

# A Fokker–Planck–Kolmogorov equation approach for the monthly affluence forecast of Betania hydropower reservoir

Efraín Domínguez and Hebert Rivera

## ABSTRACT

This paper presents a finite difference, time-layer-weighted, bidirectional algorithm that solves the Fokker–Planck–Kolmogorov (FPK) equation in order to forecast the probability density curve (PDC) of the monthly affluences to the Betania hydropower reservoir in the upper part of the Magdalena River in Colombia. First, we introduce a deterministic kernel to describe the basic dynamics of the rainfall–runoff process and show its optimisation using the  $S/\sigma_{\Delta}$  performance criterion as a goal function. Second, we introduce noisy parameters into this model, configuring a stochastic differential equation that leads to the corresponding FPK equation. We discuss the set-up of suitable initial and boundary conditions for the FPK equation and the introduction of an appropriate Courant–Friederich–Levi condition for the proposed numerical scheme that uses time-dependent drift and diffusion coefficients. A method is proposed to identify noise intensities. The suitability of the proposed numerical scheme is tested against an analytical solution and the general performance of the stochastic model is analysed using a combination of the Kolmogorov, Pearson and Smirnov statistical criteria.

**Key words** | Fokker–Planck–Kolmogorov equation, Magdalena River, stochastic forecast, time-layer-weighted bidirectional finite difference scheme

**Efraín Domínguez** (corresponding author)  
Departamento de Ecología y Territorio, Facultad de Estudios Ambientales y Rurales, Pontificia Universidad Javeriana, Ceiba – Complejidad, Transv. 4 # 42-00 8 Piso, Bogotá, Colombia  
Tel.: +571 320 8320 Ext. 4821  
E-mail: [e.dominguez@javeriana.edu.co](mailto:e.dominguez@javeriana.edu.co)

**Hebert Rivera**  
Subdirección de Hidrología, Instituto de Hidrología, Meteorología y Estudios Ambientales IDEAM, Carrera 10 No 20-30, Bogotá, Colombia

## INTRODUCTION

Hydrological variability is a major source of uncertainty for hydropower generation. In Colombia, 65% of electricity is hydraulically generated. Throughout the country, 24 hydropower reservoirs are installed and operating (Ministerio de Minas y Energía 2008). As the country's economy expands, the Mines and Energy Ministry has begun to develop an infrastructure expansion plan to ensure future power generation (Ministerio de Minas y Energía 2006). At present, the Hydrological and Powerplant Committee (HPC), which belongs to the Colombian Central Hydropower Dispatch System, requires probabilistic hydrological forecasts for short, medium and long term planning (daily, monthly and yearly timeframes). Assuming stationary hydrology, such forecasts can be built using Monte Carlo

techniques and time series modelling (AR, ARMA and ARIMA models). Due to climate change and human intervention in the river basins, the hypothesis of stationary hydrological conditions no longer holds. A technique to handle an unstable hydrological regime is therefore required. In this paper, we present a finite-difference, time-layer-weighted, bidirectional solution to the FPK equation that describes the evolution of probability density curves (PDC) of monthly affluences to hydropower reservoirs under non-stationary conditions. To take into account the physical basis of the rainfall–runoff process, we first introduce a deterministic kernel of low complexity. This kernel is then enhanced with the introduction of noisy parameters, leading to a general stochastic rainfall–runoff

differential equation that can be solved in several ways. One such solution is the numerical solution of the corresponding FPK equation. This approach keeps the deterministic kernel simple, reflecting only the essential features of the process (Samarsky & Mikhailov 1997) and accounting for indirect, non-essential factors with noisy parameters. An alternative approach is to use a more complex operator to represent each system element in order to present a holistic understanding of the process. This approach could require large vectors of input data and parameters to be optimised. The latter approach has several drawbacks associated with the spatial and temporal resolution required by its operator and the resolution of the available field data. Another drawback is the uncertainty of initial and boundary conditions due to random measurement error. Note that complex models can be very sensitive to such kinds of errors. Finally, this concurrent type of modelling cannot provide a probabilistic description of the dynamics of real systems, as is required by the Colombian Hydropower Sector.

Stochastic modelling using the FPK equation is becoming a valuable approach to simulate complex systems, including hydrological ones. Hydrology is a scientific discipline that has embraced probabilities since its scientific development (Hazen 1914; Foster 1923; Sokolovskiy 1930; Kritskiy & Menkel 1935, 1940). Presented here, the FPK equation approach can be treated as the next logical development of the field and as an extension of the works on stochastic hydrology proposed in the 1950s (Kartvelishvili 1958, 1969). The foundations of the method presented here were proposed by V. V. Kovalenko at the Russian State Hydrometeorological University (Kovalenko 1986, 1988, 1993) and were based on the fundamental work of A. N. Kolmogorov (1931). Kolmogorov's work and the theory of stochastic processes have been fruitful in different scientific domains, including physics, mechanical, astronautical and civil engineering, signal processing and nonlinear filtering. During the past two decades, random response predictions, stochastic stability and bifurcation, the first passage problem and nonlinear control problems have all been solved using different approximate solutions of the FPK equation (Stratonovich 1967; Sveshnikov 1968a, 2007; Gardiner 1985; Friedrich & Uhl 1996; Siegert *et al.* 1998; Ulyanov *et al.* 1998a,b; Friedrich *et al.* 2000; Zhu & Cai 2002; Frank *et al.* 2004). Now, this approach is bringing

new tools into hydrology to solve modern problems that relate to efficient water management under heavy human use, adaptation to and mitigation of the ongoing global change process, and the hydrological effects of global warming (Lambin *et al.* 2001; Labat *et al.* 2004; Piao *et al.* 2007). All of these new challenges reject the hypothesis of stationary hydrology, requiring not only non-stationary approaches but also new methods to predict the influence of various driving factors on runoff variability at different time scales (hourly, daily, weekly, annual and long-term runoff). Recently, different works on this subject have been published. Some of these use an intensive analytical approach to describe the dynamics of changing statistical characteristics of hydrological systems (Fujita & Kudo 1995; Lee *et al.* 2001; Naidenov & Shveikina 2002; Dolgonosov & Korchagin 2007), whereas others look for pseudo-stationary solutions to describe the long-term variations and stability of annual runoff probabilistic patterns (ASCE 1993; Khaustov 1999; Frolov 2006). Previous work on numerical solutions to the FPK equation has been done in the field of probabilistic affluence forecasting, all of which uses numerical schemes with one-directional drift (Kovalenko 1993; Shevnina 2001). Since pseudo-stationary solutions and one-directional numerical schemes are not suitable for understanding the system dynamics in terms of conditioned PDCs, a stable, bidirectional numerical solution to the FPK equation and its practical application will therefore be presented and discussed below (Domínguez 2004).

## METHODS

We present a stochastic rainfall–runoff model that works under non-stationary conditions, its corresponding FPK equation, the applied numerical schemes, initial and boundary conditions, and details on its application and modelling results.

The rainfall–runoff process is important to engineering design, water management and flood risk prevention, where all of these tasks require the probabilistic assessment of runoff fluctuations on different time scales (long, medium and short term). First, to understand the system dynamics, we must develop a deterministic kernel to describe the essence of the rainfall–runoff process. Doing so will keep

the model as simple as possible but not overly simple. The process can thus be represented by an ordinary differential equation (ODE) of the type

$$\tau \frac{dQ(t)}{dt} + \frac{1}{k} Q(t) = X(t) \quad (1)$$

where

- $\tau$ —relaxation time (coefficient) [T];
- $Q$ —affluences [ $L^3/T$ ];
- $t$ —time coordinate [T];
- $k$ —runoff coefficient [];
- $X$ —rainfall [ $L^3/T$ ].

Equation (1) can be obtained from the two-dimensional Saint Venant equation. With different assumptions, we can obtain Equation (1) or a higher-order ordinary differential equation (Kovalenko *et al.* 2005, pp 54–57). The differential operator has been widely applied in hydrology to simulate the rainfall–runoff process. General lumped models represented by systems of ODE can be found, for example (Kuchment 1972; Whitehead *et al.* 1979; McCann & Singh 1980; Pingoud 1982, 1983; Sharma & Murthy 1996; Wang & Chen 1996; Kovalenko *et al.* 2005). Equation (1) is linear and has been chosen just for convenience; its suitability will be established in terms of the  $S/\sigma_\Delta$  performance criterion, which is presented later. Although this equation offers good performance, we believe that it will not provide a perfect forecast or simulation, but rather will always involve some degree of error. This error may be associated with model incompleteness, initial condition uncertainties or observed data inexactness. To take into account all of these uncertainties, say that  $-(1/k\tau \equiv c)$  and  $X/\tau \equiv N$ . Then,  $c$  represents the basin's internal properties and  $N$  the external influence over the basin domain. Both  $c(t)$  and  $N(t)$  can be represented as a composition of structural and random components as  $c = \bar{c}(t) + \tilde{c}(t)$  and  $N = \bar{N}(t) + \tilde{N}(t)$ , where  $\tilde{c}$  and  $\tilde{N}$  are white noise processes (or processes without memory) with  $G_{\tilde{c}}$ ,  $G_{\tilde{N}}$  and  $G_{\tilde{c}\tilde{N}}$  noise intensities. Then, Equation (1) becomes

$$\frac{dQ(t)}{dt} = (\bar{c}(t) + \tilde{c}(t))Q(t) + (\bar{N}(t) + \tilde{N}(t)) \quad (2)$$

Integrating Equation (2) from  $t$  to  $t + \Delta t$ , we set

$$Q(t + \Delta t) = Q(t) + \int_t^{t+\Delta t} (\bar{c}(t) + \tilde{c}(t))Q(t)dt + \int_t^{t+\Delta t} (\bar{N}(t) + \tilde{N}(t))dt \quad (3)$$

Equation (3) is a generalised stochastic rainfall–runoff model (Kovalenko 1993). From (3), we deduce that  $Q(t + \Delta t)$  depends just on the value of  $Q(t)$ . Since  $\tilde{c}$  and  $\tilde{N}$  are white noise, we conclude that  $p(Q_k|Q_{k-1}, Q_{k-2}, \dots, Q_{k-3}) = p(Q_k|Q_{k-1})$ . From this, it follows that (3) fulfils all of the conditions for a simple Markov process for which we can obtain the following partial differential equation (Kolmogorov 1931; Pawula 1967; Kovalenko 1993):

$$\frac{\partial p(t, Q)}{\partial t} + \frac{\partial [A(t, Q)p(t, Q)]}{\partial Q} - \frac{1}{2} \frac{\partial^2 [B(t, Q)p(t, Q)]}{\partial Q^2} = 0 \quad (4)$$

In this equation,  $p(t, Q)$  represents the probabilistic density of  $Q$  at the moment  $t$ , and  $A(Q, t)$  and  $B(Q, t)$  are the drift and diffusion coefficients. These deterministic functions determine all the particularities of our Markov process. The analytic forms of the drift and diffusion coefficients depend on the structure of the selected deterministic kernel and on the types of noises introduced to build the stochastic rainfall–runoff model. In fact, these coefficients are defined as (Gardiner 1985; Sveshnikov 1968a):

$$A(t, Q) = \lim_{\Delta t \rightarrow 0} \frac{E[\Delta Q|Q]}{\Delta t} \quad (5)$$

$$B(t, Q) = \lim_{\Delta t \rightarrow 0} \frac{E[\Delta Q^2|Q]}{\Delta t} \quad (6)$$

Then, the drift  $A(t, Q)$  is the instantaneous rate of change of the mean of the process given that  $Q(t) = Q$ . Similarly,  $B(t, Q)$  denotes the instantaneous rate of change of the squared fluctuations of the process given that  $Q(t) = Q$ . Kovalenko (1993) has shown that, from Equation (2), the following analytic forms for the drift and diffusion coefficients can be derived:

$$A(t, Q) = -(\bar{c} - 0.5G_{\tilde{c}})Q(t) - 0.5G_{\tilde{c}\tilde{N}} + \bar{N}(t) \quad (7)$$

$$B(t, Q) = G_{\tilde{c}}Q(t)^2 - 2G_{\tilde{c}\tilde{N}}Q(t) + G_{\tilde{N}} \quad (8)$$

Provided that we can set initial and boundary conditions for Equation (4), we can solve it analytically or numerically. The stationary solution for the FPK equation can be analytically determined for a wide range of problems. In the nonstationary case, an analytical solution can be found for problems with strong restrictions on the analytical types of the drift and diffusion coefficients. Approximate analytical solutions for the transient FPK equation have been found by different authors in different fields (Stratonovich 1967; Sveshnikov 1968*a,b*, 2007; Gardiner 1985; Mitropol'skii & Nguen 1991; Mitropol'skii 1995; Ulyanov *et al.* 1998*a,b*; Guo-Kang 1999; Di Paola & Sofi 2002; Haiwu *et al.* 2003). Numerical solutions of the FPK equation solve a wide range of general problems. Successful applications of the finite difference method have been reported by different authors using explicit or implicit schemes and different finite difference approximations for the drift term (Vanaja 1992; Challa & Faruqi 1996; Wojtkiewicz *et al.* 1997; Kumar & Narayanan 2006; Schmidt & Lamarque 2007). To enable two-directional drift, some authors have recommended approximating the drift term of the FPK equation with central differences (Challa & Faruqi 1996; Wojtkiewicz *et al.* 1997; Kumar & Narayanan 2006) but, in implementing our numerical scheme, we found that central differences were always unstable. A theoretical explanation for the unconditioned instability of central differences is presented in the classical work of Potter (1973). Instead of central differences, we have implemented a bidirectional approach for the drift term of Equation (4). Assuming that the affluences  $Q$  are in the interval  $[\alpha, \beta]$ , we set an uniform mesh with  $Q \times t$  nodes defined as:  $Q_j = \alpha + j\Delta Q$  and  $t_i = t_0 + i\Delta t$  with  $j = 0, 1, \dots, n$ ;  $n = (\beta - \alpha)/\Delta Q$  and  $(i = 0, 1, \dots)$ . Then we can write the following numerical time-weighted bidirectional finite-difference approximation for the FPK equation:

$$\begin{aligned} \frac{p_j^{i+1} - p_j^i}{\Delta t} = & - \left\{ \sigma \left[ \frac{\varphi_L(A_{j+1}^{i+1}p_{j+1}^{i+1} - A_j^{i+1}p_j^{i+1})}{\Delta Q} + \frac{\varphi_R(A_j^{i+1}p_j^{i+1} - A_{j-1}^{i+1}p_{j-1}^{i+1})}{\Delta Q} \right] \right. \\ & + (1 - \sigma) \left[ \frac{\varphi_L(A_{j+1}^i p_{j+1}^i - A_j^i p_j^i)}{\Delta Q} + \frac{\varphi_R(A_j^i p_j^i - A_{j-1}^i p_{j-1}^i)}{\Delta Q} \right] \left. \right\} \\ & + \left\{ \sigma \left[ \frac{1}{2} \frac{(B_{j+1}^{i+1}p_{j+1}^{i+1} - 2B_j^{i+1}p_j^{i+1} + B_{j-1}^{i+1}p_{j-1}^{i+1})}{\Delta Q^2} \right] \right. \\ & \left. + (1 - \sigma) \left[ \frac{1}{2} \frac{(B_{j+1}^i p_{j+1}^i - 2B_j^i p_j^i + B_{j-1}^i p_{j-1}^i)}{\Delta Q^2} \right] \right\} \end{aligned} \quad (9)$$

where  $\varphi_R$  and  $\varphi_L$  are directional coefficients enabling bi-directional drift and  $\sigma$  is a weighting coefficient for time layers. Here, if  $A_j^i < 0$ , then  $\varphi_R = 0$  and  $\varphi_L = 1$ . For  $A_j^i \geq 0$ ,  $\varphi_R = 1$  and  $\varphi_L = 0$ . If  $\sigma = 1$ , the numeric scheme (9) becomes a totally implicit scheme. When  $\sigma = 0$ , we instead have a totally explicit scheme. To solve Equation (9) explicitly, it is necessary to fulfil a Courant–Frederich–Levi stability condition:

$$\max(|B(t, Q)|) \frac{\Delta t}{\Delta Q^2} < \frac{1}{2} \quad (10)$$

Boundary conditions for (9) may be either reflecting:

$$\left( A(t, Q)p(t, Q) - \frac{1}{2} \frac{\partial^2 [B(t, Q)p(t, Q)]}{\partial Q^2} \right)_{Q=\alpha} = 0 \quad (11)$$

$$Q = \beta$$

or absorbing:

$$p(t, Q)_{Q=\alpha} = 0 \quad (12)$$

$$Q = \beta$$

The explicit case for the FPK equation, discarding all right-hand terms with time index  $i + 1$  in (9), can be easily implemented using any programming language. The efficiency of the code will depend on condition (10) only, and no difficult programming issues will be faced. Solving Equation (9) in full requires more effort. This equation must be rewritten as

$$\xi_j^{i+1} p_{j-1}^{i+1} + \psi_j^{i+1} p_j^{i+1} + \gamma_j^{i+1} p_{j+1}^{i+1} = R_j^i \quad (13)$$

where

$$\xi_j^{i+1} = -\omega_2 A_{j-1}^{i+1} - \omega_5 B_{j-1}^{i+1} \quad (14)$$

$$\gamma_j^{i+1} = \omega_1 A_{j+1}^{i+1} - \omega_5 B_{j+1}^{i+1} \quad (15)$$

$$\psi_j^{i+1} = \omega_2 A_j^{i+1} - \omega_1 A_j^{i+1} + 2\omega_5 B_j^{i+1} + 1 \quad (16)$$

$$\begin{aligned} R_j^i = & -p_j^i + \omega_3 A_{j+1}^i p_{j+1}^i - \omega_3 A_j^i p_j^i + \omega_4 A_j^i p_j^i - \\ & \omega_4 A_{j-1}^i p_{j-1}^i - \omega_6 B_{j+1}^i p_{j+1}^i + 2\omega_6 B_j^i p_j^i - \\ & \omega_6 B_{j-1}^i p_{j-1}^i \end{aligned} \quad (17)$$

with

$$\omega_1 = \frac{\sigma\varphi_L\Delta t}{\Delta Q} \quad (18)$$

$$\omega_2 = \frac{\sigma\varphi_R\Delta t}{\Delta Q} \quad (19)$$

$$\omega_3 = \frac{(1-\sigma)\varphi_L\Delta t}{\Delta Q} \quad (20)$$

$$\omega_4 = \frac{(1-\sigma)\varphi_R\Delta t}{\Delta Q} \quad (21)$$

$$\omega_5 = \frac{\sigma\Delta t}{2\Delta Q^2} \quad (22)$$

$$\omega_6 = \frac{(1-\sigma)\Delta t}{2\Delta Q^2} \quad (23)$$

Equation (17) leads to a three-diagonal algebraic equation system that can be solved efficiently using the Thomas factorisation method (Potter 1973; Akai 1994).

## APPLICATION AND MODELLING RESULTS

The above numerical scheme was applied to set up the forecast of the PDCs of monthly affluences to the Betanias hydropower reservoir. This reservoir is located in the upper part of the Magdalena River basin, receiving affluences from an area of 13,600 km<sup>2</sup> with a mean streamflow of 430 m<sup>3</sup>/s. The hydrometeorological network has 64 hydrometric stations and 200 rainfall gauges. This network started to operate in 1960. Some of the observation nodes are linked through a satellite transmission system and report data hourly. The hydrometric stations Paicol (coded as 2104701 in the Network Catalogue) and Puente Balseadero (coded as 2105706) record 93% of the total streamflow to this reservoir, so the sum of the streamflows registered by these stations was used to approximate the total reservoir inflow. The rainfall stations used to determine the input precipitation were selected after a correlation analysis. A correlation matrix of size 200 × 40 was built to determine the best streamflow predictors among all the precipitation gauges. As a result, we have selected the total precipitation measured by the rainfall stations with codes 2101013,

2101014, 2103006, 2103506, 2105031, 2601005, 2601007 and 4401010. The last station is located outside the basin domain, but it still measures precipitation patterns that influence the streamflow at the upper part of the Magdalena River. The correlation coefficient between total rainfall and streamflow was 0.85.

The above rainfall–runoff information was used to determine the optimal values for the parameters  $\tau$  and  $k$  (Equation (1)). To solve this inverse problem, we developed the following numerical scheme for Equation (1):

$$Q_{t+\Delta t} = k \left[ X_{t+\Delta t} + \frac{\tau}{\Delta t} \left( \frac{Q_t}{X_t} - Q_t \right) \right] \quad (24)$$

The explanation of symbols is the same as for Equation (1).

We use the ratio  $S/\sigma_\Delta$  as a goal function. If we denote  $Q_i^{\text{obs}}$  and  $Q_i^f$  as observed and forecasted affluences then, to evaluate the  $S/\sigma_\Delta$  criterion (Popov 1968; Appolov *et al.* 1974), we must use the following expressions:

$$\Delta_i = Q_i^{\text{obs}} - Q_{i-T}^{\text{obs}} \quad (25)$$

$$\bar{\Delta} = \frac{1}{n} \sum_{i=1}^n \Delta_i \quad (26)$$

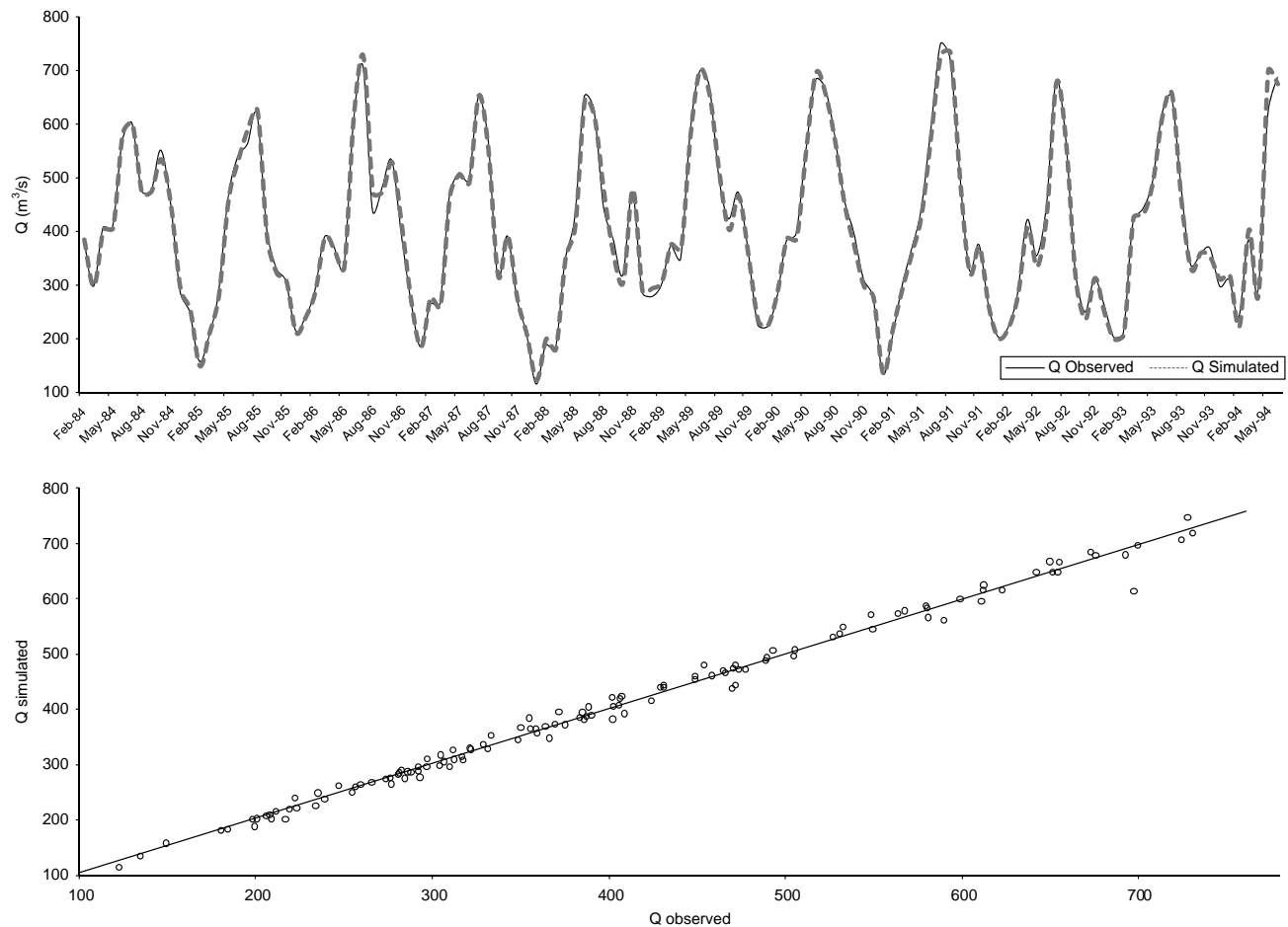
$$\sigma_\Delta = \sqrt{\frac{\sum_{i=1}^n (\Delta_i - \bar{\Delta})^2}{n-1}} \quad (27)$$

$$S = \sqrt{\frac{\sum_{i=1}^n (Q_i^{\text{obs}} - Q_i^f)^2}{n-1}} \quad (28)$$

A value of  $S/\sigma_\Delta = 0$  provides the perfect forecast/simulation. A model calibration is satisfactory if  $S/\sigma_\Delta \leq 0.8$ . In our numerical experiments, we found  $S/\sigma_\Delta \leq 0.10$  for the calibration period and  $S/\sigma_\Delta \leq 0.20$  for the blind validation test (see Figure 1). The test was performed by issuing 125 forecasts.

To apply the numerical scheme (9), we designed two kinds of modelling set-ups. The first assumes that the standard deviation of monthly affluence values comes from the typical error produced by the monitoring system, so that a normal distribution was used to characterise monthly affluences (we call these initial conditions type 1).





**Figure 1** | Visual control of validation results for Equation (1).

The second set-up type assumes that the deviation of monthly streamflows to the Betania reservoir can be characterised by the fluctuations of daily affluences to the reservoir within each month. For this set-up type, we selected, prior to simulations, the time period where daily streamflows behaved randomly inside their respective months. A sign test (Druzhinin & Sikan 2001) with a significance level of 5% was applied to determine the randomness of daily mean flows inside each month. We found that, for January, February, March, September and November, daily flows could be considered independent and identically distributed random magnitudes. Then, for the daily discharge values of the above months, we used the Pearson, Smirnov and Kolmogorov goodness-of-fit criteria (Mitropolsky 1971; Rozhdentvenskiy & Chevotariov 1974; Haan 1977; Lindley & Scott 1995; Tomas *et al.* 2002) to fit

$\gamma$  distributions that were used as initial conditions for the second type (we call these initial conditions type 2). Within the first set-up, we fit 132 normal distributions that were used as initial conditions and also as observed distributions to check the performance of the stochastic model. For the second type set-up, we adjusted only 66  $\gamma$  distributions because not all months passed the randomness test. The adjusted  $\gamma$  distributions were used as initial conditions for the FPK equation and as observed distributions in the second set-up type in order to check the stochastic model performance when simulating asymmetrical distributions. Figure 2 shows the dynamics for the observed normal and  $\gamma$  distributions in the 1991 calendar year.

To identify the internal ( $c = \bar{c}(t) + \tilde{c}(t)$ ) and external ( $N = \bar{N}(t) + \tilde{N}(t)$ ) system parameters and their noise intensities  $G_{\bar{c}}$ ,  $G_{\tilde{N}}$  and  $G_{\tilde{c}\tilde{N}}$ , we applied a pseudo-stationary

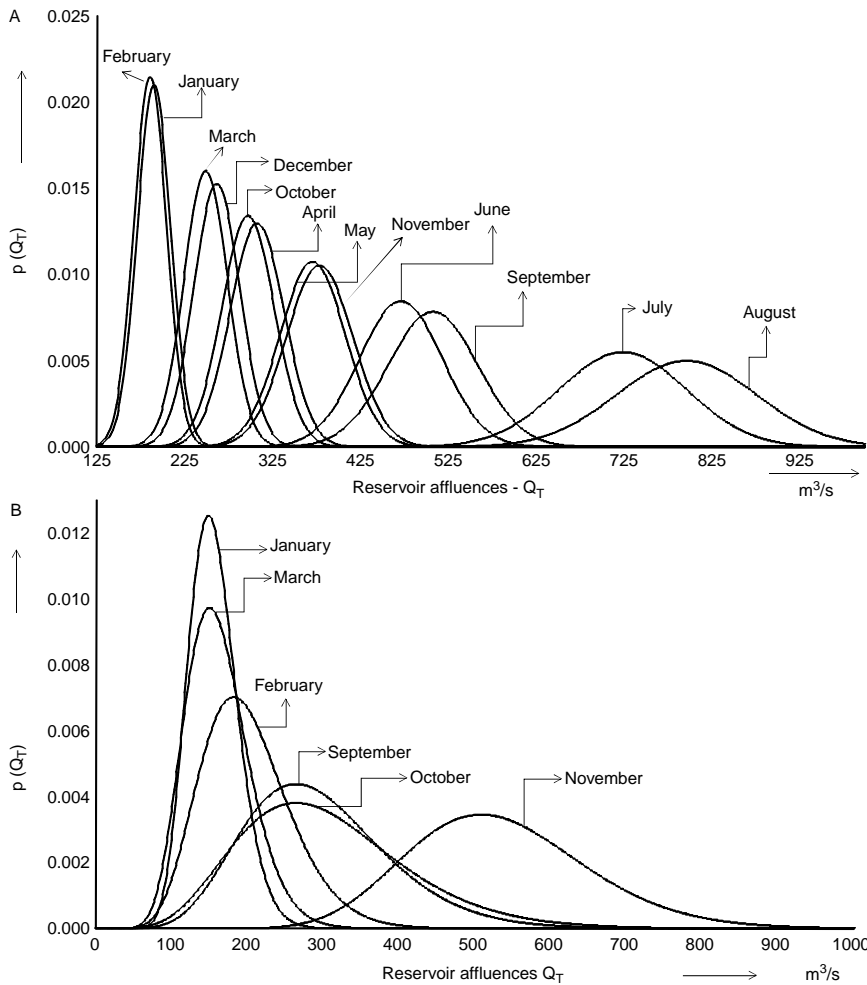


Figure 2 | Dynamics for normal (A) and  $\gamma$  distributions (B) of monthly affluences to Betania's reservoir.

solution for the FPK equation (Kovalenko 1993). Assuming  $\partial p(t, Q) / \partial t = 0$  and introducing the following notation:

$$a = \frac{G_{\bar{c}\bar{N}} + 2\bar{N}}{2\bar{c} + G_{\bar{c}}} \tag{29}$$

$$b_0 = \frac{-G_{\bar{N}}}{2\bar{c} + G_{\bar{c}}} \tag{30}$$

$$b_1 = \frac{G_{\bar{c}\bar{N}}}{2\bar{c} + G_{\bar{c}}} \tag{31}$$

$$b_2 = \frac{-G_{\bar{N}}}{2\bar{c} + G_{\bar{c}}} \tag{32}$$

Equation (4) becomes

$$\frac{dp}{dQ} = \frac{Q - a}{b_0 + b_1Q + b_2Q^2} p \tag{33}$$

Equation (33) represents the Pearson family of density curves that is widely applied in hydrology (Rozhdenstvenskiy & Chevotariov 1974). From Equation (33), regrouping, multiplying by  $Q^n$  and integrating, we can deduce an equation that links the parameters  $[a, b_0, b_1, b_2]$  with the noncentred statistical moments of order  $n$  ( $\alpha_n$ ):

$$nb_0\alpha_{n-1} + [(n+1)b_1 - a]\alpha_n + [(n+2)b_2 + 1]\alpha_{n+1} = 0 \tag{34}$$

Given  $n = 0, \dots, 3$ , we must read the system:

$$\begin{aligned} 2b_2\alpha_1 + b_1 - a &= -\alpha_1; \\ 3b_2\alpha_2 + 2b_1\alpha_1 - b_0 - a\alpha_1 &= -\alpha_2; \\ 4b_2\alpha_3 + 3b_1\alpha_2 + 2b_0\alpha_1 - a\alpha_2 &= -\alpha_3; \\ 5b_2\alpha_4 + 4b_1\alpha_3 + 3b_0\alpha_2 - a\alpha_3 &= -\alpha_4 \end{aligned} \tag{35}$$

Following from (34):

$$\begin{aligned} a &= 0.5(-\alpha_3 - 4\alpha_1^3 + 5\alpha_1\alpha_2)/(\alpha_2 - \alpha_1^2); \\ b_0 &= 0.5(-2\alpha_2^2 + \alpha_2\alpha_1^2 + \alpha_1\alpha_3)/(\alpha_2 - \alpha_1^2); \\ b_1 &= 0.5(3\alpha_1\alpha_2 - 2\alpha_1^3 - \alpha_3)/(\alpha_2 - \alpha_1^2) \end{aligned} \tag{36}$$

If the statistical moments  $[\alpha_1, \alpha_2, \alpha_3]$  are evaluated directly from hydrological records, then combining Equations (29)–(32) with the system (36), we obtain

$$\bar{c} = \frac{\bar{N}}{(a - b_1/2)} \tag{37}$$

$$G_{\bar{c}\bar{N}} = \frac{\bar{N}b_1}{(a - b_1/2)} \tag{38}$$

$$G_{\bar{N}} = \frac{-2\bar{N}b_0}{(a - b_1/2)} \tag{39}$$

Equations (37)–(39) close the inverse problem for the FPK equation parameters. The ultimate solution of this problem depends on the type of initial conditions used. For the first type, the normality of the density distributions leads to  $b_1 = b_2 = 0$ ,  $G_{\bar{c}\bar{N}} = 0$  and  $G_{\bar{c}} \cong 0$ . From this it follows that

$$a = \alpha_1 \tag{40}$$

$$\bar{c} = \frac{\bar{N}}{\alpha_1} \tag{41}$$

$$b_0 = \frac{-G_{\bar{N}}}{2\bar{c}} \tag{42}$$

For the second type, assuming the initial conditions have asymmetric distributions, setting  $G_{\bar{c}\bar{N}} \neq 0$  and keeping  $G_{\bar{c}} \approx 0$ , we obtain

$$a = \alpha_1 + b_1 \tag{43}$$

$$b_0 = \alpha_1^2 - \alpha_1 b_1 - \alpha_2 \tag{44}$$

$$b_1 = 3\alpha_1\alpha_2 - \alpha_3 - 2\alpha_1^3 \tag{45}$$

Hence the vector  $(\alpha_1, \alpha_2, \alpha_3)$  can be established directly from observed daily affluence data. For each selected month, we fitted a  $\gamma$  distribution theoretic curve  $p(Q)$  to daily affluences (Rozhdenstvenskiy & Chevotariov 1974; Haan 1977). This analytical curve was used to evaluate the vector  $[\alpha_1, \alpha_2, \alpha_3]$  as follows:

$$\alpha_k = \sum_{i=1}^n p_i Q_i^k \tag{46}$$

where  $k$  represents the order of the statistical moment. In our case  $k = (1, 2, 3)$ .

These numerical schemes were developed as a MS Windows application. This application was programmed using the Object Oriented Pascal language within the Borland Rapid Application Development Environment “Delphi 7”. As testing platforms, we used Scilab and MS Excel. This application implements absorbing boundary conditions only. To avoid the loss of probability density through the boundaries, we set an interval  $[\alpha, \beta]$  for the  $Q$  ordinates that was wider than necessary given the observed affluences. We performed 198 numerical experiments using initial conditions of types 1 and 2. Before performing the numerical simulations, we analysed the numerical scheme suitability, the model sensibility and the modelling performance of the proposed numerical solution of the FPK equation given errors in the input ( $N = \bar{N}(t) + \tilde{N}(t)$ ) and model parameters ( $G_{\bar{N}}, G_{\bar{c}}, G_{\bar{c}\bar{N}}$ ).

In order to assess the numerical scheme suitability, we compared the numerical solution of a simple set-up case to its analytic solution. Assume that  $G_{\bar{c}} = 0$ . Then the drift and diffusion coefficients take the form

$$A(t, Q) = -\bar{c}Q(t) - 0.5G_{\bar{c}\bar{N}} + \bar{N}(t) \tag{47}$$

$$B(t, Q) = -2G_{\bar{c}\bar{N}}Q(t) + G_{\bar{N}} \tag{48}$$

We introduce the following coefficients:

$$\alpha_0 = -0.5G_{\bar{c}\bar{N}} + \bar{N}(t) \tag{49}$$

$$\alpha_1 = -\bar{c} \tag{50}$$



$$\beta_0 = G_{\tilde{N}} \tag{51}$$

$$\beta_1 = -2G_{\tilde{c}\tilde{N}} \tag{52}$$

We set up an FPK equation set-up for which an analytic solution can be found as in (Sveshnikov 1968b):

$$\bar{Q}(t) = \frac{\alpha_0}{\alpha_1}(e^{\alpha_1 t} - 1) + Qe^{\alpha_1 t} \tag{53}$$

$$\sigma_Q^2 = \frac{\beta_0}{2\alpha_1}(e^{\alpha_1 t} - 1) \tag{54}$$

where  $\bar{Q}(t)$  and  $\sigma_Q^2$  represent the mathematical expectation and variance of the process. Setting up the initial conditions

$\bar{Q}(t) = 900[m^3/S]$  and  $\sigma_Q^2 = 0$  and the absorbing type boundary conditions (12), we obtained the analytic and numerical solutions presented in Figures 3 and 4. These figures show a good concordance between the analytic and numerical solutions.

The modelling performance and sensitivity analysis of the presented numerical solution to the FPK equation was done using the observed input and output and defined by Equations (29)–(31) noise intensities. Here we used initial conditions of type 1. We randomly inserted additive errors of different levels (5, 10, 15, 20, 30, 40 and 50%) around the measured values for input. Then the modelling performance and sensibility of the modelling parameters were assessed

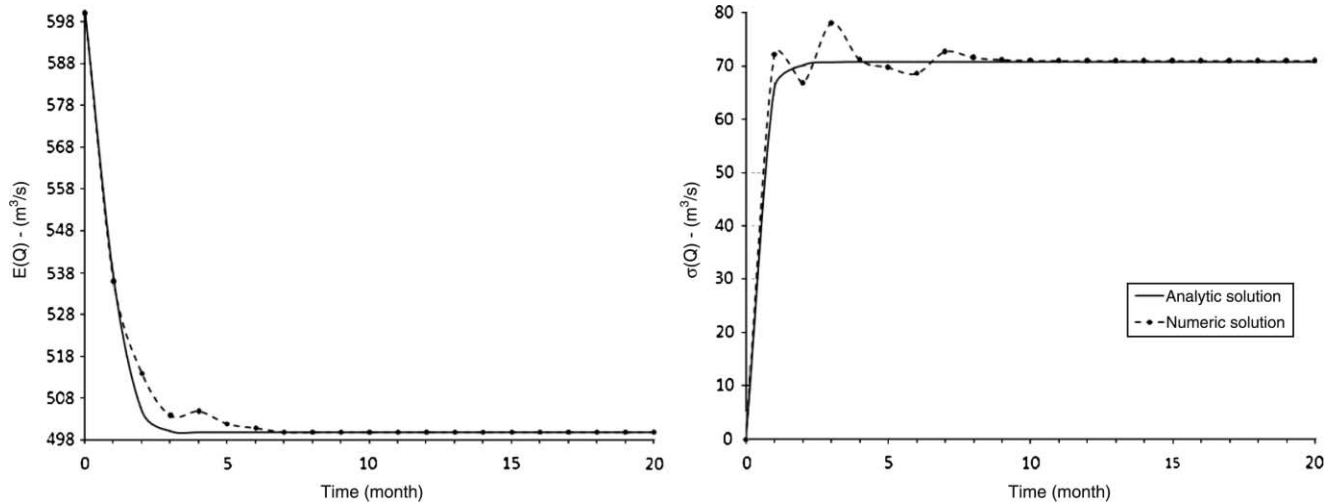


Figure 3 | Comparison of analytical and numerical solutions of the FPK equation with totally explicit schemes.

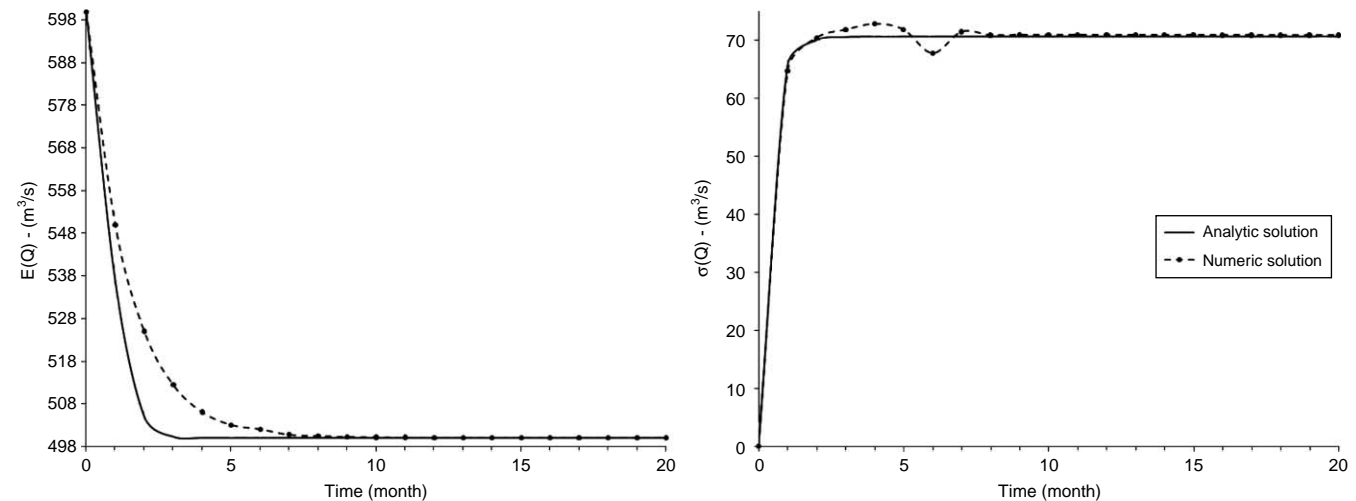


Figure 4 | Comparison of analytical and numerical solutions of the FPK equation with totally implicit schemes.

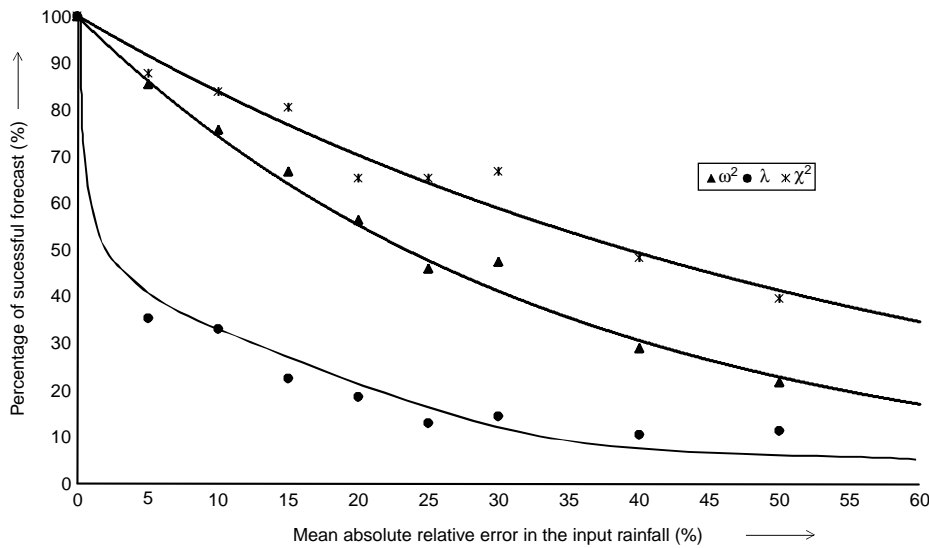


Figure 5 | FPK equation modelling performance analysis.

for each level of model input error. To do so, we established the percentage of successful forecasts for each numerical experiment. A forecast was considered successful if the forecasted PDC verified the null hypothesis about the coincidence of forecasted and observed PDCs. We required a 60% success rate. The goodness of fit was tested using Kolmogorov ( $\lambda$ ), Smirnov ( $\omega^2$ ) and Pearson ( $\chi^2$ ) criteria. This testing was completed for significance levels of 1, 5 and 10%. The sensitivity analysis showed that, on average, we can reach a satisfactory level of success if the input rainfall (observed or forecasted) has a mean absolute relative error

of 15% with a standard deviation of 10% (Figure 5). Analysing Figure 5, we note that the Kolmogorov criterion rejects the null hypothesis more frequently than the Smirnov and Pearson criteria do. It is known that the Kolmogorov criterion tends to “over-rejection of true hypothesis” (Rozhdenstvenskiy & Chevotariov 1974). If we follow the Smirnov and Pearson criteria only, we find that the mean allowable absolute error in precipitation is 20%, with a standard deviation of 15%. Figure 6 shows how sensitive the parameter  $G_{\bar{N}}$  and the criteria  $S/\sigma_{\Delta}$  are to errors induced in the precipitation input. Here we see that

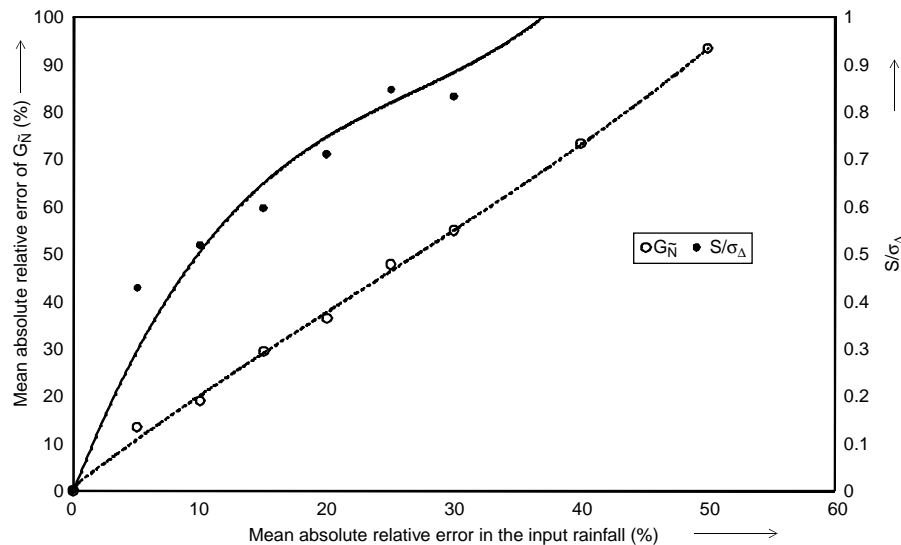


Figure 6 | FPK equation sensitivity analysis.

the error in  $G_{\bar{N}}$  increases linearly with error in input precipitation. It seems that the error in  $G_{\bar{N}}$  will be almost twice that of input rainfall. On the other hand, the criterion  $S/\sigma_{\Delta}$  behaves in a nonlinear manner and has a major sensitivity to error in rainfall from 0 to 20%. The value of  $S/\sigma_{\Delta}$  increases more slowly when the rainfall error is greater than 20%. Further, for the deterministic kernel (1), rainfall with an error level not greater than 25% has enough precision for a good deterministic forecasting performance ( $S/\sigma_{\Delta} \leq 0.80$ ). This shows that the information requirements for the stochastic forecast using the FPK equation are greater than for the deterministic forecast case.

When we forecast the PDC dynamically, we find two types of error patterns. The first pattern is related to an incorrect drift and the second to a flawed flattening or sharpening of the forecasted PDC. An incorrect drift leads to no coincidence in the modal values of observed and forecast PDCs. Incorrect drift in highly sharpened PDCs (see, for example, the January PDC in Figure 2) increases the  $\lambda$  values for the Kolmogorov criteria (Figure 7), leading to more frequent rejection of the null hypothesis. Figure 7 shows that very flat PDCs are less sensible to drift errors. In general, point-by-point error will be greater for sharper

PDCs, so cumulative criteria (such as Pearson or Smirnov) will also increase their null hypothesis rejection rates.

Finally, we performed numerical experiments varying the degree of freedom ( $\nu$ ), rainfall input and parameters used by the FPK equation in the following ways:

Type I numerical experiments—use initial conditions of type 1 and

- actual information about precipitation ( $X$ ), runoff coefficient ( $k$ ) and external noise intensity ( $G_{\bar{N}}$ );
- lag one values for  $X$ ,  $k$  and  $G_{\bar{N}}$ ;
- actual values for  $X$  and forecast values for  $k$  and  $G_{\bar{N}}$ ;
- monthly averages for  $X$  and forecast values for  $k$  and  $G_{\bar{N}}$

Type II numeric experiments—use initial conditions of type 2 and:

- actual information about precipitation ( $X$ ), runoff coefficient ( $k$ ) and external noise intensity ( $G_{\bar{N}}$ );
- monthly averages for  $X$ ,  $k$  and  $G_{\bar{N}}$  values.

Summarising the performance analysis we have performed for the results of numerical experiments, we found that forecast performance decreases as the liberty degree  $\nu$

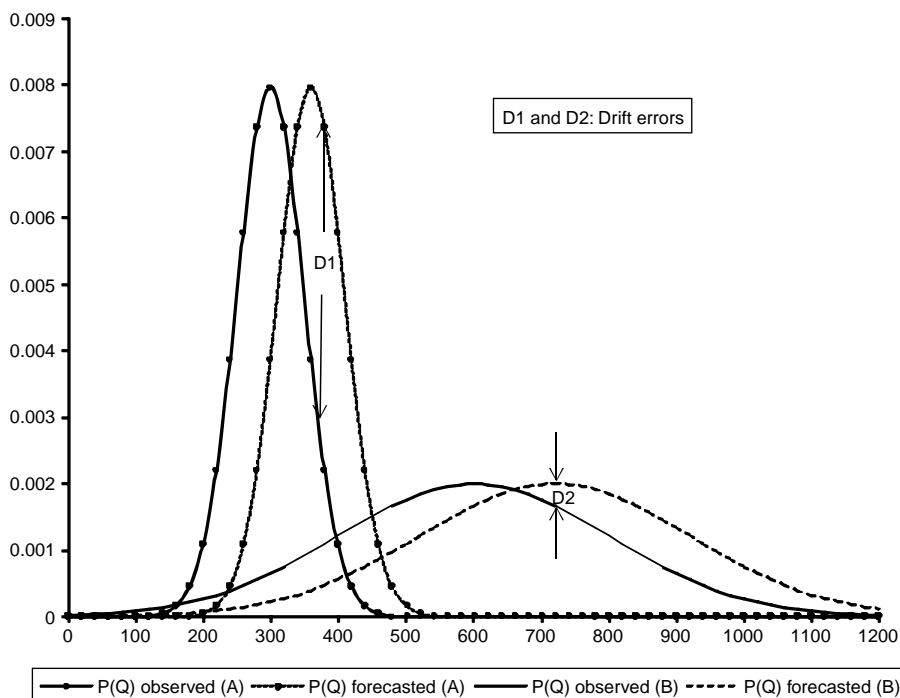


Figure 7 | Drift error effects on sharpened (A) and flattened (B) PDCs.

**Table 1** | Success of the monthly PDC forecast using initial conditions of type 1 and different degrees of freedom ( $\nu$ )

$\nu$	Rainfall and FPK equation parameters	Successfulness (%)								
		Smirnov $\omega^2$			Kolmogorov $\lambda$			Pearson $\chi^2$		
		Significance level $\alpha$								
		1	5	10	1	5	10	1	5	10
15	Actual values for $X, k, G_{\bar{N}}$	100	100	100	100	100	100	100	100	100
	Lagged values for $X, k, G_{\bar{N}}$	56	42	38	15	14	14	65	64	64
	Actual values for $X$ and optimal $k, G_{\bar{N}}$	93	84	81	40	36	36	90	90	90
	Monthly average for $X$ and optimal $k, G_{\bar{N}}$	65	56	49	19	17	17	75	70	68
30	Actual values for $X, k, G_{\bar{N}}$	100	100	100	100	100	100	100	100	100
	Lagged values for $X, k, G_{\bar{N}}$	40	31	28	12	11	11	45	42	40
	Actual values for $X$ and optimal $k, G_{\bar{N}}$	61	52	46	18	17	17	57	53	51
	Monthly average for $X$ and optimal $k, G_{\bar{N}}$	45	35	28	15	15	15	41	39	38
40	Actual values for $X, k, G_{\bar{N}}$	100	100	100	100	100	100	100	100	100
	Lagged values for $X, k, G_{\bar{N}}$	0	0	0	7	6	7	8	8	8
	Actual values for $X$ and optimal $k, G_{\bar{N}}$	57	47	41	15	15	15	56	52	52
	Monthly average for $X$ and optimal $k, G_{\bar{N}}$	39	28	21	14	13	13	41	39	38

increase (Tables 1 and 2). In the experiments that used factual information about rainfall and FPK equation parameters, we obtained a 100% success rate, demonstrating that the identification of the vector  $(G_{\bar{c}}, G_{\bar{c}\bar{N}}, G_{\bar{N}})$  can be done properly (Figure 8). Focusing our attention on the Smirnov and Pearson criteria, we note that, for type I numerical experiments (Table 1), when we use accurate precipitation forecasts and optimised  $k$  and  $G_{\bar{N}}$ , we can produce satisfactory results (more than 60% acceptance of the null hypothesis) at a 1% significance level. For the type II numerical experiments, we find that, using monthly averaged values for  $X, G_{\bar{c}}, G_{\bar{N}}$  and  $G_{\bar{c}\bar{N}}$ , it is possible to obtain an acceptance level greater than 60% even at a 10%

significance level with a liberty degree of  $\nu = 15$  (Table 2). As in the case of type I numerical experiments, this success rate decreases as the liberty degree  $\nu$  increases, but it remains acceptable at the 5% significance level with  $\nu = 40$ . Thus, we realise a better performance for the type II set-up. We assume this result is because this initial conditions set-up better corresponds with the observed asymmetry of hydrological statistical datasets and because asymmetric distributions are less sensitive to the wrong drift and diffusion of forecasted PDCs.

The performance assessment of the stochastic model deserves independent research. For the purpose of this paper we have applied criteria that usually are used to

**Table 2** | Success of the monthly PDC forecast using initial conditions of type 2 and different liberty degrees ( $\nu$ )

$\nu$	Rainfall and FPK equation parameters	Successfulness, %								
		Smirnov $\omega^2$			Kolmogorov $\lambda$			Pearson $\chi^2$		
		Significance level $\alpha$								
		1	5	10	1	5	10	1	5	10
15	Actual values for $X, k, G_{\bar{N}}$ y $G_{\bar{c}\bar{N}}$	99	99	99	99	99	99	99	99	99
	Monthly averages for $X, k, G_{\bar{N}}$ y $G_{\bar{c}\bar{N}}$	93	71	68	39	36	36	99	99	96
30	Actual values for $X, k, G_{\bar{N}}$ y $G_{\bar{c}\bar{N}}$	99	99	99	96	93	93	96	96	96
	Monthly averages for $X, k, G_{\bar{N}}$ y $G_{\bar{c}\bar{N}}$	68	61	50	21	21	21	93	82	61
40	Actual values for $X, k, G_{\bar{N}}$ y $G_{\bar{c}\bar{N}}$	99	99	99	93	93	93	96	96	96
	Monthly averages for $X, k, G_{\bar{N}}$ y $G_{\bar{c}\bar{N}}$	64	50	39	21	18	18	71	64	46

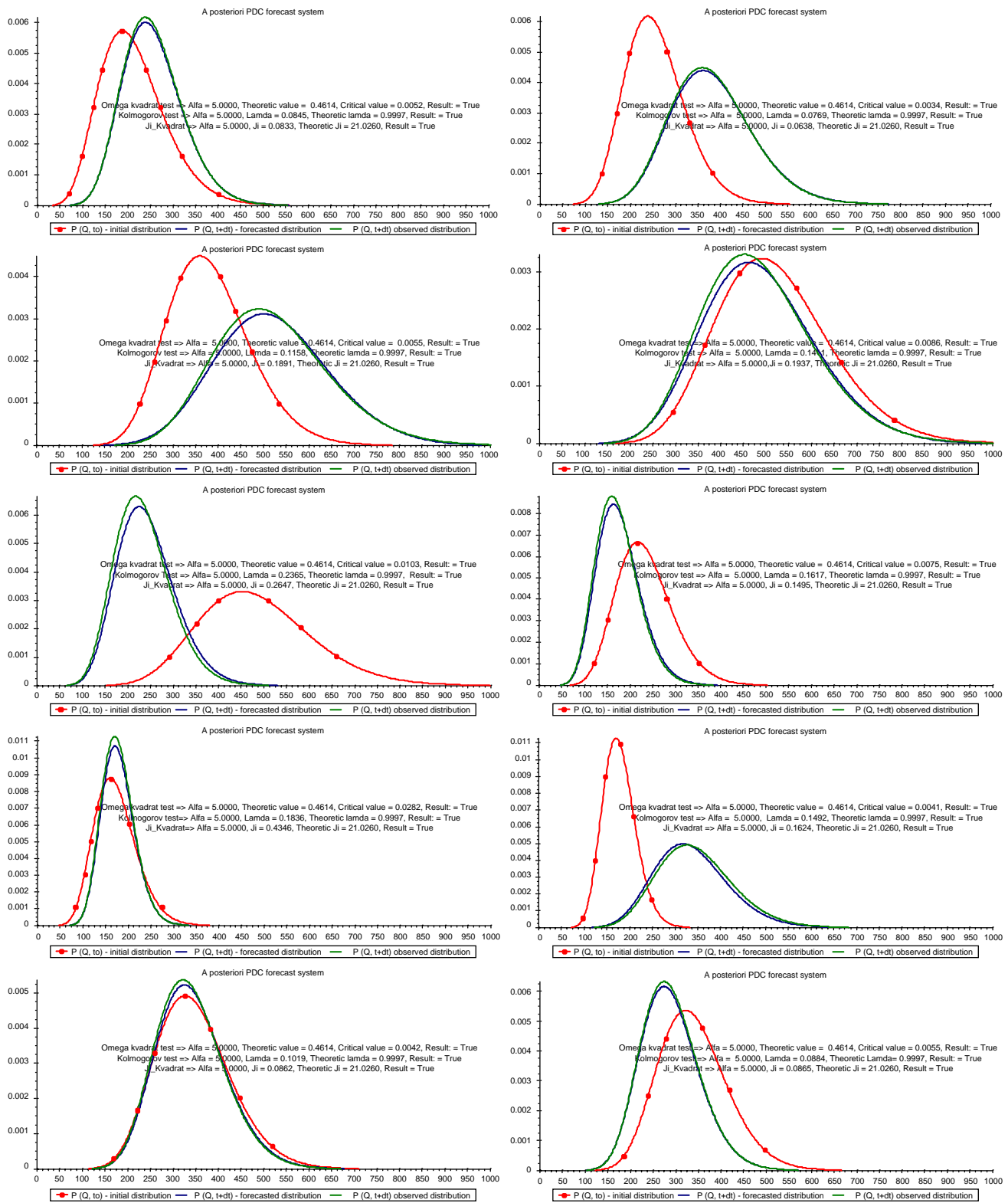


Figure 8 | Some examples of PDC forecasts of monthly affluences to Betania hydropower reservoir using actual values for rainfall input and FPK equation parameters.

compare empirical probability distributions against theoretical distribution functions in order to select the theoretical curve that better represents the empirical data. In our case we are comparing the forecasted PDC against the theoretical curve that has been adjusted to observed data, hence two theoretical sets are being compared: then a lack of proper criteria emerges. The authors did not find any work regarding this issue and the proposed assessment can be understood as a first approach to this matter.

## CONCLUSIONS

The Numerical Time-Weighted Bidirectional Scheme (NTBS) presented here solves a wide spectrum of complex set-ups for the FPK equation. This scheme allows time-dependent nonlinear drift and diffusion coefficients and can work in a totally explicit, implicit or weighted manner. For the explicit solution, where  $\sigma = 0$  in Equation (9), a stability CFL condition was proposed as in Equation (10). Even for the explicit solution, the computational time has proved to be acceptable (no more than minutes for  $\Delta t \leq 10^{-6}$ ). The proposed scheme enables a two-directional drift, overcoming the instability of centred finite differences and guaranteeing the exit to a Dirac  $\delta$  function when noise intensities tend to zero. For the linear drift and diffusion coefficients presented here, the numerical solution of the FPK equation agrees very well with the analytical solution. The stability condition for the diffusive term is stronger than the same condition for the drift term. Because its stability condition (10) is stronger, we allow numerical diffusivity to take place within the solution to the numerical FPK equation. We found that with a real problem set-up (PDC forecast for affluences to the Betania hydropower reservoir), this numerical diffusion could be handled by optimising the noise intensities. We therefore reached a satisfactory success rate for the operative PDC forecast (less than 40% of null hypothesis rejection). We used trial and error to optimise the time dynamics of the noise intensities. We suppose that more sophisticated algorithms, such as gradient solvers, could offer better results.

A better performance was realised using the initial conditions of the type II set-up rather than those of the type I set-up. We ensured that this was because an asymmetric PDC was less sensitive to incorrect drift and diffusion in the

forecasted PDC and because asymmetric initial conditions correspond better to the natural asymmetry of hydrological datasets. However, we still think that the performance of such nonstationary PDC forecasts must be studied deeper. In fact, for deterministic models there is a long tradition and very well-established performance criteria (Dawson *et al.* 2007), but for stochastic models this could still be considered an open question.

The technique presented here for solving the FPK equation allows hydrological risk assessment under nonstationary conditions and can be used as a model operator to solve the inverse modelling problem for water resource management. This tool is valuable for the climate change process and can be useful for establishing measures of adaptability to future climate conditions. For the hydropower sector, using this approach can already be considered mandatory. Finally, the statistical moments of streamflows can be used to indicate water availability. The sensibility of such indicators to the climate change process and even to human pressure on river basins can be established using this method to assess the dynamics of statistical moments in response to changes to the input and system parameters.

## ACKNOWLEDGEMENTS

This work was supported by Pontificia Universidad Javeriana, the Instituto de Hidrología, Meteorología y Estudios Ambientales (IDEAM) and the World Meteorological Organization (WMO). Special acknowledgements must be expressed to the power company EMGESA S.A., who has provided detailed information about the Betania hydropower reservoir. The authors also wish to acknowledge the reviewers for their helpful comments, which led to the improvement of this work.

## REFERENCES

- Akai, T. J. 1994 *Applied Numerical Methods for Engineers*. John Wiley & Sons, New York.
- Appolov, B., Kalinin, G. & Komarov, V. 1974 *Hydrological Forecasting Course*. Guidrometeoizdat, Leningrad.
- ASCE 1993 *Criteria for evaluating watershed models*. *J. Irrig. Drain. Eng.* **119** (3), 429–442.



- Challa, S. & Faruqi, F. A. 1996 Non-linear filtering using probability density evolutions. In: *Proc. International Symposium on Signal Processing and its Applications*. Signal Processing Research Centre, Gold Coast, Australia, pp. 873–876.
- Dawson, C. W., Abraham, R. J. & See, L. M. 2007 **HydroTest: a web-based toolbox of evaluation metrics for the standardised assessment of hydrological forecasts**. *Environ. Modell. Softw.* **22**, 1034–1052.
- Di Paola, M. & Sofi, A. 2002 Approximate solution of the Fokker–Planck–Kolmogorov equation. *Probab. Eng. Mech.* **17** (4), 369–384.
- Dolgonosov, B. M. & Korchagin, K. A. 2007 **A non-linear stochastic model describing the formation of daily and mean monthly water flow in river basins**. *Water Res.* **34** (6), 624–634.
- Domínguez, E. 2004 *Stochastic Forecasting of Streamflow to Colombian Hydropower Reservoirs*. PhD Thesis, Russian State Hydrometeorological University, St Petersburg.
- Druzhinin, V. C. & Sikan, A. B. 2001 *Statistical Methods for the Treatment of Hydrometeorological Information*. RGGMU, St Petersburg.
- Foster, A. 1923 Theoretical frequency curves and their application to engineering problems. *Trans. Am. Soc. Civ. Eng.* **87**, 142.
- Frank, T. D., Beek, P. J. & Friedrich, R. 2004 **Identifying noise sources of time-delayed feedback systems**. *Phys. Lett. A* **328** (2–3), 219–224.
- Friedrich, R. & Uhl, C. 1996 **Spatio-temporal analysis of human electroencephalograms: Petit-mal epilepsy**. *Physica D Nonlin. Phenom.* **98** (1), 171–182.
- Friedrich, R., Siegert, S., Peinke, J., Lück, S., Siefert, M., Lindemann, M., Raethjen, J., Deuschl, G. & Pfister, G. 2000 **Extracting model equations from experimental data**. *Phys. Lett. A* **271** (3), 217–222.
- Frolov, A. V. 2006 **Dynamic-stochastic modelling of long-term variations in river runoff**. *Water Res.* **33** (5), 483–493.
- Fujita, M. & Kudo, M. 1995 **Stochastic response of a storage function model for flood runoff estimation of higher-order moments**. *Environ. Int.* **21** (5), 523–531.
- Gardiner, C. W. 1985 *Handbook of Stochastic Methods*. Springer-Verlag, Berlin.
- Guo-Kang, E. 1999 **A consistent method for the solution to reduced FPK equation in statistical mechanics**. *Physica A* **262**, 118–128.
- Haan, C. T. 1977 *Statistical Methods in Hydrology*. Iowa State University Press, IA.
- Haiwu, R., Xiangdong, W., Guang, M., Wei, X. & Tong, F. 2003 **Approximation closure method of FPK equations**. *J. Sound Vib.* **266**, 919–925.
- Hazen, A. 1914 **The storage to be provided in impounding reservoir for municipal water supply**. *Trans. Am. Soc. Civ. Eng.* **77**, 1539–1640.
- Kartvelishvili, N. A. 1958 **A contribution to the theory of long-term regulation of streamflow**. *Izv. VNIIG* **58**.
- Kartvelishvili, N. A. 1969 **Theory of stochastic processes in hydrology and river regulation**. *Izv. VNIIG* **58**.
- Khaustov, B. A. 1999 *Ustoichivost vieroyatnostnij jarakteristik godovova stoka k antropogennim vozdieystviam*. Doctoral dissertation Thesis, Russian State Hydrometeorological University, St Petersburg.
- Kolmogorov, A. N. 1931 **On analytical methods in probability theory**. *Math. Ann.* **104**, 415–458.
- Kovalenko, V. 1986 *Hydrometrical Assessment of Streamflow with a Stochastic Approach*. Leningradskiy Politekhnikcheskiy Institut, Leningrad.
- Kovalenko, V. 1988 *Dynamic and Stochastic Models of Hydrological Cycle*. Leningradskiy Politekhnikcheskiy Institut, Leningrad.
- Kovalenko, V. 1993 *Modelling of Hydrological Processes*. Guidrometeoizdat, St Petersburg.
- Kovalenko, V., Viktorova, N. V. & Gaidukova, E. 2005 *Modelling of Hydrological Processes*. Guidrometeoizdat, St Petersburg.
- Kritskiy, S. N. & Menkel, M. F. 1935 **Long-term streamflow regulation**. *Gidrotekhn. Stroit.* **2**, 3–10.
- Kritskiy, S. N. & Menkel, M. F. 1940 **A generalized approach to streamflow regulation**. *Gidrotekhn. Stroit.* **2**, 19–24.
- Kuchent, L. S. 1972 *Mathematical Modelling of River Runoff*. Guidrometeoizdat, Leningrad.
- Kumar, P. & Narayanan, S. 2006 **Solution of Fokker–Planck equation by finite element and finite difference methods for nonlinear systems**. *Sadhana Acad. Proc. Eng. Sci.* **31** (4), 445–461.
- Labat, D., Godd ris, Y., Probst, J. L. & Guyot, J. L. 2004 **Evidence for global runoff increase related to global warming**. *Adv. Water Res.* **27**, 631–642.
- Lambin, E. F., Turner, B. L., Geist, H. J., Agbola, S. B., Angelsen, A., Bruce, J. W., Coomes, O. T., Dirzo, R., Fischer, G., Folke, C., George, P. S., Homewood, K., Imbernon, J., Leemans, R., Li, X., Moran, E. F., Mortimore, M., Ramakrishnan, P. S., Richards, J. F., Sk nes, H., Steffen, W., Stone, G. D., Svedin, U., Veldkamp, T. A., Vogel, C. & Xu, J. 2001 **The causes of landuse and land-cover change: moving beyond the myths**. *Global Environ. Change* **11**, 261–269.
- Lee, C.-C., Tan, Y.-C., Chen, C.-H. & Jim Yeh, T. C. 2001 **Stochastic series lumped rainfall–runoff model for a watershed in Taiwan**. *J. Hydrol.* **249** (1–4), 30–45.
- Lindley, D. V. & Scott, W. F. 1995 *Cambridge Statistical Tables*. Cambridge University Press, Cambridge.
- McCann, R. C. & Singh, V. P. 1980 **An analysis of the Chow–Kulandaiswamy GHS model**. *Adv. Water Res.* **3** (4), 173–180.
- Ministerio de Minas y Energ a 2006 *Plan de Expansi n de Referencia, Generaci n Transmisi n 2006–2020*. UPME, Bogota.
- Ministerio de Minas y Energ a 2008 *Centrales hidroel ctricas de Colombia*. Notas de Energ a, Bogot . Available at: [http://www.minminas.gov.co/minminas/sectores.nsf/pages/notas\\_energia](http://www.minminas.gov.co/minminas/sectores.nsf/pages/notas_energia)
- Mitropol’skii, Y. A. 1995 **Approximate solution of the Fokker–Planck–Kolmogorov equation**. *Ukr. Math. J.* **47** (3), 408–419.
- Mitropol’skii, Y. A. & Nguen, T. K. 1991 **A class of nonlinear oscillational systems admitting exact solution of the Fokker–Planck–Kolmogorov equation**. *Ukr. Math. J.* **43** (10), 1383–1388.

- Mitropolsky, A. K. 1971 *Technique of Statistical Assessment. Physical-Mathematical Engineer's Library*. Nauka, Moscow.
- Naidenov, V. I. & Shveikina, V. I. 2002 A nonlinear model of level variations in the Caspian Sea. *Water Res.* **29** (2), 160–167.
- Pawula, R. F. 1967 Generalization and extensions of the Fokker–Planck–Kolmogorov equations. *IEEE Trans. Inf. Theory* **IT-13** (1), 33–41.
- Piao, S., Friedlingstein, P., Ciais, P., de Noblet-Ducoudré, N., Labat, D. & Zaehle, S. 2007 Changes in climate and land use have a larger direct impact than rising CO<sub>2</sub> on global runoff trends. *Proc. Nat. Acad. Sci.* **104** (39), 15242–15247.
- Pingoud, K. 1982 A lumped-parameter model for infiltration. *J. Hydrol.* **57** (1–2), 175–185.
- Pingoud, K. 1985 A dynamical model for the overland flow on an infiltrating catchment. *Appl. Math. Modell.* **7** (2), 128–134.
- Popov, E. G. 1968 *Fundamentals of Hydrological Forecasting. Guidrometeorologicheskoe izdatielztvo*, Leningrad.
- Potter, D. 1973 *Computational Physics*. Wiley, New York.
- Rozhdenstvenskiy, A. V. & Chevotariov, A. I. 1974 *Statistical Methods in Hydrology. Guidrometeoizdat*, Leningrad.
- Samarsky, A. A. & Mikhailov, A. P. 1997 *Mathematical Modelling: Ideas, Methods and Examples*. Nauka, Moscow.
- Schmidt, F. & Lamarque, C. H. 2007 Computation of the solutions of the Fokker–Planck equation for one and two DOF systems. *Commun. Nonlin. Sci. Numer. Simul.* **14** (2), 529–542.
- Sharma, K. D. & Murthy, J. S. R. 1996 Ephemeral flow modelling in arid regions. *J. Arid Environ.* **33** (2), 161–178.
- Shevnina, E. V. 2001 *Application of Deterministic and Stochastic Models to Forecast Monthly Streamflow to a Hydropower Reservoir (Volkhovskiy Hydropower Reservoir Study Case)*. PhD Thesis, Russian State Hydrometeorological University, St Petersburg.
- Siegert, S., Friedrich, R. & Peinke, J. 1998 Analysis of data sets of stochastic systems. *Phys. Lett. A* **243** (5–6), 275–280.
- Sokolovskiy, D. L. 1930 *Distribution Curve Application Defining Probabilistic Yearly Runoff Fluctuations for European Rivers of the USSR. Gostekhizdat*, Leningrad.
- Stratonovich, R. L. 1967 *Topics in the Theory of Random Noise, 1*. Gordon & Breach, New York.
- Sveshnikov, A. A. 1968a *Applied Methods of the Theory of Random Functions*. Nauka, Moskva.
- Sveshnikov, A. A. 1968b *Problems in Probability Theory, Mathematical Statistics and Theory of Random Functions*. W. H. Saunders, Philadelphia.
- Sveshnikov, A. A. 2007 *Applied Methods for the Theory of Markov Processes*. Lan, St Petersburg.
- Tomas, X., González, L., Fernandez, L. & Cuadros, J. 2002 *Tablas Estadísticas*. IQS, Sarria.
- Ulyanov, S. V., Feng, M., Ulyanov, V. S. & Yamafuji, K. 1998a Stochastic analysis of time-variant nonlinear dynamic systems. Part 2: methods of statistical moments, statistical linearization and the FPK equation. *Probab. Eng. Mech.* **13** (3), 205–226.
- Ulyanov, S. V., Feng, M., Ulyanov, V. S., Yamafuji, K., Fukuda, T. & Arai, F. 1998b Stochastic analysis of time-variant nonlinear dynamic systems. Part 1: the Fokker–Planck–Kolmogorov equation approach in stochastic mechanics. *Probab. Eng. Mech.* **13** (3), 183–203.
- Vanaja, V. 1992 Numerical solution of a simple Fokker–Planck equation. *Appl. Numer. Math.* **9**, 533–540.
- Wang, G.-T. & Chen, S. 1996 A linear spatially distributed model for a surface rainfall–runoff system. *J. Hydrol.* **185** (1–4), 183–198.
- Whitehead, P., Young, P. & Hornberger, G. 1979 A systems model of stream flow and water quality in the bedford-ouse river-1. Stream flow modelling. *Water Res.* **13** (12), 1155–1169.
- Wojtkiewicz, S. F., Bergman, L. A. & Spencer, B. F. 1997 High fidelity numerical solutions of the Fokker–Planck equation. In *ICOSSAR'97, Kyoto, Japan* (ed. N. Shiraiishi, M. Shinozuka & Y. K. Wen), pp. 933–940.
- Zhu, W. & Cai, G. 2002 Nonlinear stochastic dynamics: a survey of recent developments. *Acta Mech. Sin (English Ser.)* **18** (6), 551–566.

First received 14 November 2008; accepted in revised form 22 June 2009. Available Online 2 February 2010