision à court, moyen et long terme. Les recommandations que nous effectuons varient aussi avec le niveau d'information *a priori*. Par exemple, lorsque l'on dispose de séries chronologiques stationnaires de longue durée, nous recommandons l'emploi de la méthode non paramétrique des  $k$  plus proches voisins pour les prévisions à court et moyen terme. Au contraire, pour la prévision à plus long terme à partir d'un nombre restreint d'observations, nous suggérons l'emploi d'un modèle conceptuel couplé à un modèle météorologique basé sur l'historique. Bien que l'emphase soit mise sur le problème de la prévision des débits, une grande partie de cette revue, principalement celle traitant des modèles empiriques, est aussi pertinente pour la prévision d'autres variables.

**Mots-clés :** prévision, débit, modèle stochastique, modèle conceptuel, réseau de neurones artificiels, théorie des sous-ensembles flous.

## 1 – INTRODUCTION

Il est important, au début de cette revue bibliographique, de faire la distinction entre la prévision et la prédiction. La prévision concerne l'estimation des conditions futures à une date donnée, ou durant un intervalle de temps spécifique, alors que la prédiction représente l'estimation des conditions futures, sans référence à un temps spécifique (LETTENMAIER et WOOD, 1993). Ainsi, on fait une prévision du débit de la rivière Rouge en amont de la chute McNeil pour le lendemain à 16 h 00, mais on fait une prédiction du débit centenaire (de période de retour 100 ans) au même site. La prévision des débits horaires, journaliers, hebdomadaires et mensuels d'un cours d'eau est une tâche nécessaire pour une gestion efficace des ressources en eaux de surface, et pour permettre une réaction adéquate et rapide dans les situations de crise, comme les inondations. Il existe différentes approches pour effectuer ces prévisions. Nous les présenterons brièvement, en soulignant leurs avantages et inconvénients pour les problèmes différents de la prévision à court terme (débits horaires et journaliers), moyen terme (débits hebdomadaires et mensuels) et long terme (tendance pour la prochaine année). Nous tenterons aussi d'en proposer une classification.

Nous distinguerons d'abord les modèles conceptuels et empiriques. Dans la première catégorie se retrouveront les modèles basés sur des connaissances physiques du bassin versant, alors que la seconde catégorie contiendra les

modèles qui visent à inférer à partir des données la relation structurelle entre les observations des variables d'entrée et de sortie (le débit) du bassin versant. Parmi les modèles empiriques, nous distinguerons les modèles stochastiques paramétriques et non paramétriques, les modèles à base de réseaux de neurones artificiels, les méthodes de régression floue, ainsi que les modèles de systèmes chaotiques. Des modèles conceptuels nous retiendrons les modèles physiques, les modèles à base de règles floues et les machines hydrauliques (c'est ainsi que nous nommerons les modèles conceptuels qui proposent de représenter un bassin versant par un réseau de réservoirs et de conduites). Nous procéderons ensuite à la comparaison des approches et formulerons des recommandations pour l'utilisation pratique des modèles de prévision du débit. Mais au préalable nous précisons le problème de la prévision des débits d'un cours d'eau.

## 1 – LA PRÉVISION DES DÉBITS : FORMULATION DU PROBLÈME

Par prévision des débits, on entend le fait d'estimer le débit  $Q_\tau$  à un temps futur  $\tau$  en n'utilisant que l'information disponible au temps  $t < \tau$ , correspondant normalement au temps présent. Cette information comprend les valeurs mesurées du débit lui-même (si elles sont disponibles), depuis un temps reculé  $t - r$  jusqu'au temps  $t$ , ainsi que d'autres séries temporelles  $X_1, X_2, \dots, X_K$ . Il s'agit généralement d'information météorologique, par exemple les précipitations et la température. On note  $\mathbf{x}_t = (x_{1,t}, x_{2,t}, \dots, x_{K,t})'$  la transposée du vecteur des valeurs prises par chacune de ces variables au temps  $t$ .

### 1.1 Utilisation de prévisions météorologiques

Une prévision du débit doit généralement tenir compte non seulement de l'information hydrologique et météorologique passée, mais aussi des prévisions météorologiques. Formellement, il s'agit d'information incertaine sur la valeur des variables  $X_1, X_2, \dots, X_K$  à des temps postérieurs à l'instant présent. Les prévisions météorologiques doivent être traitées comme des variables explicatives distinctes, puisque leurs caractéristiques statistiques sont fort différentes de celles des observations météorologiques. Nous représenterons par  $\hat{\mathbf{x}}_t$  les prévisions météorologiques au temps  $t$ .

### 1.2 Calibration d'un modèle de prévision

Pour obtenir une prévision du débit à partir des observations passées et des prévisions météorologiques, on suppose l'existence d'un modèle  $f$  reliant ces observations et ces prévisions au débit  $Q$  au temps  $\tau$  :

$$Q_\tau = f(Q_t, Q_{t-1}, \dots, Q_{t-r}, \hat{\mathbf{x}}_{\tau-1}, \hat{\mathbf{x}}_{\tau-2}, \dots, \mathbf{x}_t, \mathbf{x}_{t-1}, \dots, \mathbf{x}_{t-r}, \varepsilon_\tau | \theta) \quad (1)$$

où  $\varepsilon_\tau$  est un terme d'erreur et  $\theta = (\theta_1, \theta_2, \dots, \theta_m)'$  est un ensemble de  $m$  coefficients d'ajustement ne variant pas dans le temps. Il est alors possible d'obtenir une prévision  $\hat{Q}_\tau$  de  $Q_\tau$  en estimant les coefficients d'ajustements et le terme d'erreur :

$$\hat{Q}_\tau = f(Q_\tau, Q_{\tau-1}, \dots, Q_{\tau-r}, \hat{x}_{\tau-1}, \hat{x}_{\tau-2}, \dots, x_1, x_{\tau-1}, \dots, x_{\tau-r}, \hat{\varepsilon}_\tau | \hat{\theta}) \quad (2)$$

où  $\hat{\theta} = (\hat{\theta}_1, \hat{\theta}_2, \dots, \hat{\theta}_m)'$  et  $\hat{\varepsilon}_\tau$  sont respectivement des estimateurs de  $\theta$  et  $\varepsilon_\tau$ . Dans la plupart des cas, l'erreur est supposée nulle (*i.e.*  $\hat{\varepsilon}_\tau = 0$ ), mais son estimation est parfois intrinsèquement reliée à la prévision du débit.

Un modèle de prévision doit être ajusté à l'aide d'observations passées. De cette période de calibration, le modèle doit extraire l'information reliée au comportement du système modélisé, en éliminant le bruit. Si la souplesse du modèle de prévision est trop grande, le bruit risque d'être modélisé comme s'il s'agissait d'un comportement prévisible : sur la période de calibration le modèle pourrait alors donner un très bon ajustement, puisqu'il reproduirait même les variations aléatoires du débit. Par contre les prévisions effectuées par un tel modèle seraient alors, en comparaison, particulièrement mauvaises. On doit donc en vertu du principe de parcimonie chercher à minimiser le nombre de paramètres d'un modèle de prévision, tout en permettant qu'il soit assez souple pour modéliser convenablement le système.

Une façon de vérifier si la modélisation permet d'extraire convenablement l'information disponible consiste à observer l'auto-corrélation des écarts entre les débits observés et modélisés. Une forte auto-corrélation indique que ce que l'on a considéré comme un bruit aléatoire comporte une information pouvant être incorporée dans le modèle.

### 1.3 Validation d'un modèle de prévision

Une étape essentielle de la construction d'un modèle de prévision consiste à évaluer la précision des prévisions. Il faut réserver une portion des données disponibles pour cette phase de validation. Ces données ne seront pas utilisées lors de la phase de calibration. Lors de la validation, on peut calculer différentes statistiques des prévisions, par exemple leur variance et leur biais. La période de validation doit être assez grande pour que ces statistiques soient suffisamment précises pour être utiles. Ainsi, il importe de déterminer les distributions d'échantillonnage des statistiques utilisées comme critères de qualité d'un modèle de prévision.

Pour construire un modèle de prévision robuste respectant le principe de parcimonie, il est possible d'optimiser le nombre de paramètres du modèle en observant l'évolution de ces critères de qualité, toujours calculés sur la période de validation, en fonction de la complexité du modèle. On peut alors déterminer à partir de quel moment l'augmentation du nombre de paramètres entraîne une détérioration des critères de validation au lieu d'une amélioration des prévisions.

### 1.4 Prévision multivariée

Il faut noter que nous étudierons dans le cadre de ce travail la prévision du débit à un seul site. Le problème plus complexe de la prévision simultanée à plusieurs sites, ou prévision multivariée, ne sera pas considéré ici. Évidemment on peut obtenir de telles prévisions en appliquant indépendamment plusieurs modèles aux différents sites ; cependant une telle approche n'assure pas une cohérence entre les prévisions à plusieurs sites (*i.e.* ne reproduit pas les corrélations spatiales).

## 2 – MODÈLES EMPIRIQUES

La caractéristique principale d'un modèle empirique est qu'on ne demande pas à sa structure de ressembler à celle du système modélisé, dans la mesure où le modèle permet de bien reproduire le comportement du système. Un modèle empirique peut être une boîte noire : dans la mesure où il produit des prévisions précises, il importe peu de comprendre sa structure interne. Cependant, dans les cas les plus simples, il peut être possible d'interpréter de façon physique les paramètres et les fonctions de transfert d'un modèle empirique. Nous distinguerons dans cette catégorie quatre grands types de modèles : les modèles stochastiques, les modèles à base de réseaux de neurones artificiels, les modèles à base de régression floue et les modèles de systèmes chaotiques.

### 2.1 Modèles stochastiques

La plupart des modèles empiriques utilisés de nos jours sont des modèles stochastiques. Notons que nous utilisons ici ce terme au sens large, tel que défini par FELLER (1968) : il s'agit de tous les modèles utilisant la théorie des probabilités pour représenter l'incertitude. Nous détaillerons deux approches paramétriques : la régression linéaire et le modèle de BOX et JENKINS (1976), ainsi qu'une approche nonparamétrique : la méthode des  $k$  voisins les plus proches.

#### 2.1.1 Régression linéaire multiple

La régression linéaire multiple (WETHERILL *et al.*, 1986) est un outil de modélisation très puissant et bien connu, mais parfois mal utilisé. En effet, le modèle de régression linéaire multiple repose sur un ensemble d'hypothèses qui ne sont pas toujours vérifiées en pratique. Le modèle peut s'écrire ainsi pour la prévision des débits :

$$Q_t = \sum_{t^*=t-r}^t (\varphi_{t^*} Q_{t^*} + \lambda'_{t^*} x_{t^*}) + \varepsilon_t \quad (3)$$

où  $\varphi_{t^*}$  est le coefficient d'ajustement associé à la variable  $Q_{t^*}$ ,  $\lambda_{t^*} = (\lambda_{1,t^*}, \lambda_{2,t^*}, \dots, \lambda_{K,t^*})'$  le vecteur de coefficients d'ajustement associé au vecteur de variables explicatives  $x_{t^*} = (x_{1,t^*}, x_{2,t^*}, \dots, x_{K,t^*})'$ , et  $\varepsilon_t$  un bruit normal  $N(0, \sigma)$ . Pour estimer les coefficients d'ajustement d'une régression linéaire statistique, on doit disposer d'un échantillon d'observations pour toutes les variables explicatives mesurées à chaque instant  $t^*$  de l'horizon  $[t-r, t]$ . En pratique, pour les séries temporelles comme le débit d'une rivière, cela revient à supposer que le système modélisé est périodique :  $Q_{t+T}$  doit avoir la même distribution statistique que  $Q_t$ , où  $T$  est la période du système, normalement une année). Si cette hypothèse peut tenir pour les débits annuels et mensuels, à la limite hebdomadaires, il en est autrement des débits journaliers ou horaires. De plus notons que les coefficients d'ajustement du modèle, dont le nombre peut atteindre  $(r+1) \cdot (K+1)$ , ne sont utiles que pour estimer  $Q_t$  pour un  $t$  fixé : ils doivent être estimés à nouveau pour chaque instant  $t = 1, 2, \dots, T$ . Le nombre total de paramètres peut donc atteindre  $T \cdot (r+1) \cdot (K+1)$ .

Un avantage important de la régression linéaire est qu'elle permet de quantifier aisément l'erreur de prévision à l'aide d'un intervalle de confiance, lorsque les hypothèses sous-jacentes au modèle sont vérifiées. Une de ces hypothèses est

l'indépendance stochastique des variables explicatives. À l'échelle journalière et à l'échelle horaire,  $X_{k,t}$  ( $k = 1, 2, \dots, K$ ) est généralement corrélé positivement avec  $X_{k,t-1}$ , et  $Q_t$  est fortement corrélé avec  $Q_{t-1}$ . Conséquemment, il risque fort d'être nécessaire, pour pouvoir utiliser la régression statistique, de se limiter à l'utilisation d'un seul temps pour chaque variable explicative  $X_1, X_2, \dots, X_K$  et pour le débit. Par exemple, on pourrait considérer le modèle de régression simplifié suivant :

$$Q_t = \varphi_t \cdot Q_t + \sum_{i=1}^K \lambda_{i,t-b(i)} \cdot X_{i,t-b(i)} + \varepsilon_t \quad (4)$$

où  $b(i)$  est le délai nécessaire pour que la variable explicative  $X_i$  affecte le débit. S'il subsiste alors une corrélation entre les variables, une estimation biaisée des paramètres, par exemple en utilisant la ridge-régression (HOERL et KENNARD, 1970) peut être nécessaire. Les autres hypothèses de la régression linéaire, qui sont l'homogénéité de la variance  $\sigma$  et la normalité des erreurs  $\varepsilon_t$ , peuvent être vérifiées *a posteriori*. S'il s'avère qu'elles ne sont pas acceptables, il est possible d'effectuer des transformations sur les variables explicatives ou sur le débit pour obtenir un modèle valide de régression (WETHERILL *et al.*, 1986).

La régression est un outil relativement simple d'utilisation, et approprié pour la prévision à moyen terme (débits mensuels). Cependant, les hypothèses du modèle de régression sont trop restrictives pour la prévision à court terme (débits journaliers et horaires). Pour les débits hebdomadaires, un modèle de régression risque d'être très limité dans sa capacité prévisionnelle en raison de son grand nombre de paramètres. Dans la grande majorité des cas, une autre méthode sera plus appropriée, peut-être la méthodologie de Box-Jenkins, qui est une approche similaire à la régression, mais spécialement adaptée pour les séries temporelles.

### 2.1.2 Méthodologie de Box-Jenkins

La famille des modèles de Box-Jenkins, dans sa formulation la plus générale, permet de représenter un vaste ensemble de séries temporelles. Il s'agit d'un modèle linéaire, tout comme le modèle de régression présenté précédemment ; il ne doit donc pas y avoir d'interaction entre les entrées du modèle. Une hypothèse souvent plus contraignante que la linéarité est la stationnarité : les modèles de Box-Jenkins ne permettent que de modéliser des séries dont les propriétés statistiques n'évoluent pas dans le temps. Cependant, il est en général possible d'effectuer une transformation sur les séries de données pour les rendre stationnaires. Par exemple, la transformation  $z_t = Q_t - Q_{t-1}$ , qui correspond à la première différence du débit est souvent satisfaisante. Lorsque la première différence n'est pas stationnaire, on peut considérer la deuxième différence, c'est-à-dire utiliser la transformation  $z_t = (Q_t - Q_{t-1}) - (Q_{t-1} - Q_{t-2}) = Q_t - 2Q_{t-1} + Q_{t-2}$ .

Pour les séries périodiques comme le débit, il est aussi possible de désaisonnaliser les observations pour obtenir une série stationnaire. Une transformation possible consiste à standardiser les observations en soustrayant de chaque observation  $Q_t$  la moyenne  $\bar{Q}_t$  observée pour cette période et en divisant par l'écart-type  $s_t$  observé pour cette période :  $z_t = (Q_t - \bar{Q}_t) / s_t$ . Ceci ajoute cependant  $2T$  paramètres au modèle, ce qui est prohibitif pour des données journalières. De plus, on doit supposer que les moments d'ordre supérieur à deux (en particulier les coefficients d'asymétrie et d'aplatissement) sont constants tout au cours de l'année, ce qui est difficile à vérifier. D'autres transformations peuvent être envisagées, mais toutes ajoutent des paramètres au modèle, ce qui peut avoir des conséquences désastreuses sur la qualité des prévisions.



Un modèle de Box-Jenkins comprend deux parties : une composante autorégressive (AR) et une composante à moyenne mobile (MA). Nous présenterons d'abord les cas particuliers des modèles AR et MA, pour ensuite exposer le modèle général (ARIMA). Nous discuterons ensuite des modèles périodiques (PARMA), des fonctions de transfert de type ARMAX utilisées pour ajouter des variables explicatives au modèle, et nous terminerons par une discussion des mérites de ce type de modèle pour la prévision des débits.

### Modèle autorégressif (AR)

Un modèle autorégressif linéaire (AR) d'une transformation  $z_t$  du débit  $Q_t$  relie la valeur de  $z_t$  au temps  $t$  à ses valeurs précédentes de la façon suivante :

$$z_t = \gamma + \varphi_1 z_{t-1} + \varphi_2 z_{t-2} + \dots + \varphi_p z_{t-p} + \varepsilon_t = \gamma + \sum_{i=1}^p \varphi_i z_{t-i} + \varepsilon_t \quad (5)$$

où  $\gamma$ ,  $\varphi_i$  ( $i = 1, 2, \dots, p$ ) et  $p$  sont des coefficients d'ajustement, et  $\varepsilon_t$  est un bruit normal  $N(0, \sigma)$ . Un modèle comportant  $p$  termes (en plus de  $\gamma$  et  $\varepsilon_t$ ), est dénoté AR( $p$ ). La valeur de  $p$ , que l'on nomme l'ordre du modèle, doit être fixée *a priori*, alors que les autres paramètres sont estimés par calibration. En général on doit essayer plusieurs valeurs de  $p$  et observer la qualité des ajustements ainsi obtenus avant de faire un choix final. Il existe plusieurs méthodes permettant de choisir une valeur de  $p$  donnant un ajustement adéquat et respectant le principe de parcimonie. On peut utiliser des tests statistiques (BOX et JENKINS, 1976), ou encore un critère basé sur la quantité d'information extraite (AKAIKE, 1972).

### Modèle à moyenne mobile (MA)

Un modèle à moyenne mobile (MA) d'une transformation  $z_t$  du débit  $Q_t$  relie la valeur de  $z_t$  au temps  $t$  aux erreurs  $\varepsilon_t$  précédentes de la façon suivante :

$$z_t = \gamma + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \dots + \theta_q \varepsilon_{t-q} + \varepsilon_t = \gamma + \sum_{i=1}^q \theta_i \varepsilon_{t-i} + \varepsilon_t \quad (6)$$

où  $\gamma$ ,  $\theta_i$  ( $i = 1, 2, \dots, q$ ) et  $q$  sont des coefficients d'ajustement, et  $\varepsilon_t$  est un bruit normal  $N(0, \sigma)$ . Un modèle comportant  $q$  termes (en plus  $\gamma$  et de  $\varepsilon_t$ ), est dénoté MA( $q$ ). La valeur de  $q$ , que l'on nomme l'ordre du modèle, doit être fixé *a priori*, alors que les autres paramètres sont estimés par calibration. Comme dans le cas d'un modèle AR, différentes valeurs de  $q$  peuvent être essayées avant d'effectuer un choix final.

### Modèles ARMA, ARIMA et PARMA

Un modèle ARIMA ( $p, q$ ) est une généralisation des modèles AR ( $p$ ) et MA ( $q$ ). Il peut être représenté par l'équation suivante (SALAS *et al.*, 1988) :

$$\begin{aligned} z_t &= \gamma + \varphi_1 z_{t-1} + \varphi_2 z_{t-2} + \dots + \varphi_p z_{t-p} + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \dots + \theta_q \varepsilon_{t-q} + \varepsilon_t \\ &= \gamma + \sum_{i=1}^p \varphi_i z_{t-i} + \sum_{i=1}^q \theta_i \varepsilon_{t-i} + \varepsilon_t \end{aligned} \quad (7)$$

Un cas particulier du modèle ARIMA fréquemment rencontré est le modèle ARMA, qui correspond au cas où il n'est pas nécessaire de rendre la série stationnaire, *i.e.*  $z_t = Q_t$ . Lorsque la transformation choisie est  $z_t = (Q_t - \bar{Q}_t) / s_t$ , on parle alors d'un modèle ARMA périodique, ou PARMA.

Pour condenser la représentation mathématique, on utilise deux opérateurs. D'abord, l'opérateur  $B$  de décalage arrière (backshift) :

$$B^n(z_t) = z_{t-n} \tag{8}$$

On utilise également l'opérateur de différences  $\nabla^n = (1 - B)^n$ . Par exemple, pour  $n = 2$  on obtient  $\nabla^2(z_t) = (1 - 2B + B^2)(z_t) = z_t - 2z_{t-1} + z_{t-2}$ , qui correspond à la seconde différence de  $z_t$ . La  $n$ ème différence de  $Q_t$  s'écrit alors  $z_t = \nabla^n Q_t$ , et le modèle ARIMA peut être résumé ainsi :

$$\varphi_p(B) \nabla^n Q_t = \gamma + \theta_q(B) \varepsilon_t \tag{9}$$

où  $\varphi_p(B) = 1 - \varphi_1 - \varphi_2 - \dots - \varphi_p$  et  $\theta_q(B) = 1 - \theta_1 - \theta_2 - \dots - \theta_q$ .

Malgré la souplesse des modèles ARIMA ( $p, q$ ), ceux-ci n'utilisent comme entrées que les valeurs précédentes de la variable  $z_t$  et des erreurs  $\varepsilon_t$ . Il est souvent intéressant d'introduire d'autres séries temporelles explicatives, comme les précipitations et la température. Dans la méthodologie de Box-Jenkins, ceci est réalisé à l'aide des fonctions de transfert.

**Fonctions de transfert de type ARMAX**

Lorsque l'on désire introduire une variable explicative dans un modèle de Box-Jenkins, on peut utiliser des fonctions de transfert. Il en existe plusieurs types (LJUNG, 1987) ; nous ne présenterons ici que le modèle ARMAX, (pour ARMA avec entrées auxiliaires). On pourrait par exemple introduire une mesure  $x_t$  des précipitations au temps  $t$  dans le modèle ARMA de la transformation  $z_t$  du débit  $Q_t$ . La première étape lors de la construction de ce type de modèle consiste à rendre  $x_t$  stationnaire par une transformation  $z_t^{(x)}$ , par exemple en prenant  $z_t^{(x)} = x_t - x_{t-1}$ . Ensuite, on ajoute au modèle ARMA une composante additive qui est fonction de  $z_t^{(x)}$  :

$$z_t = \gamma + \sum_{i=1}^p \varphi_i z_{t-i} + \sum_{i=1}^q \theta_i \eta_{t-i} + \eta_t + C \left[ z_{t-b}^{(x)} + \sum_{i=1}^s w_i z_{t-b-i}^{(x)} \right] \tag{10}$$

où  $\gamma, \varphi_i, \theta_i, w_i$  et  $C$  sont des coefficients à estimer, et  $\eta_t$  est un bruit. L'ordre d'un tel modèle est donné par  $p, q, s$  et  $b$ . Le paramètre  $s$  représente le nombre de termes précédents de  $z_t^{(x)}$  que l'on doit conserver pour calculer  $z_t$ . Le paramètre  $b$  indique le délai (en unités de temps) que prend une valeur  $z_t^{(x)}$  pour affecter  $z_t$ . De façon générale, le bruit  $\eta_t$  n'a pas une distribution normale  $N(0, \sigma)$ . Il est parfois nécessaire de le modéliser lui même à l'aide d'un modèle ARMA :

$$\eta_t = \gamma + \sum_{i=1}^{p_\eta} \varphi_{\eta,i} \eta_{t-i} + \sum_{i=1}^{q_\eta} \theta_{\eta,i} \varepsilon_{t-i} + \varepsilon_t \tag{11}$$

Par le même procédé, d'autres variables explicatives peuvent être intégrées au modèle. Cependant, chaque ajout accroît le nombre de paramètres et la complexité du modèle. Cependant, HIPEL *et al.* (1977) et LJUNG (1987) ont proposé des méthodes pratiques d'estimation et d'identification du modèle.

**Mérites de la méthodologie de Box-Jenkins**

La méthodologie de Box-Jenkins est couramment utilisée en prévision, et permet effectivement de représenter un grand nombre de séries temporelles. Pour le problème de la prévision à court terme, où la valeur de  $Q_t$  dépend principalement des valeurs précédentes de cette même variable, il s'agit d'une approche à privi-

légier, car elle permet l'obtention d'intervalles de confiance pour les prévisions. Cependant, pour la prévision à moyen terme il devient plus difficile d'inclure toutes les autres variables explicatives dans le modèle. De plus, les composantes non linéaires du système réel, qui ne peuvent être modélisées, prennent alors de l'importance. Il peut alors être préférable de considérer l'utilisation d'un réseau de neurones artificiels, qui constitue en fait une généralisation du modèle ARMA (CONNOR *et al.*, 1994), ou encore la méthode non paramétrique des  $k$  voisins les plus proches.

### 2.1.3 Méthode des $k$ plus proches voisins

La méthode des  $k$  plus proches voisins est une approche non paramétrique intuitive, mais ayant tout de même des bases théoriques solides dans le domaine de la prévision des séries temporelles (YAKOWITZ, 1987). Comme les modèles ARMA et NARMA (voir plus bas), elle s'applique cependant aux séries stationnaires. Pour pouvoir estimer la sortie  $z$  reliée à un ensemble d'entrées  $\mathbf{y}_\tau = (z_\tau, z_{\tau-1}, z_{\tau-2}, \dots, z_{\tau-r}, \mathbf{x}_\tau, \mathbf{x}_{\tau-1}, \dots, \mathbf{x}_{\tau-r})'$ , on détermine simplement dans tout l'historique les  $k$  voisins les plus proches de la situation actuelle, selon une distance euclidienne pondérée. L'ensemble des observations de la sortie du système correspondant à ces  $k$  ensembles de paramètres d'entrées voisins de la situation actuelle forment alors une distribution prévisionnelle de  $z$ . Celle-ci peut être utilisée pour obtenir une estimation ponctuelle (en prenant la moyenne) ou un intervalle de confiance.

Par exemple, si l'on cherche à prévoir le débit hebdomadaire, on pourrait chercher dans l'historique des situations présentant une tendance semblable du débit, des prévisions météorologiques semblables, un stock de neige semblable (pour l'hiver) ou des conditions d'humidité du sol semblables (pour l'été). On pourrait alors prédire le débit à partir de ce qui a été observé dans des conditions similaires.

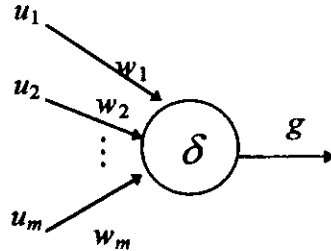
Dans cette méthode, la valeur de  $k$  ainsi que le poids associé à chaque entrée du vecteur  $\mathbf{y}_\tau$  sont des coefficients qui doivent être ajustés. Il s'agit donc d'une méthode essentiellement basée sur les données, puisque pour chaque prévision on doit chercher dans tout l'historique. Cette méthode a été appliquée avec succès par KARLSSON et YAKOWITZ (1987) pour la prévision des débits journaliers à partir des précipitations. GALEATI (1990) montre que la méthode des  $k$  plus proches voisins donne des prévisions de qualité comparable à celles d'un modèle ARMAX pour la prévision du débit, tout en étant plus simple à implanter. SHAMSELDIN et O'CONNOR (1996) discutent en détail de l'utilisation pratique de la méthode des  $k$  plus proches voisins pour la prévision des débits de façon opérationnelle. Notons finalement que le temps nécessaire pour la calibration du modèle et le calcul des prévisions peut s'avérer considérable si l'historique est important. Or, la méthode ne donnera des résultats satisfaisants précisément que dans ce cas. Selon

## 2.2 Modèles à base de réseaux de neurones artificiels

Les réseaux de neurones artificiels sont issus de travaux réalisés dans le domaine de l'intelligence artificielle pour modéliser le fonctionnement du cerveau humain (MC CULLOCH et PITTS, 1943). De ce fait, ce type de modèle a acquis une notoriété particulière. Pourtant, il ne s'agit ni plus ni moins que d'un modèle empirique non linéaire (FORTIN *et al.*, 1997). Un réseau de neurones artificiels est composé de neurones artificiels semblables, chacun étant en fait une fonction non linéaire  $g(u_1, u_2, \dots, u_m)$  de  $m$  entrées  $u_1, u_2, \dots, u_m$  de la forme suivante :

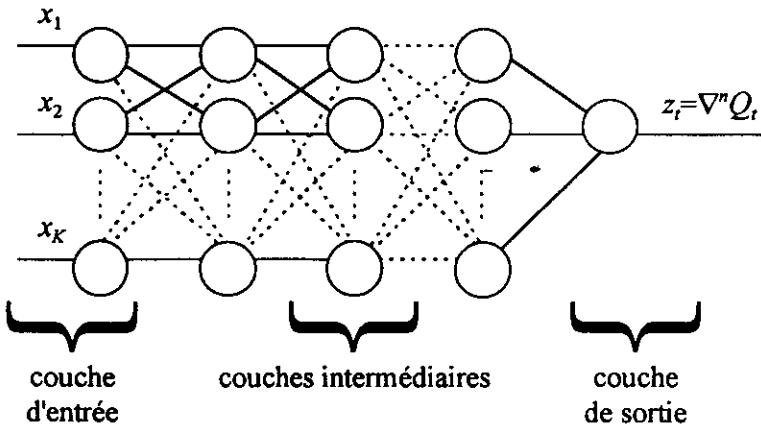
$$g(u_1, u_2, \dots, u_m) = \delta\left(\sum_{j=1}^m w_j u_j\right) \tag{12}$$

où  $w_j$  est le poids accordé à l'entrée  $u_j$  et  $\delta$  est une fonction quelconque, nommée fonction d'activation. En pratique, on retrouve trois fonctions d'activation différentes : la fonction identité  $\delta(z) = z$ , la fonction binaire  $\delta(z) = 0$  si  $z < h$ ,  $\delta(z) = 1$  si  $z > h$  (où  $h$  est le seuil d'activation), et la fonction sigmoïde  $\delta(z) = \{1 + \exp(-z + h)\}^{-1}$ . La figure 1 montre une représentation graphique d'un neurone artificiel.



**Figure 1** Représentation graphique d'un neurone artificiel.  
Graphical representation of an artificial neuron.

Il existe une grande variété d'agencements possibles de neurones artificiels (LIPPMANN, 1987), mais le type de réseau le plus utilisé en prévision est le perceptron multicouche (multilayer feed-forward network). Ce réseau comporte une couche de neurones artificiels pour capter les entrées, une ou plusieurs couches intermédiaires, ainsi qu'une couche de neurones artificiels pour émettre les sorties du modèle. Dans le cas de la prévision des débits, la couche de sortie ne comportera qu'un seul neurone. La figure 2 illustre la configuration d'un tel réseau.



**Figure 2** Configuration d'un réseau de neurones perceptron multicouche.  
Configuration of a multi-layer feed-forward artificial neural network.

Un réseau de neurones artificiels permettra ainsi de modéliser une transformation  $z_t$  du débit, usuellement la  $n$ ième différence, représentée ici sous forme compacte par  $z_t = \nabla^n Q_t$ . Le nombre de couches intermédiaires ainsi que le nombre de neurones dans chacune de ces couches doit être fixé *a priori*. Le choix de la configuration d'un réseau de neurone a une influence sur la performance, en particulier en prévision (BEBIS et GEORGIPOULOS, 1994). En général, on n'utilise qu'une seule couche intermédiaire et un nombre de neurones dans cette couche du même ordre de grandeur que le nombre de neurones dans la couche d'entrée. Ceci vient du fait qu'il a été prouvé que les réseaux de neurones artificiels ainsi configurés forment une classe d'approximateurs universels, à condition que les fonctions d'activation des neurones de la couche intermédiaire ne soient pas des polynômes, soient bornées et continues par morceaux (LESHNO *et al.*, 1993). Ceci signifie qu'un tel réseau peut approximer n'importe quelle fonction à n'importe quel degré de précision (HORNIK *et al.*, 1989). Cette propriété est partagée par les polynômes, et de plus elle ne garantit pas qu'un réseau de neurones artificiels saura modéliser une fonction quelconque avec un nombre raisonnable de neurones intermédiaires, ni qu'il sera possible de déterminer ce nombre. Même s'il est suffisant de considérer les réseaux à une seule couche intermédiaire, d'autres configurations pourraient permettre de réduire le nombre total de paramètres du modèle (BEBIS et GEORGIPOULOS, 1994).

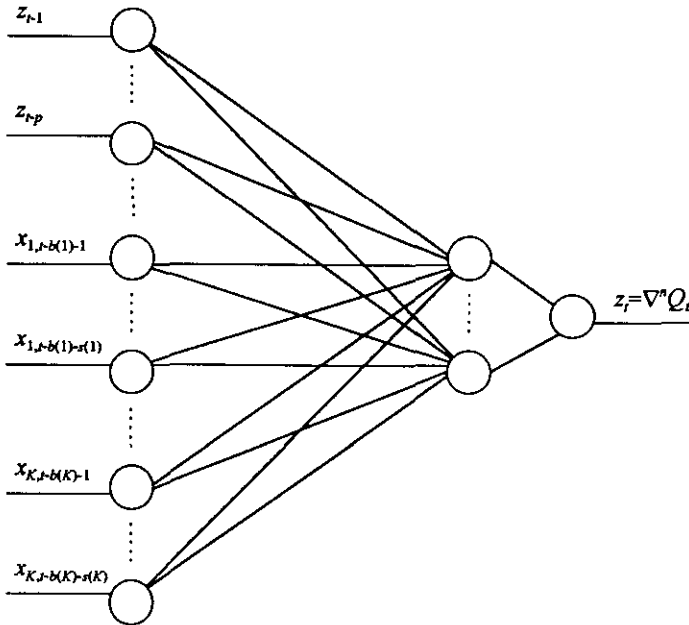
Les réseaux de neurones artificiels ont déjà été appliqués à la prévision de séries temporelles. En particulier, ils ont eu un succès retentissant pour prédire l'évolution des prix des actions en bourse (REFENES *et al.*, 1994). Dans le domaine des sciences de l'eau, on retrouve des applications à la prévision de la demande en eau (CUBERO, 1991), des précipitations (FRENCH *et al.*, 1992), de la qualité de l'eau (MAIER et DANDY, 1996) et du débit (KARUNANITHI *et al.*, 1994 ; HSU *et al.*, 1995 ; SMITH et ELI, 1995 ; et ZHANG et TRIMBLE, 1996) ainsi que pour la gestion d'un réservoir en temps réel (SAKAKIMA *et al.*, 1992). Pour toutes ces applications, un réseau de neurones perceptron à une couche intermédiaire a été utilisé. LAPEDES et FARBER (1987) ont montré que ce type de réseau constitue un modèle autorégressif non linéaire (NAR). Des architectures plus complexes, que l'on nomme réseaux récurrents (WILLIAMS, 1990), constituent une classe de modèles autorégressifs à moyenne mobile non linéaires (NARMA). Nous présenterons ces deux types de réseaux plus en détails pour ensuite discuter de leur calibration.

### 2.2.1 Réseau de neurones artificiels autorégressif non linéaire (NAR)

LAPEDES et FARBER (1987) ont montré que le réseau perceptron à une couche intermédiaire constitue un modèle autorégressif non linéaire (NAR). CHAKRABORTY *et al.* (1992) ont montré que pour des séries réelles cette approche pouvait s'avérer supérieure à l'utilisation de modèles linéaires AR et ARMA dans le cas multivarié. La figure 3 montre l'architecture d'un réseau NAR multivarié :  $b(i)$  représente le délai nécessaire avant que la variable  $X_i$  influence le débit, et  $s(i)$  représente le nombre de termes de cette même variable ayant une influence sur le débit. Ce type de modèle peut être généralisé en y ajoutant une composante à moyenne mobile. On obtient alors un réseau récurrent.

### 2.2.2 Réseau de neurones artificiels récurrents (NARMA)

Il existe une grande variété de réseaux de neurones récurrents (WILLIAMS, 1990). Nous ne présenterons dans ce rapport qu'une seule configuration, le



**Figure 3** Architecture d'un réseau de neurones artificiels autorégressif multivarié.  
*Configuration of a multivariate autoregressive artificial neural network.*

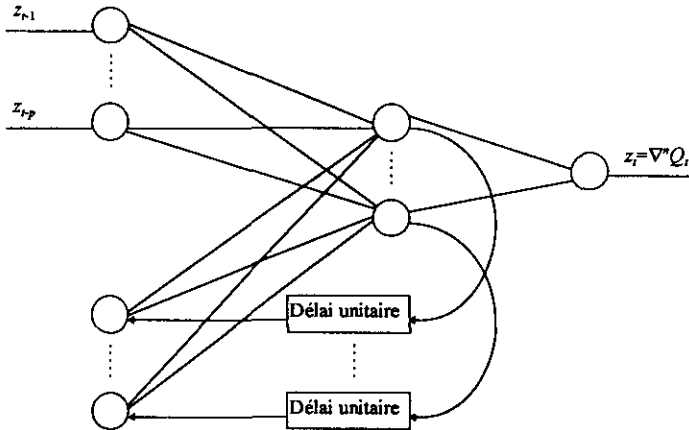
réseau complètement récurrent (CONNOR *et al.*, 1994). Il s'agit simplement de réintroduire les sorties des neurones intermédiaires dans des neurones d'entrée, après leur avoir fait subir un délai d'une unité de temps, comme le montre la figure 4. Ce type de réseau récurrent forme une classe de modèles non linéaires autorégressifs avec moyenne mobile (NARMA). Ce modèle n'a à notre connaissance pas encore été appliqué au cas multivarié, mais sa généralisation (pour obtenir un modèle NARMAX) est possible.

### 2.2.3 Calibration d'un réseau de neurones artificiels

La calibration d'un réseau de neurones artificiels est une tâche ardue. Un grand nombre d'algorithmes d'optimisation des poids et des seuils d'activation des neurones existent. Le plus utilisé demeure la méthode de la descente la plus rapide (*steepest descent* en anglais, CAUCHY 1847), nommé propagation arrière (backpropagation) dans le jargon de l'intelligence artificielle (LIPPMANN, 1987). Parmi les autres méthodes on retrouve l'adoucissement simulé (simulated annealing, AARTS et KORST, 1989), les algorithmes génétiques (HOLLAND, 1992) et la programmation évolutionnaire (FOGEL *et al.*, 1989). Cette dernière technique est très intéressante pour les réseaux récurrents, et elle permet non seulement d'établir les valeurs des paramètres du réseau, mais aussi sa configuration (MC DONNELL et WAAGEN, 1994).

### 2.2.4 Mérites des réseaux de neurones artificiels

Les réseaux de neurones artificiels sont plus souples que les modèles ARMA, mais ont plusieurs désavantages :



**Figure 4** Architecture d'un réseau de neurones artificiels complètement récurrent.  
*Configuration of a fully recurrent artificial neural network.*

1. la sélection des variables d'entrées est difficile ;
2. la configuration du réseau est une tâche délicate ;
3. l'estimation des paramètres est une tâche ardue ;
4. le modèle ne fournit pas de mesure de l'incertitude sur les prévisions.

De plus, tout comme les modèles ARMA, les réseaux de neurones ne peuvent s'appliquer qu'à des séries stationnaires (ou transformées pour l'être). Cependant lorsque l'application d'un modèle ARMA ne donne pas les résultats escomptés, l'utilisation d'un réseau de neurones peut s'avérer une alternative intéressante et améliorer la qualité des prévisions en modélisant les aspects non linéaires du système. Pour cette raison, les modèles à base de réseaux de neurones artificiels reçoivent de plus en plus d'attention dans le domaine de la prévision des phénomènes hydrologiques.

### 2.3 Régression linéaire floue

La régression floue est un outil récent issu de la théorie des sous-ensembles flous (ZADEH, 1965 ; KAUFMANN, 1973). Proposée initialement par TANAKA *et al.* (1982), la régression floue a été étudiée plus en profondeur par BÂRDOSSY (1990) ainsi que BÂRDOSSY *et al.* (1990). Il s'agit d'un outil utile lorsque la régression statistique ne peut s'appliquer soit parce que (1) il est connu *a priori* que des hypothèses du modèle de régression statistique ne sont pas respectées, ou parce que (2) le nombre d'observations est insuffisant pour permettre de valider les hypothèses du modèle *a posteriori*. La régression floue repose sur l'idée de nombre flou (DUBOIS et PRADE, 1988). C'est un concept qui permet de représenter l'incertitude sur une quantité réelle  $X$ . Il s'agit d'une fonction  $\mu : \mathbb{R} \rightarrow [0, 1]$  qui associe à chaque nombre réel  $x$  un degré d'appartenance  $\mu(x)$  compris entre 0 et 1, avec la contrainte que  $\mu(x) = 1$  pour au moins une valeur de  $x$ . Ce degré d'appartenance reflète le niveau auquel la valeur  $x$  fait partie du nombre flou. DUBOIS et PRADE (1988) montrent que sous certaines conditions la fonction  $\mu(x)$  permet de dériver une borne supérieure pour la distribution de probabilité de  $X$ , que l'on nomme dis-

tribution de possibilité. La possibilité  $\Pi(A)$  d'un événement A quelconque se calcule par :

$$\Pi(A) = \sup_{x \in A} \mu(x) \geq \Pr[A] \quad (13)$$

Parce que  $\Pi(A)$  est uniquement une borne supérieure de  $\Pr[A]$ , il est plus facile de dériver à partir de peu d'information une distribution de possibilité qu'une distribution de probabilité : en effet une surestimation de  $\Pi(A)$  n'a aucun effet sur la validité de la distribution de possibilité. C'est pourquoi, en cas d'ignorance totale sur la distribution de probabilité de X, on peut supposer chaque valeur x entièrement possible ( $\Pi(\{x\}) = \mu(x) = 1$ ). Une modélisation probabiliste adéquate de l'incertitude sur X en cas d'ignorance totale est beaucoup plus compliquée dans cette situation (BOX et TIAO, 1973).

La régression linéaire floue repose sur l'hypothèse d'un modèle linéaire dont les coefficients sont flous :

$$Y = A_0 + \sum_{k=1}^K A_k (x_k - \bar{x}_k) \quad (14)$$

où  $A_k$ ,  $k = 0, 1, \dots, K$  sont des nombres flous et  $\bar{x} = \{\bar{x}_1, \bar{x}_2, \dots, \bar{x}_k\}$  est un point de l'espace choisi *a priori* pour lequel on estime que le modèle est le plus précis. L'estimation des paramètres d'une régression linéaire floue est un problème de programmation linéaire. Les avantages du modèle de régression linéaire floue proviennent du fait qu'il est basé sur des hypothèses moins restrictives que celles de la régression linéaire classique : (1) la corrélation entre les observations d'une même variable explicative  $x_k$  n'influence pas la validité du modèle et (2) la variance n'a pas besoin d'être homogène ; en fait, on peut modéliser explicitement une structure particulière d'erreur de mesure. La régression linéaire floue est cependant très sensible aux valeurs extrêmes, qui doivent donc être validées avec précaution. Un autre désavantage de la régression floue est la complexité de l'algorithme d'estimation des paramètres. En effet, il est nécessaire de résoudre  $2^K$  systèmes d'inéquations à  $2(K+1)$  inconnues et  $2n$  contraintes, où  $n$  est le nombre d'observations. Il est donc clair que la méthode devient impraticable lorsque  $K$  et  $n$  sont très grands. Cependant, la régression floue est particulièrement avantageuse précisément lorsque  $n$  est petit, puisqu'il n'est alors pas possible, en utilisant la régression statistique classique, de valider *a posteriori* les hypothèses (1) et (2) de façon satisfaisante.

Il faut mentionner la controverse importante entourant la théorie des sous-ensembles flous. Il s'agit d'une théorie relativement récente dont les fondements mêmes sont remis en question par un noyau non négligeable de scientifiques, par exemple ELKAN (1994) et LINDLEY (1994). Dans la mesure où la théorie des sous-ensembles flous est utilisée comme base d'une théorie des possibilités, son utilisation semble justifiable. Cependant, la théorie des possibilités impose une condition stricte sur les distributions de possibilités en demandant de respecter l'axiome (13). Cette contrainte sur la probabilité supérieure d'un événement peut être réduite dans le cadre de la théorie de Dempster-Shafer (SHAFER, 1976) et éliminée si l'on préfère la théorie des prévisions inférieures (WALLEY, 1991). Cependant, le concept de la régression n'a pas encore été étendu à ces modèles d'incertitude.



## 2.4 Modèles de systèmes chaotiques

Par définition, un système chaotique est un système dynamique complexe non périodique. Dans la mesure où l'on accepte l'hypothèse d'un changement climatique, un bassin versant répond parfaitement à cette définition. Il existe plusieurs outils pour extraire la structure d'un système chaotique. Nous ne ferons que mentionner l'algorithme MARS (FRIEDMAN, 1991), qui a été appliqué avec succès dans le domaine des sciences de l'eau. La méthode est particulièrement complexe : le lecteur averti pourra consulter l'article extensif de Friedman, ou encore une application à la prévision du niveau d'eau du Grand Lac Salé (LALL *et al.*, 1996), qui montre l'intérêt de cette méthode pour la prévision à moyen terme (bihebdomadaire) et à long terme (plusieurs années). L'approche semble particulièrement intéressante pour étudier l'évolution des grands systèmes sous l'hypothèse d'un changement climatique, puisque l'avantage principal de ce type de modèle est qu'il permet de traiter directement les séries non stationnaires.

## 3 – MODÈLES CONCEPTUELS

Par opposition aux modèles empiriques, les modèles conceptuels reflètent, de par leur structure interne, une conception du système modélisé. En pratique, cette structure peut être vue comme un ensemble de contraintes qui limitent le champ d'investigation lors de la calibration. Ainsi, moins de données sont nécessaires pour obtenir des prévisions précises. Cependant, la spécification incorrecte du modèle peut causer un biais qui ne pourra être corrigé par l'acquisition de données supplémentaires. Nous distinguerons trois catégories de modèles conceptuels : (1) les modèles physiques, basés uniquement sur les lois physiques fondamentales telles que la conservation de la masse et de l'énergie, (2) les machines hydrauliques qui assimilent un bassin versant à un ensemble de conduites et de réservoirs, et (3) les modèles à base de règles floues qui permettent d'utiliser des connaissances imprécises. Nous discuterons aussi brièvement de la combinaison des modèles physiques et des machines hydrauliques, ainsi que du couplage possible entre les modèles conceptuels et les modèles empiriques.

### 3.1 Modèles physiques

Les modèles physiques sont basés sur une ou plusieurs lois physiques fondamentales : conservation de l'énergie, de l'entropie, de la masse, et loi de la gravitation. En pratique, leur application est réalisée en divisant le bassin en unités élémentaires, sur lesquelles les équations décrivant ces lois physiques sont assez simples et peuvent être résolues facilement. Les deux principales approches, que nous ne détaillerons pas ici, sont la méthode des différences finies et celle des éléments finis. Ces méthodes sont couramment utilisées en hydrogéologie, en hydraulique et en modélisation hydrodynamique (LECLERC *et al.*, 1992), mais encore assez peu en hydrologie de surface.

Un modèle physique peut être particulièrement utile pour évaluer le débit à un site non jaugeé, surtout lorsque la pente du cours d'eau est faible. Pour la prévision des débits à très court terme (débit horaire, par exemple), cette approche

peut s'avérer utile parce que le débit prévu dépend généralement, pour un bassin versant de taille moyenne, des conditions existant en amont et beaucoup moins des précipitations qui pourraient survenir (LETTENMAIER et WOOD, 1993). Si le débit varie de façon graduelle par rapport à l'échelle de temps, il peut être suffisant d'appliquer l'équation de continuité du débit (DOOGE, 1973). Autrement, il est nécessaire d'utiliser les équations différentielles dites de Saint-Venant (LAMBERTI et PILATI, 1996). LETTENMAIER et WOOD (1993) présentent une revue assez exhaustive des méthodes de prévisions des débits qui sont basées sur la modélisation de la transformation précipitations-débits, et des méthodes basées sur le laminage des volumes d'eau (onde dynamique, équation de diffusion, réponse impulsionnelle, Muskingum, etc.). L'intention de la présente étude n'est pas de reproduire la discussion des méthodes présentées par LETTENMAIER et WOOD (1993) mais de présenter une revue complémentaire tenant compte en particulier des résultats de recherche les plus récents. Une exigence importante des modèles physiques est que le réseau hydrographique doit être modélisé dans le détail. L'application de techniques fractales permet de construire efficacement un modèle du réseau à partir d'un modèle d'élévation de terrain (ICHOKU *et al.*, 1996).

La limite principale de ces approches, outre leur complexité, est le nombre de paramètres à considérer. En effet, on retrouve sur chaque unité élémentaire un ensemble de paramètres qui doivent être mesurés ou estimés. De plus, tous les processus hydrologiques impliqués ne sont pas assez bien compris pour pouvoir être décrits par un ensemble d'équations. Même pour de petits bassins très homogènes, une modélisation purement physique, bien que désirable, reste un objectif éloigné. Cependant, il serait possible de coupler un modèle physique utilisé pour une composante du bassin, par exemple l'eau souterraine, avec un autre type de modèle. Les modèles physiques ont été introduits ici principalement pour les distinguer des modèles basés sur les machines hydrauliques, souvent nommés à tort modèles physiques.

### 3.2 Machines hydrauliques

Une machine hydraulique est un ensemble de canaux, de conduites et de réservoir, dont le comportement est de nos jours simulé par ordinateur. Il ne s'agit pas d'un modèle physique, mais d'une représentation de la réalité basée sur des simplifications importantes du système, et des approximations souvent grossières, en particulier en ce qui concerne l'eau souterraine. Cependant de telles machines, simulées sur ordinateur, permettent souvent de modéliser adéquatement un bassin versant. La problématique associée aux machines hydrauliques et leur application en prévision est discutée par JAKEMAN *et al.* (1994).

Un désavantage important des machines hydrauliques est que leur calibration nécessite une quantité importante de données sur l'ensemble des paramètres météorologiques ayant une influence sur le débit. En pratique, on arrive souvent à effectuer une modélisation convenable en utilisant uniquement les mesures de précipitation et de température (JAKEMAN *et al.*, 1994).

### 3.3 Combinaison des modèles physiques et des machines hydrauliques

Il est important de noter qu'il est souvent difficile de distinguer les modèles physiques des modèles basés sur les machines hydrauliques à cause du cheveu-

chement de ces deux types de modèles tant dans les théories qui les régissent que dans la pratique de leurs applications. La majorité des modèles conceptuels de prévision des débits qui ont été présentés dans la littérature combinent un ou plusieurs éléments de ces deux types de modèles.

### 3.4 Couplage de modèles conceptuels et de modèles empiriques

Une avenue intéressante pour la prévision des débits consiste à utiliser plusieurs types de modèles pour les différentes composantes du système modélisé. Par exemple, il est possible d'utiliser un modèle stochastique pour générer des scénarios météorologiques qui peuvent être utilisés comme intrants dans un modèle conceptuel. C'est l'approche utilisée dans les modèles ESP (signifiant *extended streamflow prediction*), qui ont été présentés par TWEDT *et al.* (1977), LETTENMAIER (1984), DAY (1985), et LETTENMAIER et WOOD (1993). Les approches du type ESP sont basées sur des modèles de précipitations-apports spatialement agrégés, et utilisant comme intrants les valeurs observées et prévues des données météorologiques. Une application pratique de ces modèles, dans le cas de la province du Québec, a été proposée par BISSON (1995). Ce modèle conceptuel consiste à dériver des prévisions d'apports en utilisant un modèle global météo-apports spatialement agrégé (machine hydraulique préalablement calibrée). Le modèle est utilisé avec les données météorologiques (de précipitations et de température) observées jusqu'au moment de la prévision ainsi que la prévision météorologique pour les quatre prochains jours. Il est important de noter que LETTENMAIER (1984) a indiqué qu'il est essentiel de remettre à jour continuellement les paramètres desquels les prévisions ESP dépendent pour assurer une utilisation convenable de la méthode. En plus la distribution statistique des prévisions obtenues avec la méthode ESP reste différente de celle de l'historique quand le délai de prévision augmente. La procédure heuristique utilisée à Hydro-Québec pour garantir la convergence (BISSON, 1995) consiste à combiner linéairement les séries d'apports observés et les séries d'apports simulés en utilisant une fonction de transition. L'approche ESP, qui répond en général assez bien aux besoins de prévision, pourrait être formalisée en utilisant l'analyse bayésienne.

### 3.5 Modèles à base de règles floues

Les modèles à base de règles floues sont couramment utilisés dans le domaine du contrôle (KOSKO, 1992a), mais leur application aux systèmes ouverts est plus récente (BÁRDOSSY et DUCKSTEIN, 1995). Un tel modèle consiste en un ensemble de règles SI-ALORS qui relient les entrées du modèle à ses sorties. Voici deux exemples de règles pour la prévision du débit :

SI il y a une pluie *importante* en un temps *court*  
ALORS il y aura une *forte* augmentation du débit.

SI il y a une pluie *faible* en un temps *long*  
ALORS il y aura une *faible* augmentation du débit.

Les termes en italiques dénotent des quantificateurs imprécis, qui montrent que chaque règle est incertaine. Ceux-ci sont représentés dans ce type de modèle par des nombres flous. Lorsque le modèle est utilisé, on cherche d'abord toutes les règles qui peuvent s'appliquer à la situation (il y en a généralement plusieurs) et on combine les conséquences prédites par ces diverses règles pour obtenir une réponse unique. Cette réponse est en général aussi un nombre flou, dont

on prend souvent le centre de gravité pour obtenir une prévision précise. Évidemment, le nombre flou obtenu donne une idée de l'incertitude sur la valeur précise.

Les modèles à base de règles floues sont très utiles lorsque l'on dispose de connaissances générales sur les processus physiques du système à modéliser, sans que ces connaissances soient sous la forme d'équations précises. Ce type de modèle permet de donner un ordre de grandeur pour la prévision, mais peut aussi conduire à un modèle plus précis, puisqu'il a été montré que les modèles à base de règles floues forment aussi une classe d'approximateurs universels, tout comme les réseaux de neurones et les polynômes (KOSKO, 1992b). D'ailleurs, cette approche a été utilisée avec succès pour quantifier la prévision de la demande en eau en tenant compte des prévisions météorologiques, et pour prévoir la température moyenne journalière (BÁRDOSSY et DUCKSTEIN, 1995).

#### **4 – AMÉLIORATION DES MÉTHODES DE PRÉVISION À L'AIDE DU FILTRE DE KALMAN**

En pratique, un modèle de prévision des débits sera généralement utilisé en temps réel et sur une période de temps plus ou moins longue pour effectuer une séquence de prévisions avant d'être réévalué et calibré de nouveau. Pendant cette période, on peut réaliser que les erreurs d'estimation du modèle sont significativement corrélées, ce qui est une indication que ces dernières comportent encore une quantité d'information utile qui pourrait être modélisée. S'il serait préférable à ce point de modifier le modèle pour tenir compte de cette information, il peut s'agir d'une solution impraticable. Une solution plus simple que de modifier le modèle consiste à utiliser un filtre de Kalman pour éliminer l'auto-corrélation et améliorer les prévisions (KALMAN, 1960). BERGMAN et DELLEUR (1985) discutent de l'application du filtre de Kalman pour la prévision du débit journalier à l'aide d'un modèle autorégressif (AR). Cette technique a été utilisée avec succès entre autres pour la prévision à l'aide d'un modèle ARMAX du débit journalier sur la rivière des Outaouais (BENNIS et RASSAM, 1991) et du niveau mensuel du Lac Érié (BENNIS et ASSAF, 1994).

L'application du filtre de Kalman à un modèle complexe requiert cependant des efforts considérables, et peut mener à une amélioration limitée des performances pour les systèmes hydrologiques (O'CONNELL et CLARKE, 1981). Il peut être préférable de songer à employer un algorithme de calibration en temps réel qui serait appliqué en continu (AMBRUS, 1980).

#### **5 – DISCUSSION**

Il est difficile de proposer un classement absolu des modèles de prévisions des débits. En fait, l'ensemble des modèles présentés dans ce travail peut être appliqué à la prévision des débits. Chacun a ses forces et ses faiblesses, que

nous tenterons de résumer ici. Nous utiliserons plusieurs critères de comparaison pour chaque méthode :

1. la validité : les hypothèses sous-jacente sont-elles raisonnables ?
2. la facilité d'utilisation : peut-on aisément concevoir un tel modèle et le calibrer ?
3. la complexité des calculs : peut-on calculer rapidement et efficacement les prévisions ?
4. la modélisation de l'incertitude : peut-on évaluer la précision des prévisions ?
5. les besoins en information : quel type de données faut-il, combien en faut-il ?
6. l'adaptabilité : est-il facile de mettre à jour le modèle ?

Le tableau 1 discute de ces différents critères pour chaque méthode, et résume en fait les caractéristiques, présentées ci-haut, de chaque méthode.

**Tableau 1** Comparaison des méthodes de prévision des débits.

**Table 1** Comparison of streamflow forecasting methods.

Approche	Caractéristiques
Régression linéaire statistique	<ol style="list-style-type: none"> <li>1. Les hypothèses du modèle limitent son application aux prévisions des débits mensuels et annuels ; il s'agit d'un modèle linéaire, qui peut s'avérer insuffisant ;</li> <li>2. d'excellentes méthodes existent pour concevoir et calibrer un modèle de régression linéaire ;</li> <li>3. le calcul des prévisions est très simple ;</li> <li>4. on peut facilement obtenir un intervalle de confiance pour les prévisions ;</li> <li>5. on doit disposer de suffisamment de données pour pouvoir valider les hypothèses <i>a posteriori</i> ;</li> <li>6. le modèle doit être calibré à nouveau pour ajuster les coefficients d'ajustement.</li> </ol>
Régression linéaire floue	<ol style="list-style-type: none"> <li>1. Il s'agit d'un modèle linéaire ;</li> <li>2. la calibration du modèle est un problème de programmation linéaire dont la complexité augmente rapidement avec le nombre d'observations ;</li> <li>3. le calcul des prévisions est simple ;</li> <li>4. l'incertitude sur les prévisions est représentée par un nombre flou ;</li> <li>5. peut s'appliquer avec très peu de données ; la complexité de la calibration augmente rapidement avec la quantité de données ;</li> <li>6. le modèle doit être calibré à nouveau pour ajuster les coefficients d'ajustement.</li> </ol>
Box-Jenkins	<ol style="list-style-type: none"> <li>1. Les hypothèses de ce modèle permettent son application aux prévisions à court et moyen terme ; il s'agit cependant d'un modèle linéaire, qui peut s'avérer insuffisant – en particulier pour la prévision à moyen terme ; le système étudié doit être stationnaire, ou alors il faut appliquer une transformation pour rendre les séries stationnaires ;</li> <li>2. d'excellentes méthodes existent pour concevoir et calibrer un modèle de Box-Jenkins ;</li> <li>3. le calcul des prévisions est assez simple ;</li> <li>4. on peut obtenir un intervalle de confiance pour les prévisions ;</li> <li>5. on doit disposer de suffisamment de données pour pouvoir valider les hypothèses <i>a posteriori</i> ; il est possible de n'utiliser que le débit (ARIMA) ou d'ajouter d'autres variables (ARMAX) ;</li> <li>6. il existe des méthodes pour modifier les coefficients d'ajustement de façon continue.</li> </ol>
Réseau de neurones artificiels	<ol style="list-style-type: none"> <li>1. Modèle non linéaire généralisant le modèle de Box-Jenkins ;</li> <li>2. la conception et la validation d'un réseau de neurones est une tâche complexe ;</li> <li>3. le calcul des prévisions est complexe, mais relativement facile à automatiser ;</li> <li>4. on ne dispose pas de méthodes pour obtenir un intervalle de confiance pour les prévisions ;</li> <li>5. on doit disposer d'une base de données importante pour obtenir des résultats précis ;</li> <li>6. le modèle doit être calibré à nouveau pour ajuster les coefficients d'ajustement.</li> </ol>

**Tableau 1** Comparaison des méthodes de prévision des débits.**Table 1** Comparison of streamflow forecasting methods.

Méthode des k plus proches voisins	<ol style="list-style-type: none"> <li>1. Modèle non paramétrique s'appliquant à tous les systèmes stationnaires ou transformés pour l'être ;</li> <li>2. l'ajustement des coefficients de calibration est une étape délicate ;</li> <li>3. le calcul des prévisions est très lourd, car il implique une recherche dans tout l'historique ;</li> <li>4. on peut facilement obtenir un intervalle de confiance approximatif pour les prévisions ;</li> <li>5. on doit disposer d'un historique important pour que les résultats soient précis ;</li> <li>6. le modèle prend automatiquement en compte les nouvelles données dans la recherche des voisins, mais doit être calibré à nouveau pour ajuster les coefficients d'ajustement.</li> </ol>
Algorithme MARS pour systèmes chaotiques	<ol style="list-style-type: none"> <li>1. Approche nonparamétrique permettant de modéliser les systèmes non stationnaires ; particulièrement utile pour les prévisions à long terme ;</li> <li>2. la conception et la calibration sont des étapes ardues car le modèle est très complexe ;</li> <li>3. le calcul des prévisions est assez simple une fois le modèle déterminé ;</li> <li>4. on ne dispose pas d'intervalles de confiance pour les prévisions ;</li> <li>5. on doit disposer d'un historique important pour obtenir de bonnes prévisions ;</li> <li>6. le modèle doit être calibré à nouveau pour ajuster les coefficients d'ajustement.</li> </ol>
Modèles physiques	<ol style="list-style-type: none"> <li>1. Approche basée sur les lois physiques de l'écoulement, principalement la conservation de la masse ; pour le moment applicable uniquement que lorsque le débit peut être prévu à partir des conditions d'écoulement observées en amont, en négligeant les précipitations, la fonte et l'infiltration ; conséquemment applicable pour la prévision à très court terme, par exemple horaire ;</li> <li>2. la conception et la calibration sont des étapes ardues ;</li> <li>3. le calcul des prévisions est lourd ;</li> <li>4. on ne dispose pas d'intervalles de confiance pour les prévisions, mais celle-ci est négligeable si le niveau de détail du modèle est suffisant et que les hypothèses sont vérifiées (prévision à très court terme) ;</li> <li>5. ne nécessite pas d'observations passées du débit, sauf pour s'assurer que celui-ci dépend bien des conditions en amont ; cependant, un relevé topographique détaillé du bassin versant et une connaissance de la géologie sont essentielles ;</li> <li>6. le modèle doit être calibré à nouveau pour ajuster les coefficients d'ajustement (s'il y a lieu – une modélisation complète à l'aide des équations de Saint-Venant peut ne pas comporter de coefficients d'ajustement si toutes les conditions limites sont mesurées).</li> </ol>
Machines hydrauliques	<ol style="list-style-type: none"> <li>1. Les hypothèses reliées à l'approximation d'un bassin versant par un ensemble de canaux et de réservoirs sont souvent raisonnables ;</li> <li>2. la conception et la calibration d'un tel modèle est en général une longue démarche ;</li> <li>3. le calcul des prévisions est assez lourd ;</li> <li>4. en couplant une machine hydraulique à un modèle stochastique de génération de séries météorologiques, on peut obtenir des intervalles de confiance pour les prévisions par simulation Monte-Carlo ;</li> <li>5. quelques années de mesure peuvent suffire pour calibrer le modèle, mais il est essentiel de disposer d'un réseau de mesure dense pour les entrées, en particulier les précipitations ;</li> <li>6. le modèle doit être calibré à nouveau pour ajuster les coefficients d'ajustement.</li> </ol>
Modèles à base de règles floues	<ol style="list-style-type: none"> <li>1. Approche non paramétrique très générale – plus appropriée pour la prévision à moyen et long terme, puisqu'elle fournit principalement l'ordre de grandeur des prévisions ;</li> <li>2. la conception et la calibration sont des étapes assez simples ;</li> <li>3. le calcul des prévisions est simple ;</li> <li>4. on obtient des nombres flous comme mesure de l'incertitude des prévisions ;</li> <li>5. un historique important est nécessaire ;</li> <li>6. le modèle doit être calibré à nouveau pour ajuster les coefficients d'ajustement.</li> </ol>

## 5.1 Classement proposé pour la prévision à court terme

D'une prévision à court terme on demande qu'elle soit très précise. En effet, de telles prévisions sont utilisées uniquement pour une gestion très serrée de la

ressource, qui demande des données de qualité pour être efficace. On peut donc exclure à notre avis les méthodes ne fournissant pas de réponses précises (modèles à base de règles floues) et celles ne permettant pas de quantifier l'incertitude des prévisions (réseaux de neurones artificiels, modèle MARS), à l'exception des modèles physiques puisque l'erreur de prévision d'un modèle physique peut être contrôlée par le niveau de détail utilisé pour la résolution des équations différentielles. Évidemment, une méthode peut être très précise même si cette précision ne peut être quantifiée, mais le risque est alors difficile à évaluer. Or une procédure de gestion n'est optimale que si elle prend en compte ce risque d'erreur (DAVIS *et al.*, 1979).

Il faut aussi éliminer les méthodes de régression, qui ne permettent pas de modéliser la dépendance entre les observations successives des variables explicatives, alors qu'elles sont très importantes pour la prévision à court terme. La forte corrélation entre les observations successives à cette échelle de mesure favorise les modèles stochastiques tels que la méthode des  $k$  plus proches voisins et la méthodologie de Box-Jenkins, dont le fonctionnement repose précisément sur l'exploitation de cette persistance. Comme la base des données passées disponibles pour la prévision à court terme devient rapidement importante, ceci favorise encore davantage la méthode des  $k$  plus proches voisins, qui exploite directement cette information.

Au niveau des modèles physiques, il peut être suffisant de respecter l'équation de conservation de la masse au lieu de solutionner les équations de Saint-Venant pour la prévision à court terme. Un modèle physique est à privilégier si les conditions d'application sont réunies, c'est-à-dire lorsque le débit en aval dépend presque uniquement des conditions existant en amont. Une machine hydraulique, couplée à un modèle météorologique stochastique basé sur les prévisions peut aussi être employé, mais demande un effort considérable.

Du plus simple à appliquer au plus complexe, nous proposons donc l'ordre suivant pour les modèles de prévision à court terme (débits horaires ou journaliers) :

1. méthode des  $k$  plus proches voisins ;
2. modèle de Box-Jenkins (ARIMA ou ARMAX, selon le type de données disponibles) ;
3. modèle physique basé sur l'équation de conservation de la masse (si applicable) ;
4. modèle physique basé sur les équations différentielles de Saint-Venant (si applicable) ;
5. machine hydraulique couplée à un modèle météorologique stochastique basé sur les prévisions.

Notons encore une fois que la méthode des  $k$  plus proches voisins ainsi que le modèle de Box-Jenkins nécessite des séries de données stationnaires (ou transformées pour l'être). Il faut donc que le passé soit garant de l'avenir, ce dont on peut douter étant donné les changements climatiques manifestes. Il faut aussi s'assurer que l'aménagement du bassin par l'homme n'a pas modifié sa réponse. Si c'est le cas, il faut alors ne considérer que les données représentatives de l'état actuel du bassin.

## 5.2 Classement proposé pour la prévision à moyen terme

Pour la prévision à moyen terme (débits hebdomadaires ou mensuels), une plus grande variété de modèles peuvent être utilisés avec succès. Il n'est pas

nécessaire par exemple qu'un modèle permette de quantifier précisément l'incertitude : à cette échelle de temps, une analyse de sensibilité peut être effectuée sur chaque modèle pour évaluer l'ordre de grandeur de la variabilité des prévisions. L'effort mis à déterminer les variables explicatives importantes et à structurer le modèle peut avoir à cette échelle plus d'influence que le choix du modèle, dans la mesure où celui-ci est assez souple. Cet effort peut cependant être considérablement réduit s'il existe une base de données historiques (stationnaires) importante : la méthode non paramétrique des  $k$  plus proches voisins peut alors être appliquée efficacement.

Il est possible pour la prévision à moyen terme de considérer un modèle de régression pour chaque période (52 pour les prévisions hebdomadaires, 12 pour les prévisions mensuelles). Cependant, le grand nombre de paramètres devant être estimés, de même que l'effet des corrélations entre les débits font qu'à cette échelle l'approche n'est pas recommandée. Il faut aussi éliminer les modèles physiques, s'ils ne sont pas couplés à un modèle de prévision météorologique et à un modèle hydrogéologique : l'influence des précipitations et de l'infiltration se font sentir à cette échelle (il faudrait aussi ajouter un modèle physique de fonte de neige au printemps).

Pour l'ensemble des autres modèles, on peut proposer un classement qui est fonction du type et de la quantité d'information disponible (tabl. 2). Ceci ne permet pas d'identifier le « meilleur » modèle pour la prévision à moyen terme, mais plutôt d'identifier le plus simple à appliquer en fonction de l'information disponible.

**Tableau 2** Classement des modèles de prévision à moyen terme en fonction de l'information disponible.

**Table 2** *Ranking of medium-term forecasting models with respect to the available information.*

Type et quantité d'information disponible	Modèle le plus simple pouvant être utilisé
Séries stationnaires pour plusieurs dizaines d'années	Méthode des $k$ plus proches voisins
Séries stationnaires plus courtes	Méthode de Box-Jenkins
Séries stationnaires courtes, mais information subjective ou qualitative disponible d'un expert	Modèle à base de règles floues
Séries courtes (quelques années)	Machine hydraulique couplée à un modèle météorologique basé sur l'historique et les prévisions météorologiques à long terme

### 5.3 Classement proposé pour la prévision à long terme

Pour la prévision à long terme (débits annuels, tendance pour les prochaines années), il faut tenir compte des tendances climatiques, autrement la prévision se résume à une moyenne des observations. De façon empirique, ceci peut se faire par une régression en fonction du temps, à condition de pouvoir identifier clairement la tendance, à l'aide d'un modèle de Box-Jenkins en travaillant sur les différences pour stationariser les séries, ou à l'aide de l'algorithme MARS, qui permet de travailler avec des séries non stationnaires, à condition que les changements soient graduels. Il serait aussi possible d'employer une machine hydraulique couplée à un modèle météorologique à long terme. Cette dernière



entreprise peut cependant demander des ressources énormes, bien que le potentiel de cette approche soit important au niveau de la compréhension de l'effet des changements climatiques. De la plus simple à la plus complexe, on peut donc proposer le classement suivant :

1. régression statistique ou floue, selon la quantité d'information disponible ;
2. modèle de Box-Jenkins (ARIMA ou ARMAX, selon le type de données disponibles) ;
3. l'algorithme MARS ;
4. une machine hydraulique couplée à un modèle de prévision météorologique à long terme.

Le choix d'un modèle, spécialement pour la prévision à long terme, ne devrait cependant pas être effectué en considérant d'abord la complexité d'application des différentes approches, mais plutôt en fonction des hypothèses qu'il est possible d'avancer et de vérifier. En particulier, la seule approche empirique mentionnée qui permet véritablement de prendre en compte un système non stationnaire est l'algorithme MARS.

## CONCLUSION

Un nombre important de méthodes empiriques et conceptuelles de prévisions des débits ont été présentées et comparées pour la prévision à court, moyen et long terme. Il en est ressorti que les modèles empiriques pouvaient fournir des solutions plus simples, mais que ceux-ci reposent sur une hypothèse forte : que les observations passées sont représentatives des conditions futures. Lorsque cette hypothèse est vérifiée et qu'une grande quantité d'information est disponible sur le comportement passé du débit et des variables explicatives du débit, la méthode des  $k$  plus proches voisins nous apparaît très prometteuse pour la prévision à court et moyen terme, puisqu'elle n'impose pas *a priori* de contraintes sur la forme du modèle. De plus, un modèle construit pour un site à l'aide de cette méthode peut rapidement être adapté à un autre site, puisque cette approche repose uniquement sur la consultation d'une base de données.

De façon générale, on ne devrait pas utiliser un modèle empirique pour faire la prévision d'événements exceptionnels, puisque par définition on ne dispose alors que de peu d'observations de phénomènes semblables. Un modèle basé sur des hypothèses physiques réalistes, même s'il est simple, pourrait alors être beaucoup plus utile.

S'il est important de choisir un bon modèle en fonction du problème et de l'information disponible, il ne faut pas non plus négliger la possibilité de combiner plusieurs modèles pour améliorer la prévision. Cette idée, d'abord suggérée par CAVADIAS et MORIN (1985), a été appliquée avec succès à la prévision en temps réel des crues de la Garonne (LOUMAGNE *et al.*, 1995). Une application à des cas réels des différentes approches considérées dans ce rapport pourrait permettre d'établir des lignes directrices plus précises pour la sélection d'un modèle de prévision des débits.

## RÉFÉRENCES BIBLIOGRAPHIQUES

- AARTS, E. et J. KORST (1989). *Simulated Annealing and Boltzmann Machines: A Stochastic Approach to Combinatorial Optimization and Neural Computing*. Wiley.
- AKAIKE, H. (1972). Use of an Information Theoretic Quantity for Statistical Model Identification. *Proc. 5th Hawaii Int. Conf. System Sciences*, Western periodicals, p. 249-250.
- AMBRUS, S. (1980). Real time forecasting of discharges on the River Danube using self-tuning predictor algorithms. *Hydrological forecasting*. Proceedings of the Oxford Symposium, April 1980, IAHS Publ. 129.
- BARDOSSY, Á. (1990). Note on fuzzy regression. *Fuzzy Sets and Systems*, 37 : 65-75.
- BARDOSSY, Á., BOGÁRDI, I., DUCKSTEIN, L. (1990). Fuzzy Regression in Hydrology. *Water Resources Research*, 26 (7) : 1497-1508.
- BARDOSSY, Á., DUCKSTEIN, L. (1995). *Fuzzy Rule-Based Modeling with Applications to Geophysical, Biological and Engineering Systems*. CRC Press, Boca Raton, USA.
- BEBIS, G., GEORGIPOULOS, M. (1994). Feed-forward neural networks - why network size is so important. *IEEE Potentials*. October/November 1994 : 27-31.
- BENNIS, S., ASSAF, G.J. (1994). Adaptive forecast of multi-month lake level elevation. *Canadian Journal of Civil Engineering*, 21 : 778-788.
- BENNIS, S., RASSAM, J.-C. (1991). Utilisation d'un modèle ARMAX pour la prévision du débit à Carillon. *Canadian Journal of Civil Engineering*, 18 : 864-870.
- BERGMAN, M.J., DELLEUR, J.W. (1985). Kalman Filter Estimation and Prediction of Daily Stream Flows: Review, Algorithm and Simulation Experiments. *Water Resources Bulletin*, 21 (5) : 815-825.
- BESSON, J.-L. (1995). Génération d'apports prévisionnels. *Méthodes statistiques et bayésiennes en hydrologie*. Conférence internationale en l'honneur de Jacques Bernier. UNESCO, Paris, 11-13 septembre 1995.
- BOX, G.E.P., JENKINS, G.M. (1976). *Time Series Analysis: Forecasting and Control*, 2<sup>e</sup> éd., Holden-Day, San-Francisco.
- BOX, G.E.P., TIAO, G.C. (1973). *Bayesian Inference in Statistical Analysis*, Addison-Wesley.
- CAUCHY, A.L. (1847). Méthode générale pour la résolution des systèmes d'équations simultanées. *C. R. Acad. Science*, Paris, 25 : 536-538.
- CAVADIAS, G., MORIN G. (1985). Amélioration des performances des modèles hydrologiques par combinaison des débits simulés. *Revue internationale des sciences de l'eau*, 1 : 29-35.
- CHAKRABORTY, K., MEHROTRA, K., MOHAN, C.K., RANKA S. (1992). Forecasting the Behavior of Multivariate Time Series Using Neural Networks. *Neural Networks*, 5 : 961-970.
- CONNOR, J.T., MARTIN, R.D., ATLAS, L.E. (1994). Recurrent Neural Networks and Robust Time Series Prediction. *IEEE Transactions on Neural Networks*, 5 (2) : 240-254.
- CUBERO, R.G. (1991). Neural Networks for Water Demand Time Series Forecasting. In: *Artificial Neural Networks, Proceedings of the International Workshop IWANN'91*, A. Prieto Éd., Springer-Verlag, Germany, 453-460.
- DAVIS, D.R., DUCKSTEIN, L., KRZYSZTOFOWICZ, R. (1979). The worth of hydrological data for nonoptimal decision making. *Water Resources Research*, 15 (6) : 1733-1742.
- DAY, G.N. (1985). Extended Streamflow Forecasting Using NWSRFS. *J. Water Resour. Plann. & Manage.*, vol. 111, n° WR2, p. 157-170.
- DOOGE, J.C.I. (1973). *Linear Theory of Hydrologic Systems*. Technical bulletin no. 1468, U.S. Department of Agriculture, Washington, D.C.
- DUBOIS, D., PRADÉ, H. (1988). *Théorie des possibilités : Applications à la représentation des connaissances en informatique*, 2<sup>e</sup> éd., Masson, Paris.
- ELKAN, C. (1994). The Paradoxical Success of Fuzzy Logic. *IEEE Expert*, Août 1994, 3-49.

- FELLER, W. (1968). *An Introduction to Probability Theory and Its Applications, Volume I*, 3<sup>e</sup> éd., Wiley.
- FOGEL, L.J., OWENS, A.J., WALSH, M.J. (1989). *Artificial Intelligence Through Simulated Evolution*. Wiley.
- FORTIN, V., OUARDA, T.B.M.J., BOBÉE, B. (1997). Comment on « The use of artificial neural networks for the prediction of water quality parameters » by H.R. Maier and G.C. Dandy. *Water Resources Research*, 33 (10) : 2423-2424.
- FRENCH, M.N., KRAJEWSKI, W.F., CUYKENDALL, R.R. (1992). Rainfall forecasting in space and time using a neural network. *Journal of Hydrology*, 131 : 1-31.
- FRIEDMAN, J.H. (1991). Multivariate Adaptive Regression Splines. *Annals of Stat.*, 19 (1) : 1-141.
- GALEATI, G. (1990). A comparison of parametric and non-parametric methods for runoff forecasting. *Hydrol. Sci. J.*, 35 (1) : 79-94.
- HIPPEL, K.W., MCLEOD, A.I., LENNOX, W.C. (1977). Advances in Box-Jenkins Modeling - 1. Model Construction. *Water Resources Research*, 13 (3) : 567-575.
- HOERL, KENNARD (1970). Ridge Regression: Applications to Nonorthogonal Problems. *Technometrics*, 12 (1) : 55-67.
- HOLLAND, J.H. (1992). *Adaptation in Natural and Artificial Systems*, 2<sup>nd</sup> ed., MIT Press, Cambridge.
- HORNIK, K., STINCHCOMBE, M., WHITE, H. (1989). Multilayer Feedforward Networks are Universal Approximators. *Neural Networks*, 2 : 359-366.
- HSU, K.-I., GUPTA, V., SOROOSHIAN, S. (1995). Artificial neural network modeling of the rainfall-runoff process. *Water Resources Research*, 31 (10) : 2517-2530.
- ICHOKU, C., KARNIELI, A., VERCHOVSKY, I. (1996). Application of fractal techniques to the comparative evaluation of two methods of extracting channel networks from digital elevation models. *Water Resources Research*, 32 (2) : 389-399.
- JAKEMAN, A.J., POST, D.A., SCHREIDER, S. YU, YE, W. (1994). Modelling environmental systems: partitioning the water balance at different catchment scales. Dans : *Computer Techniques in Environmental Studies V vol II : Environmental Systems*, Computational Mechanics Publications, Southampton, P. Zannetti éd., p. 157-170.
- KALMAN, R.E. (1960). A new approach to linear filtering and prediction problems. *Journal of Basic Engineering* (Transactions of the ASME, Series D), 82 : 34-45.
- KARLSSON, M., YAKOWITZ, S. (1987). Nearest-Neighbor Methods for Nonparametric Rainfall-Runoff Forecasting. *Water Resources Research*, 23 (7) : 1300-1308.
- KARUNANITHI, N. et al. (1994). Neural networks for river flow prediction, *Journal of computing in civil engineering*, 8 (201).
- KAUFMANN, A. (1973). *Introduction à la théorie des sous-ensembles flous à l'usage des ingénieurs (Fuzzy Sets Theory) 1. Éléments théoriques de base*, Masson, Paris.
- KOSKO, B. (1992a). Neural networks and fuzzy systems: a dynamical systems approach to machine intelligence, Prentice Hall, New Jersey, 449 p.
- KOSKO, B. (1992b). Fuzzy Systems as Universal Approximators, *FUZZ-IEEE* : 1153-1162.
- LALL, U., SANGOYOMI, T., ABARBANEL, H.D.I. (1994). Nonlinear dynamics of the Great Salt Lake: Nonparametric short term forecasting. *Water Resources Research*, 32 (4) : 975-985.
- LAMBERTI, P., PILATI, S. (1996). Flood propagation models for real-time forecasting. *Journal of Hydrology*, 175 : 239-265.
- LAPEDES, A., FARBER, R. (1987). *Nonlinear signal processing using neural networks: Prediction and Modeling*. Technical Report, LA-UR87-2662, Los Alamos National Laboratory, Los Alamos, New Mexico.
- LECLERC, M., BOUDREAU, P., BOUDREAU, P., FORTIN, G.R., CLEARY, L. (1992). *Modélisation hydrodynamique des écoulements en eau libre du tronçon Tracy - lac Saint-Pierre*. Rapport de recherche 317a, INRS-Eau.
- LESHNO, M., LIN, V.Y., PINKUS, A., SCHOCKEN, S. (1993). Multilayer Feedforward Networks With a Nonpolynomial Activation Function Can Approximate Any Function. *Neural Networks*, 6 : 861-867.
- LETTENMAIER, D.P. (1984). Limitations on seasonal snowmelt forecast accuracy, *J.*

- Water Resour. Plann. & Manage.*, vol. 108, n° WR3, p. 255-269.
- LETTENMAIER, D.P., WOOD, E.F. (1993). Hydrologic Forecasting. In: *Handbook of hydrology*, D.R. Maidment ed., McGraw-Hill.
- LINDLEY, D. (1994). Foundations. Dans : *Subjective Probability*, Ayton, P. et G. Wright édés., Wiley, 3-15.
- LIPPMANN, R.P. (1987). An Introduction to Computing with Neural Nets. *IEEE ASSP Magazine*, April 1987 : 4-22.
- LJUNG, L. (1987). *System Identification: Theory for the user*. Prentice-Hall, New Jersey.
- LOUMAGNE, C., VIDAL, J.J., FELIU, C., TERTEROTOT, J.P., ROCHE, P.A. (1995). Procédure de décision multimodèle pour une prévision des crues en temps réel : application au bassin supérieur de la Garonne. *Revue des sciences de l'eau*, 8 : 539-561.
- MAIER, H.R., DANDY, G.C. (1996). The use of artificial neural network for the prediction of water quality parameters. *Water Resources Research*, 32 (4) : 1013-1022.
- MCCULLOCH, W.S., PITTS, W. (1943). A Logical Calculus of the Ideas Imminent in Nervous Activity. *Bulletin of Mathematical Biophysics*, 5 : 115-133.
- MCDONNELL, J.R., WAAGEN, D. (1994). Evolving Recurrent Perceptrons for Time-Series Modeling. *IEEE Transactions on Neural Networks*, 5 (1) : 24-38.
- O'CONNELL, P.E., CLARKE, R.T. (1981). Adaptive hydrological forecasting - a review. *Hydrological Sciences Bulletin*, 26 (2) : 179-205.
- REFENES, A.N., ZAPRANIS, A., FRANCIS, G. (1994). Stock Performance Modeling Using Neural Networks: A Comparative Study With Regression Models. *Neural Networks*, 7 (2) : 375-388.
- SAKAKIMA, S., KOJIRI, T., ITOH, K. (1992). Real-Time Reservoir Operation with Neural Nets Concept. In: *Applications of Artificial Intelligence in Engineering VII*, Computational Mechanics Publications, Southampton, p. 502-514.
- SALAS, J.D., DELLEUR, J.W., YEVJEVICH, V., LANE, W.L. (1988). *Applied Modeling of Hydrologic Time Series*. Water Resources Publication, Littleton, Colorado.
- SHAFER, G. (1976). *A Mathematical Theory of Evidence*, Princeton University Press, New Jersey.
- SHAMSELDIN, A.Y., O'CONNOR, K.M. (1996). A nearest neighbor linear perturbation model for river flow forecasting. *Journal of hydrology*, 179 : 353-375.
- SMITH, J., ELI, R.N. (1995). Neural networks models of rainfall-runoff process, *Journal of water resour. Plann. & Manage.*, 121 (56).
- TANAKA, H., UEJIMA, S., ASAI, K. (1982). Linear regression analysis with fuzzy model, *IEEE Trans. Syst. Man Cybern.*, SMC-12 : 903-907.
- TWEDT, T. M., SCHAAKE, J.C., PECK, Jr., E.L. (1977) National Weather Service Extended Streamflow Prediction, *Proc. 45th Western Snow Conference*, Albuquerque, N. Mex., p. 52-57.
- WALLEY, P. (1991). *Statistical Reasoning with Imprecise probabilities*, Chapman and Hall, London.
- WETHERILL, G.B., DUNCOMBE, P., KENWARD, M., KÖLLERSTRÖM, J., PAUL, S.R., VOWDEN, B.J. (1986). *Regression Analysis with Applications*, Chapman and Hall, London.
- WILLIAMS, R.J. (1990). Adaptive Sate Representation and Estimation using Recurrent Connectionist Networks. In: *Neural Networks for Control*, W.T. Moller, R.S. Sutton et P.J. Werbos Eds, MIT Press, Cambridge.
- YAKOWITZ, S. (1987). Nearest neighbor methods for time series analysis, *J. Time Ser. Anal.*, 8 (2) : 235-247.
- ZADEH, L.A. (1965). Fuzzy sets. *Information and Control*, 8 : 338-353.
- ZHANG, E.Y., TRIMBLE, P. (1996). Forecasting water availability by applying neural networks with global and solar indices. *Proceedings of the sixteenth annual American Geophysical Union Hydrology Days*. H. J. Morel-Seytoux Ed., Hydrology Days Publications, p. 559-570.