Expertise des travaux effectués par le BRGM pour le compte de la DIREN Poitou-Charentes et annexés au rapport : RP-54569 FR-V2 :
" Recherche d'indicateurs piézométriques pour la gestion des prélèvements en nappe (Phase 1: bassins de la Dive-Nord, du Clain, de la Sèvre Niortaise, de la Boutonne et de la Seudre)"

Préambule

Cette expertise est réalisée pour le compte de la DIREN Poitou-Charentes demandant la formulation d'un avis sur : "le choix et la pertinence du logiciel TEMPO de traitement et de modélisation de séries piézométriques retenu par le BRGM dans le cadre d'une étude confiée à cet établissement public". Il nous a semblé judicieux de décomposer notre réponse selon les deux points principaux suivants :

1- le bien fondé d'une approche de modélisation via les concepts inclus dans le logiciel TEMPO du BRGM,
2- quelques réflexions à propos des résultats obtenus par application de cette modélisation aux bassins versants de la Dive-Nord, du Clain, de la Sèvre Niortaise, de la Boutonne et de la Seudre (rapport d'étude BRGM RP-54569 FR-V2).

Il faut d'abord rappeler rapidement quels étaient les attendus de la DIREN Poitou-Charentes, mandataire de l'étude confiée au BRGM. En première simplification, la préoccupation principale est de comprendre dans son ensemble le fonctionnement hydrologique des bassins versants sus-nommés afin de mieux apprécier comment évaluer le potentiel d'exploitation en eau et ses conséquences sur le débit des rivières, en particulier en cas de régime récessif. Ceci dit, on comprendra que l'objectif n'est pas d'obtenir une modélisation fine du couplage entre la dynamique climatique, le comportement hydraulique des nappes souterraines et les eaux de surface. Ce genre d'approche est très long à mettre en place et s'il offre de nombreux intérêts en particulier dans l'analyse des incidences locales d'effets anthropiques ou d'aménagements, il n'est pas certain qu'au final le modèle soit plus précis que d'autres approches bien plus simples. La raison en est pour l'essentiel, qu'un modèle à physique fine demande une masse considérable d'information et de surcroît bien distribuée sur le domaine d'étude pour son conditionnement. Cette information est souvent manquante, auquel cas, il faut avoir recours à des simulations prospectives non-triviales évaluant le potentiel prédictif du modèle.

Le contre-pied à cette approche rigoureuse mais très lourde est de s'engager dans une démarche que l'on pourrait qualifier d'analyse pragmatique des relations de causes à effets. Par exemple, s'il pleut sur un bassin, on s'attend à ce que le niveau des nappes puisse monter, ainsi que le niveau des rivières. De cette simple constatation, on adopte une approche qualifiée de <u>systémique</u> qui ne voit pas explicitement la complexité physique du système hydrologique et ses couplages, mais se contente d'analyser les relations entre <u>entrée</u> (e.g., la pluie) et <u>sortie</u> (e.g., le débit d'une rivière). Bien évidemment, cette approche systémique se coupe *a priori* de tout sens physique donné aux résultats, encore que le propos puisse être modulé, on verra cela plus loin. En contre-partie, la modélisation reste "simple" de mise en œuvre, ne nécessite pas une masse impressionnante d'information. Elle fournit des résultats (et des simulations) dont l'interprétation est aisée, généralement en accord avec la philosophie

qui consiste à regarder le système hydrologique "en grand" et y rechercher des indicateurs de gestion. Le logiciel TEMPO et l'étude du BRGM discutés dans cette expertise sont basés sur cette approche systémique. En ce sens, il y a adéquation parfaite entre les outils utilisés et la demande formulée par la DIREN. Le fonctionnement des bassins hydrologiques est ausculté selon deux aspects principaux :

1- les relations sur un bassin entre les pluies entrantes et un indicateur du niveau des nappes. Cet indicateur (niveau piézométrique) choisi sur un ouvrage doit être le moins perturbé possible par des forçages externes qui seraient fugaces, néanmoins il peut (et il est difficile d'y échapper) intégrer les effets des prélèvements par pompage (AEP, Agricole, Industriel,…);
2- connaissant la réponse des nappes et sachant que des relations d'échanges de flux entre nappes et rivières sont possibles, on analyse ensuite la réponse en débit des rivières (exutoire du bassin versant) aux sollicitations climatiques (pluie, pour simplifier) et au comportement de la nappe, le niveau piézométrique choisi en (1) faisant foi d'indicateur du stock d'eau souterraine. De plus, le calcul est assorti d'une évaluation des contributions des composantes respectives du débit : alimentation rapide ou lente par la nappe, ruissellement des eaux pluviales.

Le problème étant de principe bien posé, l'expertise est décomposée en deux parties. La première est rédigée sous la forme d'un texte informel, non directement dédié à l'étude spécifiquement menée par le BRGM mais discutant de l'approche systémique en général. On en rappelle les grands principes et les points cruciaux où le modèle mérite quelques attentions. Ce faisant, on donne ensuite quelques éclaircissements sur les fonctionnalités proposées par TEMPO, notamment celles qui peuvent être utilisées pour l'étude concernée. De fait, il est hors de question de prétendre dans cet exposé général à l'exhaustivité, mais bien de fournir les quelques explications qui, selon nous, justifient de l'assertion consignée plus haut à savoir l'adéquation entre l'outil et les attendus de l'étude. La seconde partie regarde et commente les résultats de l'utilisation de TEMPO sur divers bassins versants et constituant le corps du rapport BRGM RP-54569 FR-V2. Cette partie n'entend pas reprendre point par point les termes du rapport, d'autant que ce dernier est pour beaucoup constitué de figures et tableaux qui synthétisent les résultats de modélisation. Partant du principe que l'outil est bon (voir plus haut) il s'agira plus de mettre en avant quelques éléments qui mériteraient éclaircissement ou discussion que d'épingler des lacunes rédhibitoires.

Partie I : L'approche systémique

Les grands principes de l'approche

Rappelons que modélisation dite systémique aborde, comme son nom l'indique, le comportement global d'un objet complexe sous la forme d'un système, sans chercher à identifier explicitement la complexité interne de l'objet. Sans que cela ait une quelconque connotation péjorative, cette approche prend parfois le nom de méthode "Boite Noire", là encore, terme explicite pour signifier qu'on se contente de regarder la complexité d'un objet sous une forme générique permettant d'établir, sans disséquer les processus, des relations de causes à effets. Bien qu'il existe plusieurs formes d'approches systémiques (c'est par exemple un des chevaux de bataille de l'automatique et de la robotique), dans le cas présent l'essentiel du problème est de traiter de séries temporelles de mesures (chroniques) et voir

ainsi comment une ou plusieurs séries considérées comme des entrées du système influencent (ou permettent de mimer) un signal de sortie. Dans toute modélisation systémique, les étapes et objectifs incontournables sont les suivants :

1- on "éduque" la "boite"noire" en modifiant ses réglages de telle sorte que, jouant sur le signal d'entrée, la boite est capable de mimer le signal de sortie. On considère le modèle satisfaisant quand la sortie qu'il calcule est comparable aux données disponibles. Par exemple, on simule le niveau piézométrique d'une nappe en fonction des chroniques mesurées de pluies sur un bassin, puis on compare ce niveau piézométrique à une (ou plusieurs) chroniques de niveaux mesurés sur des puits;
2- lorsque le modèle est jugé satisfaisant (éduqué), on peut parfois redonner un peu de sens physique à la modélisation effectuée dans la mesure où on est capable d'analyser comment la boite noire a déformé le signal d'entrée : modification de son amplitude ; effet de retard entre l'entrée et la sortie, une entrée à un instant t ne pouvant influencer la sortie qu'au temps t', t' > t ; s'il y a plusieurs signaux d'entrées, comment ces derniers sont modulés et contribuent au signal de sortie ; etc. Notons que ce sens physique redonné à l'approche reste général et analyse le système dans sa globalité. Ce sens physique reste toujours très inféodé aux éléments mathématiques que l'on met dans la boite noire (voir plus loin, la notion de réponse impulsionnelle);
3- enfin, un des intérêt du modèle éduqué est de prédire. En première simplification et sous couvert que l'objet modélisé ne change pas (ou peu) avec les formes et amplitudes des signaux d'entrée, on simule un signal de sortie en réponse à des signaux d'entrées fictifs, généralement ceux que l'on peut calculer, estimer (inventer !) et qui correspondraient à des scénarios de tests. Par exemple : simuler le niveau piézométrique d'une nappe en réponse à une année (fictive) de pluies particulièrement intenses/faibles.

L'architecture générale d'un modèle

Attendu que l'approche systémique se rapproche d'une analyse de traitement du signal quel que soit l'objet étudié, l'architecture de base d'un modèle est toujours sensiblement la même. Cette architecture est décrite en Figure 1. On désigne par $Eb_i(t)$, les chroniques de mesures correspondant aux entrées brutes, par exemple, la pluie journalière, la température au sol, le degré hygrométrique etc… Ces entrées peuvent influencer directement la sortie ou nécessiter un pré-traitement. Ce pré-traitement est un point extrêmement important de l'approche systémique et très souvent, il conditionne la qualité des résultats, pour ne pas dire la capacité du modèle à simuler le signal de sortie. Ce pré-traitement est rapporté sur la Figure 1 sous le module Compilateur d'Entrées (CE). Il faut prendre le terme au sens le plus large. Parfois la compilation de données ne consistera qu'en un lissage des données brutes et l'établissement de chroniques à pas de temps constant. Dans d'autres cas, le CE effectuera une analyse statistique plus ou moins fouillée des chroniques, en essayant d'extraire de l'information sur la corrélation en temps des données, que cette corrélation intéresse un seul type de mesures (auto-corrélation) ou deux chroniques différentes (corrélation croisée). Enfin, le CE peut aussi utiliser les données brutes $Eb_i(t)$ pour calculer des chroniques $E_i(t)$ qui correspondront aux entrées réelles dans le système. Par exemple, à partir de la pluie, de l'hygrométrie, de la température au sol, on calculera un bilan d'énergie radiative, les effets évaporatoires et on déduira la pluie "efficace" s'infiltrant réellement

dans la nappe. Ces entrées $E_i(t)$ (Fig. 1) sont finalement les chroniques que l'opérateur mathématique FT (Fonction de Transfert) va manipuler pour calculer la sortie du système S(t). Notons que l'importance du CE dans la chaîne de traitement d'un modèle systémique est d'autant plus fondamentale qu'on peut lui adjoindre des outils et méthodes de simulations d'entrées fictives. Comme on l'a vu plus haut, ces simulations d'entrées permettent le test de scénarios prospectifs. A ce propos, le logiciel TEMPO est particulièrement bien doté, en l'occurrence sur les aspects simulations climatiques : pluie, bilan radiatif, calcul de l'évaporation etc. Ce sont ces simulations qui permettent d'évaluer ensuite l'évolution de la nappe et le débit des rivières en réponse à une succession d'années sèches.

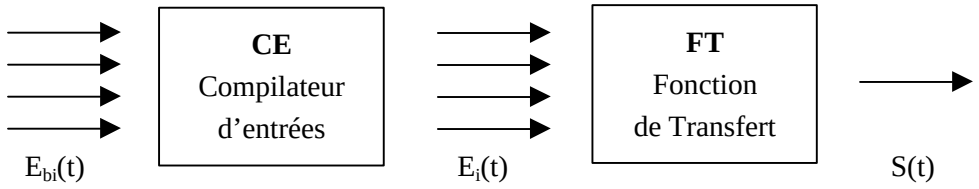

Figure 1 : Architecture générale d'un modèle systémique

Le module Fonction de Transfert FT est le cœur du calcul d'un modèle systémique. Le principe est d'identifier ce qu'on appelle communément la réponse impulsionnelle de l'objet étudié. En d'autres termes, comment un signal unitaire de durée infinitésimale (une impulsion) en entrée se transforme sur le signal de sortie : modification de son amplitude, effet de retard, étalement dans le temps… Bien évidemment s'il doit y avoir plusieurs entrées au système, la réponse impulsionnelle quantifie la transformation d'une impulsion sur chaque type d'entrée. Plusieurs méthodes conceptuelles peuvent être employées pour transformer (Transférer) E(t) en S(t). On peut projeter le signal sur une base vectorielle, le décomposer en fonctions monômes, le projeter dans un espace de fonctions de transformation (Laplace, Fourier, Hankel, …) etc. Dans le cas de TEMPO et de ce qui nous intéresse plus particulièrement, on procède par convolution. De principe, on pose une relation entre la sortie S(t) et l'entrée E(t) sous la forme :

$$S(t) = \int_0^t E(\tau)g(t-\tau)\,d\tau$$

Cette expression traduit que la sortie à l'instant t est l'addition (le cumul) des effets des entrées à des temps antérieurs ($E(\tau)$) pondérées par un opérateur de transformation $g(t - \tau)$. Cet opérateur module l'amplitude du signal E(t) en fonction du décalage en temps $t - \tau$. Ainsi une entrée à un temps τ_1 peut par exemple influencer la sortie au temps t ($t > \tau_1$) selon un facteur scalaire $g(t - \tau_1) = a$, alors que cette même entrée au temps τ_2 influencera la sortie au temps t ($t > \tau_2$) selon un facteur $g(t - \tau_2) = b$. Dans cet exemple trivial et en se limitant à deux termes on a alors $S(t) = a\times E(\tau_1) + b\times E(\tau_2)$. On peut aisément comprendre que g(t) est le cœur du problème de "transfert" entre l'entrée et la sortie. C'est en fait la réponse impulsionnelle du système. Pour les "érudits", si $E(t) = \delta(t)$ c'est à dire la fonction Kronecker (1 en t = 0, 0 ailleurs), δ étant une impulsion unitaire au temps 0 mais aussi l'invariant du produit de convolution :

$$S(t) = \int_0^t \delta(\tau)g(t-\tau)\,d\tau = g(t)$$

g(t) est bien la réponse du système à une impulsion en entrée. A partir de ce principe, on peut augmenter la complexité de l'approche. Par exemple dans les cas de systèmes à plusieurs entrées, on peut considérer la sortie comme la somme de produits de convolution :

$$S(t) = \sum_j \left(\int_0^t E_j(\tau) g_j(t-\tau) d\tau \right)$$

chaque fonction g_j étant la réponse à une impulsion sur l'entrée de type E_j. C'est l'option choisie par TEMPO pour les systèmes à plusieurs entrées. Il existe d'autres formes plus complexes comme les convolutions "multiples" :

$$S(t) = \iiint_{\tau_i} F_1(\tau_1) F_2(\tau_2 - \tau_1) .... F_i(\tau_i - \tau_{i-1}) .... F_n(\tau_n - \tau_{n-1}) d\tau_1 ... d\tau_n$$

où les fonctionnelles F_i sont elles-mêmes des produits entre les Entrées $E_i(t)$ et des fonctions élémentaires $g_i(t)$. Cependant, aussi élégant soit-il, ce noyau de convolutions multiples conduit le plus souvent à un système mal posé pour lequel on ne trouve pas de solution unique au jeu de fonctions $\{g_i(t)\}$.

Quelques variantes intéressantes s'avèrent bien souvent pertinentes pour les problèmes de modélisation des systèmes hydrologiques. La première de ces variantes est la convolution "auto-régressive". C'est un modèle où la sortie elle-même intervient comme fonction d'entrée (système en boucle, voir Figure 2). Il s'agit d'une manière élégante de formuler les problèmes dans lesquels on suppose que l'état du système à un instant t peut en partie être modifié par l'état dans lequel il était auparavant. Ainsi, calculer la saturation en eau d'un sol en fonction d'évènements pluvieux suppose que cette saturation intègre les teneurs en eau du sol (la saturation) aux temps antérieurs. La raison en est que les perméabilités pour un écoulement à deux phases eau – air dépendent des teneurs en eau et qu'au delà ou en deçà de certains seuils, des écoulement préférentiels peuvent se produire. On a alors une sortie S qui s'écrit sous la forme d'une fonctionnelle de E et de S: S(t) = F(E(t), S(t)) se traduisant dans la convolution par une expression du type :

$$S(t) = \int_0^t E(\tau) g(t-\tau) d\tau + \int_0^{t'} S(\tau) g'(t-\tau) d\tau; \quad (t' < t)$$

Dans le cas de modélisation hydrologique des relations pluie – piézométrie, pluie – débit de rivière, cette auto-régression est un moyen simple d'intégrer ce qu'il est commun et intuitif d'appeler l'état hydrique du système. En d'autres termes, un système "sec" ne répond pas de la même façon qu'un système "saturé".

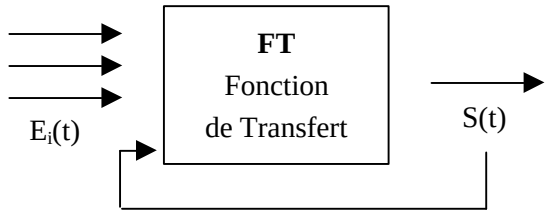

Figure 2 : Convolution avec calcul auto-régressif

Une autre variante intéressante est de proposer des systèmes convolués en séquence (Fig. 3) où la sortie $S_1(t)$ d'un modèle peut devenir une entrée d'un autre noyau de convolution pour la simulation d'une sortie $S_2(t)$:

$$S_1(t) = \int_0^t E_1(\tau) g_1(t-\tau) d\tau$$

$$S_2(t) = \int_0^t E_2(\tau) g_2(t-\tau) d\tau + \int_0^t S_1(\tau) g_2^{'}(t-\tau) d\tau$$

Bien évidemment, cette organisation séquentielle qu'elle intéresse la modélisation d'un système en cascades ou d'un système décomposable en entités plus petites (e.g., sous-bassin dans un grand bassin versant) doit s'appuyer sur quelques considérations donnant du sens aux séquences utilisées (notamment une corrélation établie entre $S_1(t)$ et $S_2(t)$). Notons que dans l'étude qui nous intéresse ici, ce mode séquentiel est employé selon le schéma suivant : 1- relation pluie → piézométrie de la nappe; 2 – relation (pluie + piézométrie de la nappe) → débit des rivières.

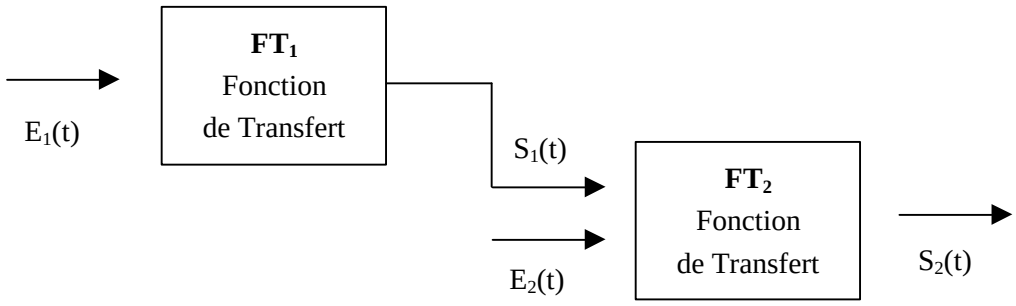

Figure 3 : Convolution avec calcul séquentiel

On notera que ces deux variantes : calcul auto-régressif, systèmes en séquence, sont implémentées sous TEMPO et utilisées par le BRGM pour l'étude des bassins versants du Poitou-Charentes

Quelques points clefs de la modélisation par approche systémique

On l'aura compris, pour simple qu'elle puisse paraître, l'approche systémique évacue cependant en grande partie la physique du problème posé et on pourrait tenter de simuler avec le même noyau de convolution une relation pluie – débit et une relation niveau d'ensoleillement – production primaire de phytoplancton lacustre ! Il s'agit bien évidemment d'un exemple peu orthodoxe dont la vocation est de poser clairement les choses et qu'il ne faut pas prendre pour autant au pied de la lettre. Cela étant, deux niveaux d'intervention permettent de tempérer cette assertion et par conséquent essayer de retrouver une connotation physique (ou pour le moins logique) à l'approche systémique.

La première intervention intéresse la fonction de transfert g(t). Il est possible si on l'exprime sous forme paramétrique (c'est à dire un modèle de fonction déclaré *a priori*) de donner du sens physique au module Fonction de Transfert (FT). Par exemple, on sait que dans les nappes, l'écoulement Darcéen se traduit par une équation locale de diffusion de pression fluide (ou de charge hydraulique). On peut alors envisager que chaque condition de forçage externe sur la nappe (e.g., les pluies) voit son signal

convolué par un noyau diffusif dont la fonction élémentaire est une distribution gaussienne des temps de transfert. Cette fonction pourrait alors s'écrire :

$$g(t) = \frac{1}{a}\exp(-b\,t^2)$$

avec a et b les deux paramètres scalaires de la fonction qu'on essaye d'optimiser afin que la convolution $S(t) = \int_0^t \frac{E(\tau)}{a}\exp\left(-b\left(t-\tau\right)^2\right)$ soit comparable à l'observé h(t) (piézométrie relevée sur un puits). L'avantage d'une fonction paramétrique (plutôt qu'une fonction discrète faite d'une succession de valeurs scalaires) est de ré-introduire du sens physique au transfert entre l'entrée et la sortie. Lorsque le choix est judicieux, on peut s'attendre à un meilleur ajustement du modèle et une capacité accrue de prédiction. En contre-partie ce choix préliminaire peut s'avérer non pertinent voire erroné. Ainsi, une fonction de transfert g(t) gaussienne ne peut pas s'appliquer pour un problème pluie → piézométrie si de forts écoulements hypodermiques se produisent dans les sols (au lieu d'une infiltration diffuse). La fonction g(t) devient alors généralement une Erf : $g(t) \propto \int \exp(-\tau^2)\,d\tau$ ou une fonction Gamma : $g(t) \propto \int \tau^{a-1}\exp\left(-\tau\right)d\tau$. En se trompant de fonction, il faudra ensuite "tordre" le modèle afin qu'il ajuste les données d'observation. Cet ajustement reste parfois réalisable mais dès qu'on relaxe les contraintes sur le modèle, notamment en le faisant travailler en contexte prédictif, on peut obtenir des comportements pour le moins anodins.

Le second point stratégique se situe à l'amont au niveau du module Compilateur d'Entrées (CE, Fig. 1). Bon nombre de problèmes supposent de retravailler les données brutes, qu'il s'agisse simplement d'analyser les caractéristiques de leur distribution statistique, leur auto-corrélation, ou plus généralement de recalculer des variables pertinentes pour le problème considéré. C'est généralement à ce niveau de modélisation que l'utilisateur peut au mieux "mettre les mains" dans le système et y insuffler son savoir-faire, son expérience du problème traité, voire une intuition. En jargon mathématique, le CE est le lieu du conditionnement *a priori* du modèle, par opposition au conditionnement a posteriori, ce dernier étant l'ajustement de la sortie simulée sur les données équivalentes d'observation. En théorie du problème inverse (en première simplification, l'ajustement des paramètres d'un modèle), ce conditionnement *a priori* est crucial pour les raisons suivantes :

- il permet de trier ce qu'il faut évacuer ou garder comme éléments pertinents de la modélisation;
- il permet d'argumenter autrement que par de simples considérations mathématiques les raisons du calcul;
- il donne la possibilité de se ménager plusieurs scénarios de traitement des données et par conséquent de tester des hypothèses (ou des intuitions);
- enfin en se plaçant dans un contexte stochastique, il permet d'analyser la sensibilité générale du modèle aux données d'entrées. Autrement dit, une petite perturbation sur les données d'entrée génère-t-elle une modification importante des paramètres d'ajustement ?

Plus spécifiquement dans le cas de l'étude considérée aujourd'hui, ce sont essentiellement les simulations prospectives qui exploitent et bénéficient des fonctionnalités du CE. Nous allons voir à ce propos que le <u>logiciel TEMPO est doté de tous les outils nécessaires à une bonne assimilation puis exploitation des données disponibles.</u>

Les principales fonctionnalités proposées par TEMPO

L'objectif n'est pas de passer en revue l'ensemble des capacités de traitement afférentes au logiciel. Il est juste question de mettre en avant quelques éléments généraux qui argumentent de la bonne adéquation entre l'outil de simulation et l'étude menée.

L'ensemble du logiciel est architecturé autour de quatre grands items que nous décrirons rapidement :

- les variables et traitements;
- les réponses impulsionnelles (noyaux de convolution);
- les méthodes de travail;
- les simulateurs d'entrées.

Notons avant même de discuter de ces items que le logiciel TEMPO intègre des concepts connus de longue date mais fait aussi montre d'innovations, ayant d'ailleurs conduit à la publications de plusieurs articles scientifiques dans de grandes revues d'audience internationale. Ces éléments novateurs sont pour l'essentiel :

- le calcul de réponses impulsionnelles sous contraintes permettant d'honorer le principe de conservation (de la masse pour un problème de transport de concentrations, des flux hydriques pour des problèmes d'écoulement);
- la régularisation des réponses impulsionnelles. Elle intéresse les fonctions g(t) non paramétriques, c'est à dire les fonctions discrètes définies comme une suite de valeurs scalaires pondérant les données d'entrée. Cette méthode permet de diminuer le nombre de degrés de liberté du problème d'optimisation (trouver la meilleure fonction g(t)). Les fonctions prennent la forme de trapèzes obéissant à des contraintes "géométriques" qui en première approximation les rendent relativement "lisses" (régulières, régularisées) pour que la sortie S(t) convoluée ne soit pas erratique. Cette régularisation permet d'obtenir des solutions à des problèmes sous-contraints;
- enfin, on dispose de simulateurs performants de chroniques de pluies qui intègrent bien évidemment la distribution statistique des données mais aussi qui permettent de respecter des contraintes de corrélation temporelles (observées ou définies *a priori*) entre pluie et Evapo-Transpiration Potentielle (ETP), entre pluie et température au sol. Ces corrélations sont honorées en arrangeant par permutation les évènements pluies d'une chronique jusqu'à ce que leur structure (évolution dans le temps) obéisse à certaines lois. Il est affirmé que préserver ces corrélations est un moyen d'intégrer dans une simulation prospective les effets de dérive climatique, voire les corrélations spatiales pluie – ETP sur les grands bassins. Selon nous c'est certainement vrai et utile pour des analyses demandant une simulation sur de longues périodes (de l'ordre du siècle) ou des bassins du gabarit de celui de grands fleuves. Dans le cas de l'étude sur les bassins versants de Poitou-Charentes et l'établissement de simulations prospectives sur un an, il est vraisemblable que ces corrélations, si elles existent, aient finalement assez peu d'influence sur les résultats.

Au chapitre "Variables et Traitements", on retiendra que TEMPO dispose de tout l'arsenal permettant de faire de l'analyse corrélatoire et spectrale sur les chroniques des données d'entrée. On peut donc regarder comment une entrée E(t) a ses valeurs locales (en temps) $E(t_1)$, $E(t_2)$,… $E(t_n)$ corrélées entre

elles en fonction des écarts $t_1 - t_2$, …, $t_1 - t_n$. On peut aussi croiser cette analyse de corrélation entre deux types d'entrées différentes $E_1(t)$, $E_2(t)$. Mathématiquement, on définit l'auto-covariance d'une entrée et la covariance croisée de deux entrées par :

$$C_{E-E}(\tau)=\langle E(t)E(t+\tau)\rangle \ ; \ C_{E_1-E_2}(\tau)=\langle E_1(t)E_2(t+\tau)\rangle$$

avec $\langle\ \rangle$, l'opérateur Espérance Mathématique (prise de moyenne). C'est la forme de ce type de fonction qui permet d'affirmer si oui ou non les entrées sont temporellement corrélées et sur quelle période de temps τ. Ce sont ces fonctions qui sont utilisées comme éléments de conditionnement des chroniques simulées de pluies utilisées à des fins prospectives. L'analyse spectrale repose sur le même principe mais calcule la corrélation sur la transformée de Fourier de l'entrée. Le passage en domaine spectral (fréquences) permet de mieux apprécier la périodicité éventuelle des données, les effets temps courts (hautes fréquences) et les effets temps longs (basses fréquences).

Concernant les "Réponses Impulsionnelles", le logiciel propose une batterie de fonctions paramétriques sélectionnées en fonction des problèmes posés. Pour la simulation de transferts de masse, ces fonctions sont les densités de probabilités des temps d'arrivée d'un traceur issues de solutions analytiques simples. Il ressort que le sens physique inhérent au transport (e.g., par convection dispersion) est intégré à la réponse g(t). De la même manière, la fonction g(t) peut prendre des formes gaussiennes ou exponentielles décroissantes dans des relations pluie → piézométrie ou (pluie + piézométrie)→ débit. En ce sens, les fonctions assurent une cohérence physique d'écoulement dans un réservoir (pour la gaussienne) ou de vidange (pour l'exponentielle). Bien que ce ne soit pas très clair à la lecture du manuel d'utilisation, TEMPO semble aussi disposer de fonctions non paramétriques. On peut par exemple penser que ce genre de fonction est utilisé sur la réponse "temps courts" dans la simulation des débits des rivières pour la convolution de l'entrée "niveau piézométrique" (voir nombreuses figures du rapport BRGM RP-54569 FR-V2). Enfin, la convolution avec calcul auto-régressif (voir plus haut), fait partie des options possibles de modélisation sous TEMPO. Comme nous l'avons déjà souligné, il s'agit d'un "plus" important permettant de considérer que le système n'est pas stationnaire (invariant au cours du temps) mais peut évoluer en fonction des conditions générales dans lesquelles il se trouve. Ces conditions sont matérialisées par (ou contractées sur) les valeurs de sortie S(t).

Au chapitre des "Méthodes de Travail", il faut finalement comprendre que les trois grands domaines d'application de TEMPO sont :

- les modèles de mélange, c'est à dire la simulation du transfert de masse dans les systèmes hydrologiques. C'est dans ce cadre que sont utilisées les réponses impulsionnelles paramétriques basées sur des solutions analytiques simples du transport;
- les modèles inverses. On peut d'ailleurs largement discuter du terme "inverse" pris ici dans un sens assez restrictif. Au sens général, inverser c'est identifier les paramètres du modèle par comparaison de ses résultats avec des données d'observation (ou un *a priori* sur ces derniers). Dans le cas présent, TEMPO restreint la définition du problème inverse à l'identification des paramètres des noyaux de convolution utilisés pour les simulations de débit des rivières et de piézométrie à partir des entrées de pluies. C'est en l'occurrence, sous ce registre que les méthodes de régularisation de la réponse impulsionnelle g(t) sont utilisées. Attendu qu'on utilise très souvent un calcul auto-régressif pour ce genre de simulation, les réponses

impulsionnelles g(t) sont alors généralement peu étalées en temps donc peu contraintes et justifient pleinement du "lissage" proposé par la régularisation;

- enfin la décomposition d'hydrogrammes. C'est en particulier dans ces approches que des fonctions paramétriques g(t) répondant à la physique d'un écoulement ou d'une vidange de réservoir sont utilisées. Ceci permet alors de distinguer dans les débits d'exutoires : - la fraction d'eau "pré-évènementielle" correspondant par exemple à l'alimentation permanente d'une rivière par une nappe ; - la fraction "évènementielle" correspondant par exemple à l'arrivée du ruissellement dans la rivière à la suite d'un épisode pluvieux. On voit clairement (et TEMPO est bien construit de ce point de vue) l'intérêt de redonner du sens physique aux convolutions via des noyaux paramétriques.

Le dernier grand volet des capacités de TEMPO concerne la "Simulation d'Entrées". On aura compris l'intérêt de tels outils dans la prospective et la simulation sur un modèle "éduqué" de scénarios fictifs. Comme cela a déjà été consigné plus haut, TEMPO fait preuve d'innovation dans le domaine de la simulation de chroniques de pluies en conditionnant ces dernières tant sur une distribution statistique que sur des lois de corrélation pluie – ETP, pluie – température. Cette corrélation est acquise en permutant les séries temporelles aléatoirement simulées jusqu'à ce qu'elles présentent la bonne fonction de corrélation. Supposons qu'une série P(t) ait un écart (quadratique) de corrélation ε avec la fonction que l'on cherche à atteindre. On permute la série P pour obtenir P'(t). Si l'écart ε' est inférieur à ε, la permutation est "bonne" puisqu'elle améliore l'objectif. A l'inverse si ε' est supérieur à ε la permutation est "néfaste". Il n'est pas dit cependant qu'une permutation néfaste ne sera pas quelques itérations plus loin la clef d'une autre permutation quant-à-elle extrêmement bénéfique. Préjugeant de cette éventualité, on peut accepter la permutation P(t) → P'(t) sous couvert d'une probabilité en loi de Boltzmann :

$$\mathrm{Prob}_{\mathrm{Accept}} \propto \exp\left(-\frac{\varepsilon'-\varepsilon}{T}\right)$$

avec T une "température de descente". En début de processus, fixant des températures élevées la probabilité d'acceptation est quasi égale à un. On accepte donc de nombreuses permutations "néfastes". En d'autres termes, on autorise la méthode à "divaguer" pour chercher la meilleure solution. En fin de processus, on est assez proche de la solution et fixant des températures basses, la probabilité d'acceptation d'une permutation néfaste tend vers zéro. Autrement dit, une fois placé sur la bonne voie, on évite par trop de s'en écarter. Hormis la relative lenteur de ce type d'algorithme (encore appelé le recuit simulé), la méthode est versatile, capable de conditionner à peu près tout ce que l'on veut. En l'occurrence, le principe de permutation sur série chronologique est opportun. Il structure en temps la séquence des évènements pluies sans modifier pour autant la distribution statistique des valeurs scalaires prises par les pluies.

En conclusion

Après avoir discuté des grands schémas et des attendus de la modélisation par approche systémique, puis avoir mis en parallèle ces attendus et les capacités de réalisation de TEMPO, on ne peut que <u>conclure à l'excellence du logiciel</u>. Bien évidemment, cet avis ne peut préjuger d'éventuelles erreurs de

programmation et/ou difficultés d'utilisation. A notre avis, il est cependant peu probable que TEMPO souffre de lacunes rédhibitoires. Dès lors, l'approche systémique qui, nous l'avons dit, s'avère judicieuse s'il est question de fournir des indicateurs de gestion ou d'analyser le comportement global de systèmes hydrologiques est parfaitement relayée par les fonctionnalités de TEMPO. Rappelons aussi que certaines méthodes ont été publiées dans de grands journaux scientifiques, ces derniers mettant au plus haut rang de leurs préoccupations le fait que la pertinence scientifique soit réelle. In fine, et pour l'étude qui nous concerne ici, TEMPO n'est l'objet d'aucune critique tant sur le plan conceptuel que numérique. C'est le modèle qui répond aux attentes de l'étude mandatée.

Partie II : Analyse de l'étude du BRGM

La seconde partie de l'expertise regarde et commente le rapport BRGM RP-54569 FR-V2 appliquant la modélisation par approche systémique (cf. Partie I) aux 5 bassins versants suivants : Dive Nord, Clain, Boutonne, Sèvre-Niortaise et Seudre.

Un commentaire préliminaire s'impose. Le rapport fait la démonstration des capacités du BRGM à compiler, synthétiser et modéliser une quantité très importante d'information concernant l'hydrologie en Poitou-Charentes. Dans son ensemble, le travail est bien construit, repose sur une connaissance enviable des systèmes hydrologiques régionaux et valorise dûment celle-ci. De fait, il n'y a pas de raison particulière de remettre en cause le bien fondé des choix effectués comme :

- le découpage en sous-bassins en fonction du contexte géologique, hydrologique et hydrogéologique;
- la sélection parmi les piézomètres disponibles des ouvrages les plus pertinents.

On a pu aussi apprécier, à la lecture du document, la connaissance approfondie du terrain et des spécificités de chaque système hydrologique. Enfin, l'ensemble des résultats est présenté sous une forme abondamment illustrée qui pour être synthétique peut parfois manquer de précision. C'est plus particulièrement sur ce point que notre avis est formulé, l'essentiel de nos commentaires cherchant à éclaircir les modalités d'utilisation de TEMPO pour l'étude pratique considérée.

Commentaires généraux

Il est apparu difficile en première lecture de comprendre l'imbrication entre les études de type pluie → piézométrie et (pluie + piézométrie) → débit des rivières. On pourrait parfaitement dissocier les deux études attendu que l'analyse pluie → piézométrie cherche des indicateurs du stock d'eau souterraine tandis que ce même indicateur servira ultérieurement de contrainte (de fonction d'entrée) à l'analyse du débit des rivières. En posant clairement le problème, on aurait pu éviter quelques propos prêtant à confusion, par exemple, pourquoi utiliser le niveau piézométrique d'une nappe captive dans la modélisation des débits d'une rivière *a priori* non connectée directement à l'aquifère. Si on comprend que ce niveau piézométrique pour l'aquifère considéré est un bon indicateur de la dynamique de l'ensemble du système "eaux souterraines", toutes nappes confondues, alors utiliser cet indicateur

n’est pas incompatible avec la modélisation systémique des eaux de surface quand bien même il proviendrait d’une nappe captive.

On a vu dans la Partie I de l’expertise que la modélisation de systèmes hydrologiques procédant par convolution entrée → sortie, pouvait s’appuyer sur plusieurs types de noyaux de convolution. Certains d’entre eux (les fonctions paramétriques) sont assortis d’une connotation à sens physique, d’autres (fonctions discrètes) sont de simples opérateurs mathématiques. Il aurait été judicieux de préciser quels types de fonctions sont utilisés selon le problème posé (par exemple, relation nappe – rivière "temps courts", relation nappe – rivière "temps longs", relation pluie – débit). En inspectant les différentes figures proposées dans le rapport, il semblerait que la réponse impulsionnelle nappe – rivière "temps courts" soit une fonction scalaire, la réponse impulsionnelle "temps longs" une fonction gaussienne ou loi Gamma. Il est impossible de se faire une idée de la réponse impulsionnelle pluie-ruissellement.

En précisant le type de noyau utilisé, il serait plus facile d’argumenter et d’asseoir la validité des résultats proposés. Avoir une bonne connaissance de l’amplitude et de la forme de la réponse impulsionnelle "ruissellement" permettrait de justifier l’écrêtage quasi-systématique par les seuils de débordement.

De la même façon, on peut s’interroger du pas de temps décadaire utilisé sur les données d’entrée alors que la réponse impulsionnelle "temps courts" semble être de forme triangulaire (trois points représentatifs). Est-ce un artefact lié au pas décadaire ou une forme justifiée notamment en conduisant quelques tests sur un pas plus petit ?

Le rapport exploite à bon escient les capacités de TEMPO à générer des chroniques de pluie bien conditionnées (cf. Partie I). Par conséquent, on peut supposer que les fonctions d’entrée dans les scénarios de simulations cherchant à valider/invalider les DOE / DCR sont des problèmes bien posés. Cependant, on ignore complètement à quoi correspond une année sèche, versus humide, dans ces scénarios. Selon nous, une année est sèche dès que la moyenne des pluies est inférieure à la moyenne des données. Une année est humide dès que la moyenne des pluies est supérieure à la moyenne des données. Une autre définition est toutefois possible. On peut simuler n chroniques équiprobables de pluies, classer ces chroniques en fonction de leur moyenne respective puis considérer par exemple qu’une simulation de p années sèches correspond au p-quantile de la distribution des moyennes de l’ensemble des chroniques. Ce genre de définition semble être celui adopté par TEMPO (sous couvert d’une bonne lecture de notre part). Parce que ces simulations conditionnent l’essentiel des résultats attendus de l’étude, il semble opportun de clairement préciser les définitions et les valeurs afférentes aux notions année sèche / humide.

On aurait souhaité, en marge des calculs réalisés, que soit fourni un bilan approximatif chiffré des contributions de chacun des compartiments : contribution eau souterraine "rapide", contribution eau souterraine "lente" et ruissellement. Ce bilan est accessible quasi-directement via l'analyse des réponses impulsionnelles.

Commentaires spécifiques

Se rattachant à la relation pluie → piézométrie, on relève, par exemple p.40 du rapport, des assertions stipulant que les meilleures corrélations possibles entre la pluie et la piézométrie sont obtenues avec les stations pluviométriques de Loudun, Marigny, Nuaillé et l'ETP de Poitiers ! Il semble qu'un argumentaire circonstancié serait nécessaire sauf à considérer que toute information locale dans une approche systémique perd ledit caractère local au profit du conditionnement global du système. Dans ces conditions, il n'est pas usurpé d'utiliser des informations d'origine géographique différente sous couvert qu'elles présentent de bonnes cohérences statistiques. Ce point aurait dû être abordé et présenté dans le rapport (au moins pour quelques exemples).

A propos des chroniques piézométriques, il manque des définitions claires et précises des seuils de débordements, des niveaux de base, et l'argumentaire des fonctions de prélèvements ajoutés aux chroniques pour moduler l'amplitude du signal piézométrique. La même remarque s'applique à la notion de tendance (page 112, Fig. 96). Par exemple, le niveau de débordement correspond-t-il au niveau topographique ou à une ligne de sources pour un aquifère libre, à la cote topographique d'une tête de tubage pour un aquifère captif ?… Avec une définition clairement exprimée, on aurait pu éviter par exemple, p.41-42, d'avoir pour un aquifère libre, un niveau de débordement à 97 mètres correspondant au lit de la rivière. Est-ce à dire qu'en environnement climatique tempéré, les nappes libres se trouvent en dessous du niveau de la rivière ? La réponse est vraisemblablement non, sauf configuration spécifique locale. Ceci n'est qu'un exemple, d'autres existent ailleurs. Il serait presque souhaitable que les seuils de débordements inclus dans la modélisation soient assortis au coup par coup d'une justification contextuelle y compris par des schémas précisant la géométrie locale et les relations nappe-rivière.

A l'usage, il apparaît peu de différences entre les temps modaux des réponses impulsionnelles "temps courts" et des réponses impulsionnelles "temps longs" (une dizaine de jours pour les temps courts, 20-60 jours pour les temps longs). Cette différence est-elle suffisante pour physiquement individualiser des écoulements karstiques du drainage général du réservoir ? On a pu vérifier que TEMPO manipule explicitement deux composantes : une rapide, l'autre lente pour les relations piézométrie – débit. Est-il judicieux ensuite d'invoquer du karst versus du drainage alors que la décomposition temps courts / temps longs est la conséquence nécessaire du mode d'analyse du logiciel TEMPO ? Cela étant, la modélisation reste valide en termes de simulation – prédiction mais pas forcément en termes de mécanismes d'écoulements.

Finalement, et concernant les tableaux récapitulatifs des volumes disponibles, comment doit-on interpréter l'ajout d'1/2 AEP alors que les prélèvements AEP sont implicitement pris en compte dans le noyau de convolution ? Est-ce parce qu'il est prévu à l'horizon 2011 une augmentation significative des prélèvements AEP ?

En conclusion

On rappelle que l'essentiel des remarques formulées dans cette partie de l'expertise ne s'attache ni à la méthodologie portée par TEMPO, ni aux fondements de son application à la recherche d'indicateurs piézométriques sur les bassins versants de Poitou-Charentes. Elles discutent de manière critique des informations complémentaires permettant d'améliorer la lecture et la compréhension générale des travaux réalisés. Dans l'état actuel des connaissances et des travaux consignés au rapport, on peut considérer que les attendus du mandataire sont honorés. L'analyse des systèmes hydrologiques proposée par le BRGM est une bonne base de réflexion-discussion pour l'élaboration des stratégies futures de gestion de la ressource en eau. Cela dit, l'outil utilisé nous fournit des tendances générales qu'il ne faudrait surtout pas prendre au premier degré comme des contraintes ou des résultats incontournables devant être respectés. En d'autres termes, une gestion pointue de la ressource sur un bassin supposerait, en guise de test opérationnel, de reconduire ce type de modélisation à une échelle temporelle (voire spatiale) plus fine (pas de temps journalier, sous-bassin de l'ordre de 50 km²).

Dans cette perspective, on pourrait comparer sur ce bassin test, les résultats obtenus par TEMPO et son approche systémique à une modélisation physique (hydrodynamique). Dans cette dernière approche, on sait "exactement" où intervenir sur les paramètres du modèle afin d'influencer l'amplitude et la dynamique transitoire des flux hydriques du bassin.

In fine, l'approche systémique reste un excellent outil de gestion de mise en œuvre autrement plus aisée que celle d'un modèle hydrodynamique. Il faut par conséquent poursuivre les travaux et les efforts consentis à propos de cette méthode pour la gestion future de la ressource en eau.

Poitiers, le 10 janvier 2007

Frederick DELAY Alain DUPUY Gilles POREL