26th International Congress on Modelling and Simulation, Adelaide, SA, Australia, 30 Nov–4 Dec 2025
mssanz.org.au/modsim2025

High-fidelity emulation of a groundwater simulator using a deep LSTM neural network architecture

Dan Pagendam [a], **Adam Siade** [b], **Stephanie Clark** [b], **Joel Dabrowski** [a], **Zhibin Li** [c], **Sreekanth Janardhanan** [b]

[a] *CSIRO Data61, Australia,* [b] *CSIRO Environment, Australia,* [c] *CSIRO Health & Biosecurity, Australia*
Email: dan.pagendam@data61.csiro.au

Abstract: Model emulation (or surrogate modelling) is the art of modelling a slow, (often mechanistic) simulator with a faster, black-box model. The emulator is trained using datasets (input-output pairs) generated from the simulator using a range of potential inputs (e.g. parameters, forcing variables etc). The fidelity of the emulator as a surrogate of the original simulator is dependent upon the volume of data available for training, and the flexibility of the black-box model to learn the specific dynamics of the original model. We demonstrate how deep LSTM architectures can be used to obtain high-fidelity emulators of groundwater level dynamics at many locations of interest using data from a MODFLOW simulator. The deep LSTM emulator provides a one-step-ahead predictor that can be called recursively to generate arbitrarily long time series predictions of groundwater levels. In our example application, we show that the emulator attains a high degree of fidelity even when recursively called hundreds of times. These emulation methods are highly valuable for modelling activities such as Bayesian inference and optimisation that may require a computationally expensive model to be run sequentially many times.

Keywords: *Emulation, surrogate modelling, groundwater, MODFLOW, dynamic model, deep learning*

1. INTRODUCTION

Many modelling tasks such as Bayesian inference, calibration, and optimisation require a simulator to be run numerous times in series. When the simulator is slow, these activities can become computationally infeasible or difficult. Model emulation, is the art of modelling a computationally expensive simulator with a more efficient black-box surrogate (Gramacy, 2020). Gaussian Processes (GPs) have a long history in their use as model emulators, particularly when the number of simulations available is not large. More recently, additional machine learning models such as Random Forests (e.g. Gladish et al., 2018) have been successfully used as emulators in hydrology. Deep neural networks also offer flexible architectures that lend themselves to the emulation of complex mechanistic processes that evolve through time.

A particularly attractive aspect of GPs as model emulators is the ease with which they can be used to also model the emulator's confidence around the predicted model output. In standard form, neural networks lack this functionality, but there are a variety of ways that they can be extended to quantify uncertainty around their outputs. One of the simplest such approaches is the use of density networks (Bishop, 1994), where the neural network outputs the parameters of a predictive distribution and is trained to minimise a negative log-likelihood loss function. Recently, Pagendam et al. (2023) showed that density networks attained excellent coverage properties when predicting groundwater levels over space and time.

2. METHODOLOGY AND RESULTS

In this work, we examine the effectiveness of neural networks based on deep Long Short-Term Memory (LSTM; Hochreiter and Schmidhuber, 1997) blocks as emulators of a complex physics-based groundwater model: MODFLOW. The emulator is set up as a density network, using a simple Gaussian distribution to model the groundwater level at the next timestep. The MODFLOW model that was the target of the emulator, was developed to investigate a scenario of onshore gas extraction. For a given time series of forcing variables (including recharge, stream losses, evapotranspiration, well extractions etc) time series of groundwater levels at 77 wells were recorded as output from MODFLOW. For building the emulator, we used approximately 2300 model runs with each run using a different time series of forcing variables.

The deep LSTM emulator is a one-step-ahead predictor, taking the forcing variables at time t and outputing the expected groundwater level and standard deviation of the prediction for time $t + 1$. In addition to the forcing variables, our emulator also takes the spatial location (easting and northing) of the well location as additional inputs, resulting in one trained emulator that is capable of predicting groundwater dynamics over the spatial domain of interest.

© MSSANZ and The Authors, CC-BY

We split the dataset generated from MODFLOW into training, valiation (hyper-parameter tuning) and test (assessing final performance) sets by splitting it by unique well locations. The training set conisted of 61 of the 77 wells, while the test and validation sets each consisted of 8 the wells. At the end of training, the set of neural network parameters that had achieved the lowest loss on the test set was restored and used for the final emulator. To assess the performance of the emulator, we studied its ability to reconstruct the time series of groundwater levels of 8 validation wells. To do this, we recursively called the neural network for the entire time series of forcing variables (approximately 400 time steps) for each of these wells and visually assessed the fidelity of the emulator to the MODFLOW model. Figure 1 shows the impressive performance of the emulator in reconstructing a diverse set of groundwater hydrographs from the forcing variables at wells in the test set. We note that our results can show some divergence between the two time series over time, as a result of the recursive approximation of the physics, but despite this fidelity to MODFLOW remains high over long time horizons.

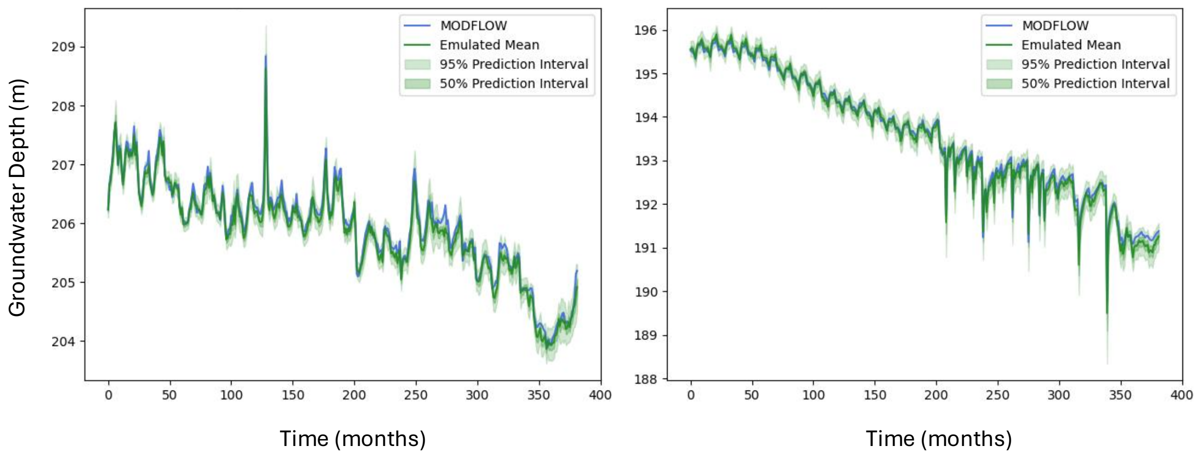

Figure 1. Two groundwater time series from the test set showing the close agreement between the emulator and the original MODFLOW model.

3. CONCLUSIONS

We have shown how a deep LSTM neural network can achieve a high degree of fidelity in replicating groundwater level dynamics from the MODFLOW simulator. When called recursively, we see that these internal state variables allow the emulator to attain a high degree of fidelity to the original MODFLOW simulator at well locations that were not included in the training dataset.

High-fidelity emulators of the type shown here are incredibly useful for improving the efficiency of common modelling tasks such as Bayesian inference, calibration and optimisation. In addition, they also allow for knowledge from complex simulators to be liberated from their high-performance computing environments and incorporated into simpler systems such as decision support tools and software dashboards, where a user may want to interactively explore various scenarios without having to endure long model runtimes.

ACKNOWLEDGEMENTS

This research was funded by CSIRO's Gas Industry Social and Environmental Research Alliance (GISERA). GISERA is a collaboration between CSIRO, Commonwealth and state governments and industry established to undertake publicly reported independent research.

REFERENCES

Bishop, C.M. (1994). Mixture Density Networks. Neural Computing Research Group Report: NCRD/94/004.

Gladish, D.W, Pagendam, D.E., Peeters, L.J.M., Kuhnert, P. & Vaze, J. (2018). Emulation engines: choice and quantification of uncertainty for complex hydrological models. Journal of Agricultural, Biological and Environmental Statistics 23(1), 39-62.

Gramacy R.B. (2020) Surrogates: Gaussian process modeling, design, and optimization for the applied sciences. Chapman & Hall/CRC, Boca Raton, FL.

Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. Neural Computation, 9(8), 1735–1780.

Pagendam, D., Janardhanan, S., Dabrowski, J. & MacKinlay, D. (2023). A log-additive neural model for spatio-temporal prediction of groundwater levels. Spatial Statistics, 55, 100740.

© MSSANZ and The Authors, CC-BY