Received: 10 May 2022 | Accepted: 13 May 2022

DOI: 10.1002/hyp.14596

INVITED COMMENTARY

HP TODAY WILEY

Coevolution of machine learning and process-based modelling to revolutionize Earth and environmental sciences: A perspective

Saman Razavi[1,2] | David M. Hannah[3] | Amin Elshorbagy[2] | Sujay Kumar[4] | Lucy Marshall[5] | Dimitri P. Solomatine[6,7,8] | Amin Dezfuli[9] | Mojtaba Sadegh[10] | James Famiglietti[1]

[1]Global Institute for Water Security, School of Environment and Sustainability, University of Saskatchewan, Saskatoon, Canada

[2]Department of Civil, Geological and Environmental Engineering, University of Saskatchewan, Saskatoon, Canada

[3]School of Geography, Earth and Environment Sciences, University of Birmingham, Birmingham, UK

[4]Hydrological Sciences Lab, NASA Goddard Space Flight Center, Greenbelt, USA

[5]School of Civil and Environmental Engineering, University of New South Wales, Kensington, Australia

[6]Department of Hydroinformatics and Socio-Technical Innovation, IHE Delft Institute for Water Education, Delft, The Netherlands

[7]Water Resources Section, Delft University of Technology, Delft, The Netherlands

[8]Water Problems Institute of RAS, Moscow, Russia

[9]Science Systems and Applications, Inc., Global Modeling and Assimilation Office, NASA Goddard Space Flight Center, Greenbelt, USA

[10]Department of Civil Engineering, Boise State University, Boise, USA

Correspondence
Saman Razavi, Global Institute for Water Security, School of Environment and Sustainability, University of Saskatchewan, 11 Innovation Blvd, Saskatoon, SK S7N 3H5, Canada.
Email: saman.razavi@usask.ca

Funding information
IAS Vanguard Fellowship; Natural Sciences and Engineering Research Council of Canada

Abstract

Machine learning (ML) applications in Earth and environmental sciences (EES) have gained incredible momentum in recent years. However, these ML applications have largely evolved in 'isolation' from the mechanistic, process-based modelling (PBM) paradigms, which have historically been the cornerstone of scientific discovery and policy support. In this perspective, we assert that the cultural barriers between the ML and PBM communities limit the potential of ML, and even its 'hybridization' with PBM, for EES applications. Fundamental, but often ignored, differences between ML and PBM are discussed as well as their strengths and weaknesses in light of three overarching modelling objectives in EES, (1) nowcasting and prediction, (2) scenario analysis, and (3) diagnostic learning. The paper ponders over a 'coevolutionary' approach to model building, shifting away from a borrowing to a co-creation culture, to develop a generation of models that leverage the unique strengths of ML such as scalability to big data and high-dimensional mapping, while remaining faithful to process-based knowledge base and principles of model explainability and interpretability, and therefore, falsifiability.

KEYWORDS
artificial intelligence, deep learning, machine learning, modelling objective, policy support, predication, process-based modelling, scenarios, scientific discovery

1 | INTRODUCTION

Machine learning (ML), particularly deep learning (DL), has achieved revolutionary performance in areas such as computer vision (Krizhevsky et al., 2012), natural language processing (Young et al., 2018), and gaming (Silver et al., 2018). Such unprecedented success has increased the momentum of ML applications in non-native fields, such as Earth and environmental sciences (EES), where process-based (also called knowledge-based) modelling has dominated to date. For example, the number of ML-related presentations at the American Geophysical Union's Fall Meetings has increased from 0.2% in 2015 to >4% in 2020 and as high as 28%, 9%, and 7.5% in the non-linear geophysics, natural hazards, and hydrology sections, respectively. ML is believed to provide processes and systems in EES with new and fertile research horizons, leveraging the boom in computational power, 'big data' sets, and novel sensing technologies (Reichstein et al., 2019).

We argue this recent momentum might lead to disappointment, akin to 'artificial intelligence (AI) winters' (Antun et al., 2022; Choi, 2021; Colbrook et al., 2022; Hendler, 2008; Strickland, 2021), if we do not properly recognize and address two inter-related grand challenges: (1) the lack of explainability and interpretability, and therefore falsifiability, of how ML models emulate underlying systems; and (2) the divorce of ML models from the knowledge base in those emulations, and therefore missing opportunities in extrapolation beyond available data. Motivated by these challenges is a recent increasing drive towards 'Explainable ML' using various numerical heuristics (e.g., Bach et al., 2015; Rudin, 2019; Samek & Müller, 2019; Toms et al., 2020), typically rooted in sensitivity analysis (see section 3.4 of Razavi et al., 2021 for a review), to explain why or how an ML model responds to inputs. In parallel, there are growing efforts to integrate physical knowledge into the fabric of ML models (e.g. Raissi et al., 2019; Champion et al., 2019; Jiang et al., 2020). An often-ignored fact, however, is that these challenges are intrinsically domain- and problem-specific and their meaning and implications in applications native to computer science (CS) can be different from those in the context of EES, where typical systems of interest are *complex*, *open*, *partially observable* and *non-stationary*.

A disciplinary focus, augmented by cultural barriers between different modelling communities, has led ML and process-based modelling (PBM) to evolve in isolation from each other and with different worldviews towards problem solving. Apparently generic concepts can mean different things to different communities. For example, computer and environmental scientists often use different lenses to view the notion of model explainability and interpretability. The former group looks more into 'numerics' and how inputs are mapped onto outputs, while the latter focuses on 'processes' and how causalities and feedback mechanisms give rise to an output conditional to an input. The differences in worldviews may also be attributed to differences in modelling objectives in those communities; for example, whether to solely gain predictive or generative power, or rather to learn why and how a system behaves as it does.

This article provides a perspective on the current status and potential directions of ML applications in EES (Figure 1). We first outline and contrast the unique features of ML and PBM, developed within an 'isolation phase' since the inception of each modelling paradigm, in light of modelling objectives in EES. We then discuss general ML-PBM coupling frameworks in EES arising during the ongoing 'hybridization phase', dating back to at least the early 2000s and recently re-promoted by Reichstein et al. (2019), and contend their compounding potential is limited because their underlying philosophies are rooted in isolation. We argue that truly integrating the two modelling paradigms has not yet occurred, but a 'coevolution phase' has started emerging which, if jointly embraced by both ML and EES communities, can provide fertile ground for transformative innovations in an age of big data and computational power.

2 | HOW IS ML DIFFERENT FROM PBM AND WHY DOES IT MATTER?

ML in most cases is rooted in *connectionism*, *hyper-flexibility* and *vigorous optimization*, which are aliens to PBM (Razavi, 2021). Connectionist ML techniques, particularly those based on DL or more broadly artificial neural networks (ANNs), stack many identical or similar algebraic operators both in parallel and series so they can *collectively* perform complicated tasks. As such, the roles and functions of different individual operators in producing the model response are not readily distinguishable, unlike the *modular* approach typically adopted in PBM where each model part is designed to represent a specific system component. Hyper-flexibility, which can be viewed as an inversion of *Occam's razor*, is characteristic of a model that can literally fit any dataset with any desired level of accuracy owing to its excessive degrees of freedom. Vigorous optimization complements hyper-flexibility by manipulating model parameters at any cost to maximize the goodness-of-fit to calibration data. As a result, considerations of *identifiability*, *equifinality*, *parsimony* and *physical consistency* of model structure and parameters are rather irrelevant in the context of ML yet are typical considerations of PBM, with either statistical models or those based on differential or other types of equations (Beven, 2006; Guillaume et al., 2019).

The above characteristics make ML uniquely suitable for complex mappings due to possible *correlations* of any form embedded in any dataset of any dimensionality and nature. Conversely, PBM is generally based on *causations*, presumptive or real, typically defined in low-dimensional spaces consistent with human cognitive capacity and bounded within the realm of existing knowledge or perceptions. As such, PBM is suitable for system identification and hypothesis development and testing to understand the processes and their interactions within a system. However, the attachment to legacy knowledge and often rigid setups can sometimes hinder the *scalability* of PBM to different datasets, which, in some cases, requires altering the model conceptualization, structure and parameterization (Nearing et al., 2021). Conversely, ML applications are typically agnostic to prior knowledge,

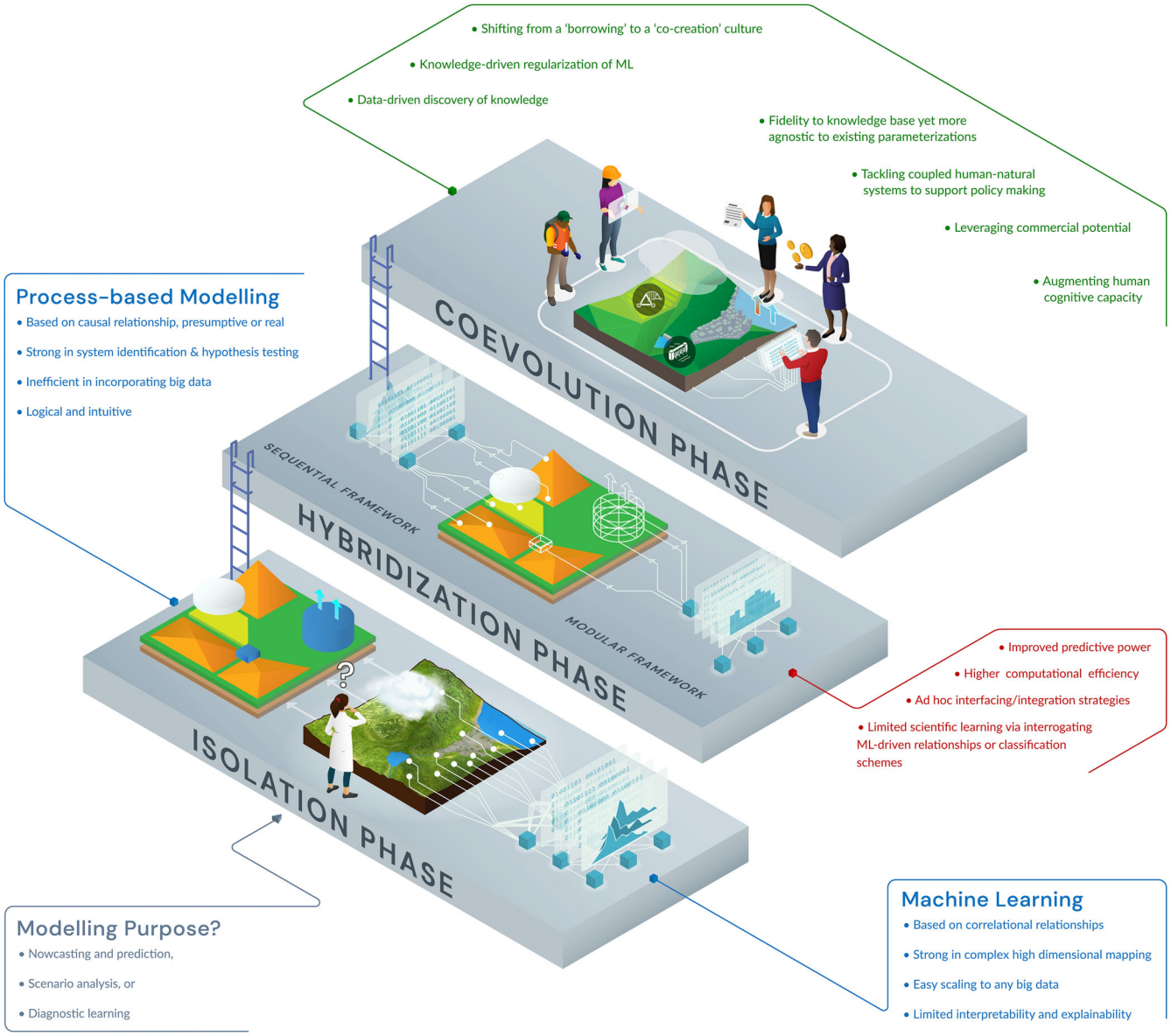

FIGURE 1 Machine learning and process-based modelling in Earth and environmental sciences: Past, present and future

assumptions, and setups and can therefore relatively easily *scale* to any data type, size and spatio-temporal resolution.

This 'agnosticism' further enables ML to be more robust to missing data or knowledge about a problem in the sense that it may still predict some variables with desired accuracy based on data that are deemed incomplete from a process-representation viewpoint. The other side of this agnosticism is that ML applications typically provide little basis for interpretability and explainability; for example, to enable the modeller to reason why a model may respond differently to different perturbations and, more importantly, to explain that to a stakeholder. In contrast, PBM applications tend to be more explainable and intuitive, even under circumstances not seen in the period of record, because they try to emulate real-world processes as observed or speculated by the modeller.

To understand the significance of the above fundamental differences, let us revisit the overarching objectives of modelling in EES, which have largely been served by PBM in past decades. In our view, they include: (1) *nowcasting and prediction*, (2) *scenario analysis* and (3) *diagnostic learning*. The first aims to look into the now or foreseeable future and predict what *will* happen, for example, in a local or regional weather system (e.g. Shi et al., 2015). This directly supports real-time operation and management at different levels, from individual citizens to local, regional or global institutions. The second takes a *what-if* view of the future and aims to determine how the system might respond under new or altered conditions, such as climate or land-use change (Maier et al., 2016). Thus, it supports long-term decision-making pertaining, for example, to adaptation to change and building resilience in human-controlled systems.

The third, diagnostic learning, is about looking backward, using models to simulate the past and present behaviour of a system to determine why it behaves as it does. This supports the development

and testing of new theories and hypotheses, thereby extending our process understanding and knowledge base. In other words, diagnostic learning is about *exploring causations* and *attributions of observed variability* to controlling factors; for example, to explain the observed sea level rise in past decades (e.g., Wada et al., 2016). Modelling for diagnostic learning is a fundamental underpinning of scientific discovery, thereby improving model fidelity.

How does ML perform with respect to the above modelling objectives? ML is expected to perform well in nowcasting and prediction, particularly when the underlying real-world system remains structurally unchanged in the forecasting horizon of interest. ML, however, can become handicapped or unreliable in scenario analysis if extrapolating well beyond the training data is needed or a potential structural change in the underlying system via human intervention or otherwise is of interest (see section 7 in the study by Razavi, 2021 for an experiment). In such extrapolations, the purely data-driven correlational structure of variables may not be transferable (Beven, 2020), while PBM might be salvaged by domain knowledge built in models.

Furthermore, most ML techniques, particularly connectionists, may arguably be of limited help for diagnostic learning because they are not essentially built for system identification and attribution. In other words, it is typically non-trivial to compartmentalize an ML model and attribute its parts to the components of the system being modelled. Thus, contributions of ML to learning have been generally limited to those based on interrogating ML-driven correlational relationships or classification schemes to improve parameterizations (Mount et al., 2016; Shen et al., 2018).

3 | THE NOTION OF HYBRIDIZATION AND ITS SHORTCOMINGS

The desire to bridge ML and PBM has a long history in the AI community (e.g. Towell & Shavlik, 1994) and has received much more attention recently (e.g. Karpatne, Atluri, et al., 2017; Raissi et al., 2019; von Rueden et al., 2019). In the context of EES, such efforts have mainly been under the notion of 'hybridization', cast in three general frameworks as articulated in Abrahart et al. (2012) and more recently reintroduced in Reichstein et al. (2019), herein referred to as *sequential framework*, *modular framework* and *surrogate modelling*.

Sequential framework refers to a hybridization framework in which an ML model receives the output of a process-based model and attempts to estimate its errors (e.g. Anctil et al., 2003; Li et al., 2021; Shamseldin & O'Connor, 2001) or their distribution as a measure of model uncertainty (e.g. Solomatine & Shrestha, 2009; Wani et al., 2017). This form of hybridization is useful to improve our overall predictive power via ML in a data-driven fashion, when PBM cannot explain some aspect(s) of observed data.

In modular framework, a (sub-)model takes charge to directly represent a process, or a set of processes combined, in the body of another larger model. This framework is intrinsically ad hoc depending on the problem at hand and data available. When serving as a submodel in PBM, ML can target processes that are not well understood theoretically to enable a process-based representation, but provide sufficient direct observations for input–output mapping (e.g. Bennett & Nijssen, 2021; Chen & Adams, 2006; Chua & Wong, 2010; Corzo et al., 2009; Mekonnen et al., 2015). Alternatively, PBM may at times be preferred within larger ML models, when it offers place-specific knowledge or proven skills in representing particular processes in the underlying system of interest (e.g. Chua & Wong, 2010; Humphrey et al., 2016; Jiang et al., 2020).

Surrogate modelling is fundamentally different from the above two frameworks as it employs ML to develop an emulator of some targeted aspects of a complex process-based model (e.g. Maskey et al., 2000; Yu et al., 2020). The main motivation for this framework is to improve computational efficiency and tractability of model-based analyses, particularly in multi-query applications (Razavi et al., 2012).

In general, these hybridization frameworks are expected to improve our computational efficiency and predictive power of any system of interest in EES, which are particularly beneficial to operations for nowcasting and prediction. However, we argue such hybridizations still treat the theories and models of the two paradigms in isolation, while integrating them through different and ad hoc interfacing procedures. Therefore, any resulting 'hybrid model' may naturally inherit limitations of the parent models, particularly when it comes to scenario analysis and diagnostic learning.

4 | TOWARDS COEVOLUTION OF ML AND PBM

The notion of hybridization may yet need to break through philosophical barriers between CS and EES communities, and shift attention from a 'borrowing culture' centered at adoption of CS tools in EES applications to a 'co-creation culture' that brings the two communities together to build new modelling paradigms engineered to address outstanding and emerging challenges in EES. Enabled with a coevolutionary approach towards model building, such modelling paradigms would: (1) leverage the unique strengths of ML such as scalability to big data and high-dimensional mapping, while (2) remain *faithful to knowns* in the knowledge domain yet more agnostic to existing mechanistic formulations that might be deficient.

This endeavour can build upon some new, yet embryonic, methodological approaches under two complementary notions, referred to herein as '*knowledge-driven regularization of ML*' and '*data-driven discovery of knowledge*'. Regularization is widely used to leash any possibly *undesired* flexibility of mathematical models with an aim to improve their generalizability (Johansen, 1997; Krogh & Hertz, 1991). A common approach of regularization is to introduce a 'penalty function' to the model development exercise, sometimes as simple as the sum of squares of the model parameters, that needs to be minimized during model calibration (Razavi & Tolson, 2011). This regularization approach can be extended to account for the knowledge base available, by designing and minimizing '*knowledge-driven training functions*' that measure violations from known physical laws and principles. For example, Karpatne, Watkins, et al. (2017) designed a training function

to enforce the monotonicity of the water density–depth relationship in an ML-based lake temperature model. Raissi et al. (2019) furthered this approach by incorporating a set of known differential equations (ordinary or partial) such as Navier–Stokes into a training function.

Knowledge-driven regularization of ML has great potential that needs to be exploited. In principle, a vast array of knowledge types can be built into ML, whether quantitative, such as those typically represented in PBM including *conservation laws*, *monotonicity and rates*, and *feedback mechanisms*, or qualitative, such as any information around the general form of a relationship. Regularizing ML with such a knowledge base, which typically comes with *known limits of validity*, provides a basis for building confidence in modelling results, even in extrapolation into parts of the problem space for which no data are available. However, a major challenge is that the design of a knowledge-driven training function is inherently a domain- and scale-specific problem and its implementation can largely be heuristic. What helps is that such knowledge representation in ML does not necessarily need to be exact or deterministic and can be effectively treated as *soft* constraints during training, with the degree of softness controlled by a weight factor in the training function.

In parallel, new fronts are emerging for data-driven discovery of knowledge, following the longstanding ambitions to discover the unknown dynamical properties of a system from data, with roots in efforts to reconstruct the state space of a system and its possible future trajectories purely from time series data (Casdagli et al., 1991). A first front revolves around using ML to transform an original variable space, where data are collected, into a new space typically with lower dimensionality that explains the governing processes in a more parsimonious and interpretable way (Champion et al., 2019; Iten et al., 2020). A second front aims to derive differential or other types of equations governing a system from data collected across spatio-temporal domains, for example, via sparse regression methodologies (Brunton et al., 2016; Rudy et al., 2017). While any set of equations derived from data may not necessarily embed causal relationships, this approach can allow for theory and hypothesis development around dynamical properties, possible feedback mechanisms, and emerging behaviours. When leveraged by big data, such techniques have potential for new scientific discoveries, particularly in poorly known systems, across a range of space and time scales. Such insights can potentially improve extrapolatability and generalizability beyond observations by constraining model behaviour within the new knowledge learned.

Data-driven discovery of knowledge should not be limited to purely natural processes. Today, EES is dealing with a wide range of human-induced or -driven processes that are still largely unknown or under-characterized. For example, modelling and predicting the impact of policy, societal response to conservation measures or the role of corruption on the management of water and other resources is too complicated. A critical question is how such processes come together and interact with natural processes to sustain, amplify or dampen outcomes like drought, groundwater depletion or streamflow withdrawals (Elsawah et al., 2020). Where data are available around such human-natural systems, either quantitative, qualitative, or a combination thereof, ML can help us find relationships that may not be readily apparent.

This potential can even lead to an inversion of views about techniques such as data assimilation (DA), which traditionally use PBM as the 'reference' with which different datasets are fused. Standard DA techniques can be deficient if signals in the data are dominated by processes that are poorly or un-represented in PBM, a common situation for many real-world systems with ubiquitous human impacts. In such situations, ML embedding physical knowledge may be used as an alternative reference for DA, as it may demonstrate improved explanatory power of the processes involved. More generally, any new knowledge gained about the underlying processes through the use of ML can be built into PBM.

5 | THE BOTTOM LINE

Current excitement in the EES and CS communities around ML provides a fertile ground for breaking through cultural barriers by developing initiatives aimed at coevolution of ML and PBM. Eventually, such initiatives may lead to the development of models that further expand our cognitive capacity to make sense of the underlying processes in higher-dimensional spaces, and across heterogeneous and non-stationary domains. Commercialization potential of ML to address Earth and environmental problems can be a catalyst in this endeavour; it has already motivated high-tech companies, which have been spearheading the transformative and widespread impacts of ML on society through smartphones and other applications, to invest in such new ideas (e.g. the ML-based flood warning system by Google [Vincent, 2020]). These interdisciplinary synergies and private investments must be embraced in sustainable ways when training the new generation of scientists and practitioners. Away from any possible 'hype' or overexcitement around ML, we need to ensure they retain curiosity about understanding Earth and environmental processes while being equipped with emerging ML technologies.

AUTHOR CONTRIBUTIONS

S.R. and D.M.H. conceived the idea. S.R. wrote a first draft in discussion with D.M.H. and coordinated a series of workshops where all coauthors provided valuable contributions and discussions. S.R. wrote the final paper with significant contributions from all the coauthors.

ACKNOWLEDGEMENTS

This Perspective is the outcome of a series of workshops administered under the University of Birmingham's Institute of Advanced Studies (IAS) and the University of Saskatchewan's Global Institute for Water Security (GIWS) in 2021. Saman Razavi is deeply grateful for receiving financial and logistical supports during his visit under the University of Birmingham's IAS Vanguard Fellowship. David Cunha graphically designed Figure 1.

ORCID

Saman Razavi https://orcid.org/0000-0003-1870-5810
David M. Hannah https://orcid.org/0000-0003-1714-1240
Mojtaba Sadegh https://orcid.org/0000-0003-1775-5445

REFERENCES

Abrahart, R. J., Anctil, F., Coulibaly, P., et al. (2012). Two decades of anarchy? Emerging themes and outstanding challenges for neural network river forecasting. *Progress in Physical Geography: Earth and Environment*, *36*(4), 480–513. https://doi.org/10.1177/0309133312444943

Anctil, F., Perrin, C., & Andréassian, V. (2003). ANN output updating of lumped conceptual rainfall/runoff forecasting models. *Journal of the American Water Resources Association*, *39*(5), 1269–1279.

Antun, V., Colbrook, M. J., & Hansen, A. C. (2022). Proving existence is not enough: Mathematical paradoxes unravel the limits of neural networks in artificial intelligence. *SIAM News*, *55*(04), 1–4.

Bach, S., Binder, A., Montavon, G., Klauschen, F., Müller, K. R., & Samek, W. (2015). On pixel-wise explanations for non-linear classifier decisions by layer-wise relevance propagation. *PLoS One*, *10*(7), e0130140.

Bennett, A., & Nijssen, B. (2021). Deep learned process parameterizations provide better representations of turbulent heat fluxes in hydrologic models. *Water Resources Research*, *57*(5), e2020WR029328.

Beven, K. (2006). A manifesto for the equifinality thesis. *Journal of Hydrology*, *320*(1–2), 18–36.

Beven, K. (2020). Deep learning, hydrological processes and the uniqueness of place. *Hydrological Processes*, *34*(16), 3608–3613.

Brunton, S. L., Proctor, J. L., & Kutz, J. N. (2016). Discovering governing equations from data by sparse identification of nonlinear dynamical systems. *Proceedings of the National Academy of Sciences*, *113*(15), 3932–3937.

Casdagli, M., Eubank, S., Farmer, J. D., & Gibson, J. (1991). State space reconstruction in the presence of noise. *Physica D: Nonlinear Phenomena*, *51*(1–3), 52–98.

Champion, K., Lusch, B., Kutz, J. N., & Brunton, S. L. (2019). Data-driven discovery of coordinates and governing equations. *Proceedings of the National Academy of Sciences*, *116*(45), 22445–22451.

Chen, J., & Adams, B. J. (2006). Integration of artificial neural networks with conceptual models in rainfall-runoff modeling. *Journal of Hydrology*, *318*(1–4), 232–249.

Choi, C. Q. (2021). 7 revealing ways AIs fail: Neural networks can be disastrously brittle, forgetful, and surprisingly bad at math. *IEEE Spectrum*, *58*(10), 42–47.

Chua, L. H., & Wong, T. S. (2010). Improving event-based rainfall–runoff modeling using a combined artificial neural network–kinematic wave approach. *Journal of Hydrology*, *390*(1–2), 92–107.

Colbrook, M. J., Antun, V., & Hansen, A. C. (2022). The difficulty of computing stable and accurate neural networks: On the barriers of deep learning and Smale's 18th problem. *Proceedings of the National Academy of Sciences*, *119*(12), e2107151119.

Corzo, G. A., Solomatine, D. P., Hidayat, de Wit, M., Werner, M., Uhlenbrook, S., & Price, R. K. (2009). Combining semi-distributed process-based and data-driven models in flow simulation: A case study of the Meuse river basin. *Hydrology and Earth System Sciences*, *13*, 1619–1634. https://doi.org/10.5194/hess-13-1619-2009

Elsawah, S., Filatova, T., Jakeman, A. J., Kettner, A. J., Zellner, M. L., Athanasiadis, I. N., Hamilton, S. H., Axtell, R. L., Brown, D. G., Gilligan, J. M., Janssen, M. A., Robinson, D. T., Rozenberg, J., Ullah, I. I. T., & Lade, S. J. (2020). Eight grand challenges in socio-environmental systems modeling. *Socio-Environmental Systems Modelling*, *2*, 16226.

Guillaume, J. H., Jakeman, J. D., Marsili-Libelli, S., Asher, M., Brunner, P., Croke, B., Hill, M. C., Jakeman, A. J., Keesman, K. J., Razavi, S., & Stigter, J. D. (2019). Introductory overview of identifiability analysis: A guide to evaluating whether you have the right type of data for your modeling purpose. *Environmental Modelling & Software*, *119*, 418–432.

Hendler, J. (2008). Avoiding another AI winter. *IEEE Intelligent Systems*, *23*(02), 2–4.

Humphrey, G. B., Gibbs, M. S., Dandy, G. C., & Maier, H. R. (2016). A hybrid approach to monthly streamflow forecasting: Integrating hydrological model outputs into a Bayesian artificial neural network. *Journal of Hydrology*, *540*, 623–640.

Iten, R., Metger, T., Wilming, H., Del Rio, L., & Renner, R. (2020). Discovering physical concepts with neural networks. *Physical Review Letters*, *124*(1), 010508.

Jiang, S., Zheng, Y., & Solomatine, D. (2020). Improving AI system awareness of geoscience knowledge: Symbiotic integration of physical approaches and deep learning. *Geophysical Research Letters*, *47*(13), e2020GL088229.

Johansen, T. A. (1997). On Tikhonov regularization, bias and variance in nonlinear system identification. *Automatica*, *33*(3), 441–446.

Karpatne, A., Atluri, G., Faghmous, J. H., et al. (2017). Theory-guided data science: A new paradigm for scientific discovery from data. *IEEE Transactions on Knowledge and Data Engineering*, *29*(10), 2318–2331. https://doi.org/10.1109/tkde.2017.2720168

Karpatne, A., Watkins, W., Read, J., & Kumar, V. (2017). Physics-guided neural networks (pgnn): An application in lake temperature modeling. *arXiv preprint*, arXiv:1710.11431.

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. In *The proceedings of the 25th international conference on neural information processing systems*. Association for Computing Machinery (pp. 1097–1105).

Krogh, A., & Hertz, J. (1991). A simple weight decay can improve generalization. *Advances in Neural Information Processing Systems*, *4*, 950–957.

Li, D., Marshall, L., Liang, Z., Sharma, A., & Zhou, Y. (2021). Characterizing distributed hydrological model residual errors using a probabilistic long short-term memory network. *Journal of Hydrology*, *603*, 126888.

Maier, H. R., Guillaume, J. H., van Delden, H., Riddell, G. A., Haasnoot, M., & Kwakkel, J. H. (2016). An uncertain future, deep uncertainty, scenarios, robustness and adaptation: How do they fit together? *Environmental Modelling & Software*, *81*, 154–164.

Maskey, S., Dibike, Y. B., Jonoski, A., & Solomatine, D. P. (2000). Groundwater model approximation with artificial neural network for selecting optimal pumping strategy for plume removal. In O. Schleider & A. Zijderveld (Eds.), *AI methods in civil engineering applications* (Vol. 2000, pp. 67–80).

Mekonnen, B. A., Nazemi, A., Mazurek, K. A., Elshorbagy, A., & Putz, G. (2015). Hybrid modelling approach to prairie hydrology: Fusing data-driven and process-based hydrological models. *Hydrological Sciences Journal*, *60*(9), 1473–1489.

Mount, N. J., Maier, H. R., Toth, E., Elshorbagy, A., Solomatine, D., Chang, F. J., & Abrahart, R. J. (2016). Data-driven modelling approaches for socio-hydrology: Opportunities and challenges within the Panta Rhei science plan. *Hydrological Sciences Journal*, *61*(7), 1192–1208.

Nearing, G. S., Kratzert, F., Sampson, A. K., Pelissier, C. S., Klotz, D., Frame, J. M., Prieto, C., & Gupta, H. V. (2021). What role does hydrological science play in the age of machine learning? *Water Resources Research*, *57*(3), e2020WR028091.

Raissi, M., Perdikaris, P., & Karniadakis, G. E. (2019). Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational Physics*, *378*, 686–707.

Razavi, S. (2021). Deep learning, explained: Fundamentals, Explainability, and Bridgeability to process-based modelling. *Environmental Modelling & Software*, *144*, 105159.

Razavi, S., Jakeman, A., Saltelli, A., Prieur, C., Iooss, B., Borgonovo, E., Plischke, E., Piano, S. L., Iwanaga, T., Becker, W., Tarantola, S., Guillaume, J. H. A., Jakeman, J., Gupta, H., Melillo, N., Rabitti, G., Chabridon, V., Duan, Q., Sun, X., … Maier, H. R. (2021). The future of sensitivity analysis: An essential discipline for systems modeling and policy support. *Environmental Modelling & Software*, *137*, 104954.

Razavi, S., & Tolson, B. A. (2011). A new formulation for feedforward neural networks. *IEEE Transactions on Neural Networks*, *22*(10), 1588–1598.

Razavi, S., Tolson, B. A., & Burn, D. H. (2012). Review of surrogate modeling in water resources. *Water Resources Research*, *48*(7), 1–32.

Reichstein, M., Camps-Valls, G., Stevens, B., Jung, M., Denzler, J., & Carvalhais, N. (2019). Deep learning and process understanding for data-driven earth system science. *Nature*, *566*(7743), 195–204.

Rudin, C. (2019). Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nature Machine Intelligence*, *1*(5), 206–215. https://doi.org/10.1038/s42256-019-0048-x

Rudy, S. H., Brunton, S. L., Proctor, J. L., & Kutz, J. N. (2017). Data-driven discovery of partial differential equations. *Science Advances*, *3*(4), e1602614.

Samek, W., & Müller, K.-R. (2019). Towards explainable artificial intelligence. In *Explainable AI: Interpreting, explaining and visualizing deep learning*. Springer (pp. 5–22). https://doi.org/10.1007/978-3-030-28954-6_1

Shamseldin, A. Y., & O'Connor, K. M. (2001). A non-linear neural network technique for updating of river flow forecasts. *Hydrology and Earth System Sciences*, *5*, 577–598. https://doi.org/10.5194/hess-5-577-2001

Shen, C., Laloy, E., Elshorbagy, A., Albert, A., Bales, J., Chang, F. J., & Tsai, W. P. (2018). HESS opinions: Incubating deep-learning-powered hydrologic science advances as a community. *Hydrology and Earth System Sciences*, *22*(11), 5639–5656.

Shi, X., Chen, Z., Wang, H., Yeung, D. Y., Wong, W. K., & Woo, W. C. (2015). Convolutional LSTM network: A machine learning approach for precipitation nowcasting. In *Advances in neural information processing systems* (pp. 802–810).

Silver, D., Hubert, T., Schrittwieser, J., Antonoglou, I., Lai, M., Guez, A., Lanctot, M., Sifre, L., Kumaran, D., Graepel, T., Lillicrap, T., Simonyan, K., Hassabis, D., & Lillicrap, T. (2018). A general reinforcement learning algorithm that masters chess, shogi, and go through self-play. *Science*, *362*(6419), 1140–1144.

Solomatine, D. P., & Shrestha, D. L. (2009). A novel method to estimate model uncertainty using machine learning techniques. *Water Resources Research*, *45*(12), 1–16.

Strickland, E. (2021). The turbulent past and uncertain future of AI: Is there a way out of AI's boom-and-bust cycle? *IEEE Spectrum*, *58*(10), 26–31.

Toms, B. A., Barnes, E. A., & Ebert-Uphoff, I. (2020). Physically interpretable neural networks for the geosciences: Applications to earth system variability. *Journal of Advances in Modeling Earth Systems*, *12*(9), e2019MS002002.

Towell, G. G., & Shavlik, J. W. (1994). Knowledge-based artificial neural networks. *Artificial Intelligence*, *70*(1–2), 119–165.

Vincent, J. (2020). Google's AI flood warnings now cover all of India and have expanded to Bangladesh. *The Verge*. https://www.theverge.com/2020/9/1/21410252/google-ai-flood-warnings-india-bangladesh-coverage-prediction

von Rueden, L., Mayer, S., Beckh, K., Georgiev, B., Giesselbach, S., Heese, R., Kirsch, B., Pfrommer, J., Pick, A., Ramamurthy, R., Walczak, M., Garcke, J., Bauckhage, C., & Schuecker, J. (2019). Informed machine learning--A taxonomy and survey of integrating knowledge into learning systems. *arXiv preprint*, arXiv:1903.12394.

Wada, Y., Lo, M. H., Yeh, P. J. F., Reager, J. T., Famiglietti, J. S., Wu, R. J., & Tseng, Y. H. (2016). Fate of water pumped from underground and contributions to sea-level rise. *Nature Climate Change*, *6*(8), 777–780.

Wani, O., Beckers, J. V., Weerts, A. H., & Solomatine, D. P. (2017). Residual uncertainty estimation using instance-based learning with applications to hydrologic forecasting. *Hydrology and Earth System Sciences*, *21*(8), 4021–4036. https://doi.org/10.5194/hess-21-4021-2017

Young, T., Hazarika, D., Poria, S., & Cambria, E. (2018). Recent trends in deep learning based natural language processing. *IEEE Computational Intelligence Magazine*, *13*(3), 55–75.

Yu, X., Cui, T., Sreekanth, J., Mangeon, S., Doble, R., Xin, P., Rassam, D., & Gilfedder, M. (2020). Deep learning emulators for groundwater contaminant transport modelling. *Journal of Hydrology*, *590*, 125351.

How to cite this article: Razavi, S., Hannah, D. M., Elshorbagy, A., Kumar, S., Marshall, L., Solomatine, D. P., Dezfuli, A., Sadegh, M., & Famiglietti, J. (2022). Coevolution of machine learning and process-based modelling to revolutionize Earth and environmental sciences: A perspective. *Hydrological Processes*, *36*(6), e14596. https://doi.org/10.1002/hyp.14596