Stochastic Environmental Resources and Risk Assessment manuscript No.
(will be inserted by the editor)

Spatio-temporal assessment of soil infiltration capacity physical-based models and geostatistical inference

Daniele Pedretti · Daniel Fernàndez-Garcia · Xavier Sanchez-Vila · Marco Barahona-Palomo · Diogo Bolster

the date of receipt and acceptance should be inserted later

Abstract Surface infiltration ponds (SP) provide an efficient and effective method to perform managed artificial recharge of aquifers whenever the soil hydraulic parameters are correctly predicted. In particular, the local-scale infiltration capacity varies in space and time and the pond-scale infiltration varies in time. Any estimation of these is uncertain due to complex clogging mechanisms and lack in the geological characterization of an infiltration sites. Here, we propose a method to combine geostatistical inference of spatial variability of soil characteristic with a temporally dependent physical model. In a field campaign we obtain experimental measurement of local-scale infiltration capacity at select locations within the SP as well pond-scale values of the global infiltration capacity during a large scale experiment. We adopt an exponential model to reproduce changes in the infiltration capacity with time. Accounting for experimental, site-dependent and model errors, we can map characteristic variables associated with the soil (e.g. soil-dependent clogging factors) with local uncertainty and the distribution of local infiltration

D. Pedretti · D. Fernàndez-Garcia · X. Sanchez-Vila
Hydrogeology Group, Department of Geotechnical Engineering and Geosciences, Technical University of Catalonia (UPC-BarcelonaTech), 08034 Barcelona, Spain,
E-mail: daniele.pedretti@upc.edu

M. Barahona-Palomo
Hydrogeology Group, Department of Geosciences, Institute of Environmental Assessment and Water Research (IDAEA-CSIC), 08034 Barcelona, Spain

D. Bolster Department of Civil Engineering and Geological Sciences, University of Notre Dame, Indiana, USA

capacity at different times during the infiltration. Assessing the spatial distribution of infiltration rates is important for the optimum scheduling of maintenance operations on the SP as well as for cost-effective designs of managed artificial recharge facilities.

Keywords artificial recharge, surface ponds, probabilistic, infiltration capacity, clogging, cokriging, satellite images

1 Introduction

Managed artificial recharge (MAR) practices include several methods that aim to recover and enhance groundwater quality and productivity of depleted aquifers [7]. Among them, the use of excavated artificial ponds, or surface ponds (SP) are common [1,31] . For SPs, the recharge is induced by flooding the excavation with available waters (e.g. reclaimed water, stormwater, river water), which percolate towards the subsurface by infiltration. In a properly designed facility evaporation and other losses are typically negligible. As the infiltration occurs through the natural ground, assessing the hydraulic properties of the soils is necessary for optimum design and management of SP facilities.

The critical variable to be determined in SPs is the infiltration capacity (I_c), which is the maximum capacity of the soil to infiltrate water at a given time. I_c depends on the hydraulic conductivity (K) and the porosity of the soil (ϕ). Several models exist to relate K and ϕ (e.g., the Kozeny-Carman relationship [23,8]). In a SP K or ϕ vary in space and change over time such that $I_c = I_c(\mathbf{x}, t)$.

To make educated management decisions regarding the operation of such facilities, two operation scales should be considered. At the pond scale and from a purely managerial perspective, the spatial average infiltration in the SP $I_c^e(t)$ is probably most important. It provides quantitative information about the total amount of actual infiltration in an integrated way. $I_c^e(t)$ is also important for qualitative purposes as it determines certain important characteristic times (such as the residence time of the water within the SP) that can influence other biological and

chemical processes. In addition, it is necessary to have the ability to predict when the average infiltration capacity falls below some predefined minimum threshold (e.g. I_c^{e*}) below which the system is no longer deemed to be efficient.

However, for optimal operation and maintenance (O&M) on the SP, the complete knowledge about the spatial and temporal distribution of the local scale infiltration capacity ($I_c(\mathbf{x}, t)$) is required [29]. Due to the occurrence of complex spatio-temporal processes in SPs, the estimation of the local infiltration capacity is uncertain; as such we suggest that a probabilistic approach to predict soil hydraulic parameters is advisable [24], which could then be integrated into a larger scale probabilistic risk assessment [33,6] or targeted information quest [15,25] .

Spatial variability of hydrogeological parameters is typically related to the ubiquity of heterogeneity in soils. Uncertainties naturally come about as it is next to impossible to completely characterize these geological heterogeneities [34].

The temporal variability of the infiltration capacity occurs due to natural and extreme events which affect the SP soils before or during operation. In a properly designed SP, accidental extreme events occur with low frequency/probability and are not easily predictable. Thus, including the occurrence of such extreme events is outside the scope of most design strategies and deemed to be external to the boundary of the problem at hand (e.g. P 103-104 [5]).

However, natural processes are a common cause of failure in SP facilities [29] and one of the dominant mechanisms that leads to operational is clogging of the porous media. Clogging leads to a reduction of the soil drainable porosity. Consequently it also reduces the hydraulic conductivity K and I_c. Clogging occurs to a variety of processes, including physical, biological and chemical mechanisms [4, 36,18]. Such processes act in tandem on the topsoil but they develop at different rates, depending on a variety of site-specific conditions and factors, such as:

- textural properties of the soil (e.g. characteristic grain size), which are normally heterogeneous [37,35];
- soil density [13];

– chemical heterogeneity [18];
– state variables such as temperature [11,12].

Although several published field and laboratory experiments exists that aim to understand and quantify the various clogging processes and their mutual interaction depending on the type of soil [30,22,19], predicting the clogging rates via cumulation of individual mechanisms remains, in the practice, very uncertain.

Often, in practice, lumped models that quantify clogging in terms of bulk parameters are used to simulate clogging, or better said, to predict I_c. A typical example is the use of exponential decay functions [20,30].

In this paper we propose a simple method to determine $I_c(\mathbf{x}, t)$ at all times, accounting for the fact that direct measurements of $I_c(\mathbf{x}, t)$ (primary data) are challenging to be collected. We combine traditional geostatistical inference on primary data with lumped physically based models and secondary information from alternative sources. As an illustrative example, we apply the method to a case study based on an experimental SP site in Spain.

2 Problem Statement

Let us consider a given artificial recharge site in which direct measurements of the local infiltration rate, $I_c(\mathbf{x}, t)$, are known at sparse locations and for a few discrete times, i.e., $I_c(\mathbf{x}_i, t_j)$ $\{i = 1, ..., m\}$, $\{j = 1, ..., n\}$. This is typically the case in most practical applications as the operation of an artificial recharge pond is typically done under flooding conditions and thereby any exhaustive characterization of the infiltration capacity is too costly and time-consuming. The support scale of the measurement is local and given, for example, by the size of an infiltrometer test. Under these conditions, we attempt to map the temporal evolution of the infiltration capacity at the site so that a better management of the artificial recharge pond can be undertaken.

As the amount of infiltration data is typically not enough to pursue any traditional geostatistical inference, most studies applied simple deterministic calcu-

lations to estimate the global infiltration capacity. Here, we take advantage that indirect (secondary) information about the infiltration capacity can be obtained from satellite images to depict the full spatial distribution of the infiltration capacity. Satellite images can provide valid information related to the soil hydraulic properties such as the moisture content of the soil, the vegetative canopy density on the ground or the infiltration capacity [10,17,26,28]. Evidently, a limitation of satellite images is that image-based methods are restricted to non flooded stages. Flooding tests [3] are normally used for large-scale assessment of the total infiltration capacity of the pond; however, these experiments do not give any information about the spatial distribution of the infiltration, which is, on the other hand, required for optimizing the O&M of an artificial recharge pond.

3 Methodology: Geostatistical Approach

We start by assuming that an extensively sampled variable which (at least partially) correlated with infiltration capacity is available at some support scale. An example would be data coming from intensity of a satellite image, which has been found to be correlated with infiltration capacity in an area dominated by clean sands [28]. Several geostatistical techniques can be used to incorporate such secondary information of an extensively sampled variable (pixel intensities) into the estimation of infiltration rates. The collocated cokriging model can be used for this purpose. Yet, in this case, its direct application is cumbersome as one needs to estimate the full time evolution of the variogram matrix as function of the clogging processes. To overcome this problem, we propose the following approach. Let us start by assuming that the temporal reduction of the infiltration capacity at every point $\mathbf{x}$ in the domain due to clogging processes follow a decaying exponential law [20,30], of the form,

$$I_c(\mathbf{x},t) = I_c(\mathbf{x},t=\infty) + R(\mathbf{x})I_c(\mathbf{x},t_0)\exp\left[-\lambda(\mathbf{x})t\right], \tag{1}$$

where $I_c(\mathbf{x}, t_0)$ is the infiltration capacity at some initial operation stage, λ is the lumped clogging coefficient; $R[-]$ is a generic, instantaneous reduction function of infiltration caused by gas production and other mechanisms [27, 7], and $I_c(\mathbf{x}, t = \infty)$ is the large-time asymptotic value of the infiltration capacity (such values should not be achieved in practical situations, since they suppose the absence of O&M in the facilities). We assume for simplicity that λ does not change with time and $R = 1$. This infiltration model is widely used in practice and well established in the literature [30, 22, 19]. The model assumes that clogging takes place in the order of tens of days. Temperature can therefore be neglected as it is acting in two temporal scales, seasonality and night-day fluctuations, that do not affect clogging occurrence. Seasonality occurs at temporal scales much longer than clogging occurrence, and night-day fluctuations are too short to affect clogging development permanently.

In Eq. 1, the temporal evolution of the infiltration capacity has been parametrized by two separate parameters: the initial infiltration capacity, $I_c(\mathbf{x}, t = 0)$, and the clogging coefficient, $\lambda(\mathbf{x})$. These parameters are spatially varying parameters. Based on this, our approach consists in cokriging these two parameters based on exhaustive secondary information stemming from satellite images. More specifically, to avoid matrix instabilities caused by densely sampled secondary data and reduce the burden of modelling all variograms and cross-variograms of a cokriging system, we employed the collocated cokriging model under the Markov model I [2, 21]. Once the spatial distributions of $I_c(\mathbf{x}, t = 0)$ and $\lambda(\mathbf{x})$ are known we can back calculate the spatio-temporal distribution of the infiltration rate, $I_c(\mathbf{x}, t)$, through Eq. 1.

4 An application

4.1 Site description: Sant Vicenç dels Horts Pilot Test

A pilot artificial pond is located in Sant Vicenç dels Horts, close to the city of Barcelona (Spain). The geological background is a sequence of deltaic deposits of the river Llobregat. A high resolution image was captured from Google Earth (fig. 1) in November, 15th 2007. It displays the ground conditions at the site before a flooding experiment was conducted; the resolution of the image is around 0.5 m^2 per pixel. The infiltration area (A) is approximately 100 m x 46 m. The unconfined and highly transmissive aquifer below the artificial pond is between 12 and 15 meters in thickness, and has been suffering from overexploitation and pollution in the last years due to an increasing population and industrial density in the Barcelona area [9, 14].

Two types of seepage information are required to operate on an artificial pond. While knowing the spatial distribution of local infiltration rates, $I_c(\mathbf{x}, t)$, at different times is fundamental to optimize future practical operations [29], an estimation of the total infiltration rate, I_c^e, provides worthwhile quantitative information about the maximum amount of infiltrated water, the averaged residence time of water in the artificial pond, and the time for occurrence of biological and chemical processes. The two infiltration quantities are related through

$$I_c^e(t) = \frac{1}{A} \int I_c(\mathbf{x}, t)\, d\mathbf{x} \tag{2}$$

where A is the total infiltration area.

These two types of estimates of the infiltration capacity were measured at the pilot test site. A large-scale flooding test was conducted to assess the total infiltration rate capacity of the pond. Knowing that the variations in the water level at the pond are typically small, neglecting evaporation, and considering that water infiltrates driven mostly by gravity, the infiltration capacity can be approximated

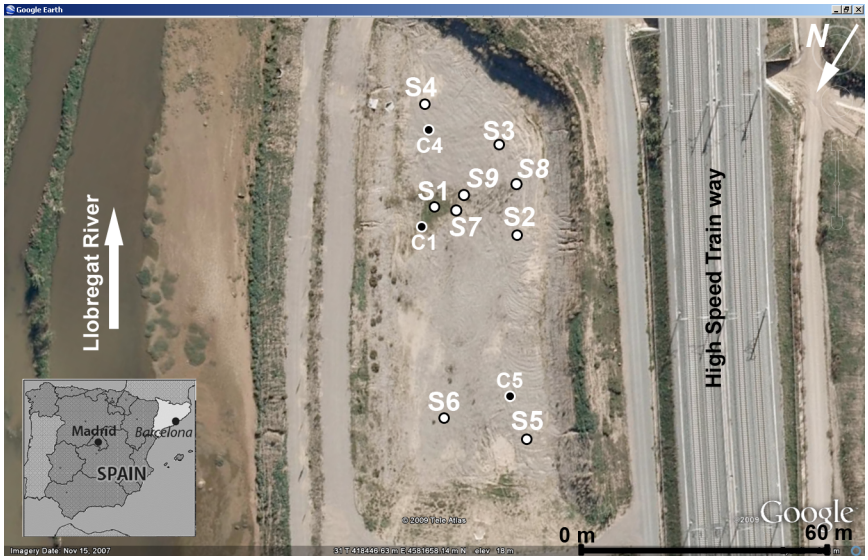

Fig. 1 Google Earth satellite image of the Sant Vicenç dels Horts EAP, close to Barcelona (property of ICC, 2007). S- tags refer to the locations of infiltrometer tests; C- tags are excavated pits to observe the local geological stratigraphy

as the ratio of the unitary discharge rate (Q_{in}/A) to the head gradient at the pond surface,

$$I_c^e(t) \approx \frac{Q_{in}(t)}{A(t)h(t)/L} \tag{3}$$

where Q_{in} is the water discharge rate, A is the infiltration area, h is the pond potential pressure head, and L is the distance from the surface to the water table.

The flooding test was performed between March and June 2009. Results are shown in Figure 2. We consider that preclogged conditions took place in February while clogging was mainly affecting the total infiltration capacity in June. Note that only infiltration related clogging is considered (no air dust clogging or extreme events modified ground conditions). The data have been been filtered out with a moving average algorithm that eliminates the noise associated with the daily temperature. Figure 2 shows that a good match between data and the infiltration model (Eq. 1) can be obtained. The solid line represents the discharge rate, while

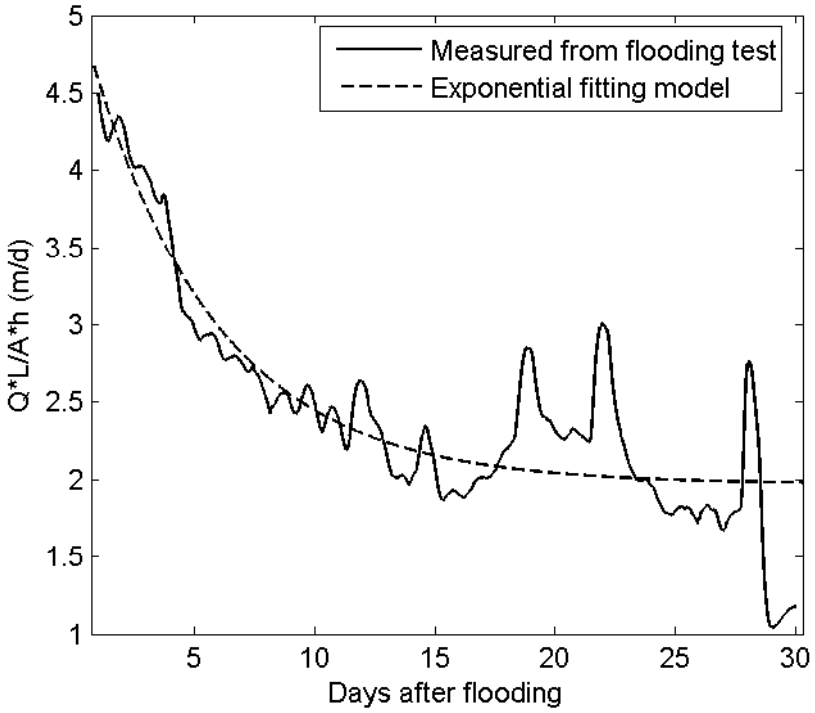

Fig. 2 Experimental observations (solid line) of $\overline{I_c}$ (expressed as in Eq. **??**) from the flooding test performed in the pilot SP between March 2009 and May 2009 . The dotted line is the exponential model. The Pearson's coefficient of the curve is $r^2 = 0.79$

the dotted line is the fitted model solution obtained with $R = 0.75$, $\lambda = 0.019\ d^{-1}$, and $I_c^e(t = \infty) = 1.98$ m/d.

In addition to this, at two different times (before and after the flooding experiment), 9 measurements of the local infiltration rate were obtained using a double-ring infiltrometer [32]. Table 1 summarizes the results of the double-ring infiltration tests. Double-ring tests provides a direct estimation of $I_c(\mathbf{x}, t)$ that is representative of a small support scale (≈ 0.12 m^2), as compared to the whole domain area (≈ 5000 m^2). Thus, in practice, we can assume them to be point values. The double-ring infiltrometer intends to reproduce the local soil transition from unsaturated to saturated conditions. Basically, two metal rings are keyed a few centimeters into the soil and filled with water (at constant or variable height).

After a standard time of two hours, the infiltration rate approaches a steady state value, identified as $I_c(\mathbf{x}, t)$ at that location $\mathbf{x}$ and time t.

The infiltration rates measured after the flooding experiment were used to determined quantitative measures of the clogging coefficient at the 9 different measurement locations by

$$\lambda(\mathbf{x}) = -\frac{1}{\Delta t} \ln \left(\frac{\Delta I_c(\mathbf{x}, t)}{R I_c(\mathbf{x}, t_0)} \right) \tag{4}$$

where $\Delta I_c(\mathbf{x}, t) = I_c(\mathbf{x}, t) - I_c(\mathbf{x}, t = 0)$ is the local maximum change of the infiltration rate at different locations in the domain, and Δt is the duration of the flooding test ($\Delta t = t - t_0$).

In fig. 1 the letters inside the symbols denote the specific locations where in situ infiltration experiments have been performed. The "S" letters denote the location of double-ring data values. The "C" letters refer to excavated pits where geology was directly observed and described. Results highlight that "S1" and "S7" are located in an area of low infiltration capacity before and after the flooding test.

The "S5" and "S6" are located initially ($t = 0$) in an area of high capacity, but after the flooding stage the infiltration observed at "S5" and "S6" locations decreased more than 90% and 50%, respectively. This difference is probably due to a small slide took place in surroundings and the soil was covered by a thin layer of very fine materials. Thus clogging of this point (and probably partially S6) was caused by a catastrophic event that doesn't follow Eq. 1.

Thus, results indicate that low infiltration values are less sensitive to clogging than other areas. Another important observation is that after flooding the infiltration capacity is more homogeneous (smaller variance) than its initial state. The effects of homogenization are somehow expected because of the soil selective clogging processes.

Table 1 Double ring test results. All values are expressed in m/d. Starred point are affected by error and are not used for the regression.

Location	February 2009	June 2009	Ratio (June/Feb)
S1	0.19	0.18	0.94
S2	2.6	2.1	0.80
S3	2.9	2.5	0.86
S4	3.3	1.1	0.33
S5	12.9	1.2	0.09*
S6	12.6	6.3	0.5
S7		0.17	
S8		3.04	
S9		0.75	
mean	5.74	2.07	0.36
variance	30.6	3.98	

4.2 Regression Models

The cokriging inference of $I_c(\mathbf{x}, t = 0)$ and $\lambda(\mathbf{x})$ based on exhaustive secondary information, i.e. the pixel intensity (P_v) obtained from satellite images, requires to establish two linear regression models. One that relates the pixel intensity with $I_c(\mathbf{x}, t = 0)$ an another that relates the pixel intensity with $\lambda(\mathbf{x})$. Following [28], an adequate regression model between pixel values of a Google Earth image and the infiltration capacity can be determined at the Sant Vicenç dels Horts test site. The highly detailed RGB image was taken before the flooding experiment and therefore corresponds to its initial state (without clogging). The volume support of the pixel value is close to the primary variables [28].

Results are plotted in Figure 3. The correlation between the red band image and the logarithm of the infiltration values is linear. Coefficients differ from February (initial conditions) to June (affected by clogging).

$$S = \ln(I_c) = a_s P_v + b_s + \varepsilon_s \tag{5}$$

where ε_s is the model error. A good linear correlation is obtained for both states of the artificial pond, with a Pearson's coefficient of $r^2 = 0.87$ in February and $r^2 = 0.89$ in June.

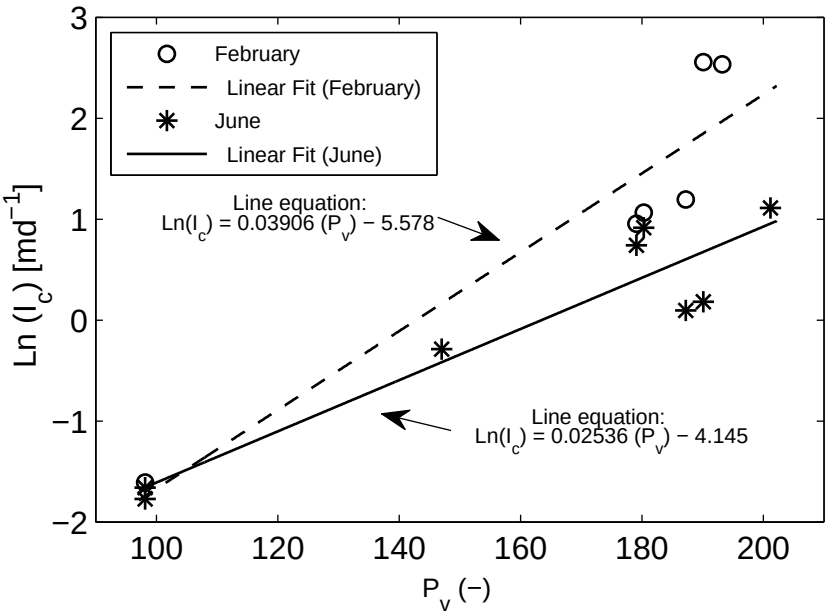

Fig. 3 Correlation between the pixel values (P_v) from the red band of the satellite image (fig. 1) and measured infiltration capacities (I_c) at selected locations in the SP at different temporal moments. The lines show the results from linear fitting to the experimental data. The Pearson's coefficient for the dotted line (February dataset) is r^2=0.8695; for the straight line (June dataset), r^2=0.8937

An adequate linear regression model between λ and P_v was also found. Results are shown in Figure 4. Formally, the linear regression model is of the following form

$$\lambda = a_\lambda P_v + b_\lambda + \varepsilon_\lambda. \tag{6}$$

The linear correlation coefficient is $r^2 = 0.90$, when excluding the λ value corresponding to the lowest P_v.

4.3 Variograms and Cross-variograms

The collocated cokriging system under the Markov model I limits the secondary variable to the data available at the estimation location, and further estimates

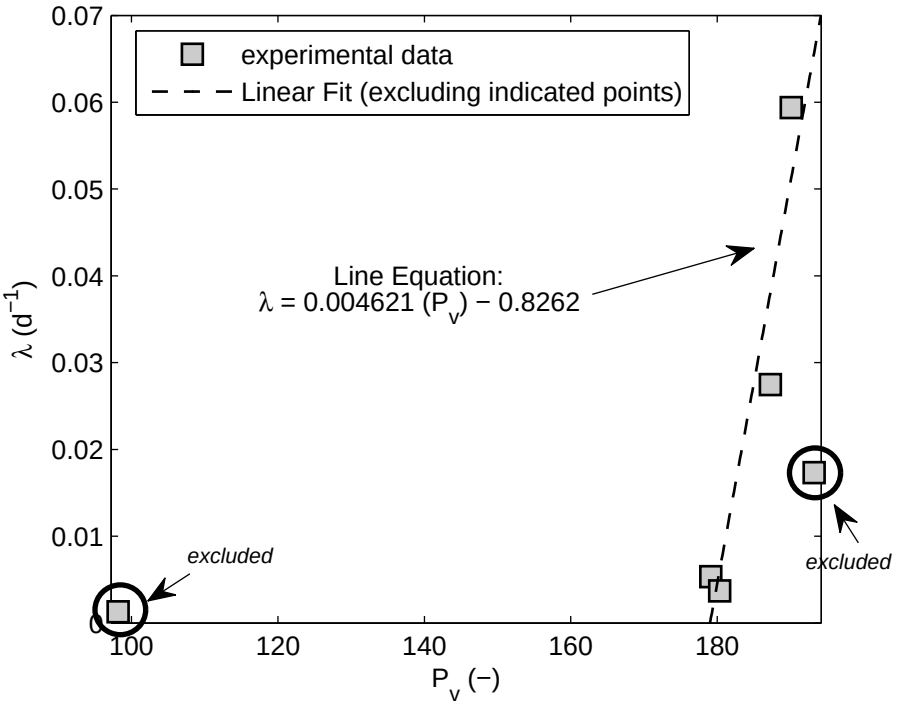

Fig. 4 Correlation between the pixel values (P_v) from the red band of the satellite image (fig. 1) and the clogging parameters (λ') from field measurements. Two points (highlighted and tagged) have been excluded from the fitting process because of the different reasons. r^2 for the straight line is 0.90

the cross-variograms by employing the underlying regression model. In this case, the cokriging model only needs the knowledge of the variogram of the primary variable, the correlation coefficient, and the variance of the secondary variable. Thus, the modelling effort is almost the same as for kriging one variable.

In our case, having too few infiltrometer experiments did not allow for a proper estimation of the auto-variogram of the infiltration capacity $I_c(\mathbf{x}, t = 0)$ or the clogging coefficient $\lambda(\mathbf{x})$. Instead, a well behaving sampled variogram was observed for the intensity of pixel values, P_v. Knowing that the primary and the secondary variables have the same support volume [28] and given the high correlation between them, we expect both correlation structures to depict similar patterns. For

this reason, and in an attempt to overcome this problem, we estimated the auto-variogram of the primary variables as

$$\gamma_S(\mathbf{h}) = a_s^2 \gamma_{P_v}(\mathbf{h}) + \gamma_{\varepsilon_\lambda}(\mathbf{h}) \tag{7}$$

$$\gamma_\lambda(\mathbf{h}) = a_\lambda^2 \gamma_{P_v}(\mathbf{h}) + \gamma_\varepsilon(\mathbf{h}) \tag{8}$$

where $\mathbf{h}$ is the lag distance between data values, and we have assumed that the regression model errors are uncorrelated with the secondary variable. The auto-variogram of the regression model error was modeled as a pure nugget effect, i.e., $\gamma_\epsilon \approx \sigma_\epsilon^2$.

Fig. 5 shows the resulting two directional variograms (N-S and E-W) from the secondary information. The fitted model variograms show three different structures, whose combined formulation is the following

$$\gamma_{P_v}(\mathbf{h}) = 297 \left[0.32 \cdot Gauss(\frac{\|\boldsymbol{h}\|}{7}) + 0.44 \cdot Sph(\frac{\|h_{NS}\|}{92}, \frac{\|h_{EW}\|}{68}) + 0.24 \cdot Exp(\frac{\|h_{NS}\|}{400}, 0) \right] \tag{9}$$

where Gauss, Exp and Sph are the standard unitary variogram models [16].

5 Results

The formulations in section 4.3 are now applied to the study area to obtain the spatial distribution of $\lambda(\mathbf{x})$ and thus to map $I_c(\mathbf{x}, t)$ at different time stages. Using Eq. 4 to the data in Table 1, we obtain six local values of the clogging factors ($\lambda(\mathbf{x})$). We used t=40 days. Such values are plotted in fig. 4 versus P_v.

Disregarding point S1 and S5 (tagged as *exluded*), we found a linear correlation between local $\lambda(\mathbf{x})$ and P_v given by Eq. 6 with a_λ = 0.004621 and b_λ =-0.8262. This correlation error has r^2=0.90. We already addressed the discussion for the anomalous behavior at the S5 areas because of failure of the system. We also

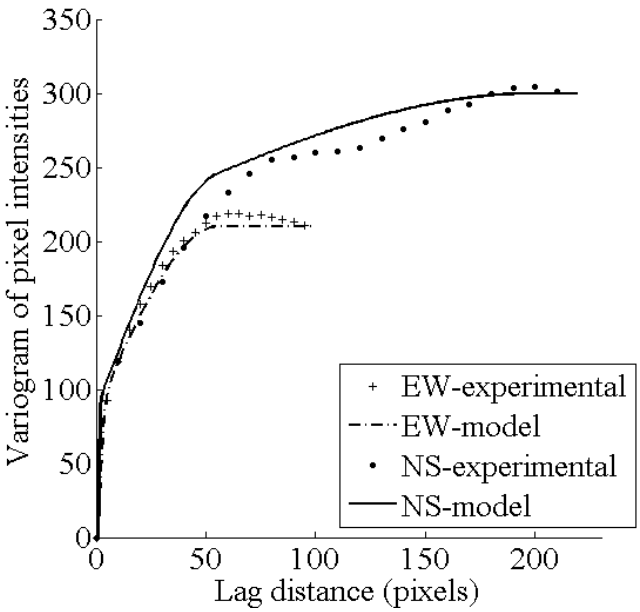

Fig. 5 Directional variograms of the pixel values in the reference image (exhaustive dataset)

suppose that λ at S1 is not representative for the statistical inference. Indeed, the difference between the two experimental datasets (see Table 1) is too small (only 0.01 m/d) and can be affected by measurements errors.

Nonetheless, a correction must be applied for the mapping process since the correlation equation would tend to negative values of λ for P_v less than 178.8 ($I_c = 0.63$ m/d). This happens because we only consider a partial subset of data for the regression. Evidently, negative values are not physically meaningful (they would let I_c increase in time, while the clogging model was defined as a negative monotonic function in Eq. 1).

For simplicity, we apply Eq. 8 to cokrige the map and then assign λ= 0.001 d^{-1} to the SP areas with P_v below 178.8. Still, since the areas below $P_v = 178.8$ have low I_c, their importance on the maintenance programs is secondary with regards to the higher I_c areas. Thus, this simplification should not influence significantly the purposes of the method. Of course, other considerations should be posed in case a larger dataset would be available for the statistical inference.

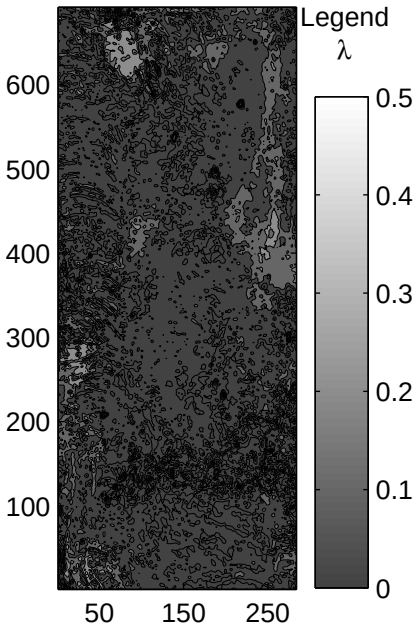

Fig. 6 Map of $\lambda(\mathbf{x})$ in the studied domain. Value are expressed in d^{-1}

The λ map is shown in fig. 6. We can see from the color bar that λ spans a range of values from 0.001 d^{-1} to 0.5 d^{-1}. Once $\lambda(\mathbf{x})$ is fully known, we can apply Eq. 1 to map the spatial distribution of $I_c(\mathbf{x}, t)$ at different temporal stages.

Three infiltration maps are shown in Fig. 7, considering t = 7, 14 and 21 days. We adopted an homogeneous R=0.75, which was found from the fitting process to the results of the flooding test (see section **??**). By simple comparison between the two extreme cases (from initial time to 21 days), we see that some zones (such as the left side of the map) darken more quickly than other areas. A more detailed analysis would suggest which areas in specific should be accounted for to optimize the use of the SP, and which time plan should be scheduled as O&M. Such decision can be combined with other factors economic, social and practical parameters to manage SP facilities in an efficient way.

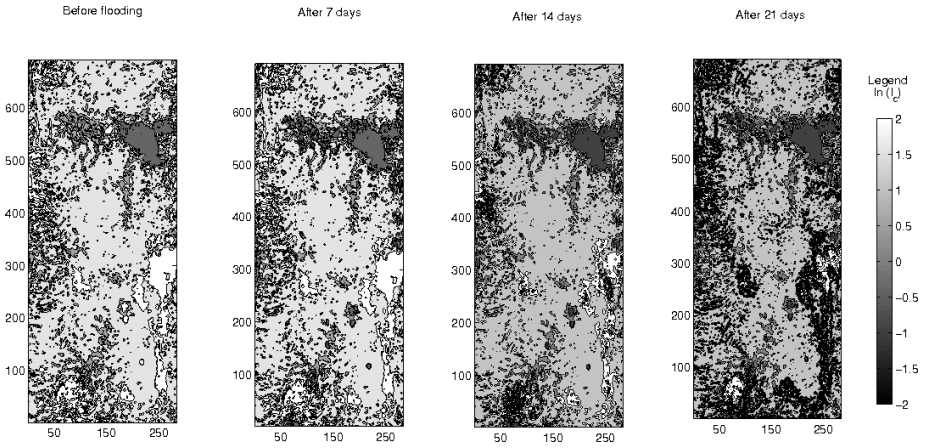

Fig. 7 Map of $I_c(\mathbf{x}, t)$ at the initial stage (t= 0 days, left), t = 7 days (center-left), t = 14 days (center-right) and t = 42 days (right) after a flooding event in the domain. Values are expressed in m/d

6 Conclusions

To take optimum, effective and efficient decisions in many branches of the groundwater hydrology, hydraulic variables governing the problem should be completely known in space and time. Under complex processes and in heterogeneous geological context, the estimation of such parameters is uncertain. This is mainly due to scarce primary information, to the high costs of direct field observation, to limited accessibility to the sites or even to the partially knowledge of the physical processes involved. Therefore, lumped models are often preferable to work with due to their simplicity and versatility. However, the spatial distribution of the characteristic lumped temporal factors is fundamental to estimate the spatio-temporal variability of the hydraulic parameters and to take decisions that might mitigate the risk in O&M operations by attacking specific locations (a subset of the total domain).

In this paper, we studied the spatial and temporal variability of the infiltration capacity of the soils (I_c) during managed artificial recharge of aquifers via excavated surface ponds. We coupled a simple exponential clogging model with the available initial distribution of local measurements about I_c; moreover, we added

knowledge from ancillary or secondary information in order to corroborate limited measurements about primary information. This method was applied on a pilot area in Spain. Despite some premises and simplifications were introduced in this illustrative example, we showed that we were able to to map the spatial distribution of the lumped clogging factor by means of a small dataset of measured λ and by means of secondary information from satellite image. With such information, we obtain different infiltration maps at selected time interval, which can be used to find the original strategy for clean-up operations that can minimize the effect of clogging without having to operate the full facility.

Acknowledgements This project was financially by the Spanish Ministry of Science and Innovation via the CONSOLIDER-Ingenio 2010 (CSD2009-00065), RARA-AVIS (CGL2009-11114) and HEROS (CGL2007-66748) projects. DP would like to further acknowledge the Spanish government for funding through the FPU scholarship programme (Formation of University Lecturers).

References

1. Management of aquifer recharge and subsurface storage. Netherlands National Committee - International Association of Hydrogeology, No. 4. NNCIAH Publication, 2003.
2. A. Almeida and A. Journel. Joint simulation of multiple variables with markov-type coregionalization model. *Mathematical Geology*, 26(5):565–588, 1996.
3. M. Barahona-Palomo, D. Pedretti, and X. Sanchez-Vila. Infiltration tests at the Sant Vicenç dels Horts artificial recharge experimental site. In EGU General Assembly 2010, editor, *Geophysical Research Abstracts*, volume 12 of *EGU2010-5326*, 2010.
4. P. Baveye, P. Vandevivere, B. L. Hoyle, P. C. DeLeo, and D. de Lozada-Sanchez. Environmental impact and mechanisms of the biological clogging of saturated soils and aquifer materials. *Critical Rev. Environ. Sci. Tech.*, 28(2):123–191, 1998.
5. T. Bedford and R. Cook. *Probabilistic Risk Analysis: Fundations and methods.* Cambridge University Press, New York, 2001.
6. D. Bolster, M. Barahona-Palomo, M. Dentz, D. Fernàndez Garcia, X. Sanchez-Vila, P. Trinchero, C. Valhondo, and D. M. Tartakovsky. Probabilistic risk assessment applied to contamination scenarios in porous media. *Water Resour. Res.*, 45:W06413, doi:10.1029/2008WR007551, 2009.

7. H. Bouwer. Artificial recharge of groundwater: hydrogeology and engineering. *Hydrogeol. J.*, 2002.
8. P.C. Carman. The determination of the specific surface of powders. *J. Soc. Chem. Ind. Trans*, 57:225, 1938.
9. J. Carrera, E. Vazquez-Sune, E. Abarca, B. Capino, D. Gamez, A. Simo, J.M. Ninerola, and E. Queralt. *El baix Llobregat. Historia i actualitat ambiental d'un riu*, chapter Les aigues subterranies del Baix Llobregat, pages 72–92. 2005.
10. M. Chica-Olmo and F. Abarca-Hernandez. Computing geostatistical image texture for remotely sensed data classification. *Computers & Geosciences*, 26:373–383, 2000.
11. F. Civan. Temperature effect on power for particle detachment from pore wall described by an Arrhenius-type equation. *Trans. Porous Med.*, 67(2):329–334, 2007.
12. F. Civan. Non-isothermal permeability impairment by fines migration and deposition in porous media including dispersive transport. *Transp. Porous Med.*, pages doi:10.1007/s11242–010–9557–0, 2010.
13. T. P. Clement, B. S. Hooker, and R. S. Skeen. Macroscopic models for predicting changes in saturated porous media properties caused by microbial growth. *Groundwater*, 34(5):934–942, 1996.
14. E. Custodio. Aquifer overexploitation: what does it mean? *Hydrogeology Journal*, 10:254–277, 2002. 10.1007/s10040-002-0188-6.
15. F. P. J. de Barros and Y. Rubin. A risk-driven approach for subsurface site characterization. *Water Resour. Res.*, 44:W01414, 2008.
16. C. Deutsch and A. Journel. Gslib: Geostatistical software library and user's guide. Technical report, Oford University PressNew York, 340 p., 1998.
17. R.J. Granger. Satellite-derived estimates of evapotranspiration in the gediz basin. *Journal of Hydrology*, 229:70–76, 2000.
18. J. Greskowiak, H. Prommer, G. Massmann, C. D. Johnston, G. Nützmann, and A. Pekdeger. The impact of variable saturated conditions on the hydrochemistry during artificial recharge of groundwater – A field study. *Appl. Geochem.*, 20:1409–1426, 2005.
19. A. Hoffmann and G. Gunkel. Bank filtration in the sandy littoral zone of lake tegel (berlin): Structure and dynamics of the biological active filter zone and clogging processes. *Limnologica - Ecology and Management of Inland Waters*, In Press, Corrected Proof:–, 2010.
20. T. Iwasaki. Some notes on sand filtration. *J. Am. Water Works Assn.*, 29:15971602, 1937.
21. A. Journel. Markov models for crosscovariances. *Mathematical Geology*, 31(8):955–964, 1999.

22. Jung-Woo Kim, Heechul Choi, and Yakov A. Pachepsky. Biofilm morphology as related to the porous media clogging. *Water Research*, 44(4):1193 – 1201, 2010. Transport and Fate of Colloids and Microbes in Granular Aqueous Environments.

23. J. Kozeny. Uber kapillare leitung des wassers im boden. *Sitzungsber. Akad. Wiss. Wien*, 136:271–306, 1927.

24. R. Krzysztofowicz. The case for probabilistic forecasting in hydrology. *Journal of Hydrology*, 249:2–9, 2001.

25. M. Masetti, S. Sterlacchini, C. Ballabio, A. Sorichetta, and S. Poli. Influence of threshold value in the use of statistical methods for groundwater vulnerability assessment. *Science of The Total Environment*, 407(12):3836 – 3846, 2009. Thematic Issue - BioMicroWorld Conference.

26. A. Milewskia, M. Sultana, E. Yanb, R. Beckera, A. Abdeldayemc, F. Solimand, and K. Gelil. A remote sensing solution for estimating runoff and recharge in arid environments. *Journal of Hydrology*, 373(1-2):1–14, 2009.

27. T.N. Olsthoorn. The clogging of recharge wells, main subjects. Technical report, Working group on recharge wells. Riswijk, The Netherlands. 136 pp., 1982.

28. D. Pedretti, M. Barahona-Palomo, D. Bolster, D. Fernàndez-Garcia, and X. Sanchez-Vila. Spatial assessment of infiltration capacity of soils for artificial recharge practices using Google Earth images. In *GeoENV Conference, Ghent University, Ghent (Belgium)*, 2010.

29. D. Pedretti, M. Barahona-Palomo, D. Bolster, X. Sanchez-Vila, D. Fernàndez-Garcia, and D.M. Tartakovsky. Probabilistic analysis of maintenance and operation of artificial recharge ponds. *Advances in Water Resources*, under review.

30. A. Perez-Paricio and J. Carrera. Clogging handbook. Technical report, Final report, EU project on Artificial Recharge of Groundwater., 1999.

31. B. R. Scanlon, K. E. Keese, A. L. Flint, L. E. Flint, G. B. Gaye, W. M. Edmunds, and I. Simmers. Global synthesis of groundwater recharge in semiarid and arid regions. *Hydrol. Processes*, 20:3335–3370, 2006.

32. RE Smith. The infiltration envelope: results from a theoretical infiltrometer. *Journal of Hydrology*, 17:1–21, 1972.

33. D. M. Tartakovsky. Probabilistic risk analysis in subsurface hydrology. *Geophys. Res. Lett.*, 34:L05404, doi:10.1029/2007GL029245, 2007.

34. D. M. Tartakovsky and C. L. Winter. Uncertain future of hydrogeology. *ASCE J. 394 Hydrologic Engrg.*, 13(1):37–39, 2008.

35. C. Tien and A. C. Payatakes. Advances in deep bed filtration. *AIChE J.*, 25(5):737–759, 1979.

427 36. P. Vandevivere, P. Baveye, D. S. de Lozada, and P. DeLeo. Microbial clogging of saturated
428 soils and aquifer materials: Evaluation of mathematical models. *Water Resour. Res.*,
429 31(9):2173–2180, 1995.
430 37. A. Zamani and B. Maini. Flow of dispersed particles through porous media – Deep bed
431 filtration. *J. Petrol. Sci. Eng.*, 69:71–88, 2009.