International Conference on Hydroinformatics

2014

# Fast Neural Network Surrogates For Complex Groundwater Flow Models

Niels Schütze

Tirthankar Roy

11th International Conference on Hydroinformatics
HIC 2014, New York City, USA

# FAST NEURAL NETWORK SURROGATES FOR COMPLEX GROUNDWATER FLOW MODELS

NIELS SCHÜTZE (1), TIRTHANKAR ROY (2), …
*(1): Institute of Hydrology and Meteorology, Technische Universität Dresden, 01062 Dresden, German*
*(2): The University of Arizona, Tucson, USA*

Surrogate modeling approach has been adopted in the study to replace computationally expensive physical-based numerical flow and transport model. Two approximate surrogate models namely, Artificial Neural Network (ANN) and Gaussian Process Model (GPM) are developed individually using a scenario database generated from the density dependent numerical flow and transport model OpenGeoSys (OGS). The state-space surrogates have the flexibility to move freely from one point to another within a time frame of decades and also to allow for moderate extrapolation in the case of extreme abstractions. The performance of the GPM was better in many cases with a little compromise on the computational time. A comparison of the performances of both the surrogates shows that they are reliable enough to be used in the management frameworks for decision making.

## INTRODUCTION

Increasing water demand in different consumer sectors has attributed to higher water stresses in the aquifers in various parts of the world (WBCSD [1]). In particular, arid coastal regions face mainly two types of problems as a result of increased water stress, namely the water table drawdown and seawater intrusion (Roy [2])]. As more water is abstracted from the aquifer, e.g. for agricultural use, the water table lowers down at the pumping stations. Seawater intrudes into the aquifer subsequent to the negative pressure gradients created because of high abstractions. Sea water contaminates the fresh water regions of the aquifer which results into the scarcity of fresh water in the aquifer. Robust management strategies are the indispensable needs for coastal aquifers to ensure both profit from agriculture as well as sustainability of aquifer.

In recent practice, management models couple a numerical groundwater model with an optimization algorithm to reach to the optimal solutions of the control variables (Das and Datta, [3], Nicklow et al., [4]). For solving such simulation-optimization problems evolutionary algorithms (EA's) are the most popular for solving groundwater management problems in arid coastal regions. A major drawback of EA's is the computational effort required for finding global optimal solutions (up to 2.5 Mio. model evaluations in Bhattacharjya and Datta [5]). In order to reduce the computational burden, surrogate modeling techniques are adopted (Bhattacharjya and Datta [5], Dhar and Datta [6], Papadopoulou et al. [7]). A surrogate model is an approximation to a more detailed and rigorous numerical model that, while used, reduces the computational effort significantly with a little compromise with the performance.

## METHODS

The present study aims to explore the potentialities of two different surrogates, namely ANN and GPM in the context of water management in arid coastal regions. The models are trained to be suitable for variable simulation time, meaning, they are time dependent (see Fig. 1).

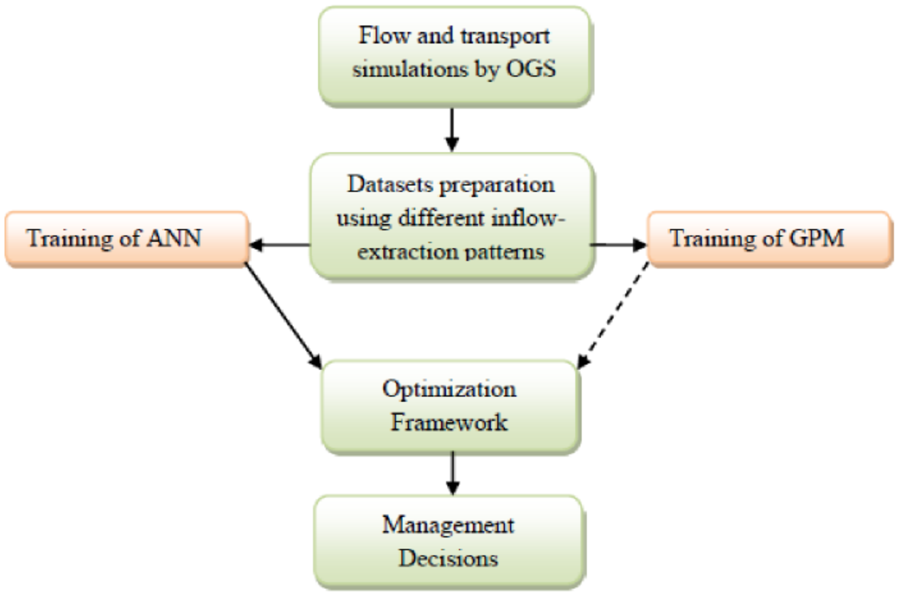


Figure 1. Basic framework of the surrogate optimization framework

This makes the models state-space and avoids the necessity of additional training when the simulation time is altered. The performance of ANN and GPM is carried out in a comparative study for a real life groundwater management problem (Fig 1).

### The state-space surrogate modeling framework

Two different surrogate models are developed in the study based on ANN and GPM to replace the computationally expensive numerical groundwater model OGS in the long term planning and management tool (Roy [2], Grundmann et al. [8]). Therefore, the surrogate state-space models are expected to provide the flexibility to shift from one point to another in the total time frame. Since the output variables of interest are a subset of the state variables of the numerical groundwater model the surrogate models can be seen as “state-space black-box” models and formulated as:

$$\mathbf{y}(t+1) = \mathbf{f}\left(\mathbf{y}(t), \mathbf{u}(t)\right) \quad \text{for} \quad t = 0 \ldots T \tag{1}$$

where $\mathbf{y}(t)$ denotes the output variables, i.e. groundwater levels $\mathbf{h}$ and salinity concentrations $\mathbf{S}$ at specific observation nodes, $\mathbf{f}$ the numerical groundwater model and $\mathbf{u}(t)$ the known external input, i.e. groundwater abstraction at different sites (see Fig 2). The surrogate models are used to emulate the numerical model $\mathbf{f}(\mathbf{y}(t), \mathbf{u}(t))$ by an approximation model $\tilde{\mathbf{f}}(\mathbf{y}(t), \mathbf{u}(t), \boldsymbol{\theta})$, with parameters $\boldsymbol{\theta}$. The parameters have to be determined by some form of training or optimization, e.g. by using a least-squares framework, minimizing the error of the differences:

$$\theta^* = \arg\min_{\boldsymbol{\theta}} \mathbf{f}_{err}\left(\mathbf{f}(\mathbf{y}(t), \mathbf{u}(t)), \tilde{\mathbf{f}}(\mathbf{y}(t), \mathbf{u}(t), \boldsymbol{\theta})\right) \tag{2}$$

using a certain error function $\mathbf{f}_{err}$.

Subsequently, the long term planning and management tool combines the trained surrogate model instead of the numerical model with a mono-criteria evolutionary optimization algorithm CMA-ES (Hansen [9]) to solve the groundwater management problem (see Fig. 1). Replacement of the numerical model with an approximate surrogate reduces considerable computational time and thus, improves the computational performance of the tool enormously.

## EXAMPLE APPLICATION

### Study area

The current study is carried out in the Barka area of Al-Batinah region situated at the north of the Sultanate of Oman as shown in the Figure 2. Barka lies at the north of the middle wadi catchment out of the three catchments in Al-Batinah. Northern part of the region is the downstream side with sea as the zero pressure boundary and the southern part is the upstream side with Hadjar mountain acting as the inflow boundary. It has a width of almost 20 km from the sea to the inside mountains and it stretches up to a length of 30 km along the coast. Agricultural farms are situated within a distance of 6 km from the coastline.

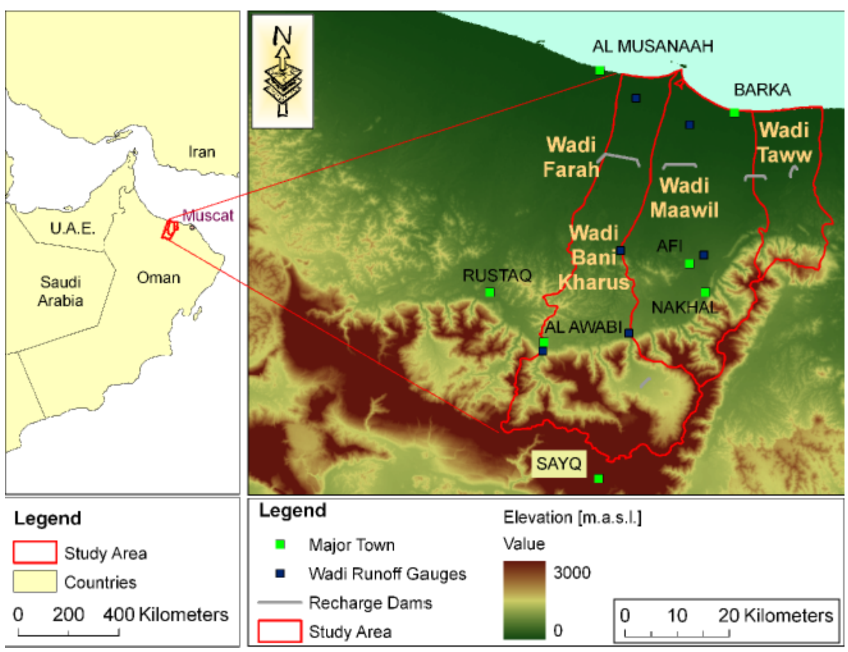


Figure 2. Study area: Al-Batinah

Most of the irrigation techniques till date are inefficient which leads to excess water pumping resulting large drawdown in water table. The increasing trend of extraction was found in the region since 1960s although the amount of the total extraction was comparatively lower at the beginning. Extraction rate from Barka region was reported to be 30 Mm3/year in 1970, whereas, now the rate is almost 120 Mm3/year (Al-Shoukri [10]).

## RESULTS

Both the selected models from the previous sections were applied to a real life problem. The 60 years simulation runs were started with the initial conditions of the aquifer (Roy [2]. Spatial variation of the profiles of the water table and salinity for a lower (55 Mm3/year) and a higher (595 Mm3/year) extraction cases are plotted over space in Figure 3 and Figure 4, respectively.

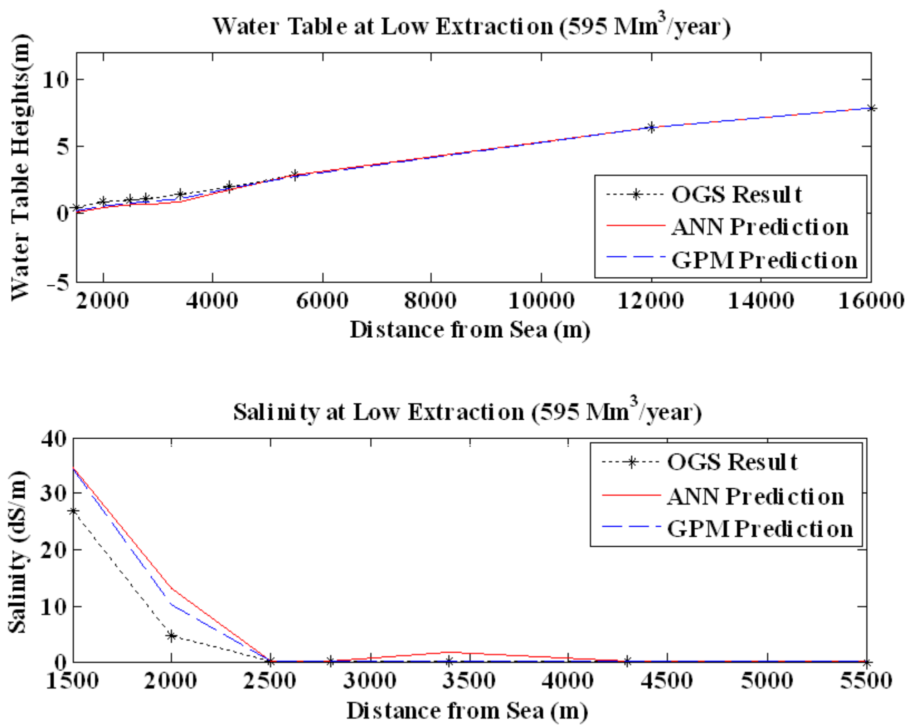


Figure 3. Comparison of OGS, ANN and GPM results (Lower extractions)

The performance of both models was compared with the results of actual OGS model. It is seen that in low extraction case, the prediction of GPM is better than ANN for both variables water table and salt concentration. In high extraction case, the water table is well approximated by GPM but there is a little bit of under prediction in the salt concentration. ANN was almost 200 to 300 times faster than the numerical model OGS. However, the computational time of GPM was higher than of ANN.

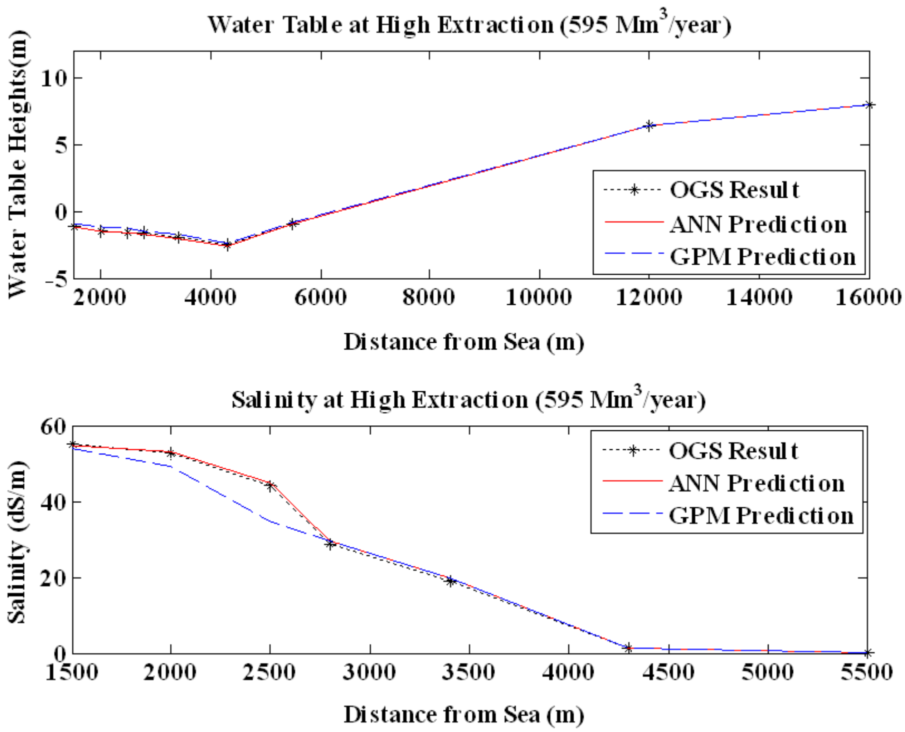


Figure 4. Comparison of OGS, ANN and GPM results (Higher extractions)

The results clearly substantiate the applicability of surrogate models in the management framework of coastal aquifers. If properly modeled, the surrogates can be used quite efficiently as they mimic the physically-based flow and transport model quite efficiently.

## CONCLUSIONS

Two surrogate models namely, ANN and GPM were trained with the data generated from the physics based groundwater model OGS. For data preparation, a 2D transient density dependent groundwater flow and salt transport model was developed in OGS with an assumption of two distinguishing layers in the aquifer. An adaptive pumping scheme was introduced to make the model even more realistic for the problem domain. The surrogates are prepared as state-space models that make it possible to start and stop the simulations at any point in the time frame and the models can also run beyond the time frame depending on the extreme values of abstractions. The performances of both ANN and GPM were compared with the results of OGS.

ANNs are quite familiar in the field of hydrologic modeling. However, GPM is an entirely new approach and till date, there have been very few research works reported on this topic in the field of hydrologic modeling. As seen from the results, GPM has high potential to outperform ANNs in many cases and so, it is quite promising for the near future. The only drawback of GPM lies in its relatively higher computational time. However, further researches can be carried out on improving the computational time of GPM in order to make it faster. The surrogate models can be used for undertaking management decisions for the arid coastal aquifers. The applications of the models are not only limited to the coastal regions, rather, they can be extended to other areas as well.

**REFERENCES**


[1] WBCSD. Facts and trends:water. Technical report, World Business Council for Sustainable Development, 2006.

[2] T. Roy. Optimal water management in arid coastal regions with surrogate models: A simulation-optimization approach. Master's thesis, Indian Institute of Technology Kanpur, 2012.

[3] A. Das and B. Datta. Application of optimisation techniques in groundwater quantity and quality management. Sadhana - Academy Proceedings in Engineering Sciences, 26(Part 4):293–316, 2001. doi: 10.1007/BF02703402.

[4] J. Nicklow, P. Reed, D. Savic, T. Dessalegne, L. Harrell, A. Chan-Hilton, M. Karamouz, B. Minsker, A. Ostfeld, A. Singh, E. Zechman, and ASCE Task Comm Evolutionary. State of the Art for Genetic Algorithms and Beyond in Water Resources Planning and Management. Journal of Water Resources Planning and Management–ASCE, 136(4):412–432, 2010. doi: 10.1061/ (ASCE)WR. 1943-5452.0000053.

[5] R. K. Bhattacharjya and B. Datta. ANN-GA-Based Model for Multiple Objective Management of Coastal Aquifers. Journal of Water Resources Planning and Management–ASCE, 135(5): 314–322, 2009. doi: 10.1061/(ASCE)0733-9496(2009)135:5(314).

[6] A. Dhar and B. Datta. Saltwater Intrusion Management of Coastal Aquifers. I: Linked Simulation-Optimization. Journal of Hydrologic Engineering, 14(12): 1263–1272, 2009. doi: 10.1061/ (ASCE) HE. 1943-5584.0000097.

[7] M. P. Papadopoulou, I. K. Nikolos, and G. P. Karatzas. Computational benefits using artificial intelligent methodologies for the solution of an environmental design problem: saltwater intrusion. Water Science and Technology, 62(7):1479–1490, 2010. ISSN 0273-1223. doi: 10.2166/wst.2010.442.

[8] J. Grundmann, N. Schütze, G. H. Schmitz, and S. Al-Shaqsi. Towards an integrated arid zone water management using simulation-based optimisation. Environmental Earth Sciences, 65(5): 1381–1394, 2012. ISSN 1866-6280. doi: 10.1007/s12665-011-1253-z.

[9] N. Hansen. Towards a new evolutionary computation. Advances on estimation of distribution algorithms: The CMA evolution strategy: a comparing review. Springer, 2006.

[10] S. S. Al-Shoukri. Modeling of Groundwater Flow in Wadi Ma 'awil Catchment, Barka in Sultanate of Oman. PhD thesis, Arabian Gulf University, Bahrain, 2008.