# A coupled stochastic inverse-management framework for dealing with nonpoint agriculture pollution under groundwater parameter uncertainty

Llopis-Albert, Carlos[a]; Palacios-Marques, D.[b]; Merigó, J.M.[c]

[a] Universitat Politècnica de València, Camí de Vera s/n, Spain, 46022; email: cllopisa@gmail.com

[b] Universitat Politècnica de València, Camí de Vera s/n, Spain, 46022, email: dapamar@doe.upv.es

[c] University of Manchester, Booth Street West, M15 6PB Manchester, United Kingdom, email: jose.merigolindahl@mbs.ac.uk

## Abstract

In this paper a methodology for the stochastic management of groundwater quality problems is presented, which can be used to provide agricultural advisory services. A stochastic algorithm to solve the coupled flow and mass transport inverse problem is combined with a stochastic management approach to develop methods for integrating uncertainty; thus obtaining more reliable policies on groundwater nitrate pollution control from agriculture. The stochastic inverse model allows identifying non-Gaussian parameters and reducing uncertainty in heterogeneous aquifers by constraining stochastic simulations to data. The management model determines the spatial and temporal distribution of fertilizer application rates that maximizes net benefits in agriculture constrained by quality requirements in groundwater at various control sites. The quality constraints can be taken, for instance, by those given by water laws such as the EU Water Framework Directive (WFD). Furthermore, the methodology allows providing the trade-off between higher economic returns and reliability in meeting the

environmental standards. Therefore, this new technology can help stakeholders in the decision-making process under an uncertainty environment. The methodology has been successfully applied to a 2D synthetic aquifer, where an uncertainty assessment has been carried out by means of Monte Carlo simulation techniques.

**Keywords:** Stochastic inversion; Gradual deformation; Non-Gaussian; Nitrate pollution; Fertilizer standards; Optimization

## 1. Introduction

Groundwater is the ultimate source of freshwater to sustain many important agricultural production areas when surface water sources have been depleted. Furthermore, irrigation is the most important water use accounting for about 70% of the global freshwater withdrawals and 90% of consumptive water uses. Although the development of an intensive agriculture represents one of the main factors in the current economic development of many regions, it has also become an important environmental issue in recent years. This is because it poses many impacts and threats to groundwater bodies, such as overdrafting, aquifer pollution, impacts on downstream demands or impacts on Groundwater Dependent Ecosystems (GDEs). Water laws and polices around the world try to deal with such problems. For instance, the EU Water Framework Directive (EC, 2000) stipulates that groundwater bodies must achieve a good chemical and quantitative status by a set deadline.

However, the decision-making process in groundwater management protection is complex because of heterogeneous stakeholder interests, multiple objectives, key drivers influencing the agricultural market and farmer’s decisions, land-use/crop pattern evolution and uncertain outcomes. A wide range of stakeholders play an active role in water resources management. They range from irrigation communities, government, river basin authority, Non Governmental Organisations (NGO's), agri-business industries, farmers to electric power industries (because of groundwater abstraction costs). Moreover, integrated water resources management incorporates technical, scientific, political, legislative and organizational aspects of a water system. Because of that, stakeholders need new technologies and tools to help them in the decision-making process. This links with the main goal of this paper, which is to present a hydro-economic modeling framework for agricultural advisory services. Specifically, this work is intended to analyze the influence of uncertainty in the physical parameters of a heterogeneous groundwater diffuse pollution problem on the results of management strategies, and to introduce methods that integrate uncertainty and reliability in order to obtain strategies of spatial allocation of fertilizer use in agriculture.

The methodology is based on the coupling of a stochastic inverse model to identify non-Gaussian parameters and to reduce uncertainty in heterogeneous aquifers with a groundwater quality management model for dealing with non-point agriculture pollution. It should be mentioned that a small number of papers in the literature have developed a similar approach as that here presented (e.g., Bark el al., 2003).

The stochastic inverse model allows identifying non-Gaussian parameters and reducing uncertainty in flow and mass transport predictions by constraining stochastic simulations to data, while the optimization management model determines the spatial and temporal distribution of fertilizer application rates that maximizes net benefits in agriculture constrained by quality requirements in groundwater at various control sites.

Inverse modelling has become an important and necessary step in hydrogeological studies (e.g., Poeter and Hill, 1997). This is because the inability to characterize aquifer heterogeneity properly, which makes predictions of contaminant concentration highly uncertain. Consequently, the predictions of management models based on groundwater quality standards are also uncertain. The literature on groundwater inverse modelling mostly focuses on the estimation of parameters and its underlying uncertainty. This is because they are the most relevant factors affecting mass transport predictions (Smith and Schwartz, 1981) and because conceptual uncertainties are difficult to be formalized in a rigorous mathematical framework (Renard, 2007). Regarding the different groundwater parameters we have focused on the hydraulic conductivity, owing to the fact that it is the paramount parameter controlling the flow and solute transport in groundwater. In fact, it can vary spatially by several orders of magnitude. For instance, the aquifer at the Columbus Air Force Base in Mississippi, commonly known as the Macrodispersion Experiment (MADE) site, is a strongly heterogeneous system with a variance of the natural logarithm of K of nearly 4.5 (e.g., Rehfeldt et al., 1992).

Eventually, once the groundwater parameter uncertainty has been strongly reduced by the inverse model, more reliable policies can be defined using the hydro-economic model. It explicitly integrates nitrate leaching and fate and transport in groundwater with the economic impacts of nitrogen fertilizer restrictions in agriculture.

The remaining of the paper is organized as follows: firstly, a background of the stochastic inverse model and the management model is presented; secondly, the methodology has been verified on a 2D synthetic case. Finally, we have highlighted the advantages of using the methodology for providing agricultural advisory services to policy-makers.

## 2. Modeling framework

The methodology is based on the coupling of a stochastic inverse model to identify non-Gaussian parameters and to reduce uncertainty in heterogeneous aquifers with a groundwater quality management model for dealing with non-point agriculture pollution. An explanation of both models is provided below:

### 2.1. Stochastic inverse model (the GC method)

The GC method is a stochastic inverse modeling technique for the simulation of conductivity (*K*) fields in aquifers which has been developed to overcome several of the limitations found in the already existing techniques (Llopis-Albert, 2008; Capilla and Llopis-Albert, 2009). The method was exhaustively verified on a 2D synthetic aquifer (Llopis-Albert and Capilla, 2009). In addition, a 3D application to the Macrodispersion Experiment (MADE-2) site, on a highly heterogeneous aquifer at Columbus Air Force Base in Mississippi (USA) was presented by Llopis-Albert and Capilla (2009a); and also on a complex real-world case study in a fractured rock site (Llopis-Albert and Capilla, 2010). Furthermore, it was extended to deal with independent stochastic structures belonging to independent *K* statistical populations (SP) of fracture families and the rock matrix, each one with its own statistical properties (Llopis-Albert and Capilla, 2010a).

The method uses an iterative optimization procedure to simulate *K* fields honoring *K* measurements, secondary information obtained from expert judgment or geophysical surveys, transient piezometric head (*h*) data and concentration (*c*) measurements. Travel time data can also be considered by means of a backward-in-time probabilistic model (Neupauer and Wilson, 1999), which extends the applications of the method to the characterization of sources of groundwater contamination. The formulation of the method does not require assuming the classical multi-Gaussian hypothesis allowing the reproduction of strings of extreme values of *K* that often take place in nature, being these formation features crucial in order to obtain realistic and safe estimations of mass transport predictions. The method has been developed using a modified version of the gradual deformation technique (Hu, 2000), and applying a Lagrangian approach to solve the mass transport equation. This allows avoiding numerical dispersion usually found in Eulerian approaches. The algorithm has been implemented for 3D transient flow problems under variable density flow conditions, considering the dispersion as a tensorial magnitude, and a first-order mass transfer approach. Performing a Bernoulli trial on the appropriate phase transition probabilities, the particle distribution between the mobile domain and the immobile domain can be determined (Salamon et al., 2006).

The iterative optimization process for constraining stochastic simulations to data is carried out by doing non-linear combinations of seed conditional realizations. These seed conductivity (*K*) fields are already conditional to *K* and secondary data, and are generated by sequential indicator simulation. The a priori stochastic structure of these *K* seed fields is defined by means of the local conditional cumulative density functions (ccdf's) and the indicator variograms, thus allowing the GC method to adopt any Random Function (*RF*) model. As a first step, the GC method builds linear sequential combinations of non multiGaussian *K* fields that honour *K* data:

$$K^m = \alpha_1 K^{m-1} + \alpha_2 K_{2m} + \alpha_3 K_{2m+1} \quad \text{with} \quad K^0 = K_1 \tag{1}$$

where subscripts stand for seed fields and superscripts for conditional fields resulting from a previous linear combination That is, at $m$ iteration, the field $K^{m-1}$, from the previous iteration, is combined with two new independent realizations $K_{2m}$ and $K_{2m+1}$. The procedure requires combining at least three conditional realizations at a time to ensure the preservation of mean, variance, variogram and *K* data in the linearly combined field. The coefficients have also to fulfill the constraints in Eq. (2):

$$\begin{cases} \alpha_1 + \alpha_2 + \alpha_3 = 1 \\ (\alpha_1)^2 + (\alpha_2)^2 + (\alpha_3)^2 = 1 \end{cases} \tag{2}$$

being the parameterization of $\alpha_i$ given by Eq. (3):

$$\begin{cases} \alpha_1 = \frac{1}{3} + \frac{2}{3}\cos\theta \\ \alpha_2 = \frac{1}{3} + \frac{2}{3}\sin(-\frac{\pi}{6} + \theta) \quad \text{with} \quad \theta \in [-\pi, \pi] \\ \alpha_3 = \frac{1}{3} + \frac{2}{3}\sin(-\frac{\pi}{6} - \theta) \end{cases} \tag{3}$$

The $\alpha_i$ coefficients are different in every iteration *m*, and correspond to a unique parameter θ; note the one to one correspondence between the parameter and the combined realization $K^m$.

Because the linear combination of independent non-Gaussian random functions does not preserve the non-Gaussian distribution, although the variogram is preserved, a transformation between Gaussian to the non-Gaussian fields (and vice versa) is required. This transformation is performed through the probability fields (see Capilla and Llopis-Albert 2009 for more details). Finally, at each iteration *m* of the method the parameter θ is determined by minimizing an objective function that penalizes deviations among computed and measured data. As aforementioned this way of operating has been successfully applied in both synthetic and real cases.

The penalty function to be minimized is made up by the weighted sum of three terms:

$$p^k(\theta) = p_h^k(\theta) + \Phi^k p_c^k(\theta) + \Psi^k p_\tau^k(\theta) \tag{4}$$

where $p_h^k(\theta)$ is the weighted sum of square differences between observed and calculated values for piezometric heads, concentrations and travel times, respectively. These terms are function of the parameter $\theta$, for every time step *t* and measurement location *i*. The terms $\Phi^k$ and $\Psi^k$ are trade-offs coefficients between the different conditioning data (see Capilla and Llopis-Albert, 2009).

## 2.2. Hydro-economic model

The method is based on previous developments of the hydro-economic modeling framework presented by Peña-Haro et al. (2009; 2011). The model was also applied to a real-complex case study (Peña-Haro et al., 2010) and further extended to assess different sources of uncertainty on the suggested control policies and the resulting economic and environmental impacts (Llopis-Albert el al., 2014).

The stochastic optimization approach determines the spatial and temporal fertilizer application rate that maximizes the net benefits in agriculture constrained by the quality requirements in groundwater at various control sites. It quantifies the relationship between emissions (fertilizer applied) and groundwater quality impacts or nitrate concentration measured at regulatory control sites. The regulation of nitrate pollution is examined within a cost-effectiveness way, in which the objective is to maximize the sum of the net benefits from agricultural production while meeting the environmental standards. The management model for groundwater pollution control is formulated as follows, where the benefits in agriculture are determined by means of crop prices and crop production functions:

$$Max\,\Pi = \sum_{s}\sum_{y}\frac{1}{\left(1+r\right)^{y}} A_s\left(p_s\cdot Y_{s,y} - p_n\cdot N_{s,y} - p_w\cdot W_{s,y} - C_s + S_s\right) \tag{5}$$

subject to:

$$\sum_{s}\mathbf{RM}_{c,t\times s,y}\cdot\mathbf{cr}_{s,y} \le \mathbf{q}_{c,t} \quad \forall\, c,t,y \tag{6}$$

where $\Pi$ is the objective function to be maximized and represents the present value of the net benefit from agricultural production (€) defined as crop revenues minus fertilizer and water variable costs; $A_s$ is the area cultivated for crop located at source $s$; $p_s$ is the crop price (€/kg); $Y_{s,y}$ is the production yield of crop located at source s at planning year y (kg/ha), that depends on the nitrogen fertilizer and irrigation water applied; $p_n$ is the nitrogen price (€/kg); $N_{s,y}$ is the fertilizer applied to crop located at source *s* at year *y* (kg/ha), $p_w$ is the price of water (€/$m^3$), and $W_{s,y}$ is the water applied to crop located at source s at each planning year *y* ($m^3$); $C_s$ is the aggregation of the remaining per hectare cost for crop located at source *s* (€/ha); $S_s$ are the subsidies for the crop located at source *s* (€/ha); r is the annual discount rate, $\boldsymbol{RM}$ is the unitary pollutant concentration response matrix where each column is the nitrate concentration for each crop area (*s*) times de number of years within the planning horizon (*y*), the number of rows equals the number of control sites (*c*) times the number of simulated time steps (*t*) in the frame of the problem; $\boldsymbol{q}$ is a vector of water quality standard imposed at the control sites over the simulation time (kg/$m^3$); $\boldsymbol{c_r}$ is a vector representing the nitrate concentration recharge (kg/$m^3$) reaching groundwater from a crop located at source *s*, which is obtained dividing the nitrate leached over the water that recharges the aquifer. $C_s$ is the aggregation of the remaining per hectare costs for crop located at sources (€/ha); Ss are the subsidies for the crop located at source *s* (€/ha).

The response matrix (***RM***) describes the influence of pollutant sources upon concentrations at the control sites over time. This is carried out by means of numerical simulation models based on the flow and solute transport governing equations. Specifically, in order to ensemble the pollutant response matrix we have used MODFLOW (McDonald and Harbough, 1988), a 3D finite difference groundwater flow model, and MT3DMS (Zheng and Wang, 1999), a 3D solute transport model. The hydro-economic framework takes into account different processes governing nitrate in groundwater in both the saturated and unsaturated zone of the aquifer. Then the agronomic, and the flow and solute transport model consider processes ranging from mineralization, nitrification, denitrification, volatization, immobilization, and plant uptake in the unsaturated zone to advection, dispersion, diffusion and biodegradation in the saturated zone.

The simulated time horizon corresponds to the time for the solute to pass all the control sites, and it is independent of the length of the planning period. Once the field of groundwater velocities is computed using the calibrated groundwater flow model, it is used by the calibrated mass transport model to compute the nitrate concentrations over time each control site (breakthrough curves) at resulting from unit nitrate concentration recharges at each pollution source. These concentration values are assembled as columns to conform the pollutant concentration ***RM***, which is a rectangular ($m$ $x$ $n$) matrix. The number of columns, $n$, equals the number of crop areas (pollution sources) times the number of years within the planning horizon. The number of rows, $m$, equals the number of control sites times the number of simulated time steps in the frame of the problem. The integration of the response matrix in the constraints of the management model allows simulating by superposition the evolution of groundwater nitrate concentration over time at different points of interest throughout the aquifer resulting from multiple pollutant sources distributed over time and space. The assumption of linearity of the system is required in order to apply superposition. The advective and dispersive transport depends on concentration and groundwater flow velocity. Because concentration is unknown, the use of both velocity and concentration as decision variables would create a non-linearity. As a result the underlying assumption is that the irrigation rate at each source is not a decision variable and has a known influence upon the velocity field.

Both nitrate leached and crop production are represented by polynomial regression equations depending on the water and fertilizer use (Peña-Haro et al., 2009). These equations are derived from the results of the agronomic simulation model GEPIC (Liu et al., 2007), a GIS-based crop growth model integrating a bio-physical EPIC model (Environmental Policy Integrated Climate) (Williams, 1995) with a GIS to simulate the spatial and temporal dynamics of the major processes of the soil-crop-atmosphere-management systems. The GEPIC package simulates crop growth using local conditions on climate, soil, irrigation water, tillage and other operations. The crop production function was introduced into the management model as follows:

$$Y_{s,y} = a + b \cdot W_{s,y} + c \cdot W_{s,y}^2 + d \cdot N_{s,y} + e \cdot N_{s,y}^2 + f \cdot W_{s,y} \cdot N_{s,y} \tag{7}$$

where $Y_{s,y}$ is the crop yield located at source s for a year *y* (kg/ha), $W_{s,y}$ is the water applied to the crop located at source *s* ($m^3$/ha) and $N_{s,y}$ is the fertilizer applied to the crop located at source *s* (kg/ha) within the year y. The nitrogen leached is defined as follows:

$$L_{s,y} = g + h \cdot W_{s,y} + i \cdot W_{s,y}^2 + j \cdot N_{s,y} + k \cdot N_{s,y}^2 + l \cdot W_{s,y} \cdot N_{s,y} \tag{8}$$

where $L_{s,y}$ is the nitrogen leached (kg/ha), $W_{s,y}$ is the water applied to the crop located at source *s* ($m^3$/ha) with in the year *y*, and $N_{s,y}$ is the fertilizer applied to the crop located at source *s* (kg/ha).

The non-linear optimization problem was coded in GAMS, a high-level modeling system for mathematical programming problems (GAMS, 2012), while the solver used was CONOPT (Drud, 1985). It is based on the Generalized Reduced Gradient algorithm designed for large programming problems.

## 3. Application to a 2D synthetic case

A two-dimensional synthetic aquifer was selected to verify the methodology. It is based on the configuration presented by Peña-Haro et al. (2009), which apply a deterministic formulation to a 2D homogeneous synthetic aquifer. In this case, however, we consider heterogeneous hydraulic conductivity. The aquifer has impermeable boundaries and steady-state flow from top to bottom of the domain (Fig. 1). The aquifer domain was divided in square cells of 500 x 500 m, with a grid made up of 58 rows and 40 columns. A confined aquifer has been modeled with a saturated thickness of 10 m, effective porosity of 0.2, and longitudinal dispersivity of 10 m. The natural annual recharge is 500 $m^3$/ha. A temporal discretization of 70 stress periods was used, each of one-year duration. In addition, seven different crop areas (which are the pollution sources in our model formulation) with five different crops are considered. For each crop a quadratic production function and a leaching function have been defined using Eq. (7) and (8). The calibrated coefficients of those quadratic functions can be found in Peña-Haro et al. (2009). The relationship between each source and the crop sown is depicted in Table 1, which also shows the irrigation rates needed by each crop. A recovery time horizon of the aquifer by the year 2015 is defined, which entails nitrate concentration lower than 50 mg/l for all control sites (three control wells are defined), as established by the EU water legislation.

Two scenarios have been simulated in order to compare the groundwater nitrate concentrations and net profits. In scenario 1 (S1) the fertilizer use is not constrained by groundwater nitrate concentration standards at the control wells. It represents the fertilizer application rates that return the maximum net benefits at each crop.

In scenario 2 (S2) the fertilizer used is constrained by the groundwater nitrate concentration standards (50 mg/l) at the control wells. Finally, a planning horizon of forty years was considered for each scenario with a constant annual fertilizer application for the whole time period.

An ensemble of a hundred conditional seed *K* fields is generated by means of sequential indicator simulation (SIS) by means of the code ISIM3D (Gómez-Hernández and Srivastava, 1990). The conditional fields represent the Case 1 (C1). In addition, an ensemble of unconditional *K* fields has also been generated in order to compare results (Case 2, C2). This geostatistical tool allows controlling the bivariate (2-point) statistics imposed on the simulated field instead of controlling a mere covariance model. As aforementioned, these seed *K* fields are subsequently used by the inverse model to constraint stochastic simulations to the available data. It should be mentioned that all seed *K* fields honour the *K* data, while during the stochastic optimization procedure carried out by the inverse model, the *K* fields are gradually modified to honour the flow and mass transport data. Eighty-four *K* data were defined as conditioning data. They have been selected to be homogeneously and spatially distributed over the whole aquifer domain. Moreover, they differ in several orders of magnitude to obtain highly heterogeneous aquifers. With the information provided by the *K* data we have defined nine deciles of the cumulative distribution to give a reasonable discretization of the local distribution and transformed to the corresponding indicator categories. The SIS uses an indicator kriging (Deutsch and Journal, 1997) to build up a discrete conditional cumulative density function (ccdf) for the individual categories at each case and the node is assigned a category selected at random from this discrete ccdf. For a continuous variable such as conductivity, indicator variables are built by comparing data values to a set of thresholds, $z_k$:

$$i(\mathbf{u}_{\boldsymbol{\alpha}};k) = \begin{Bmatrix} 1 \; if \; z(\mathbf{u}_{\boldsymbol{\alpha}}) \leq z_k \\ 0 \quad otherwise \end{Bmatrix} \tag{9}$$

Spatial continuity for the different thresholds was then evaluated using the standardized indicator semivariogram (e.g., Goovaerts, 1997):

$$\frac{\gamma_I(\boldsymbol{h}; z_k)}{\sigma_i^2} \approx \frac{1}{2N(\boldsymbol{h})}_{\boldsymbol{u}_1-\boldsymbol{u}_2=\boldsymbol{h}\pm\Delta\boldsymbol{h}} \sum [i(\boldsymbol{u}_1; z_k) - i(\boldsymbol{u}_2; z_k)]^2 \tag{10}$$

where $z_k$ are the thresholds values; $\sigma_i^2$ is the indicator variance given as $\sigma_i^2 = F(z_k)[1 - F(z_k)]$, and $F(z_k)$ is the marginal cumulative distribution function; $N(\boldsymbol{h})$ is the number of data pairs within the class of distance and direction; $\boldsymbol{h}$ is the separation vector; $z(\boldsymbol{u}_{1,2})$ denotes a measurement with $\boldsymbol{u}_{1,2}$ being the vector of spatial coordinates of the individual 1 or 2; and $\Delta\mathbf{h}$ is a tolerance vector. Finally, a mosaic variography was chosen for the generation of the seed *K* fields with the defined spatial continuity. It is spherical, with equal ranges in all directions of 40 m, 0.04 of nugget effect, and sill of 0.22. The generated seed fields are equally likely realizations, thus being plausible representations of reality since they display the same degree of spatial variability.

Note that we have focused in order to generate the heterogeneous fields on the hydraulic conductivity (*K*), since it is usually the parameter with the most significant spatial variation. One of these seed *K* fields has been chosen to be the true *K* field (i.e., it represents the actual heterogeneity in the aquifer). For the true field the hydraulic heads (*h*) are obtained and used as conditioning data for the inverse model. The spatial location of the 84 piezometric head data is the same than those defined for the *K* data.

The a priori conditional cumulative density function (ccdf) have been defined displaying a highly asymmetrical distribution with a long lower tail; thus assigning higher probabilities of occurrence to high *K* values (i.e., it could represent fracture structures). This is how the GC method allows integrating the available hard data and also the geological information. Note that GC method honors the a priori ccdf's during the whole conditioning process, while other inverse modeling techniques deal with secondary information incorporating it in initially generated fields to be perturbed, not having any implemented constraint to keep honoring these data. Furthermore, GC method tends to preserve the local ccdf's during the whole perturbation process of seed *K* fields to obtain conditional simulations see Llopis-Albert and Capilla (2009). This means that if there are zones with ccdf's belonging to independent stochastic processes they are still preserved. In addition, conductivities can vary, due to the deformation process, several orders of magnitude, and because of how the non-Gaussian feature is introduced in the inverse model by means of the probability fields allows the reproduction of strings of extreme values of *K* or preferential flow paths (Llopis-Albert and Capilla, 2009). Moreover, as many authors have pointed out (e.g., Gómez-Hernández and Wen, 1998), preferential flow pathways may play a crucial role for tracer transport and may reflect some geological settings, e.g., channelling.

Eventually, for each one of the calibrated *K* fields obtained with the inverse model the pollutant concentration response matrix is calculated, and the hydro-economic model run.

## 4. Results and discussion

Fig. 3 shows the Cumulative Density Functions (CFD's) of the maximum benefits obtained in the aquifer for both groundwater quality scenarios (S1 and S2) and both conditional (to *K* and *h* data) and unconditional conductivity realizations (i.e., cases C1 and C2). Note that a forty year planning period is considered and a recovery time horizon for the aquifer by the year 2015 has been defined, as established by the WFD. As logical, this figure shows that when no groundwater quality restrictions are applied in the optimization management model (i.e., for scenario S1) the same benefit is obtained for all realizations and both cases, i.e., conditional and unconditional *K* fields.

Moreover, it represents the maximum benefit that can be achieved in the aquifer with the defined set of parameters and variables. This is because farmers can applied as much as fertilizer as required by each crop to maximize their yields as defined in Eq. (7). Then, for scenario S1, the maximum benefit takes the value of 21.06 M€/year for all realizations and both cases, so that the standard deviations of such CDF's are zero. Contrary, for scenario S2, when groundwater quality constraints are applied (i.e., maximum nitrate concentrations of 50 mg/l are allowed at control wells) the benefits are reduced for all realizations and both cases. As expected, for conditional *K* fields (C1) and scenario S2, the uncertainty in the maximum benefits achieved in the aquifer is strongly reduced if compared with the unconditional realizations (C2). In fact, the CDF of the maximum benefits has a mean of 20.91 M€/year and a standard deviation of 0.074 M€/year for case C1, while it has a mean of 20.79 M€/year and a standard deviation of 0.22 M€/year for case C2. This proves the worth of the methodology to provide more reliable policies since it reduces the hydrogeological parameter uncertainty, and therefore, the uncertainty in the decision variables of the management model.

Similar results are obtained for the fertilizer applied to the aquifer as shown in Fig. 3. It depicts the Cumulative Density Functions (CFD's) of the total fertilizer applied to the aquifer for both groundwater quality scenarios and both conditional and unconditional conductivity fields. Then, the maximum fertilizer applied to the aquifer is obtained for S1. It has for both cases a mean of 3741.3 (ton/year) and a standard deviation of zero For S2, case C1 has a mean of 3502.03 (ton/year) and a standard deviation of 129.51 (ton/year), while C2 takes the values of 3546.34 and 57.01, respectively. Again, there is an important reduction in the uncertainty of the fertilizer applied to the aquifer.

Furthermore, the reduction of uncertainty is also depicted in Fig. 4, which shows the breakthrough curves for control well #1 and scenario S2 (with groundwater quality constraints of 50 mg/l) of all conditional conductivity realizations (C1). This figure shows how close are the time series of nitrate concentrations of all conditional realizations to the true $K$ field. Note that these concentrations are restricted by the management model to be lower, during the whole planning period, than the standard enacted by the WFD.

The trade-offs between the economic returns and the reliability in meeting the environmental standards can be compared for both scenarios. Fig. 2 and 3 show how the scenario S1 leads to higher benefits that the scenario S2. Then higher nitrate concentrations in the aquifer lead to lower benefits and vice versa. Therefore, these trade-offs have been quantified under the WFD standards. Furthermore, they have been quantified under an uncertain environment by means of the CDF of the agricultural benefits and their respective nitrate concentrations.

## 5. Conclusions

In recent years, the concern about nitrate concentrations in groundwater has increased on account of the intensive use of fertilizers in agriculture. Water legislations have dealt with such issue by establishing limits of nitrate concentrations in groundwater bodies. In Europe, the EU WFD establishes limits of 50 mg/l, and requires that groundwater bodies reach a good quantitative and chemical status by 2015. Then to control groundwater diffuse pollution is necessary to analyze and implement complex management decisions. However, the decision-making process is even more complex under uncertain environments and heterogeneous stakeholder’s interests. This uncertainty leads to different management policies with clear implications in reliability levels, costs and benefits. Therefore policy-makers need agricultural advice services to help them to come up with the best management practices. Such advices can be derived from the results provided by the tool here presented, which entails the coupling of a stochastic inverse model with a hydro-economic model. This allows reducing uncertainty by constraining stochastic simulations to data. The stochastic hydro-economic modeling framework has been verified in a 2D synthetic aquifer and its worth for agricultural advice services demonstrated. It has been proved to be a valuable tool in estimating non-Gaussian hydrogeological parameters such as the hydraulic conductivity in highly heterogeneous aquifers. This leads to reducing uncertainty in concentration distributions of contaminant plumes at control wells when reasonable amount of data is available. Finally, this is translated into a reduction of the uncertainty on the results of the hydro-economic model: maximum benefits, optimal strategies of spatial and temporal allocation of fertilizer use in agriculture and concentrations in the aquifer that meet certain groundwater quality standards. This has been carried out by means of a sensitivity analyses for conditional and unconditional *K* fields. Furthermore, the trade-offs between higher economic returns and reliability in meeting the environmental standards have been analyzed for different groundwater quality scenarios. The study of the least-cost alternative for meeting the environmental objectives is also important in order to justify potential time and objective derogation when disproportionate costs are identified (WFD, art. 4).

As a further research we could have considered different groundwater quality standards, recovery time horizons, different spatial structure of the conductivity fields, or different sets of flow and mass transport conditioning data (for instance, regarding the spatial location and/or temporal).

## References


- Bakr, M.I., te Stroet, C.B.M., Meijerink, A., 2003. Stochastic groundwater quality management: Role of spatial variability and conditioning, Water Resour. Res., 39(4), 1078, doi:10.1029/2002WR001745.

- Capilla, J.E., Llopis-Albert, C., 2009. Stochastic inverse modeling of non multiGaussian transmissivity fields conditional to flow, mass transport and secondary data. 1 Theory. Journal of Hydrology, 371, 66-74. doi:10.1016/j.jhydrol.2009.03.015.

- Deutsch, C.V., Journal, A.G., 1997. GSLIB: Geostatistical Software Library and User's GUide. New York, Oxford University Press.

- Drud, A., 1985. CONOPT: a GRG code for large sparse dynamic nonlinear optimization problems. Math. Program. 31, 153e191. doi:10.1007/BF02591747.

- EC, 1980. Council Directive 80/778/EEC relating to the quality of water intended for human consumption. Official Journal of the European Communities, L229 (30.08.1980), pp. 11-29.

- EC, 2000. Directive 2000/60/EC of the European Parliament and of the Council of October 23 2000 Establishing a Framework for Community Action n the Field of Water Policy. Official Journal of the European Communities, L327/1eL327/72. 22.12.2000.

- Gómez-Hernández, J.J., Srivastava, R.M., 1990. ISIM3D: An ANSI-C three dimensional multiple indicator conditional simulation program. Computer and Geosciences, 16(4), 395-440.

- Gómez-Hernández, J.J., Wen, X.H., 1998. To be or not to be multiGaussian? A reflection on stochastic hydrogeology. Advances in Water Resources 21 (1), 47–61.

- Hu, L.Y., 2000. Gradual deformation and iterative calibration of gaussian-related stochastic models. Mathematical Geology, Vol. 32, No 1, 2000, 87-108.

- Neupauer, R.M., Wilson, J.L., 1999. Adjoint method for obtaining backward-in-time location and travel time probabilities of a conservative groundwater contaminant. Water Resources Research, 35(11), 3389-3398.

- Llopis-Albert, C., 2008. Stochastic inverse modelling in non-multiGaussian media conditional to flow, mass transport and secondary data. PhD Thesis, Universitat Politècnica de València, 274 pp. ISBN: 978-84-691-9796-7.

- Llopis-Albert, C., Capilla, J.E., 2009. Stochastic inverse modeling of non multiGaussian transmissivity fields conditional to flow, mass transport and secondary data. 2 Demonstration on a synthetic aquifer. Journal of Hydrology, 371, 53–65. doi:10.1016/j.jhydrol.2009.03.014.

- Llopis-Albert, C., Capilla, J.E., 2009a. Gradual conditioning of non-Gaussian transmissivity fields to flow and mass transport data: 3. Application to the Macrodispersion Experiment(MADE-2) site, on Columbus Air Force Base in Mississippi (USA). Journal of Hydrology 371, 75–84. doi:10.1016/j.jhydrol.2009.03.016.

- Llopis-Albert, C., Capilla, J.E., 2010. Stochastic simulation of non-Gaussian 3D conductivity fields in a fractured medium with multiple statistical populations: case study. Journal of Hydrologic Engineering, 15, 554-566. doi: 10.1061/(ASCE)HE.1943-5584.0000214.

- Llopis-Albert, C., Capilla, J.E., 2010a. Stochastic inverse modelling of hydraulic conductivity fields taking into account independent stochastic structures: A 3D case study. Journal of Hydrology 391, 277–288. doi:10.1016/j.jhydrol.2010.07.028.

- Llopis-Albert, C., Pulido-Velazquez, M., Peña-Haro, S., 2014. Assessment of the effect of uncertainty on groundwater nitrate pollution control from agriculture. Environmental Modelling & Software, submitted.

- McDonald, M.G., Harbough, A.W., 1988. A Modular Three-Dimensional Finite-Difference Groundwater Flow Model, US Geological Survey Technical Manual of Water Resources Investigation, Book 6, US Geological Survey, Reston, Va, 586 p.

- Peña-Haro S., Pulido-Velazquez, M., Sahuquillo, A., 2009. A hydro-economic modeling framework for optimal management of groundwater nitrate pollution from agriculture. Journal of Hydrology, 373, 193–203, doi:10.1016/j.jhydrol.2009.04.024.

- Peña-Haro, S., Llopis-Albert, C., Pulido-Velazquez, M, Pulido-Velazquez, D., 2010. Fertilizer standards for controlling groundwater nitrate pollution from agriculture: El Salobral-Los Llanos case study, Spain. Journal of Hydrology, 392, 174–187.

- Peña-Haro, S., Pulido-Velazquez, M., Llopis-Albert C., 2011. Stochastic hydro-economic modeling for optimal management of agricultural groundwater nitrate pollution under hydraulic conductivity uncertainty. Environmental Modelling & Software 26, 999-1008.

- Poeter, E.P., Hill, M.C., 1997. Inverse models: a necessary next step in ground-water modeling. Ground Water;35(2):250–60.

- Rehfeldt, K.R., Boggs, J.M., Gelhar, L.W., 1992. Field study of dispersion in a heterogeneous aquifer 3. Geostatistical analysis of hydraulic conductivity. Water Resources Research, 28(12), 3309–3324.

- Renard, P., 2007. Stochastic hydrogeology: what professionals really need? Ground Water, 45(5):531–41. doi:10.1111/j.1745-6584.2007.00340.x.

- Salamon P., Fernàndez-Garcia, D., Gómez-Hernández, J.J., 2006. Modeling mass transfer processes using random walk particle tracking. Water Resources Research, 42, W11417, doi:10.1029/2006WR004927.

- Smith, L., Schwartz, F.W., 1981. Mass transport, 2. Analysis of uncertainty in prediction. Water Resources Research, 17(2):351–69.

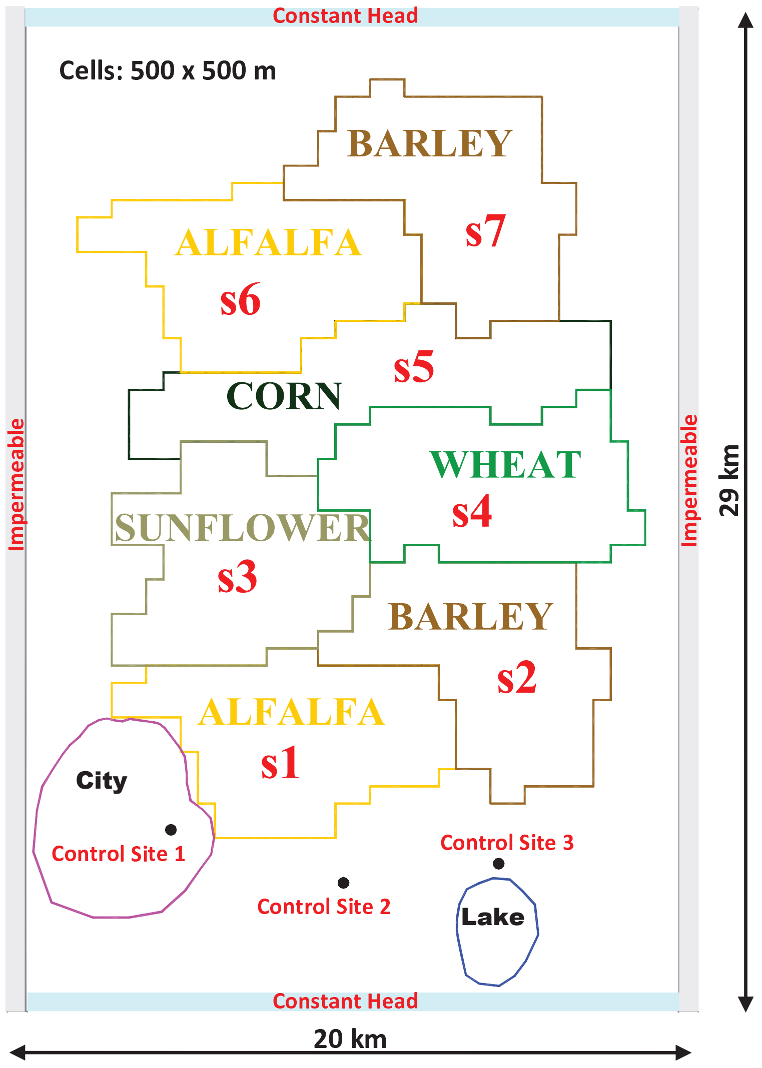
- Williams, J.R., 1995. The EPIC model. In: Singh, V.P. (Ed.) Computer Models of Watershed Hydrology, 909-1000. Water Resources Publisher.

- Zheng, C., Wang, P., 1999. MT3DMS: A Modular Three-Dimensional Multispecies Transport Model for Simulation of Advection, Dispertion and Chemical Reactions of Contaminants in Groundwater Systems; Documentation and User`s Guide.

Table 1. Sources, crops and irrigation.

| Source | Crop | Area (ha) | Water applied ($m^3$/ha) | Crop price (€/kg) |
|---|---|---|---|---|
| S1 | Alfalfa | 3600 | 950 | 0.09 |
| S2 | Barley | 3600 | 300 | 0.12 |
| S3 | Sunflower | 3600 | 400 | 0.30 |
| S4 | Wheat | 3600 | 250 | 0.13 |
| S5 | Corn | 3600 | 700 | 0.12 |

Figure1



Constant Head
Cells: 500 x 500 m
BARLEY
s7
ALFALFA
s6
s5
CORN
WHEAT
s4
SUNFLOWER
s3
BARLEY
s2
ALFALFA
s1
City
Control Site 1
Control Site 2
Control Site 3
Lake
Impermeable
Impermeable
29 km
Constant Head
20 km

Figure2

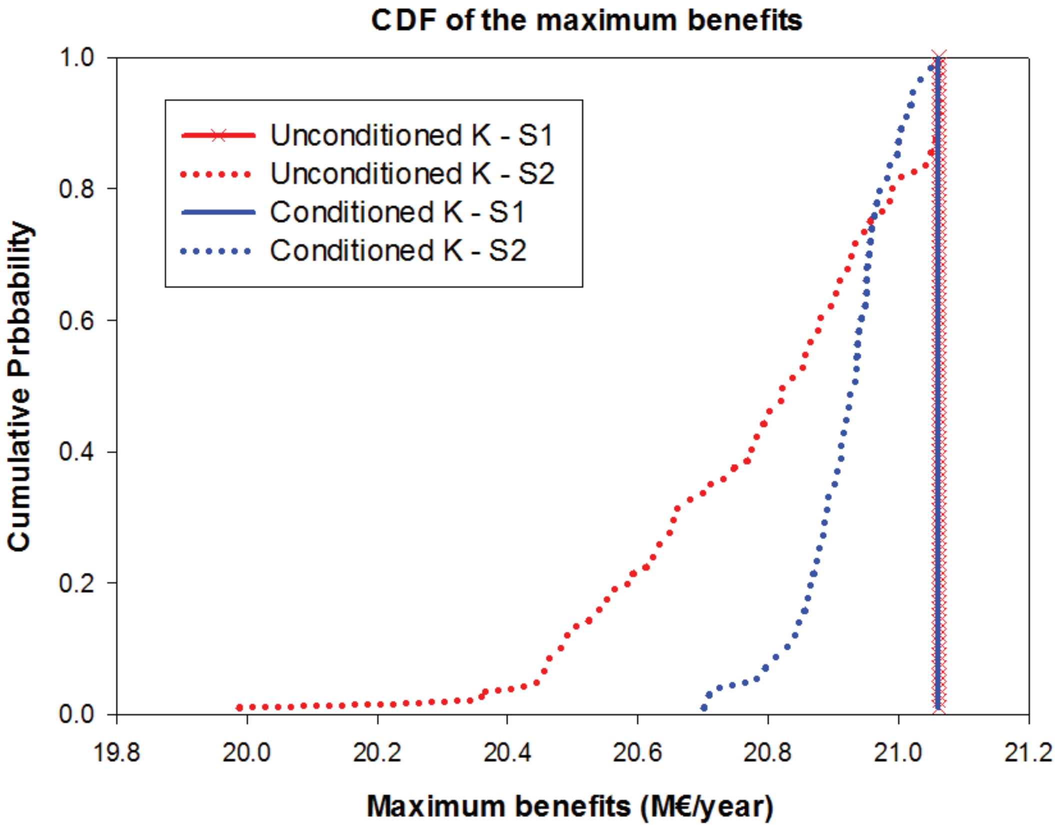
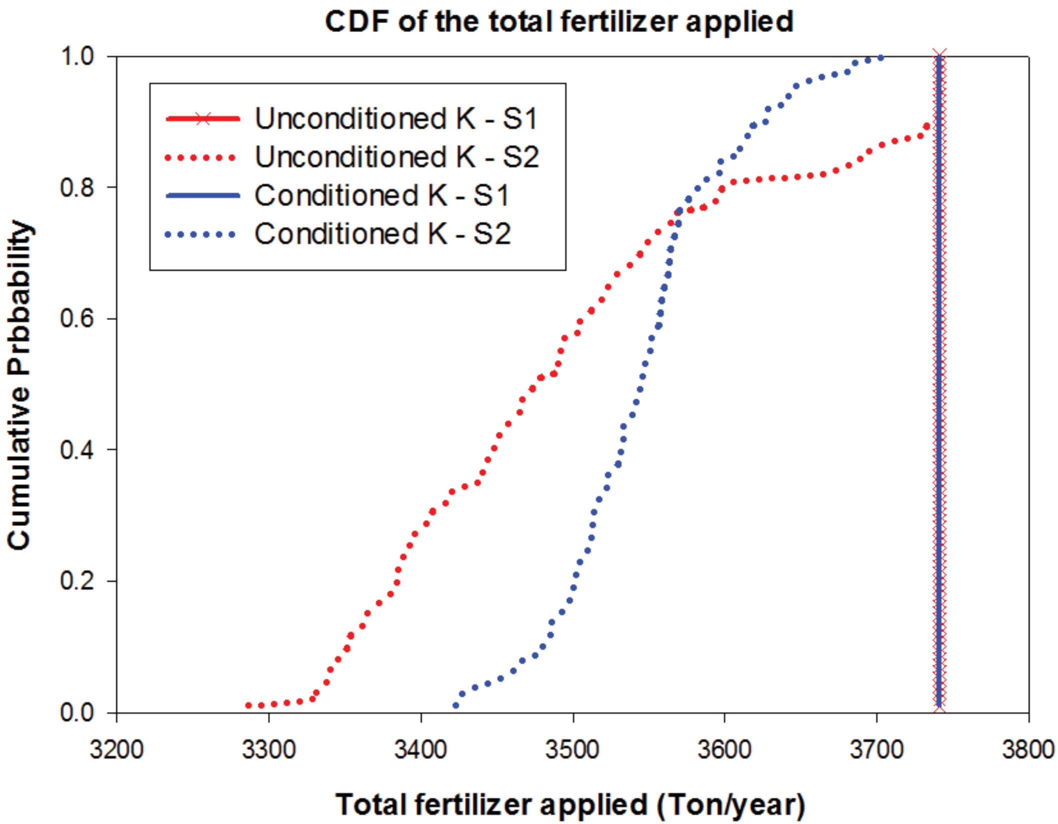
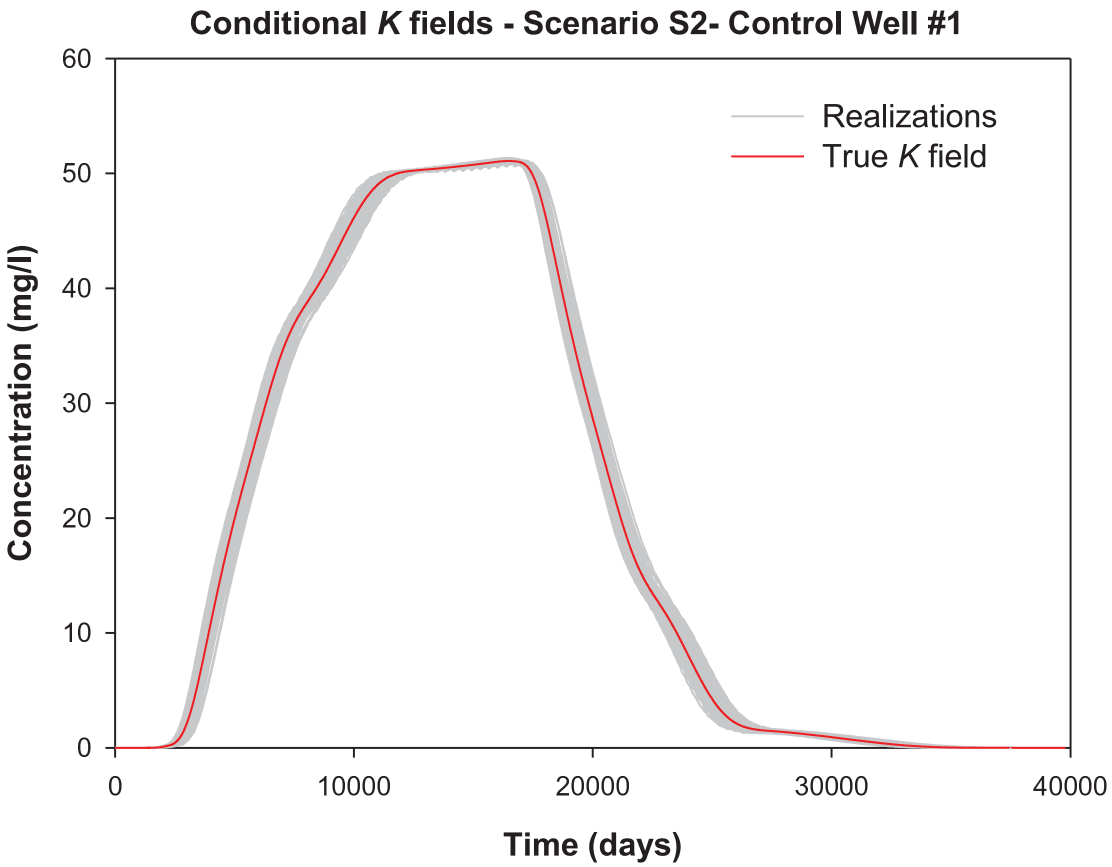

CDF of the maximum benefits
Unconditioned K - S1
Unconditioned K - S2
Conditioned K - S1
Conditioned K - S2
Cumulative Prbbability
0.0
0.2
0.4
0.6
0.8
1.0
19.8
20.0
20.2
20.4
20.6
20.8
21.0
21.2
Maximum benefits (M€/year)

Figure3



CDF of the total fertilizer applied
Unconditioned K - S1
Unconditioned K - S2
Conditioned K - S1
Conditioned K - S2
Cumulative Prbbability
1.0
0.8
0.6
0.4
0.2
0.0
3200
3300
3400
3500
3600
3700
3800
Total fertilizer applied (Ton/year)

Figure4


Conditional K fields - Scenario S2- Control Well #1
Realizations
True K field
Concentration (mg/l)
Time (days)
60
50
40
30
20
10
0
0
10000
20000
30000
40000

## Figure Captions

Figure 1. Problem domain, boundary conditions, control areas ($s_i$) and crops and spatial location of the observation sites.

Figure 2. Cumulative Density Functions (CFD's) of the maximum benefits (M€/year) for both groundwater quality scenarios and both conditional and unconditional conductivity fields.

Figure 3. Cumulative Density Functions (CFD's) of the total fertilizer applied to the aquifer (Ton/year) for both groundwater quality scenarios and both conditional and unconditional conductivity fields.

Figure 4. Breakthrough curves for control well #1 and scenario S2 (groundwater quality constraints of 50 mg/l) of the conditional conductivity fields. The figure also shows the true field. A planning period of forty years is considered.