Machine learning in Earth and environmental science requires education and research policy reforms

Leveraging advances in artificial intelligence could revolutionize the Earth and environmental sciences. We must ensure that our research funding and training choices give the next generation of geoscientists the capacity to realize this potential.

Sean W. Fleming, James R. Watson, Ashley Ellenson, Alex J. Cannon and Velimir C. Vesselinov

In Earth and environmental science (EES), quantitative prediction models gauge the state of scientific knowledge and help put it to practical use. With the emergence of big data, exponential growth in computational speed and increasing awareness of the practical limits of classical physics-based and statistical models, a new modelling approach has appeared: artificial intelligence (AI) (for a short glossary of machine learning-related jargon, see Table 1). A major component of the broader data-science tidal wave, which has been deemed the fourth industrial revolution[1] and fourth paradigm of science[2], AI can accelerate discovery and prediction thanks to its scalability, capacity to determine patterns within large datasets and wide applicability.

The problem? Every modelling framework has its own philosophy, theory, nomenclature, implementation details and culture of practice — and the proportion of geoscientists with expertise in new AI technologies and concepts, and particularly in how to use and interpret this new class of scientific tool, remains very limited. We argue that bridging this gap between the few EES researchers who specialize in AI and the rest of the EES community requires something more fundamental and comprehensive than technology transfer: reforms of policies around research funding priorities and education content will be needed for AI to more fully transition into, and find its place within, EES.

Novel applications and challenges

To illustrate, consider the application of AI to streamflow forecasting. Operational groups, primarily government agencies, use physical process-based or linear statistical hydrologic models for high-stakes practical applications such as flood and water supply forecasting[3]. For over 25 years, machine learning algorithms arising in the AI literature have been explored as an alternative river prediction paradigm by a narrow segment of the research enterprise. This led to substantial improvements in the ability to discover patterns in data, and to use those patterns to make accurate predictions. Despite this, AI has not substantially diffused into operational systems, or more generally into water resource science and engineering research and education. In our view, this boils down to the failure of AI specialists to engage the hydrology community, whose detailed theoretical and practical requirements have therefore not been met by most AI applications[4]. But there is light at the end of the tunnel: AI has begun to migrate into operations through collaborative partnerships between AI researchers and operational hydrologists, leading to the integration of a priori hydrometeorological process knowledge and pragmatic logistical needs within new, purpose-built AI technologies[4]. Nevertheless, the average hydrologist remains more likely to use AI in a smartphone application to find the quickest route to the office than in their work after they arrive.

This situation — an inability of AI to break out into the mainstream, despite a demonstrated capacity to out-compete existing quantitative methods in many respects — remains typical of EES generally.

Research policy and AI

To gain genuine and widespread acceptance among the EES community, AI cannot exist in a vacuum: it must be used in such a way as to incorporate — and ultimately, to expand — the substantial existing body of quantitative biogeophysical process knowledge.

That kind of broad-based intersection between an understanding of underlying physical processes and data-driven analysis dovetails with the current fashion among computer scientists of developing explainable AI (Table 1). Trust in AI has been limited by its nominally black-box nature; that is, the inability to explain, in terms of accepted process knowledge, why the algorithm got the answer it did. Resolving this issue is the focus of much work today[5].

But the intersection we refer to here also reflects a deeper and broader sense of 'physics-aware AI', a concept we borrow from materials science[6]. As applied to EES, physics-aware AI spans at least four different, although overlapping, categories: biogeophysical systems characterization, theory-guided or physics-informed AI[7], original knowledge discovery and emulation of process-based models (Supplementary Table 1). Alignment with domain-specific practical needs and issues, such as data limitations or computing costs, is another crucial aspect of physics-aware AI. EES examples so far have been promising (Box 1; see also Supplementary Table 1) but remain relatively rare.

The policy implication? Research funding must emphasize physics-aware AI, which will require, and in turn enable, closer ties between specialists in EES applications of AI and the rest of the EES research and practice community.

Education policy and AI

The AI learning experience for EES students remains largely one of self-teaching, short informal courses or selecting a thesis supervisor with an AI specialization—and the vast majority of EES students receive no AI training whatsoever. As a result, many geoscientists, including recent graduates, may not have the skills to work effectively with these concepts. The consequences can range from blanket dismissal of AI by senior researchers with no previous knowledge of the field to 'package-surfing' by new adopters, who draw AI capabilities from existing Python, R or Julia libraries (for

Table 1 | Data science 101

Term	Concept
Data science	Algorithm development for building information/knowledge pipelines based on data, often to enable automated predictions and actions. Data science builds on traditional statistics to include AI and tends to have a 'whatever works' philosophy.
Data mining	Extracting explanatory/predictive patterns from data, often without clear a priori expectations around specific causal linkages. The term originated in traditional statistics as a pejorative but in AI has evolved into respected techniques. Data mining is typically, but not universally, reserved for pattern identification in large/complex datasets.
Big data	Datasets satisfying three 'V's of volume (large size), velocity (fast, typically continuous, incoming data streams) and variety (spanning a wide, often unpredictable content range). YouTube is a prime example. Massive datasets are routine in some EES areas but seldom satisfy all three criteria; however, analysis methods developed for big data can be useful for such datasets as well[5]
AI	A broad science and engineering field devoted, loosely, to inventing technologies that emulate human intelligence; it in turn includes several disciplines, such as machine learning, robotics and so forth.
Machine learning (ML)	Major AI field using algorithms to identify patterns in data and applying them to make predictions. Divided into classification or regression, and unsupervised or supervised methods. Examples include various neural network types, random forests and support vector machines, and deep learning architectures such as deep, long short-term memory and convolutional neural networks. Where there is no risk of confusion, including most EES applications, AI and ML are used interchangeably.
Explainable AI	ML is typically viewed as a 'black box,' meaning that it is unclear how the AI arrived at its answers and how to explain those answers in terms of the physics of the system being studied. Explainable, interpretable or glass-box AI seeks to overcome this limitation[7]. Explainable AI connects to wider concepts of physics-aware AI[4–8].
Hyperparameters	AIs have parameters, like a neural network's neuron weights, that are optimized in training (loosely analogous to coefficients in a linear statistical model). AIs also have higher-level hyperparameters, like a neural network's learning rate, that in turn control this process of estimating parameter values and overall architecture.
Features and targets	Akin to predictors and predictands in traditional statistical modelling. Feature engineering (processing or manipulating input data to extract and select features to be passed to the AI as predictor variables) is a major element of many AI applications.
AutoML	System to automatically build the best AI for a given dataset; may include optimal hyperparameter selection and is intended to make ML easier for non-AI specialists to apply in their respective practice domains. Challenges with AutoML include the emergence of still-higher-level parameters (hyperhyperparameters) and model equifinality. AutoML is just beginning to find applications in EES[4].
Overtraining and regularization	Overtraining is an AI's memorization of the data used to fit it, compromising generalization accuracy. It happens in all models involving calibration or derivation of parameters from observational data but can be particularly acute for AI due to its flexibility. Regularization refers to technical methods that mitigate AI overtraining to the point that it is no worse than in a traditional statistical or process model.

example) without first understanding what is under the hood. Inadequate knowledge of how these prediction engines relate to the biogeophysical and statistical details of the EES problem at hand invites flawed results.

Such a lack of widespread AI technical capacity blends with deeper concerns around unevenness in the quantitative literacy of EES graduates. Not every geoscientist requires advanced quantitative skills, but, in our opinion, by failing to provide a uniformly solid grounding in mathematical and computational methods, universities are leaving students ill-prepared to effectively conduct modern science—including (but not limited to) judicious, informed use of AI to support EES problem-solving. We urge universities to rethink their policies around core curricula, so that the next generation of graduates are prepared for an AI-enabled future.

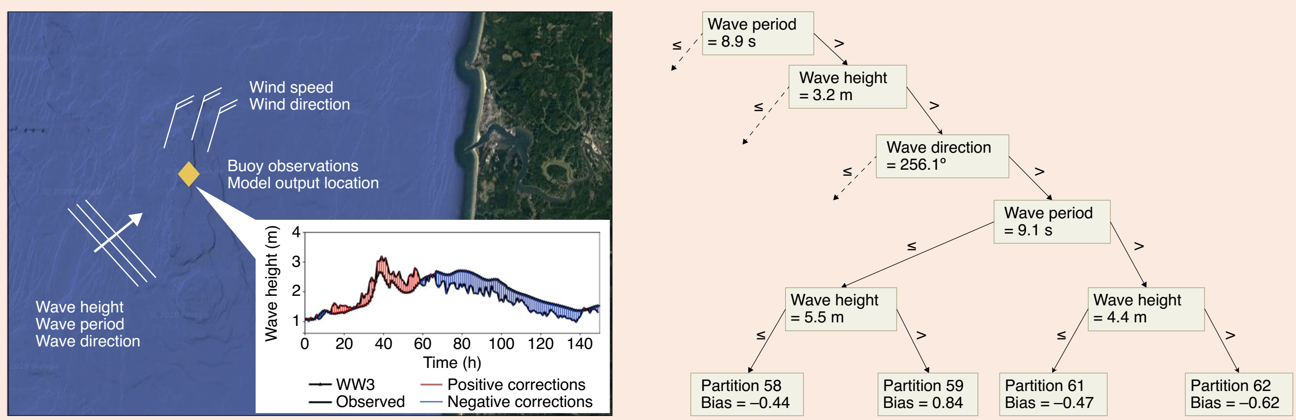
The road ahead

Which policies fit the bill? Details depend on the local context, but promising avenues can be identified. Research funding should promote novel partnerships to create transformative and widely relevant insights. Both cutting-edge AI expertise and a need for AI-facilitated practical solutions often reside in industry. Innovative funding mechanisms designed specifically to enable multisectoral academic–public–private partnerships and two-way knowledge transfers in applied research and development, such as the Mitacs programme (Canada), could therefore be a powerful policy instrument for evolving the field forwards into mainstream EES use. Another consideration is that AI interest and expertise tend to be concentrated among early-career geoscientists, who have relatively limited access to funding opportunities. This suggests that adjusting priorities at agencies such as the National Science Foundation (United States) or Natural Environment Research Council (United Kingdom) to favour high-risk/high-reward proposals integrating machine learning into EES could be helpful. Above all, funding processes that reward physics-aware forms of AI will be key to successfully bridging the gap between machine learning specialists and the broader geoscience community. As for educational policy, a mandatory EES course introducing all undergraduate geoscience students to AI, with the necessary maths, statistics and computing prerequisites, would be a good start. One step further would be the creation of joint EES/computer science undergraduate programmes that provide students with a solid grounding in both disciplines. Universities could also create dual tenure-track positions—helping promote AI expertise in geoscientists, and geoscience expertise among computer scientists. Once we start thinking seriously about such policy reforms, more directions will undoubtedly present themselves.

AI now permeates society, and it is here to stay. EES must come to terms with this world-altering shift by learning how to understand, build, use and critically evaluate AI in a way that is informed, credible, objective and powerful for answering geoscience questions. The migration of AI into the standard toolkit of EES researchers, practitioners and educators requires a three-way intersection between the established body of EES process knowledge, a mature and balanced understanding of the capabilities and limitations of current and emerging AI techniques, and innovation leading to tailored AI solutions aligned with geoscientific knowledge and practical requirements. This combination of attributes has only come together in a relative handful of EES studies, and in a spontaneous and bottom-up fashion, suggesting that a gentle nudge from above—that is, targeted policy

Box 1 | Machine learning for understanding error sources in a physics-based oceanography model

The data processing and pattern recognition abilities of machine learning can be leveraged to parse large volumes of physics-based model output and determine when and why model biases occur, improving our understanding of the underlying physics and our ability to capture it in process models. In this example[8], a wave model predicts time series of wave parameters (such as significant wave height, mean wave direction and mean wave period) at a specified location (a buoy location, yellow diamond). Imperfections in the physical parameterizations of the process model led to errors in wave height predictions compared with buoy observations (inset time series; WW3 denotes the physics-based model used (WAVEWATCH III) and its predicted wave heights). To find conditions under which this process model is likely to err, a bootstrap aggregated or 'bagged' regression tree (a representative example of which is shown on the right) is used to detect associations. The predictive inputs to the regression tree are the process model output (wave height, period and direction) and process model input (wind speed and direction), and the prediction target is the wave model error; partition numbers in the lowermost boxes in the tree in this example index different error predictions obtained via different regression tree input–space partition paths. By inspecting the architecture of the decision tree, one can determine the regions of the modelled environmental context in which over- and underestimations of wave height are likely to occur, which can in turn be used to improve the physics-based model. Map data © Google.



reform in research and education—is also required.

If these changes do not happen, the world will move on, and EES may be left behind. But where there is challenge, there is opportunity, and EES has potential to capitalize on both its existing deep process knowledge and emerging AI capabilities to push the entire discipline forwards in ways that were unimaginable a very short time ago. ❐

Sean W. Fleming[1,2,3,4] ✉, James R. Watson[2], Ashley Ellenson[2,5], Alex J. Cannon[4,6] and Velimir C. Vesselinov[7]

[1]*National Water and Climate Center, Natural Resources Conservation Service, US Department of Agriculture, Portland, OR, USA.* [2]*College of Earth, Ocean, and Atmospheric Sciences, Oregon State University, Corvallis, OR, USA.* [3]*Water Resources Graduate Program, Oregon State University, Corvallis, OR, USA.* [4]*Department of Earth, Ocean, and Atmospheric Sciences, University of British Columbia, Vancouver, British Columbia, Canada.* [5]*Department of Civil and Construction Engineering, Oregon State University, Corvallis, OR, USA.* [6]*Climate Research Division, Environment and Climate Change Canada, Vancouver, British Columbia, Canada.* [7]*Computational Earth Sciences Group, Environmental and Earth Sciences Division, Los Alamos National Laboratory, Los Alamos, NM, USA.*

✉*e-mail: sean.fleming@usda.gov*

Published online: 2 December 2021
https://doi.org/10.1038/s41561-021-00865-3

References

1. Schwab, K. *The Fourth Industrial Revolution* (Penguin Random House, 2017).
2. Hey, T., Tansley, S. & Tolle, K. *The Fourth Paradigm: Data-Intensive Scientific Discovery* (Microsoft Research, 2009).
3. Fleming, S. W. *Where the River Flows: Scientific Reflections on Earth's Waterways* (Princeton Univ. Press, 2017).
4. Fleming, S. W. & Gupta, H. V. *Phys. Today* **73**, 46–52 (2020).
5. McGovern, A. et al. *Bull. Am. Meteorol. Soc.* **100**, 2175–2199 (2019).
6. Hutchinson, M. et al. Solving industrial materials problems by using machine learning across diverse computational and experimental data. In *American Physical Society March Meeting 2018* BAPS.2018.MAR.K32.2 (American Physical Society, 2018); http://meetings.aps.org/link/BAPS.2018.MAR.K32.2
7. Karpatne, A. et al. *IEEE Trans. Knowl. Data Eng.* **29**, 2318–2331 (2017).
8. Ellenson, A., Pei, Y., Wilson, G., Özkan-Haller, H. T. & Fern, X. *Coast. Eng.* **157**, 103595 (2020).

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41561-021-00865-3.

nature portfolio

https://doi.org/10.1038/s41561-021-00865-3

Supplementary information

Machine learning in Earth and environmental science requires education and research policy reforms

In the format provided by the authors and unedited

Supplementary Materials

Supplementary materials for *Nature Geoscience* paper, "Machine Learning in Earth and Environmental Science Requires Education and Research Policy Reforms," by Fleming, Watson, Ellenson, Cannon, and Vesselinov:

- **Table S2**
- **Literature citations for Table S2**

Supplementary Table S2 Four general classes of physics-aware AI. Other classification schemes are possible, and the list of examples provided is cursory, but this illustration provides a broadly reasonable and useful representation of major directions in the intersection between EES process knowledge and AI. These directions include, but extend well beyond, current interest in explainable machine learning (type 3 below). Framing research funding calls to support physics-aware AI is one of several crucial research and education policy steps needed to successfully integrate machine learning into general EES research and practice. References cited in the examples are provided following the table.

Type	Description	Examples
1. System characterization	• Assessing the inherent predictability of an EES system, including evaluation of chaotic vs. stochastic dynamics and estimation of forecast horizons • Accuracy benchmarking and performance diagnostics for physics-based process models • Provision of direct technical or indirect knowledge support to process simulation model development, diagnosis, and refinement	Benchmarking/diagnosing complex physics-based ocean[1] and land surface[2,3] models, identifying deterministic chaos[4] or prediction horizons[5] in biological systems, guiding climate change impact modeling functional forms[6]
2. Theory-guided AI	• Using *a priori* EES process knowledge to guide development of AI prediction models; also known as 'physics-guided' or 'physics-informed' AI[11] • Spans a wide variety of approaches, from domain expert-guided features engineering, to new AI methods and metasystem algorithms created to embed and enforce biogeophysical constraints, to direct integration of AI components into physics-based process-simulation models	Nonnegativity or monotonicity constraints in AIs for water supply forecasting[7], precipitation analysis[8], and geochemical site characterization[9], hydrology-guided features engineering in flood forecast AI[10]
3. Knowledge discovery	• AI-based pattern recognition to discover physical controls/drivers and the nature of governing relationships underlying complex, nonlinear, and poorly understood phenomena • Includes graphical methods and descriptive AI • Powerful synergies with the computer science community's drive toward explainable/glass-box/white-box AI	Discovering nonlinear dynamics and physical controls for Indian monsoon strength[12], time series memory in geophysical fluid flows[13], gene-gene-environment interactions[14], and ocean beach state[15]; see [16,17] for reviews in meteorology and geoscience
4. Emulation	• Train AI on input data and corresponding simulated output data generated by a physics-based process-simulation model, and then substitute predictions made by the AI in subsequent operations • Generally for the purpose of achieving major computational efficiency gains in iterative procedures like system optimization or in severely time-constrained emergency management contexts	Surrogates for complex physics-based models in optimal aquifer remediation design[18], hurricane storm surge prediction[19], extreme hydrologic event characterization under climate change[20], long-wave radiation transfer within atmospheric models[21]

Reference list for Supplementary Table S1

[1] Ellenson A, Pei Y, Wilson G, Özkan-Haller HT, Fern X. 2020. Coastal Engineering, 157, 103595.

[2] Abramowitz G, Leuning R, Clark M, Pitman A. 2008. Journal of Climate, 21, 5468-5481.

[3] Best MJ, Abramowitz G, Johnson HR, Pitman AJ, Balsamo G, Boone A, Cuntz M, Decharme B, Dirmeyer PA, Dong J, Ek M, Guo Z, Haverd V, Van Den Hurk BJJ, Nearing GS, Pak B, Peters-Lidard C, Santanello JA Jr, Stevens L, Vuichard N. 2015. Journal of Hydrometeorology, 16, 1425-1442.

[4] Wolpert DM, Miall RC. 1990. Proceedings of the Royal Society of London B, 242, 82-86.

[5] Petchey OL, Pontarp M, Massie TM, Kéfi S, Ozgul A, Weilenmann M, Palamara GM, Altermatt F, Matthews B, Levine JM, Childs DZ, McGill BJ, Schaepman ME, Schmid B, Spaak P, Beckerman AP, Pennekamp F, Pearse IS. 2015. Ecology Letters, 18, 597-611.

[6] Solander KC, Bennett KE, Fleming SW, Middleton RS. 2019. Journal of Hydrology: Regional Studies, 26, 100642, doi.org/10.1016/j.ejrh.2019.100642.

[7] Fleming SW, Goodbody AG. 2019. IEEE Access, 7, 119943-119964.

[8] Cannon AJ. 2018. Stochastic Environmental Research and Risk Assessment, 32, 3207-3225.

[9] Vesselinov VV, Mudunuru MK, Karra S, O'Malley D, Alexandrov BS. 2019. Journal of Computational Physics, 15, 85–104.

[10] Fleming SW, Bourdin DR, Campbell D, Stull RB, Gardner T. 2015. Journal of the American Water Resources Association, 51, 502-512.

[11] Karpatne A, Atluri G, Faghmous JH, Steinback M, Banerjee A, Ganguly A, Shekhar S, Samatova N, Kumar V. 2017. IEEE Transactions on Knowledge and Data Engineering, 29, 2318-2331.

[12] Cannon AJ, McKendry IG. 2002. International Journal of Climatology, 22, 1687-1708.

[13] Fleming SW. 2007. Canadian Journal of Physics, 85, 279-294.

[14] McKinney BA, Reif DM, Ritchie MD, Moore JH. 2006. Applied Bioinformatics, 5, 77-88.

[15] Beuzen T, Splinter KD, Marshall LA, Turner IL, Harley MD, Palmsten ML. 2018. Coastal Engineering, 135, 16-30.

[16] McGovern A, Lagerquist R, Gagne DJ II, Jergensen GE, Elmore KL, Homeyer CF, Smith T. 2019. Bulletin of the American Meteorological Society, November, 2175-2199.

[17] Reichstein M, Camps-Valls G, Stevens B, Jung M, Denzler J, Carvalhais N, Prabhat. 2019. Nature, 566, 195-204.

[18] Aly AH, Peralta RC. 1999. Water Resources Research, 35, 2523-2532.

[19] Kim SW, Melby JA, Nadal-Caraballo NC, Ratcliff J. 2015. Natural Hazards, 76, 565-585.

[20] Shrestha RR, Cannon AJ, Schnorbus MA, Zwiers FW. 2017. Climatic Change, 145, 289-303.

[21] Krasnopolsky VM, Chevallier F. 2003. Neural Networks, 16, 335-348.