Journal of Contaminant Hydrology 100 (2008) 22–29

Contents lists available at ScienceDirect

Journal of Contaminant Hydrology

journal homepage: www.elsevier.com/locate/jconhyd

Stationary and non-stationary autoregressive processes with external inputs for predicting trends in water quality

Jean Louis Pinault *, Igor G. Dubus

BRGM, Water Division, 3 avenue Claude Guillemin, BP 36009, 45060 Orléans Cedex 2, France

A R T I C L E   I N F O

Article history:
Received 12 February 2007
Received in revised form 2 May 2008
Accepted 2 May 2008
Available online 16 May 2008

Keywords:
Autoregressive process
Transfer process
Decay process
Atrazine
Groundwater
Trends in water quality

A B S T R A C T

An autoregressive approach for the prediction of water quality trends in systems subject to varying meteorological conditions and short observation periods is discussed. Under these conditions, the dynamics of the system can be reliably forecast, provided their internal processes are understood and characterized independently of the external inputs. A methodology based on stationary and non-stationary autoregressive processes with external inputs (ARX) is proposed to assess and predict trends in hydrosystems which are at risk of contamination by organic and inorganic pollutants, such as pesticides or nutrients. The procedures are exemplified for the transport of atrazine and its main metabolite deethylatrazine in a small agricultural catchment in France. The approach is expected to be of particular value to assess current and future trends in water quality as part of the European Water Framework Directive and Groundwater Directives.

1. Introduction

The assessment and the prediction of trends in water quality is often a challenge when external inputs such as meteorological conditions influence the dissipation and transfer processes of pollutants, especially in situations where the observation period is short in comparison with the response time of the hydrosystems. This paper reports on the implementation of mathematical models also referred to as the autoregressive process with external inputs (ARX) that are devoted to transfer process characterization, prediction and control (Karacan, 2003). The method under consideration can be considered as a particular case of the widespread multivariate autoregressive analysis (from the Box and Jenkins approach, 1970; Berthouex et al., 1978) when the “feedback” is not observable, i.e. when the terms representing the influence of the output on the inputs are zero (Box and Jenkins, 1970).

New developments on stochastic recursions emphasize the remarkable stability of short- and long-term forecasting from autoregressive processes. Stochastic recursions appear in queuing theory (Baccelli and Brémaud, 2003), random walks in random environment or branching processes (Altman and Fiems, 2006) and the theory is being studied actively (Asmussen, 2003; de Saporta et al., 2004; de Saporta and Yao, 2005; Guivarc'h, 2006). The advantages of autoregressive processes within the context of the present study are that: i) the so-called ‘tail’ of the process, i.e. the extrapolation of the recursive process beyond the observation period, may be extended without an increase in random and systematic errors as time increases; and, ii) the calibration of the autoregressive process may be performed accurately using a time period which is short in comparison with the response time of the systems under study. This latter aspect distinguishes autoregressive processes from Kalman filtering techniques and neural network modelling.

In the present study, a “true” transfer model is presented which reports the results obtained using the ARX approach to assess and predict trends in the contamination of groundwater and surface water by inorganic or organic pollutants from monitoring data. The fact that these compounds are typically subject to chemically- or biologically-driven degradation processes means that fluxes of micropollutants in hydrological systems are typically non-stationary. Methodological developments are proposed to address the non-stationary nature of these processes and the procedures are applied to the investigation of

* Corresponding author.
E-mail addresses: jeanlouis_pinault@hotmail.fr (J.L. Pinault), i.dubus@brgm.fr (I.G. Dubus).

0169-7722/$ – see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.jconhyd.2008.05.001

trends in pesticide fluxes measured in spring water at the outlet of a small agricultural catchment in France. Two aspects of the ARX approach are investigated to characterize the natural processes of transfer and degradation of pesticides. First, the autoregressive approach is used for long-term forecasting while the initial conditions of the system have no longer influence on pesticide fluxes. Secondly, constrains are introduced into the model to account for mass conservation and the autoregressive term is weighted to take into account decay processes. The approach is demonstrated for concentrations of atrazine and its main metabolite deethylatrazine (DEA) in groundwater.

2. Materials and methods

Crop protection products are known to represent a potential risk for human and the environment (e.g. Birarder and Rayburn, 1995) and the presence of pesticides is therefore routinely monitored in environmental media such as groundwater, surface water (Gilliom et al., 2006) and, to a lesser extent, the atmosphere (Dubus et al., 2000; Shen et al., 2004; Yao et al., 2006). Most surveillance efforts are targeted towards investigating the spatial spread of the contamination of water resources by pesticides and/or detecting any positive or negative trends in concentrations (Gilliom et al., 2006). Monitoring programmes have demonstrated that the contamination of water resources by atrazine and its metabolites tends to be widespread (Thurman et al., 1991; Ritter et al., 1994; Tindall and Vencill, 1995; Goolsby et al., 2001; Kolpin et al., 1998; Clark et al., 1999; IFEN, 2006).

2.1. The context

The autoregressive approach is demonstrated for pesticide concentrations and fluxes measured in spring water for 7 years at an experimental site ca. 75 km north-west of Paris, France (Mouvet el al., 2004; Baran et al., 2005; Roulier et al., 2006; Morvan et al., 2006). The site is of particular interest because the use of atrazine on the catchment was discontinued in 1999 following concentrations of atrazine and its metabolite deethylatrazine (DEA) exceeding regulatory thresholds for drinking water. Concentrations of atrazine, DEA and major cations and anions have been monitored in the Brévilles spring at a two-week interval from December 2000 to June 2004 and on a monthly sampling rate since (Fig. 1).

2.2. Transfer and degradation processes

The release of atrazine and DEA that is observed after 1999 while atrazine was no longer used on this catchment discloses the persistence of those pollutants (Fig. 1). Pesticides that are retained into the pores of the root zone are subject to degradation. The dissolved part of pesticides migrates through the vadose zone to the watertable. During wet years, atrazine and DEA are subject to quick transfer as a result of preferential flow paths in the limestone of both the unsaturated and the saturated zones. Subsequently, quick transfer through the macropores of the vadose zone is enhanced when the capillary pressure is high enough to establish a hydraulic continuity in preferential pathways.

2.3. Autoregressive process with external inputs

The autoregressive process with external inputs (ARX) may be regarded as a mathematical representation of the hydrosystem. It therefore gives more of a functional relationship which may or may not agree with other models obtained by utilizing the knowledge of the physical, chemical and biological mechanisms of the system.

The general expression for an ARX model can be written as:

$$F^{out}(t_i) = (1-\omega)\left[\lambda_1 \Gamma_1 \cdot \mathbf{F}_1^{in} + \cdots + \lambda_q \Gamma_q \cdot \mathbf{F}_q^{in}\right] + \omega \Gamma_{out} \cdot \mathbf{F}^{out} + \varepsilon(t_i) \quad (1)$$

q is the number of external inputs

$\mathbf{F}_k^{in} = (F_k^{in}(t_i), F_k^{in}(t_{i-1}), \cdots)^T$, $\mathbf{F}^{out} = (F^{out}(t_{i-1}), F^{out}(t_{i-2}), \cdots)^T$, are $p\times 1$ column vectors of observations at time t_i representing the inputs and the output of the transfer model (p is the order)

$\mathbf{\Gamma}_k$ and $\mathbf{\Gamma}_{out}$ are $1\times p$ line vectors, the so-called impulse responses

ε is the noise of the transfer model which represents erratic, complex, and usually short-term variability of the output which is not explained by the model whatever the underlying reasons might be (e.g. lack of or inadequate description of key processes, measurement error, inadequate sampling or modelling strategy).

In the actual case which concerns us additional conditions are required since fluxes are positive functions of time constrained by the mass conservation of the system.

series F_1^{in}, ..., F_q^{in} and F^{out} are assumed to be reduced to unit average

$\mathbf{\Gamma}_k$ and $\mathbf{\Gamma}_{out}$ are unit impulse responses (area is unity) with the constraint $\Gamma(t_i)\geq 0$, $k=1,\cdots,p$

λ_k, $k=1, \cdots, q$ and ω are weighting factors such as $\lambda_1+\cdots+\lambda_q=1$

The weighting factor ω and the impulse response $\mathbf{\Gamma}_{out}$ express the 'internal' functioning of the hydrosystem.

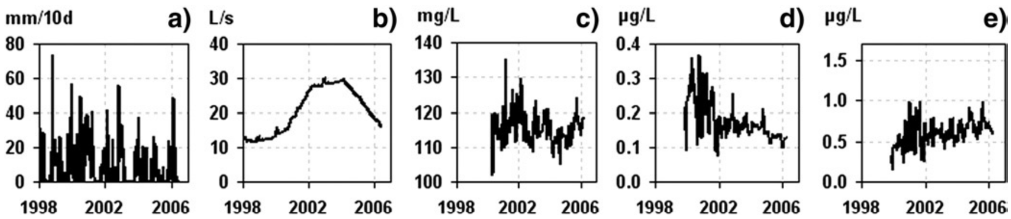

Fig. 1. Variables used in the modelling — a) effective rainfall — b) water flow at the spring — c) Ca concentrations in spring water — d) atrazine concentrations in spring water — e) deethylatrazine concentrations in spring water.

The transfer model (1) supposes the stationarity of the inputs (pesticide spraying, meteorological conditions) and the output (pesticide fluxes). Moreover, a transfer process is supposed to be ergodic, which means that statistical sampling can be performed at one instant across a group of identical processes or sampled over time on a single process with no change in the measured result.

2.4. Modelling procedure

Optimal estimates of the parameters, i.e. the calculation of the weighting factors and impulse responses, require inversion and regularization to smooth impulse responses as much as possible by minimizing the functional (Pinault et al., 2001a,b, 2005; Pinault and Schomburgk, 2006).

$$\Phi = \ln(r) + ||\Gamma_1||_2 + \cdots + ||\Gamma_q||_2 + ||\Gamma_{out}||_2 \quad (2)$$

where $||\circ||_2$ represents the l_2 norm,

$r = \sum_{j=1,N}\left(\hat{F}^{out}(t_j) - F^{out}(t_j)\right)^2$ is the residual;

$F^{out}(t_j)$ is the observed output at time t_j;

and $\hat{F}^{out}(t_j)$ is its estimator defined so that:

$$\hat{F}^{out}(t_{n+1}) = (1-\omega)\left[\sum_{j=1,q}\lambda_j \sum_{i=0,p}\Gamma_j(t_i).F_j^{in}(t_{n+1-i})\right] + \omega\sum_{i=0,p}\Gamma_{out}(t_{i+1}).F^{out}(t_{n-i}) \quad (3)$$

where n is the length of the series and p is the order of the autoregressive process.

Minimization of (2) consists in selecting solutions such that a compromise is found that reduces the residual without excessively increasing the solution norms, i.e. ensuring that the optimum order p_{opt} is the smallest as possible. The residual increases significantly when the order decreases from p_{opt} and the solution norms increase as the order increases from p_{opt} without reducing the residual significantly. In such conditions, the tail of the autoregressive process $(p,\omega,\lambda_1,\cdots,\ \Gamma_1,\cdots,\ \Gamma_{out})$ defined from Eq. (3) – in which the observed output $\mathbf{F}^{out}$ is replaced by its estimator $\hat{\mathbf{F}}^{out}$– confers on the model a reliable long-term predictive behavior. The reversibility of the linear system (1) is ascertained as long as the inputs are sufficiently uncorrelated, i.e. the absolute values of the cross-correlograms of the inputs considered two by two are significantly lower than 1 for small lags (typically <0.85).

2.5. The stability conditions of the tail

In order to specify the stability conditions of the tail, let us assume the error $|\varepsilon_k|$ associated with the prediction $\hat{F}^{out}(t_k)$ is made up of: 1) a positive constant ε that is independent of the index k and 2) a term which reflects the propagation of errors:

$$|\varepsilon_{k+1}| \le \varepsilon + \omega|\varepsilon_k|. \quad (4)$$

The recursive relationship (4) implies that the existence of an upper limit for the error, which is expressed as:

$$\omega \le \omega_{max} < 1. \quad (5)$$

In practical terms, the value $\omega_{max}=0.95$ typically ensures the numerical stability of the system.

Eq. (1) can be extended to decay processes:

$$F^{out}(t_i) = \eta^i(1-\omega)\left[\lambda_1\Gamma_1\cdot\mathbf{F}_1^{in} + \cdots + \lambda_q\Gamma_q\cdot\mathbf{F}_q^{in}\right] + \omega\Gamma_{out}\cdot F^{out} + \varepsilon \quad (6)$$

in which the decay factor η raised to the power i is such that $0<\eta<1$($\eta=1$ corresponds to stationary processes). The decaying influence is included implicitly in the output F^{out} which is defined recursively. When transient production processes occur, the term η^i should be replaced by a function $f(t_i)>0$ that converges asymptotically towards $\eta^i(0<\eta<1)$ while i increases.

Now, the statistical properties of the decay process defined from Eq. (6) vary over time and the random process has to be defined compared with a time origin.

The propagation of errors in Eq. (6) is subject to the following recursive relationship:

$$|\varepsilon_{k+1}| \le \varepsilon + \eta\omega|\varepsilon_k| \quad (7)$$

The errors as given by Eq. (7) are bounded since $\eta<1$ so long as $\omega\le 1$.

In the following applications devoted to the simulation of the trends of contaminant fluxes, our interest lies in the tails of both the stationary (1) and decay (6) models. A key problem to solve the system under study lies in the determination of the order of the autoregressive process through the minimization of the discrepancies between the observations and the tail, i.e. the capability to undertake long-term prediction of fluxes. Optimal estimates of the parameters including the order are obtained by minimizing the functional Φ in Eq. (2) where the residual r is estimated from the tail, i.e. the extrapolation of the estimator $\hat{F}^{out}(t_j)$.

3. Results and discussion

The transfer model (ARX) uses the idea of state space technique, in which we treat all the inputs and outputs as the states of the system. Since atrazine was no longer used at the start of the monitoring period, the compound and its metabolite were simulated as originating from a stock in the soil/unsaturated zone. Contaminants at the site are believed to be leached to groundwater through a soil-unsaturated zone fractured limestone sequence, and then transported towards the spring as a result of effective rainfall infiltration.

The effective rainfall, i.e. the part of rainfall that contributes to spring flow (Fig. 1a) was computed from the Buhy meteorological station located 2 km from the study site. The years 1999, 2000 and 2001 were particularly wet while the following years were significantly dryer, which is reflected in the flow of the spring (Fig. 1b) that is exclusively fed by groundwater. The response time of the groundwater system (the elapsed time required for the system to recover its initial state after a short perturbance) is long due to the slow and delayed infiltration of water and solutes through the vadose zone which is characterized by a wide distribution of micro- and macropore sizes. Rainfall patterns are reflected in the sharp increase of the spring flow in 2002 and 2003 and its subsequent slow decrease. The time series of calcium concentrations in spring water (Fig. 1c) shows that dissolution and transport of calcite was enhanced during wet years. Atrazine displayed a similar transport behavior (Fig. 1d) whereas concentrations of

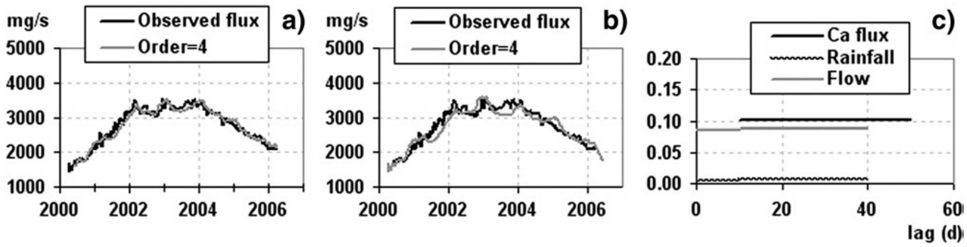

Fig. 2. Modelling of the Ca flux at the spring — a) comparison between the observed flux and the autoregressive process (8) for a forecast at one step ahead (10 days) — b) comparison between the observed flux and the autoregressive process (8) for long-term prediction (the autoregressive process is initialized at the beginning of the observed flux) — c) impulse responses, Γ_{out}, Γ_R and Γ_Q with areas proportional to the contribution of the corresponding variables to Ca flux, i.e. 0.51, 0.03 and 0.46, respectively: order = 4, η = 1.

DEA appear to increase over time although it should be noted that the signal is extremely noisy (Fig. 1e).

The stationary Eq. (1) and the non-stationary Eq. (6) are re-written so that:

$$\hat{F}^{out}(t_i) = (1-\omega)\left[\lambda_R \boldsymbol{\Gamma}_R \cdot \mathbf{R}_{eff} + \lambda_Q \boldsymbol{\Gamma}_Q \cdot \mathbf{Q}\right] + \omega \boldsymbol{\Gamma}_{out} \cdot \mathbf{F}^{out} \tag{8}$$

or

$$\hat{F}^{out}(t_i) = \eta^i (1-\omega)\left[\lambda_R \boldsymbol{\Gamma}_R \cdot \mathbf{R}_{eff} + \lambda_Q \boldsymbol{\Gamma}_Q \cdot \mathbf{Q}\right] + \omega \boldsymbol{\Gamma}_{out} \cdot \mathbf{F}^{out} \tag{9}$$

where $\mathbf{R}_{eff}$ is the effective rainfall, $\mathbf{Q}$ is the spring flow, all the positive functions $\mathbf{R}_{eff}$, $\mathbf{Q}$ and $\mathbf{F}^{out}$ being reduced to unit average.

Given the shortness of the impulse responses, the quick transfer of water and solutes to the spring is assumed to be represented by the transfer of the effective rainfall to the outlet $\boldsymbol{\Gamma}_R \cdot \mathbf{R}_{eff}$ whereas the slow and delayed transfer is considered to be represented by the spring flow $\boldsymbol{\Gamma}_Q \cdot \mathbf{Q}$ which reflects infiltration processes through the vadose zone.

The tails associated with both processes are defined from recursive relationships obtained by replacing $\mathbf{F}^{out}$ by the estimator $\hat{\mathbf{F}}^{out}$ in the second part of Eqs. (8) and (9).

3.1. Modelling of calcium flux at the spring

The presence of calcium in groundwater mainly results from the dissolution of calcite in the vadose zone (calcium is a pre-event water tracer, i.e. its concentration increases with the flow). Calcium concentrations in spring water being exclusively ruled by climatic conditions, the calibration of Eq. (8) is required to model the stationary flux of calcium at the outlet of the catchment (Fig. 2). This stationarity is not observable from the data (Fig. 1a, b, c) owing to the shortness of the series in comparison with the response time of the system but it results from both the stationarity of meteorological conditions and the genesis of calcium in groundwater.

The order of the autoregressive process obtained by minimizing the functional Φ in Eq. (2) is 4 where the residual r is estimated from the tail. The low contribution of rainfall to Ca flux in comparison with the spring flow shows that the contribution of preferential pathways to the overall transfer is not significant at the catchment scale.

3.2. Modelling of deethylatrazine flux at the spring

Fig. 3 shows the modelling results obtained for deethylatrazine (DEA), the main metabolite of atrazine. The tail of the stationary autoregressive process closely represents the observed trend in the flux. No degradation process had to be accounted for to yield a good fit to the data. The temporal variations of the observed DEA flux are probably resulting from the spatial heterogeneity of the DEA concentrations in soils and in the vadose zone, which probably reflects the spatial variability in application patterns of atrazine and in transport and potential attenuation processes. Those variations are part of the noise component of the transfer model (1).

3.3. Modelling of atrazine flux at the spring

No stationary autoregressive process could successfully explain the behavior of the atrazine flux for the whole observation period (Fig. 4) since the resolution of Eq. (8) yielded ω = 0. This can be attributed to the non-stationarity of the flux from the end of 2004 onwards, which is probably due to atrazine degradation in the vadose zone. The non-stationarity

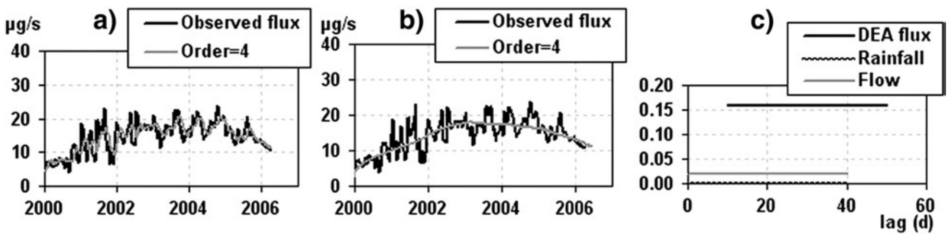

Fig. 3. Modelling of the deethylatrazine flux at the spring — a) comparison between the observed flux and the autoregressive process (8) for a forecast at one step ahead (10 days) — b) comparison between the observed flux and the autoregressive process (8) for long-term prediction — c) impulse responses Γ_{out}, Γ_R and Γ_Q (areas are 0.79, 0, 0.21): order = 4, η = 1.

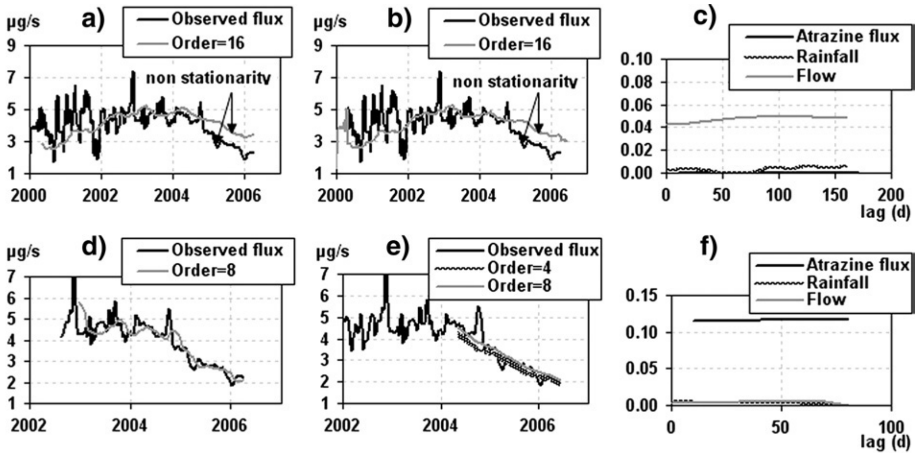

Fig. 4. Modelling of the atrazine flux at the spring with the stationary autoregressive process (8) in a, b, c and with the decaying autoregressive process (9) in d, e, f — a) comparison between the observed flux and the model for a forecast at one step ahead (10 days) — b) comparison between the observed flux and the model for long-term prediction — c) impulse responses Γ_{out}, Γ_R and Γ_Q (areas are 0, 0.03, 0.97): order = 16, — d) comparison between the observed flux and the model for a forecast at one step ahead (10 days) — e) comparison between the observed flux and the model for long-term prediction (only a part of the tail is represented) — f) impulse responses Γ_{out}, Γ_R and Γ_Q (areas are 0.93, 0.03, 0.04); order = 8, η = 0.9883.

in this period can also be detected through the reduction of the noise around atrazine concentrations over the observation period, which is a consequence of the homogenization of atrazine concentrations across the vadose zone through dispersive and degradation processes.

The modelling of the non-stationary atrazine flux is represented in Fig. 4. The inclusion of a decaying component in the system (9) allowed the successful description of time series of atrazine concentrations at the spring. The order of the non-stationary autoregressive process is 8 from Eq. (2). Moreover the residual between the observed flux and the tail does not decrease significantly as the order p increases beyond 8 (Fig. 4e). A decrease in atrazine flux was only observed from 2004, i.e. several years after the last application of atrazine on the catchment. This delay corresponds to the residence time of atrazine in soils and in sub-soils.

3.4. Long-term forecasts

Long-term forecasts in spring flow and pesticide concentrations at the spring were obtained by simulating the long-term behavior of external inputs. Rainfall and potential evapotranspiration data were synthetically generated using weather generators (Pinault et al., 2005). The precipitation and potential evapotranspiration generators use the Markov chain Monte Carlo and simulated annealing (Bardossy, 1998) procedures from the Hastings–Metropolis algorithm (Metropolis et al., 1953; Hastings, 1970).

The stochastically-generated weather series were used to simulate the long-term behavior of spring flow and fluxes according to Eqs. (8) and (9). The contaminant concentrations in spring water were then compiled as the ratio of fluxes to the spring flow:

$$\hat{C}(t_i) = \hat{F}^{\text{out}}(t_i) m_F / \hat{Q}(t_i) m_Q \tag{10}$$

where m_F and m_Q are the means of F^{out} and Q, respectively and:

$$\hat{Q}(t_i) = \boldsymbol{\Gamma}_R \cdot \mathbf{R}_{\text{eff}} \tag{11}$$

where $\mathbf{R}_{\text{eff}}$ is the effective rainfall reduced to unit average calculated from the rainfall and the potential evapotranspiration (Pinault et al., 2005) and $\boldsymbol{\Gamma}_R$ is the transfer function of effective rainfall to the outlet.

Modelling of spring flow from the transfer model (11) and associated forecasts are represented in Figs. 5 and 6. The

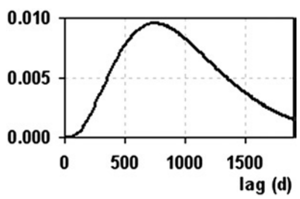

Fig. 5. The impulse response of the transfer model of spring flow (11) which represents the unit hydrograph.

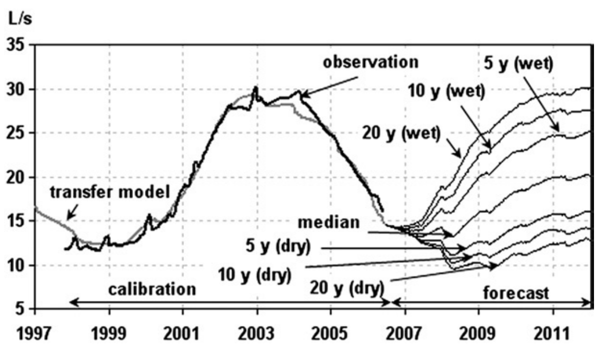

Fig. 6. Modelling and prediction of the spring flow and representation of the percentiles according to their return period. On both sides of the median, the return periods are referring to exceptionally wet or to exceptionally dry periods. The transfer model (11) is used (1000 realizations are simulated).

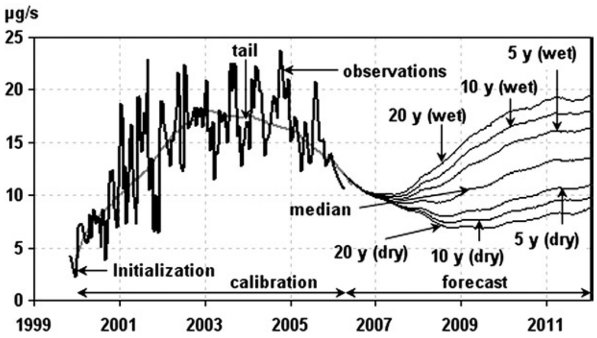

Fig. 7. Modelling and prediction of the DEA flux at the spring (order=4) and representation of the percentiles according to their return period. The stationary autoregressive process (8) is used (1000 realizations are simulated).

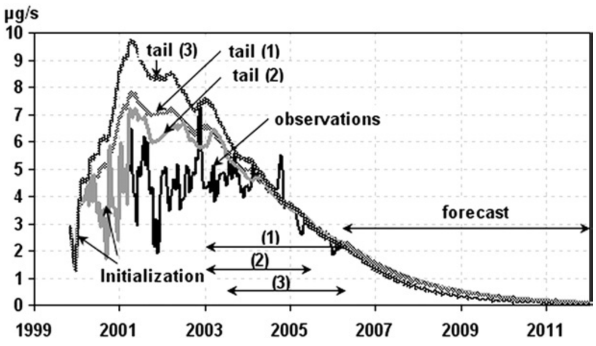

Fig. 9. Modelling and prediction of the decrease in atrazine flux at the spring from the decaying autoregressive process (9). The percentiles associated with different return periods overlap (the process becomes one-dimensional for long-term forecasts). Three time periods are used for calibration (the tails are referring to calibration regions). Region 1 (impulse response areas=0.93, 0.03, 0.04): order=8, Region 2 (impulse response areas=0.76, 0.09, 0.15): order=48, Region 3 (impulse response areas=0.90, 0.04, 0.06): order=8.

response time of the spring flow which exclusively originates from groundwater is several years (Fig. 5). Predictions of spring flow from April 2006 are presented as percentiles associated with different return periods calculated from the stochastic realizations of weather data. The percentiles are calculated from all realizations at every time step, which means that the lines are not actual projections, but prediction intervals. They appear to diverge increasingly with increasing time until 2011 where the spread of the curves stabilizes.

The modelling of the deethylatrazine (DEA) flux at the spring and the associated forecasts are presented in Fig. 7. The curves were obtained from a stationary autoregressive process (8) which was initialized at the beginning of the observed flux series. The tail of the stationary process which represents the trend of the DEA flux shows that the fluctuations of the observed data do not disclose any significant evolution around the trend. No degradation of DEA is perceptible over the observation period 2000–2006, which is a time where DEA is probably still formed from the degradation of atrazine. The hypothesis of the stationarity of the DEA flux will remain valid as long as the stock of DEA in soil and in the vadose zone does not undergo any significant degradation process. Only less than 1% of the total DEA stock is estimated to be transferred to the spring every year and the only decay processes which may significantly decrease the DEA flux are biological and geochemical degradation. It is interesting to note that the hypothesis regarding the lack of significant degradation is confirmed by the divergence of the DEA flux percentiles (Fig. 7) which are remarkably similar to those obtained for spring flow (Fig. 6). The modelling and associated forecasts for DEA concentrations in spring water, which is based on Eq. (10), is shown in Fig. 8. The curves representing the various percentiles have a narrow spread with minimum and maximum concentrations being 0.60 and 0.75 µg/L, respectively.

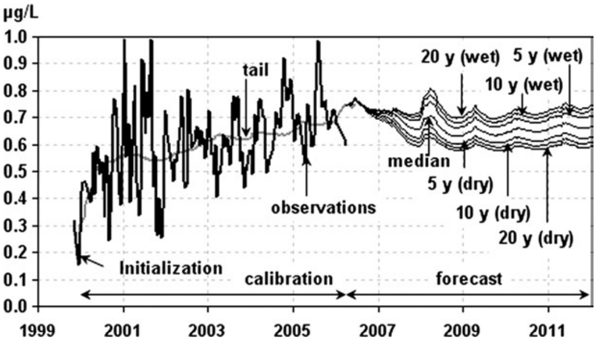

Fig. 8. Modelling and prediction of DEA concentrations at the spring from Eq. (10) and representation of the percentiles according to their return period (1000 realizations are simulated). The concentration increase observed from 2003 is interpreted as resulting from the spring flow decrease. The oscillations that are observed beyond 2008 have no physical meaning (some distortions in the model are a consequence of the noise of observed data used for calibration). The percentiles represent the DEA concentrations assuming a stationary spring flow. Deviations observed over the observation period are not reflected by the spread of the percentiles because the return period of the observed rainfall cumulated over the successive wet years 2000, 2001 and 2002 is much higher than 20 years. Due to the long response time of the system, the inter-annual correlation of rainfall has therefore to be considered for calculating the probability of occurrence of the events.

The modelling of the non-stationary atrazine flux at the spring using Eq. (9) and the associated forecasts are presented in Fig. 9. The observation dataset was split in three subsets to investigate the influence of the calibration period on the modelling and forecast results. As for the DEA stationary flux, the non-stationary autoregressive process was initialized at the beginning of the observation period and the trend of the atrazine flux is represented from the tail. The order of the process, which is selected to have the best representation of the atrazine flux from the tail, was found to be highly

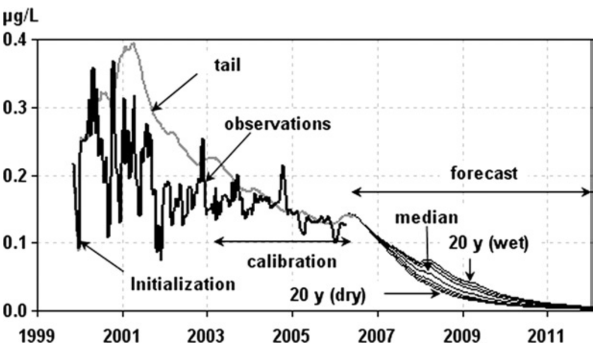

Fig. 10. Modelling of non-stationary atrazine concentration in spring water from Eq. (10) and representation of the percentiles according to their return period (return periods are the same as in Fig. 8). The model corresponding to region (1) of Fig. 9 is used (1000 realizations are simulated).

dependant on the data subset selected for calibration. Although the decrease in atrazine flux was observed from the spring of 2004, the model required a one-year overlap before this date to be calibrated accurately. Whereas region (3) begins 6 months after region (1), the associated autoregressive processes are defined with the same order (8). In contrast, region (2), which ends 8 months before the end of the observation window, requires a higher order (48) to allow the tail to adequately represent the trend of the observed flux into the calibration region (2). The tail follows accurately the observed flux outside the calibration region until the end of the observations, which may be considered as a validation of the decay model. The high order required to minimize the functional (2) proves that the order of the ARX model is not intrinsically related to the process. Moreover, all the regularized impulse responses associated with the autoregressive processes are almost constant (their value depends a little on the lag), which shows the process is equivalent to a moving average on the inputs and the output. So, the order of such a model is strongly linked to the temporal structure of the noise that is all the more related to the calibration region as this region is shorter.

Although the two-year period over which the decrease in atrazine flux is observed is short in comparison with the response time of the system, the forecasts obtained from the various calibration windows are similar. This demonstrates that the autoregressive process is calibrated adequately despite the narrowness of the time window used for calibration. The tails obtained for dates before the spring of 2004 represent the theoretical fluxes under a hypothesis of non-stationarity. These fluxes are clearly dependant of the decay factor η.

For all calibration regions, the percentiles of the predicted atrazine flux are overlapping, which means that the random term $\eta^i(1-\omega)[\lambda_R\boldsymbol{\Gamma}_R\cdot\mathbf{R}_{\mathrm{eff}}+\lambda_Q\boldsymbol{\Gamma}_Q\cdot\mathbf{Q}]$ decreases more rapidly than the autoregressive term $\hat{F}^{\mathrm{out}}(t_i)$ in Eq. (9). In this way, the decaying autoregressive process becomes one-dimensional as the time increases while the atrazine flux becomes entirely controlled by the internal functioning of the system. In other words, processes characterized by long transfer times of atrazine no longer produce any significant flux at the spring as time increases. In those instances, atrazine is being degraded and meteorological conditions have little influence on the actual transfer to the spring, which proves atrazine is being transferred through the unsaturated zone while it is practically degraded in the sub-soils. This statement has important consequences about the understanding of degradation and transfer processes of atrazine.

Predicted concentrations for atrazine in spring water as calculated from Eq. (10) are presented in Fig. 10. Due to the long response time of the system, concentrations are predicted with accuracy for one year and then the percentiles spread. The variability in atrazine concentrations in spring flow is significant until the year 2011 where atrazine concentrations are predicted to be <0.01 µg/L. This means that concentrations of atrazine can be predicted with increasing confidence as time progresses as a result of the decreasing influence of sources.

4. Conclusions

The application of an autoregressive process with external inputs to a small agricultural catchment in France was demonstrated to yield results of interest for explaining and predicting trends in the contamination of water resources by pesticides, even if the monitoring period considered was small compared to the response time of the system under study. The tail of stationary and decaying autoregressive processes was found to provide reliable long-term predictive behavior to the model and the advantages of autoregressive processes were found to be three-fold. First, the application of autocorrelation and cross-correlation analyses to the data could not lead to inferences regarding the stationarity or non-stationarity of the data due to the shortness of the observation period. In contrast, the non-stationarity of the data was clearly demonstrated through the application of the autoregressive process. Secondly, the tail of the autoregressive process was found to become one-dimensional as time increases, which means that fluxes become entirely controlled by the internal functioning of the system. Finally, the stability of the processes was found to allow reliable predictions to be made on the basis of stochastic recursions with controlling factors such as effective rainfall. The methodology outlined in the present paper is expected to be of particular interest to those involved in assessing trends in water quality within the scope of the European Water Framework Directive and Groundwater Directives.

Acknowledgements

The work reported in the present paper was undertaken as part of the TREND2 work package of the EU-funded integrated project AquaTerra (Integrated modelling of the river–sediment–soil–groundwater system; advanced tools for the management of catchment areas and river basins in the context of global change). Monitoring data for the Brévilles catchment were collected as part of the FP5 PEGASE (EVK1-CT1990-00028) and FP6 AquaTerra (GOCE 505428) projects. The funding by the European Commission of the data collection effort and the trends analysis work is gratefully acknowledged.

References

Altman, E., Fiems, D., 2006. Expected waiting time in symmetric polling systems with correlated vacations. INRIA Research Report No. 5933. June, 19 pp.
Asmussen, S., 2003. Applied Probability and Queues, second ed., vol. 51 of Applications of Mathematics (New York). Springer-Verlag, New York.
Baccelli, F., Brémaud, P., 2003. Elements of Queueing Theory. Springer-Verlag, Heidelberg. 234 pp.
Baran, N., Dubus, I.G., Morvan, X., Normand, M., Gutierrez, A., Mouvet, C., 2005. Climatic and land use data for the Brévilles experimental catchment. BRGM Contribution to Deliverable T2.2 of the TREND2 Work Package of the AquaTerra Project. 12 pp.
Bardossy, A., 1998. Generating precipitation time series using simulated annealing. Water Resour. Res. 34, 1737–1744.
Berthouex, P.M., Hunter, W.G., Pallesen, L., Shih, C.Y., 1978. Dynamic behavior of an activated sludge plant. Water Res. 12, 457–472.
Birarder, D.P., Rayburn, A.L., 1995. Flow cytogenetic analysis of whole cell clastogenicity of herbicides found in groundwater. Arch. Environ. Contam. Toxicol. 28, 13–17.
Box, G.E.P., Jenkins, G.M., 1970. Time Series Analysis: Forecasting and Control. Holden-Day, San Francisco.
Clark, G.M., Goolsby, D.A., Battaglin, W.A., 1999. Seasonal and annual load of herbicides from the Mississippi River Basin to the Gulf of Mexico. Environ. Sci. Technol. 33 (7), 981–986.
de Saporta, B., Yao, J.-F., 2005. Tail of a linear diffusion with Markov switching. Ann. Appl. Probab. 15 (1B), 992–1018.
de Saporta, B., Guivarc, ' h, Y., Le Page, E., 2004. On the multidimensional stochastic equation $y_{n+1}=a_n y_n+b_n$. C. R. Acad. Sci. 339 (7), 499–502.
Dubus, I.G., Hollis, J.M., Brown, C.D., 2000. Pesticides in rainfall in Europe. Environ. Pollut. 110 (2), 331–344.

Gilliom, R.J., Barbash, J.E., Crawford, C.G., Hamilton, P.A., Martin, J.D., Nakagaki, N., Nowell, L.H., Scott, J.C., Stackelberg, P.E., Thelin, G.P., Wolock, D.M., 2006. Pesticides in the Nation's streams and ground water, 1992–2001 — the quality of our Nation's waters. U.S. Geol. Surv. Circ. 1291 171 pp. Available at: URL http://ca.water.usgs.gov/pnsp/pubs/circ1291/.
Goolsby, D.A., Battaglin, W.A., Aulenbach, B.T., Hooper, R.P., 2001. Nitrogen input to the Gulf of Mexico. J. Environ. Qual. 30, 329–336.
Guivarc'h, Y., 2006. Heavy tail properties of stationary solutions of multidimensional stochastic recursions. IMS Lecture Notes—Monograph Series. Dynamics & Stochastics, vol. 48, pp. 85–99. doi:10.1214/074921706000000121.
Hastings, W.K., 1970. Monte Carlo sampling methods using Markov chains and their applications. Biometrika 57, 97–109.
IFEN (2006). Les pesticides dans les eaux. Données 2003 et 2004. Les dossiers de l'IFEN numéro 5, 40 pp. http://www.ifen.fr/publications/dossiers/PDF/dossier05.pdf.
Karacan, S., 2003. Application of a non-linear long range predictive control to a packed distillation column. Chem. Eng. Process 42, 943–953. doi:10.1016/S0255-2701(02)00190-3.
Kolpin, D.W., Thurman, E.M., Linhart, S.M., 1998. The Environmental Occurrence of Herbicides: The Importance of Degradates in Ground Water, 35, 3. Springer, New York, pp. 385–390. doi:10.1007/s002449900392.
Metropolis, N., Rosenbluth, A., Rosenbluth, M., Teller, A., Teller, E., 1953. Equation of state calculation by quick computing machines. J. Chem. Phys. 21, 1087–1092.
Morvan, X., Mouvet, C., Baran, N., Gutierrez, A., 2006. Pesticides in the groundwater of a spring draining a sandy aquifer: temporal variability of concentrations and fluxes. J. Contam. Hydrol. 87 (3–4), 176–190.
Mouvet, C., Albrechtsen, H.J., Baran, N., Chen, T., Clausen, L., Darsy, C., Desbionne, S., Douguet, J.M., Dubus, I.G., Esposito, A., Fialkiewicz, W., Gutierrez, A., Haverkamp, R., Herbst, M., Howles, D., Jarvis, N.J., Jørgensen, P.R., Larsbo, M., Meiwirth, K., Mermoud, A., Morvan, X., Normand, B., O' Connor, M., Ritsema, C., Roessle, S., Roulier, S., Soutter, M., Stenemo, F., Thiéry, D., Trevisan, M., Vachaud, G., Vereecken, H., Vischetti, C., 2004. PEGASE. Pesticides in European Groundwaters: detailed study of representative aquifers and simulation of possible evolution scenarios. In: Dubus, I.G., Mouvet, C. (Eds.), Final Report of the European Project #EVK1-CT1990-00028. BRGM/RP-52897-FR. www.brgm.fr/pegase/report/final_report.htm. 358 pp., 196 fig., 57 tabl., 1 app.
Pinault, J.-L., Schomburgk, S., 2006. Inverse modeling for characterizing surface water/groundwater exchanges. Water Resour. Res. 42, W08414. doi:10.1029/2005WR004587.
Pinault, J.L., Pauwels, H., Cann, C., 2001a. Inverse modeling of the hydrological and the hydrochemical behavior of hydrosystems — application to nitrate transport and denitrification. Water Resour. Res. 37, 2179–2190.
Pinault, J.L., Plagnes, V., Aquilina, L., Bakalowicz, M., 2001b. Inverse modeling of the hydrological and the hydrochemical behavior of hydrosystems — characterization of Karst system functioning. Water Resour. Res. 37, 2191–2204.
Pinault, J.L., Amraoui, N., Golaz, C., 2005. Groundwater-induced flooding in macropore-dominated hydrological system in the context of climate changes. Water Resour. Res. 41, W05001. doi:10.1029/2004WR003169.
Ritter, W.F., Scarborough, R.W., Chirnside, A.E.M., 1994. Contamination of groundwater by triazines, metolachlor and alachlor. J. Contam. Hydrol. 15, 73–92.
Roulier, S., Baran, N., Mouvet, C., Stenemo, F., Morvan, X., Albrechtsen, H.J., Clausen, L., Jarvis, N., 2006. Controls on atrazine leaching through a soil-unsaturated fractured limestone sequence at Brevilles, France. J. Contam. Hydrol. 84, 81–105.
Shen, L., Wania, F., Lei, Y.D., Teixeira, C., Muir, D.C.G., Bidleman, T.F., 2004. Hexachlorocyclohexanes in the north American atmosphere. Environ. Sci. Technol. 38 (4), 965–975 15 FEB.
Tindall, J.A., Vencill, W.K., 1995. Transport of atrazine, 2.4-D, and Dicamba through preferential flowpaths in an unsaturated claypan soil near Centralia, Missouri. J. Hydrol. 166, 37–59.
Thurman, E.M., Goolsby, D.A., Meyer, M.T., Kolpin, D.W., 1991. Herbicides in surface waters of the Midwestern United States—the effect of spring flush. Environ. Sci. Technol. 25, 1794–1796.
Yao, Y., Tuduri, L., Harner, T., Blanchard, P., Waite, D., Poissant, L., Murphy, C., Belzer, W., Aulagnier, F., Li, Y.F., Sverko, E., 2006. Spatial and temporal distribution of pesticide air concentrations in Canadian agricultural regions. Atmos. Environ. 40 (23), 4339–4351 JUL.