Towards Determining the Optimal Density of Groundwater Observation Networks under Uncertainty

Andreas Langousis, Vassilios Kaleris, Angeliki Kokosi and Georgios Mamounakis

Department of Civil Engineering University of Patras, Greece (andlag@alum.mit.edu)

1. Introduction

In most hydrologic applications, where spatial and temporal dependencies are present at various spatial scales, the stationarity assumption is a matter of the observation scale; see **Figure 1** below.

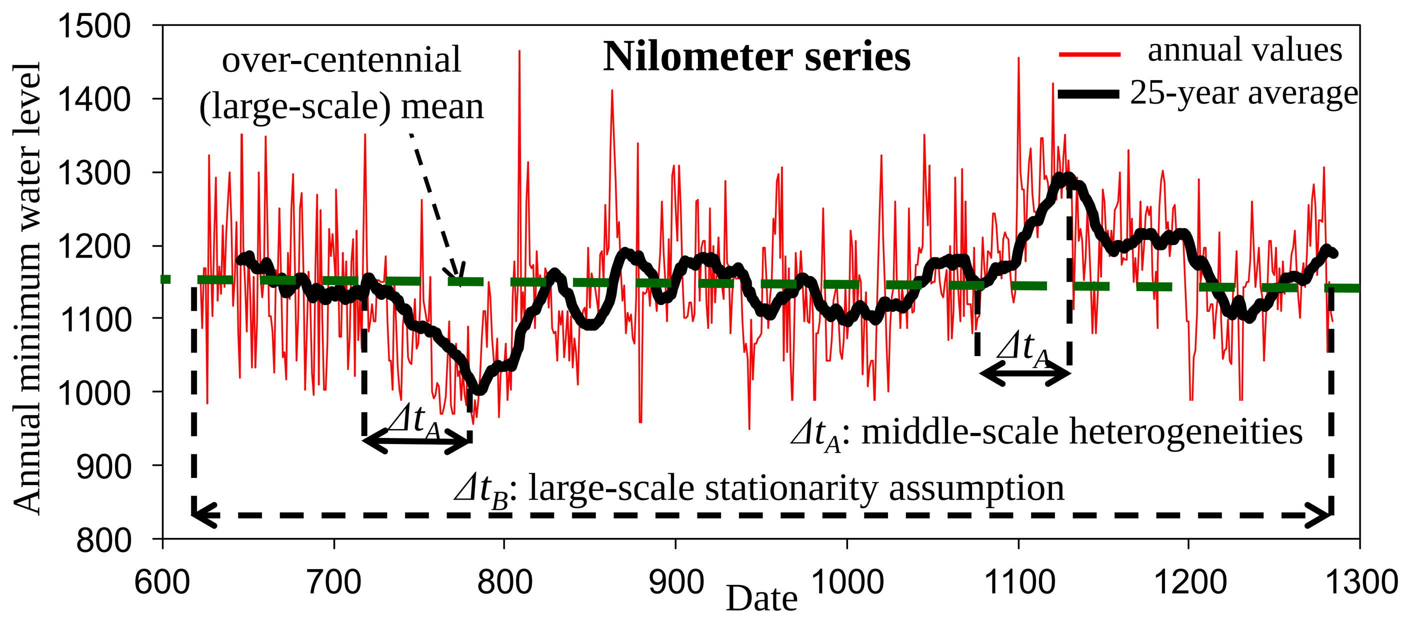

Figure 1: Annual minimum water level of the Nile River for the years 622–1284 AD (663 years; Beran, 1994).

Similarly, in ground water applications, where the spatial extent of geologic formations is unknown, one can model middle scale spatial heterogeneities using a stationary (i.e. homogeneous) field that exhibits long-range dependencies; see **Figure 2** below.

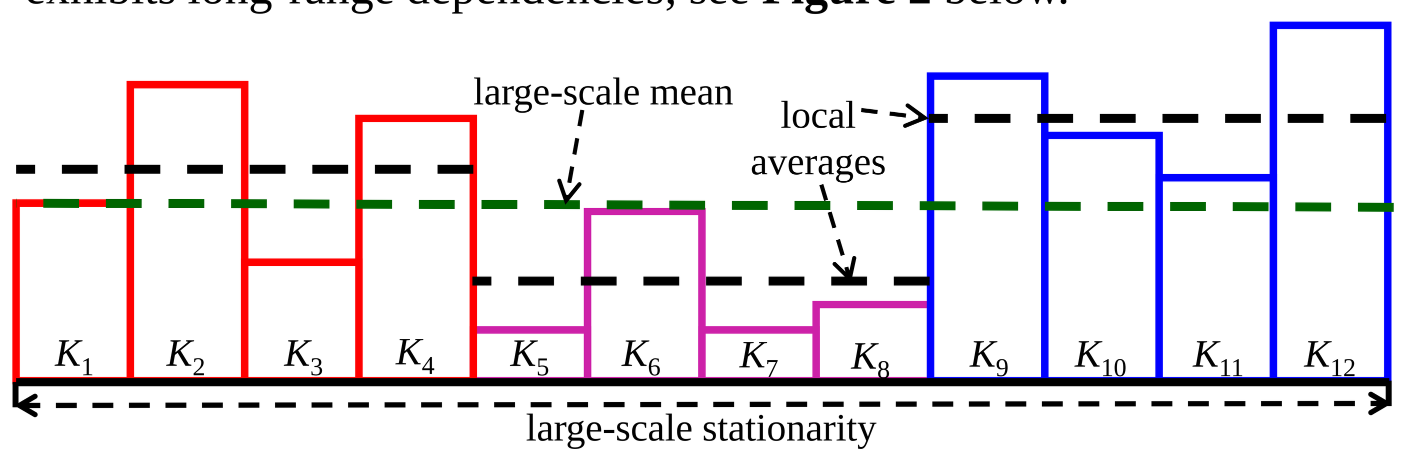

Figure 2: Schematic representation of a 1-dimensional (1D) stationary hydraulic conductivity field, demonstrating middle-scale heterogeneities (see local averages) due to long-range spatial dependencies.

➡ In what follows, we use a stationary approach, to study how spatial dependencies in geologic formations affect the error in hydraulic head estimation at ungauged locations.

2. Methodology

A. Hydraulic head estimation in confined aquifers: The case of 1D flow

Suppose an 1-dimensional (1D) confined aquifer of total length L, formed by r_{max} successive hydraulic conductivity layers of equal width $l_0 = L/r_{max}$; see **Figure 3** below.

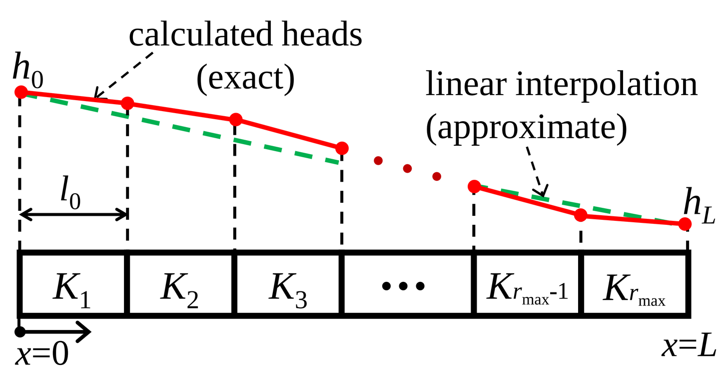

Figure 3: Schematic representation of the calculated (exact; red broken line) and linearly interpolated (green line) hydraulic heads in a 1D confined aquifer, formed by r_{max} successive hydraulic conductivity layers of equal width.

In the case when the hydraulic conductivities K_i, $i = 1, 2, \ldots, r_{max}$ are known, one can calculate the exact hydraulic head $h(x)$ at any location x in the direction of the flow (see red broken line in **Figure 3**), as:

$$h(x) = h_0 - q l_0 \left[\sum_{i=1}^{s} \frac{1}{K_i} + \frac{(x/l_0 - s)}{K_{s+1}} \right], \quad x \in [0, L]$$

where $s = \text{int}(x/l_0)$ is the integer part of the ratio x/l_0, $h_0 = h(x = 0)$, $h_L = h(x = L)$, and:

$$q = \left(\sum_{i=1}^{r_{max}} \frac{1}{K_i} \right)^{-1} \frac{h_0 - h_L}{l_0}$$

is the groundwater discharge per unit length of the aquifer span-wise (i.e. perpendicular) to the direction of the flow.

In the lack of hydraulic conductivity information, one can obtain an estimate $\hat{h}(x)$ of the standardized hydraulic head $h(x)$ by linearly interpolating between the two measuring locations:

$$\hat{h}(x) = h_0 - \frac{x}{L}(h_0 - h_L), \quad x \in [0, L]$$

For different values of r_{max} = 2, 4, and 8, **Figure 4** below shows plots of the mean value $m_{|e|}$ of the standardized absolute error:

$$|e(x)| = \left| \frac{h(x) - \hat{h}(x)}{h_0 - h_L} \right|, \quad x \in [0, L]$$

assuming that k_i ($i = 1, 2, \ldots, r_{max}$) are independent realizations drawn from a lognormal (LN) distribution with unit mean value, and coefficient of variation CV_K = 0.1, 0.5, 1, and 2.

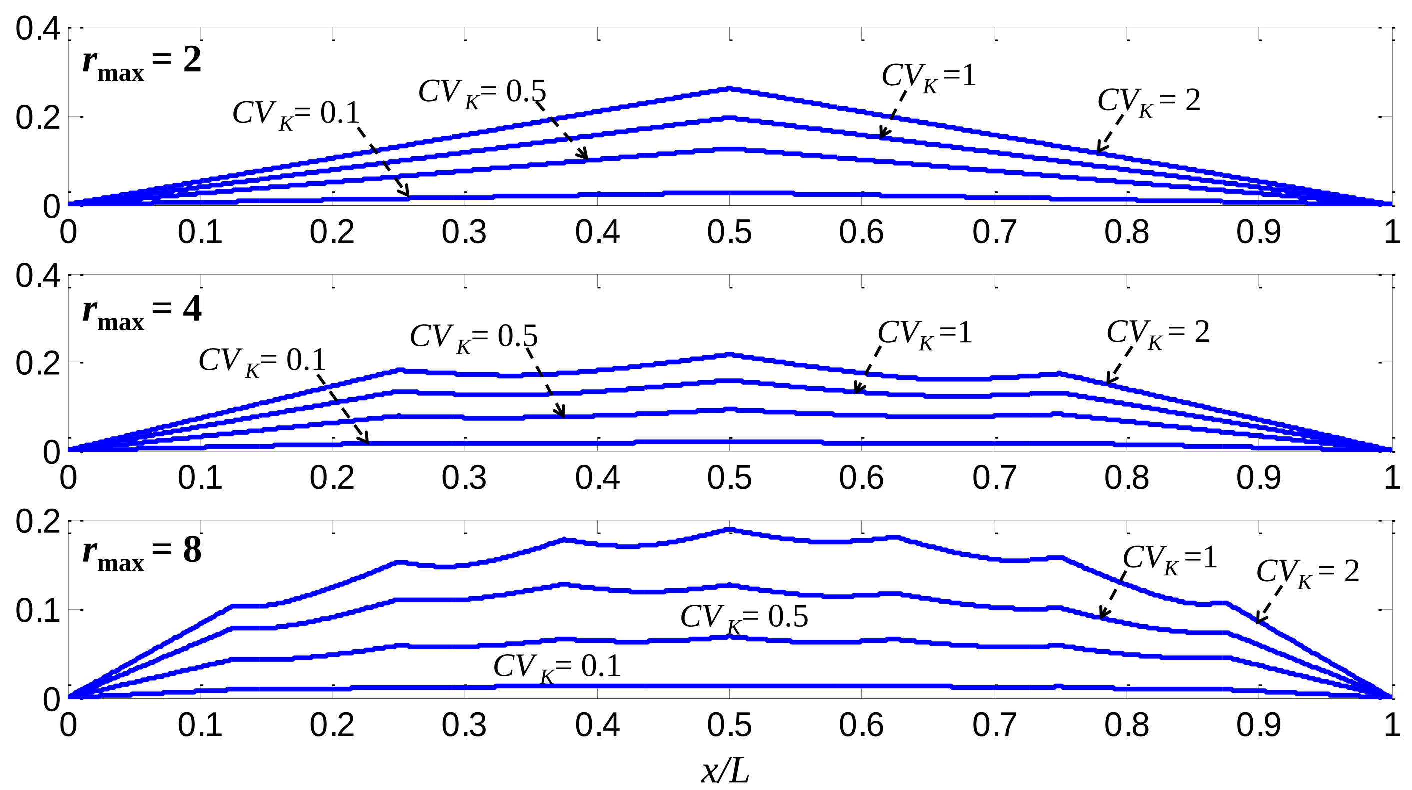

Figure 4: Mean value of the standardized absolute error $|e|$, as a function of the standardized distance x/L, for different number of hydraulic conductivity layers r_{max}= 2, 4, and 8. The corresponding curves have been obtained by ensemble averaging the results of 1000 Monte Carlo simulations, assuming that k_i (i= 1, 2, …, r_{max}) are independent realizations of a lognormal (LN) random variable with unit mean value, and coefficient of variation CV_K= 0.1, 0.5, 1, and 2.

It follows from simple mathematical arguments and considerations, that (see **Figure 4** for an illustration):

- The standardized absolute error exhibits local maxima at the interfaces of the hydraulic conductivity layers.
- In the interval [0, L], the maximum absolute error $|e|_{max}$ is located at the interface of the hydraulic conductivity layer that simultaneously maximizes the distance between the two measuring locations.

B. Introducing middle-scale heterogeneities as an attribute of spatial dependence

To study the effect of spatial dependencies as a function of scale on the statistics of the standardized maximum absolute error $|e|_{max}$, we introduce the discrete cascade representation shown below:

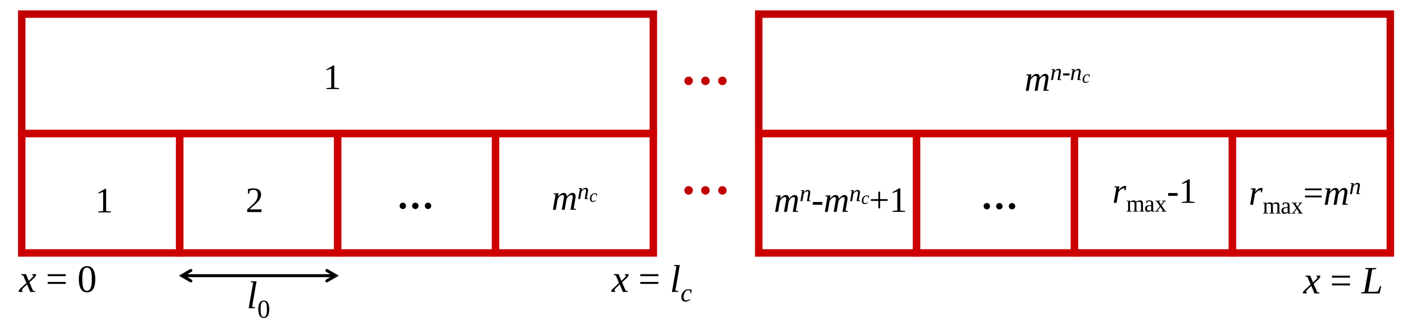

Figure 5: Discrete cascade representation of a 1D confined aquifer formed by $r_c = L/l_c = m^{n-n_c}$ independent pulses of constant length l_c and internal random structure.

In this representation, the geology of the aquifer between the two measuring locations is approximated by $r_c = L/l_c = m^{n-n_c}$ independent pulses of constant length l_c and internal random structure; $r_{max} = L/l_0 = m^n$ is the total length between the two measuring locations, m is an integer >1, and n, n_c are positive integers that satisfy n-$n_c \geq 0$.

Denote now by K_i ($i = 1, 2, \ldots, r_{max}$) the hydraulic conductivity of layer i at spatial scale l_0. The latter is the scale associated with the maximum resolution $r_{max} = L/l_0$. Many studies (see e.g. Freeze, 1975) have shown that for l_0 small (i.e. l_0 ~ 1 m used in laboratory samples) Kr_{max} can be modeled using a LN distribution, with mean value m_K and variance $(\sigma_K)^2$ that depend on the large-scale geology of the region.

Assuming stationarity of the hydraulic conductivity field at spatial scales l much larger that the inter-measuring distance L (i.e. $l >> L$), one has:

$$E[K_i] = E[K] = m_K, \ \text{Var}[K_i] = \text{Var}[K] = (\sigma_K)^2, \ i = 1, 2, \ldots, r_{max} = m^n$$

To facilitate comparisons, in what follows, we use Monte Carlo simulations to study the statistics of the maximum absolute error $|e|_{max}$, based on two limit cases:

Case 1: k_i ($i = 1, 2, \ldots r_{max}$) independent realizations of a mean-1 LN random variable K ~ LN (1, CV_K^2). ➡ **complete randomness**

Case 2: $k_{(i-1)r_c+1} = k_{(i-1)r_c+j}$, $\forall\ j = 1, 2, \ldots, r_c$, $i = 1, 2, \ldots r_{max}/r_c$, where $k_{(i-1)r_c+1}$ ($i = 1, 2, \ldots, r_{max}/r_c$) independent realizations of a mean-1 LN random variable K ~ LN (1, CV_K^2).

Case 2 corresponds to independent and spatially uniform hydraulic conductivity pulses with length $l_c = r_c\, l_0$, and random magnitude K ~ LN (1, CV_K^2). Note that Case 2 includes Case 1 as a sub-case, for r_c=1, or equivalently n_c=0.

In the proposed setting, one can parameterize the statistics of the standardized maximum absolute error $|e|_{max}$ using two dimensionless quantities (see below and **Figure 6**):

- **The coefficient of variation CV_K:** a measure for the small-scale variability of the hydraulic conductivity in the region. ⇨ **CV_K ~ 1÷2 for most cases**
- **The dependence ratio L/l_c:** a measure for the extent of apparent (i.e. observed) middle-scale heterogeneities (i.e. long-range spatial dependencies) in the aquifer. ⇨ from **geologic maps** or preliminary in-situ investigations

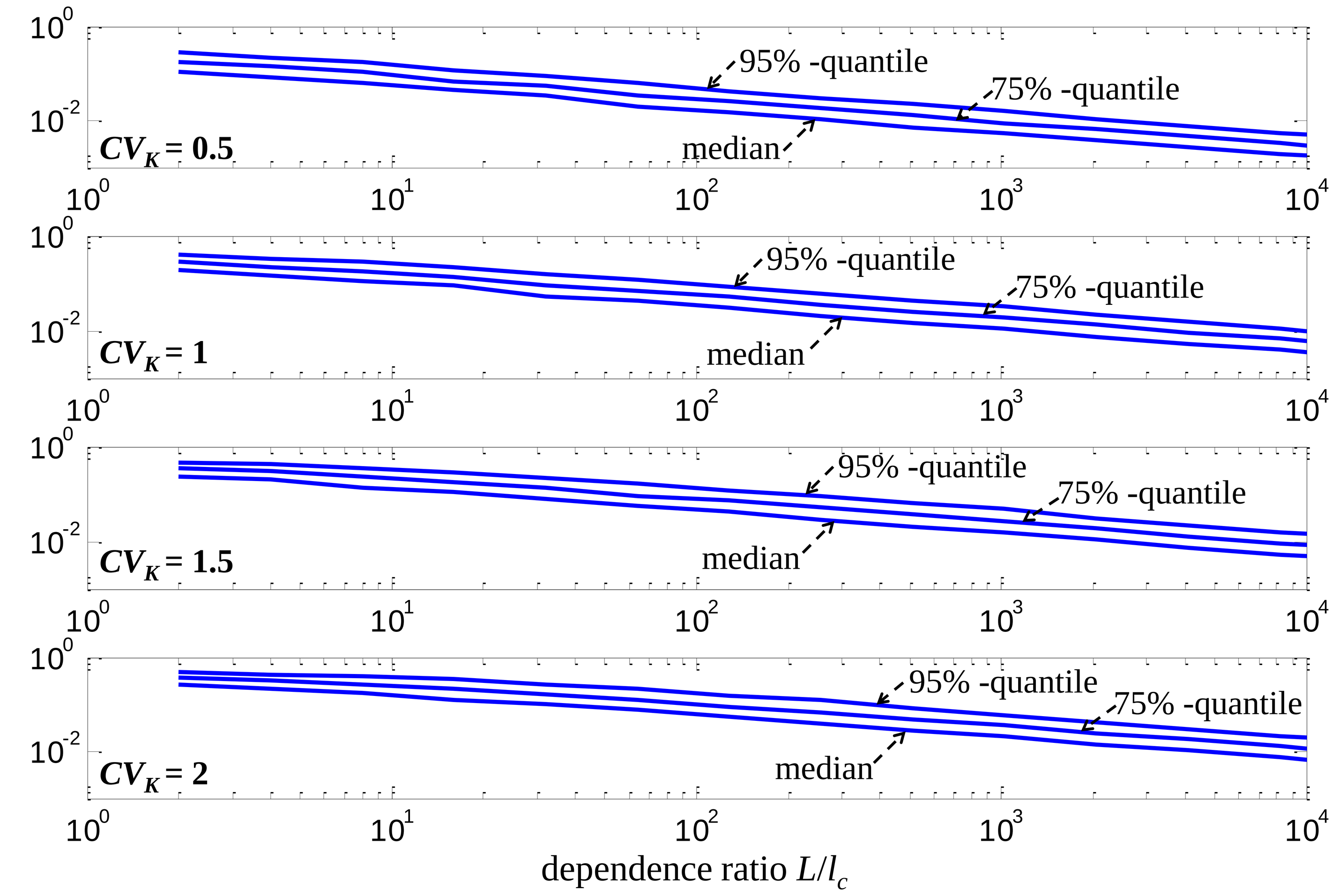

Figure 6: Quantiles of the standardized maximum error $|e|_{max}$ as a function of the ratio L/l_c for different values of CV_K = 0.5, 1, 1.5 and 2.

3. Results - Conclusions

- Rather than the small-scale variability of geologic formations, what is **most influential** in hydraulic head estimation is the **large-scale structure of the aquifer**. The latter is associated with **observed/apparent middle-scale heterogeneities** that cannot be easily dealt with ordinary regionalization approaches.
- To tackle this issue, we used the **classical notion of interior and exterior processes in an easy-to-handle stationary setting**, to **assess** the **accuracy** of **hydraulic head estimation** in aquifers that exhibit middle-scale heterogeneities.

References

Beran, J. (1994) *Statistics for Long-Memory Processes*, vol. 61 of Monographs on Statistics and Applied Probability, Chapman and Hall, New York, USA.

Freeze, R. A. (1975) A Stochastic-Conceptual Analysis of One-Dimensional Groundwater Flow in Nonuniform Homogeneous Media, *Water Resour. Res.*, **II**(5), 725-741.

Jones, P. D., K. R. Briffa, T. P Barnett and S. F. B. Tett (1998) High-resolution paleoclimatic records for the last millennium: interpretation, integration and comparison with General Circulation Model control-run temperatures, *Holocene* **8**(4), 455–471.

Acknowledgements

This work is sponsored by the Onassis Foundation under the "Special Grant and Support Program for Scholars' Association Members".