

PERSPECTIVE OPEN



AI for climate impacts: applications in flood risk

Anne Jones¹✉, Julian Kuehnert², Paolo Fraccaro¹, Ophélie Meuriot³, Tatsuya Ishikawa⁴, Blair Edwards¹, Nikola Stoyanov¹, Sekou L. Remy², Kommy Weldemariam⁵ and Solomon Assefa⁵

In recent years there has been a surge of interest in the potential of Artificial Intelligence (AI) to address the global threat of climate change. Here, we consider climate change applications, and review the ability of AI technologies to better quantify climate change-induced hazards, impacts and risks, and address key challenges in this domain. We focus on three application areas: data-driven modeling, enabling uncertainty quantification, and leveraging geospatial big data. For these, we provide examples from flood-related applications to illustrate the advantages of AI, in comparison to alternative methods, whilst also considering its limitations. We conclude that by streamlining the process of translating weather and climate data into actionable information, facilitated by a suitable technology framework, AI can play a key role in building climate change resilience.

npj Climate and Atmospheric Science (2023)6:63; <https://doi.org/10.1038/s41612-023-00388-1>

INTRODUCTION

Artificial intelligence (AI) is the science and engineering of making intelligent machines¹. Recently, with the rapid growth of machine learning (ML), a subset of AI in which models are learnt from data, there has been considerable interest in understanding how AI technologies can address climate change. The opportunities for leveraging new ML techniques for weather and climate modelling have been discussed extensively by previous authors^{2–4}. Others have considered a broader remit encompassing both climate change mitigation and adaptation^{5,6}. Here, we focus on the usefulness of AI/ML in downstream applications to quantify the risk and impacts of climate change and extreme events (for example, floods, droughts, heatwaves, wildfire). While climate risk and impact assessment considers predictive timescales of decades and beyond, we also include shorter (days to months) timescales in the scope of our discussion, since predictions on these timescales are needed for climate resilience planning by organisations across multiple sectors (e.g. energy distribution, logistics, supply chain, agriculture, and infrastructure management).

We have been building and operationalising AI-based weather/climate applications for research and commercial uses. In this paper, based on our experience, we set out to consider to what extent AI/ML can address key challenges in climate applications. This article is intended to be informative both for climate impact scientists interested in understanding what AI/ML may offer the particular challenges they face, and to the broader climate science community interested in how emerging technology can help translate climate signals into consumable, decision-relevant information.

We begin by reviewing significant challenges in climate impact quantification. We then discuss the main contributions we believe AI/ML can make to addressing these challenges, illustrated with examples from recent work by our group. While the challenges and methods are chosen to be broadly applicable across multiple impacts, the examples are taken from a range of flood-related applications since this domain often experiences the challenges most acutely, whilst also possessing a maturity of methodology, due to the history and scale of flood modeling activities (driven by the huge impact of flooding globally, and the consequent

demand for decision-relevant information). Finally, we draw conclusions and consider future directions.

CHALLENGES IN CLIMATE IMPACT AND RISK QUANTIFICATION

Transforming climate variables into local-scale impacts involves multiple scientific domains, models, and datasets. The challenges associated with this process may be grouped into the following somewhat overlapping categories: (1) system complexity; (2) uncertainty quantification; (3) localisation; (4) computational and practical constraints. We discuss each in turn below.

System complexity

Quantification of climate impacts requires the connection of models in complex workflows in order to create a quantifiable representation of the underlying physical systems. In some cases, creation of individual models may be a challenge due to missing scientific knowledge of some biophysical and earth system processes⁷ or hazard-induced damage². Integration of multiple model types (e.g. simulation, ML, rules-based) written in different programming languages, may be required, along with large spatio-temporal datasets which vary in spatial and temporal resolution, extent, projection, and file format. This process can be time-consuming and inefficient if repeated for each new study. We illustrate this complexity for the example of surface water flood in Fig. 1: the workflow contains six different models and eight different datasets to predict the impact of just one hazard. A further, related challenge is that of compound events: combinations or sequences of different hazards (e.g. droughts and heatwaves) which are climatically correlated and highly impactful. Such events are rare and difficult to model since they cannot be considered independently, or modelled from data alone. Systemic modelling of such events would require multiple, simultaneously-executed versions of Fig. 1. Previous authors⁸ highlight the need for better tools that enable “bottom up” (impact-centric) modelling to identify and better predict such events.

Uncertainty quantification

Uncertainty quantification (UQ) is a widely acknowledged challenge in climate science, addressed in climate projections by

¹IBM Research, Daresbury, UK. ²IBM Research, Nairobi, Kenya. ³Imperial College, London, UK. ⁴IBM Research, Tokyo, Japan. ⁵IBM Research, New York, NY, USA.

✉email: anne.jones@ibm.com

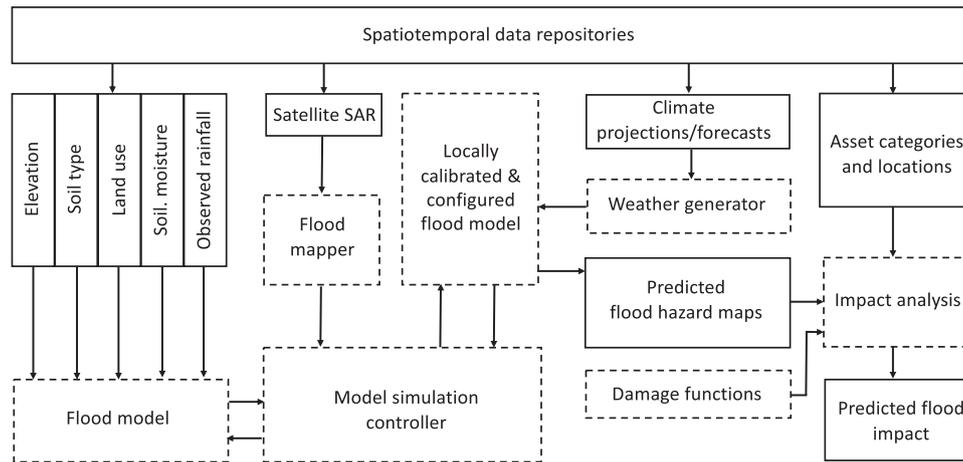


Fig. 1 Example impact modelling workflow to calculate pluvial flood impact. Solid boxes correspond to datasets, dashed boxes correspond to data transformations, consisting of simulation, ML or rules-based models, or other simple processing steps. A flood simulation model is driven by multiple spatiotemporal datasets. A separate flood mapping model is used to derive flood extents from satellite data. A simulation controller (which may be human, automated, or semi-automated) interacts with the flood simulations to calibrate and validate against the satellite-derived ground truth. The calibrated model is then deployed to map flood hazard using a weather generator to transform climate drivers into the daily rainfall inputs required by the flood model. Hazard predictions are then combined with asset locations and damage functions to predict impact in terms of percent of assets damaged or financial cost.

ensembling across multiple emission scenarios and climate models, and at shorter timescale predictions by ensembling over initial conditions. Typically, applications propagate climate ensembles through impact models, resulting in ensemble predictions of crop yield⁹, disease¹⁰ or flood¹¹. Whilst quantification of the weather/climate contribution to uncertainty in this way is often emphasised, uncertainties are introduced at every step of the modelling process: due to observational errors, availability of data for model validation, or unresolved processes. Traditional approaches to extended UQ include the creation of “super-ensembles” of multiple impact models driven by climate ensembles^{12,13}. Full UQ is considered a grand challenge in many applications¹¹.

Interestingly, climate risk assessment does not necessarily require calculation of a combined probability distribution over all contributing uncertainties: methodologies such as “robust” decision-making avoid optimisation of solutions, and instead identify risk management strategies, which minimise regret over uncertainty envelopes¹⁴. This approach is of particular interest for climate change predictions, where the considerable uncertainties mean that ensembles such as CMIP6¹⁵ cannot be used to quantify the likelihood of a particular future climate state. Related is the discrete scenario approach proposed by¹⁶, where experiments are specifically designed to capture high impact, low-probability events. This requires a flexible approach to integrated modelling, and closer collaboration across physical climate science, impact modelling, technologies, and decision-making.

Localisation

Climate impacts are experienced locally, and vary dramatically across the globe¹⁷. A common challenge in impact modelling is the mismatch between the high spatial and temporal precision required for informative local impact predictions, and those of driving climate datasets^{18,19}. This is particularly important for high-impact extreme weather events. In the medium term, this may be addressed via the development of very high-resolution global climate models²⁰, which themselves are likely to be hybrid AI/physics-based models⁴, and will require dedicated “exascale” climate computing technology which is yet to be built²¹. In the interim, dynamic or statistical downscaling is required, which still cannot completely correct the errors in current global climate models.

To provide locally-accurate predictions, it is necessary to parameterise or calibrate impact simulation models against local conditions, or, in the case of data-driven models, use a training dataset, which enables generalisation to a given region of interest. This can be a challenge for applications where ground truth data is limited or difficult to access; for example, disease outbreaks, spatial flood extents, or building damage. This is also a challenge for regions such as sub-Saharan Africa, where observations of weather/climate data, geospatial parameters and impacts themselves may be limited^{22,23}.

Computational and practical constraints

Computational constraints: the availability, speed, and cost of compute (measured in CPU-hours) affect the accuracy achievable by impact models, along with the timeliness of shorter-term predictions. Examples of applications particularly constrained in this way include complex Land Surface Models used to simulate drought impacts²⁴, economic models²⁵, and flood models²⁶. These constraints are particularly important when UQ is undertaken via ensembles of simulations.

Impact modellers and decision-makers also face practical challenges in carrying out climate risk assessment and adaptation planning. These typically center around a lack of tools or expertise to consume climate datasets, issues with data access and sharing, and the need for close interdisciplinary collaboration^{18,19}. Consequently, some authors call for a standardisation of methods, data sharing/access and tools for climate risk assessment²⁷, and current activities by the global community are working towards developing the technologies to support this (e.g.²⁸). Such tools also need to deal with the ever-increasing volume of climate data: estimated to have grown from a few thousands Petabytes in 2010 to over 100,000 in 2020, with a projection of around 350,000 Petabytes by 2030²⁹.

AI METHODS FOR CLIMATE APPLICATIONS

We now consider how AI/ML might address the challenges highlighted in the previous section. We categorise the discussion into three areas, each illustrated with an example: (1) Data-driven modelling; (2) Surrogate modelling, and (3) Leveraging geospatial big data.

AI for data-driven modelling

Data-driven modelling - using ML to derive empirical models from data alone - is a commonly-cited application of AI, which can address several of the challenges previously described: system complexity (by using data-derived models to fill gaps where scientific knowledge is missing), computational constraints (ML models are cheaper to run for inference than simulations) and localisation (ML models are particularly suited to tuning or post-processing of predictions from physically-based models to specific local conditions). Examples include real-time predictions of wildfire risk³⁰, ML-based downscaling of climate variables such as rainfall², and prediction of localised flood patterns⁶. One disadvantage of ML models is their “black-box” nature, meaning it can be difficult to interpret why a model is predicting a particular result. This has led to the field of explainable AI (XAI), which uses a variety of techniques to examine the relative importance of each predictor for specific AI model predictions. XAI is now beginning to be applied to ML modelling in meteorology, climate science, and applications^{31,32}.

Common to many climate applications of AI/ML is the constraint of lack of relevant data for model training at the location of interest. One approach is transfer learning: using models trained in one setting and then refining them with local data³³. The new paradigm of Foundation AI models (to date mainly focused on text and image applications) use this approach at scale, with self-supervised learning on “broad” unlabelled data to develop a general model which can then be refined with minimal training for a range of different applications³⁴. This also addresses the high computational cost of training complex AI models, which may outweigh their advantage in speed over simulation models when making predictions, especially if they cannot generalise to unseen case studies²⁶.

Published examples of AI/ML used for climate applications (e.g.³) are dominated by the use of older ML architectures (e.g. Decision Trees, Random Forest, Support Vector Machines, and Perceptron Neural Networks). Newer architectures, such as those used in Deep Learning (DL) - for example, Convolutional Neural Networks and Long Short-term Memory Networks - are less common, except for specific domains such as remote sensing. This partly due to the age of the technology, but also because they require specialist ML expertise and tools which may not be easily accessible or transferable to climate applications³⁰. This is starting to be addressed via application-specific libraries, for example, TorchGeo³⁵, which supports the application of computer vision models to Earth observation data.

While ML has been used extensively for mapping susceptibility to climate hazards with multivariate spatial drivers, these maps are usually derived for average climate conditions^{36,37}. In³⁸, we demonstrated how ML can also be used to map the risk corresponding to specific unseen and extreme scenarios, therefore potentially providing an alternative to expensive simulations. The flood inundation generated by the ML model for a case study event in the Humber estuary, UK, from³⁸, is compared in Fig. 2, to two alternatives: a simple static “flood fill” model, and a dynamic simulation model. More details are given in³⁸ and the Supplementary Material. The ML model took less than one minute to train on 10,000 sample points across England. Inference times for the ML model depended on the flood map resolution, and were longer than the static model, but substantially shorter than the dynamic model (Table 1). Interestingly, when compared to the flood extent recorded by the Environment Agency, all three models over-predict the inundated area, with over-prediction lowest in the ML model. This is likely to be because of the inability of the DEMs used to resolve features such as sea defenses. This poses a challenge for AI/ML: could a DL model be used to take into account more complex surface features? And if so, would it be at the expense of the fast ML training time? An alternative solution

could involve a hybrid of AI and simulation, which we will discuss in the next section.

AI surrogates of simulations

Surrogate or emulator modelling typically involves training an ML model on simulated data with the aim of replacing the simulation with a cheaper ML version, addressing the challenge of computational constraints. Surrogates may be chosen to represent climate components³⁹, or impact models, such as those for crop yield¹⁷, or hydrological processes⁴⁰. Surrogates also may be used as part of a hybrid AI/physics approach, for example by reducing simulation model resolution to improve efficiency whilst maintaining accuracy⁴¹. A different type of surrogate can be used in simulation model calibration and sensitivity testing, to map model parameter space^{42,43}. We next illustrate how an application of this second kind of surrogate modelling - to find extreme model scenarios in the combined space of climate drivers and impact model parameters - could address the challenge of UQ for climate impacts by efficiently identifying the simulations required to find the limits of the uncertainty envelope.

In this example we quantify pluvial flood risk using a simulation model⁴⁴ driven by a sub-seasonal ensemble rainfall forecast⁴⁵, for a case study event in December 2021, southwest of Bristol, UK. The flood model predicts water depth spatiotemporally, and includes spatially-varying parameters of land use and soil type, as described in Fig. 1. Calibration against past flood events resulted in parameter sets that corresponded to similarly-acceptable simulated flood. We then used an AI method, Gaussian Process (GP)-based Bayesian optimization⁴⁶, a popular technique for optimizing ML model parameters⁴⁷, to efficiently explore a combination of parameter and rainfall driver uncertainty. The optimization determines the set of parameters that, for each member of the rainfall ensemble, maximizes or minimizes an impact metric (here the spatial mean of the maximum flood depth during the event). The GP builds a surrogate function of uncertainty space, which is used to rapidly find the extremes of the metric, by guiding the choice of setting for the next simulation. More details are given in the Supplementary Material. Maps illustrating the range of flooding obtained for default and UQ-optimized parameter show considerable difference in flood extent (Fig. 3). The optimization process and predicted uncertainty envelope across climate driver and impact parameter space are shown in Fig. 4. The predicted uncertainty envelope can then be used to determine the potential minimum and maximum severity and impact of plausible events, in line with the “robust” decision-making approach previously described. Furthermore, the samples collected during optimization can directly be used to evaluate the relative importance of parameters and the climate drivers, to inform the choice of preventive or adaptive measures⁴³.

AI to leverage geospatial big data

AI/ML methods are particularly suited to gathering, structuring, and processing large and diverse geospatial datasets, for example, Earth Observation (EO) satellite images, text documents, or sensor data⁵. Applications of ML to EO data have increased substantially over the last decade, due to great advances in the field of deep learning, especially computer vision algorithms and increasing availability of high resolution satellite imagery^{48,49}, coupled with new generation compute (e.g.⁵⁰). Applications in climate impacts are relatively small in number, and include identification of desertification trends⁵¹, mapping fuel and detection of wildfires³⁰, and flood mapping^{26,49}. These applications address the challenge of system complexity by providing new data and models which can fill gaps in scientific knowledge, and localisation, by providing high-resolution data which can be used to replace or improve physical models. However, satellite data in particular poses a challenge due to its complexity, and increasing data variety and

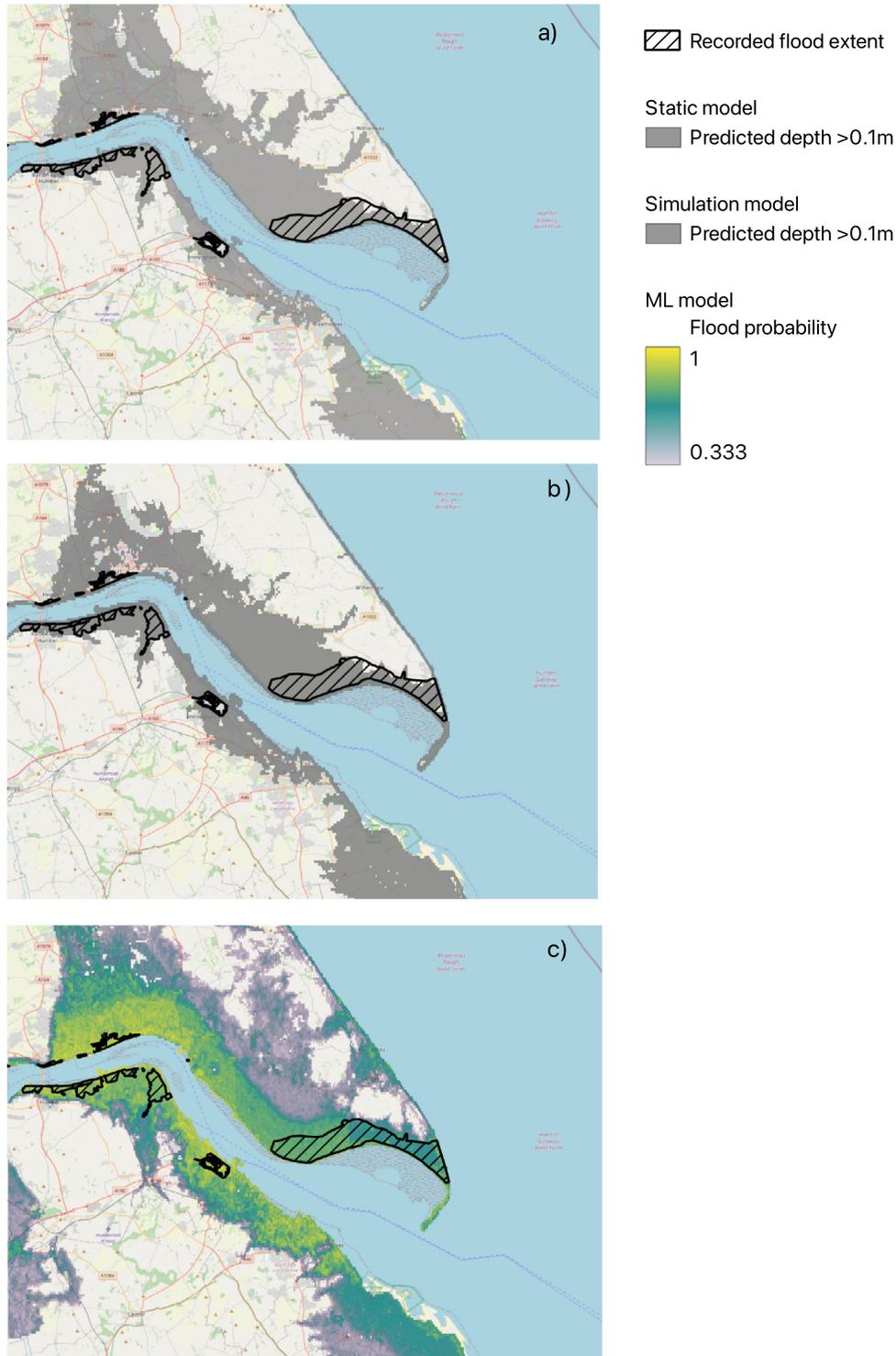


Fig. 2 Comparison of coastal flood inundation maps for the case study of Storm Xavier. Models run at 250m resolution: (a) Static model; (b) Simulation model; (c) Machine learning model. In (a) and (b), black pixels indicate the predicted maximum extent of flood of depths above 0.1m on 5 December 2013. In (c), the model predicts probability of flood inundation for the event. Flood extent recorded by the Environment Agency for 5/6 December 2013 is shown as hatched area.

volumes. For example, the Copernicus Sentinel 2 program provides multi-spectral data (13 output channels from 10 to 60 meters), in 100×100 km tiles, with a size of 700 Megabytes each. This scales to 35 Gigabytes for an individual satellite pass for a country the size of the UK, increasing to 100 Gigabytes if Synthetic Aperture Radar (SAR) imagery from the Copernicus Sentinel 1

program, is also included. Therefore, it is important that AI models are deployed within a technology framework to address these practical challenges.

In⁵², we described such a framework for operational deployment and scaling of an ML model to automatically detect flood water from Sentinel 1 satellite data. Flood extent mapping using Synthetic

Aperture Radar (SAR) satellite imagery has been traditionally undertaken using time-consuming manual or semiautomated statistical methods⁵³. More recent work has demonstrated the potential of fully-automated AI methods using labelled data for training^{54,55}. Both⁵² and⁵⁵ used a “Resnet-UNet” architecture (a state-of-the-art neural network for image segmentation tasks), and

obtained high performance for detection of flood water (0.98 and 0.99 accuracy, respectively) when validated against published quality-controlled flood extent data^{56,57}. An example is illustrated in Fig. 5. By automating the process of flood detection, AI methods can result in time savings of up to 80 percent⁵⁵. The resulting flood extents can be used as “ground truth” data to derive ML flood risk models, and to validate, calibrate and improve UQ in flood simulations, as illustrated in Fig. 1.

Table 1. Computation times for Storm Xavier coastal flood case study inference.

Resolution	Static model	ML model	Dynamic model
250m	0h01	0h03	0h37
50m	0h04	0h30	16h

Times given in CPU hours/minutes and do not include data preparation. Training time is also excluded, since, due to the use of point samples, ML model training time was less than 1 minute for both resolutions. Case study duration was 7 days and extent approximately 4400 km². CPU specification was Intel Xeon 3.3GHz with 16GB RAM.

FUTURE OUTLOOK AND SUMMARY

Quantification of climate hazards, impacts and risks forms the vital connection between climate science and the adaptation decisions which are needed for climate change resilience. This requires modelling of complex systems in the face of multiple constraints: uncertainties and inaccuracies in data and models, computational resources, and the need to connect both technically and practically across scientific communities and end users. In this paper, we have illustrated how AI/ML can begin to address some of these challenges: by improving computational efficiency of models and

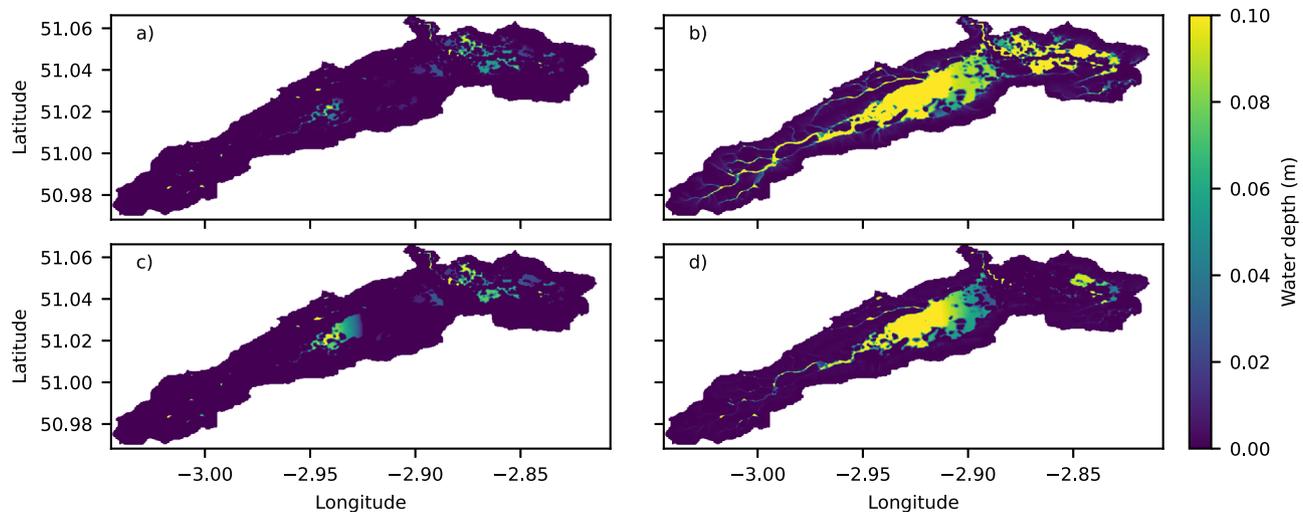


Fig. 3 Flood map for different climate driver ensemble members and parameter sets. **a, b** Use the same rainfall input, as do **(c)** and **(d)**. The default parameter sets are used in **(a)** and **(c)**, while **(b)** and **(d)** are modeled based on optimized parameters to maximize the flood metric. Snapshots taken at the last timestamp of the prediction period, on December 30, 2021.

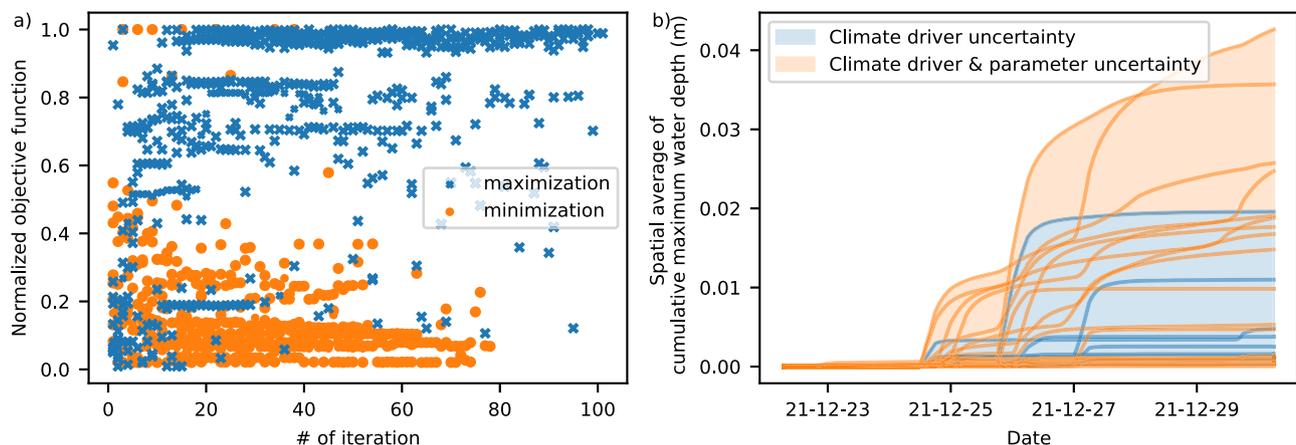


Fig. 4 Time series illustrating flood prediction uncertainty quantification. **a** Evolution of normalized metric value defined by the average of the flood depth maxima at all pixels over the entire time period. With increasing iteration number, the samples from maximization and minimization gather around higher and lower values, respectively. Samples from all 10 climate driver ensemble members are shown together. **b** Climate driver uncertainty alone and combined with parameter uncertainty as a function of prediction time, measured by the evaluation metric.

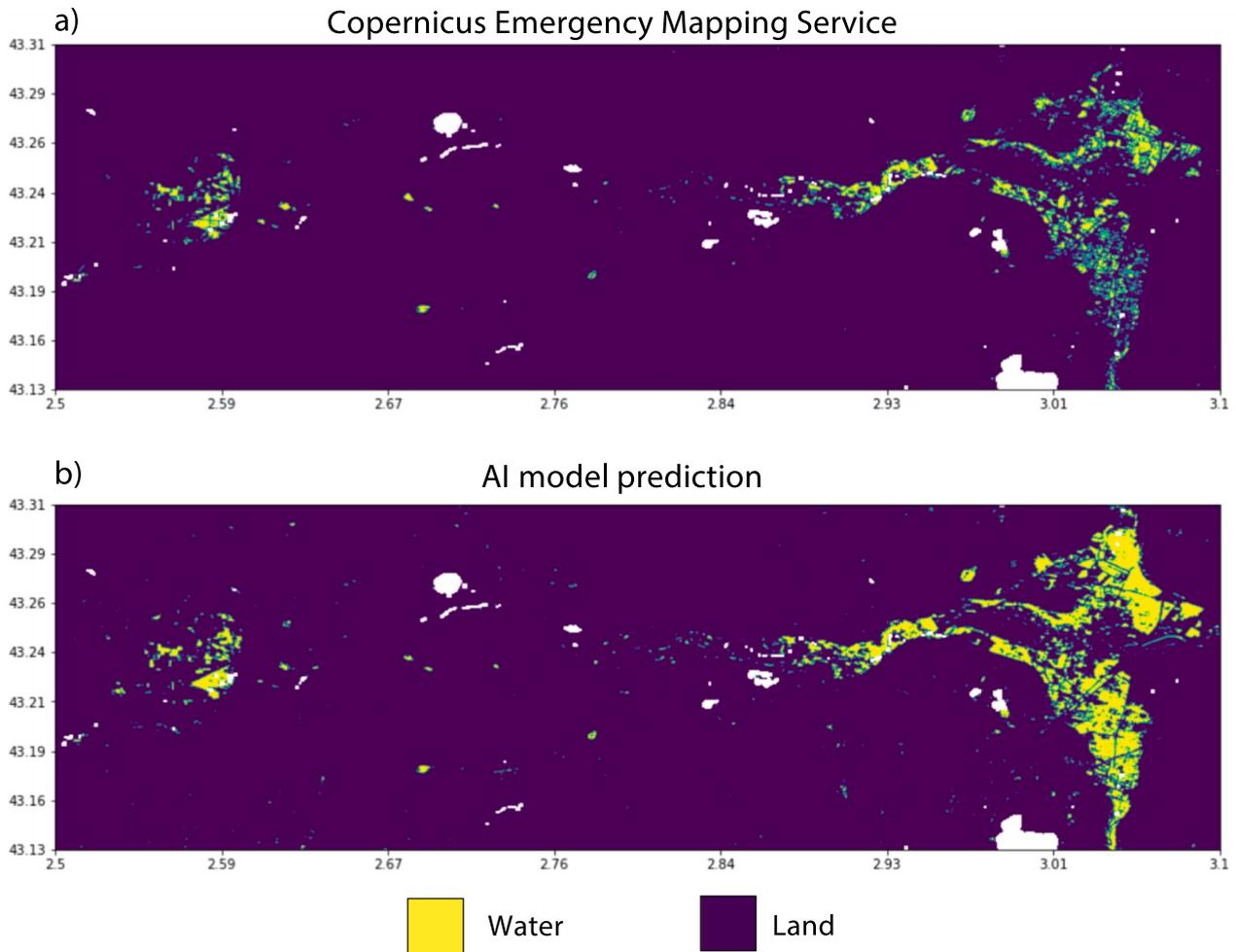


Fig. 5 Flood water detection for an event in Aude, France. Comparison between (a) ground truth data coming Copernicus and (b) AI model from⁵² for an event occurred in October 2020. Purple indicates background, while yellow indicates flood water. White areas identify permanent water areas, which have been mapped out from the plot. The AI model is able to identify more continuous flooded areas, while the Copernicus Emergency Management Service data seems to be more fragmented, possibly due to the use of a threshold-based method.

enabling uncertainty quantification, and by leveraging new geospatial datasets. In a related article⁵⁸, we outline the design of a modelling platform to enable modular, flexible development and deployment of impact modelling workflows. We envisage a future extension of the techniques described in this paper, in combination with such a platform, will enable efficient and easy development and deployment of multiple weather/climate application workflows. Such a system could then begin to address the challenge of flexible, bottom-up modelling and address acute challenges such as deep uncertainty and compound impacts, whilst enabling more streamlined collaboration between scientists and end users. Finally, we note that emerging AI methods now offer the potential to advance scientific knowledge and accelerate solutions in this multi-disciplinary science: for example, to enable data-based discovery of dynamical relationships³, or more easily draw conclusions from vast bodies of existing scientific literature⁵⁹. The full potential of AI in this context remains to be explored.

DATA AVAILABILITY

The coastal flood models used entirely open datasets: DEM data from⁶⁰, recorded flood outlines from the UK Environment Agency, available from <https://www.data.gov.uk/dataset/16e32c53-35a6-4d54-a111-ca09031eaaaf/recorded-flood-outlines>, land cover data from⁶¹ and SSH from⁶². The pluvial flood model used for the uncertainty quantification example used subseasonal rainfall forecast data⁴⁵,

available to purchase from ECMWF. The remaining datasets used to drive the model are all open: soil type from <https://soilgrids.org>, land use data from⁶³, ERA5 soil moisture data⁶⁴, available from <https://cds.climate.copernicus.eu>, and 50m elevation data available from <https://www.ordnancesurvey.co.uk/business-government/products/terrain-50>. The datasets used for the flood detection example are all open, with full details given in⁵².

CODE AVAILABILITY

The AI flood detection, coastal inundation, and pluvial flood model are proprietary models used in the examples and may be made available subject to license, please contact anne.jones@ibm.com. Bayesian optimization was carried out using the python library GPyOpt, freely available from⁴⁶.

Received: 21 September 2022; Accepted: 18 May 2023;

Published online: 08 June 2023

REFERENCES

- McCarthy, J. What is artificial intelligence? (2004). <http://faculty.otterbein.edu/dstucki/inst4200/whatisai.pdf>, Last accessed on 2022-09-12.
- Knüsel, B. et al. Applying big data beyond small problems in climate research. *Nat. Clim. Chan.* **9**, 196–202 (2019).
- Huntingford, C. et al. Machine learning and artificial intelligence to aid climate change research and preparedness. *Env. Res. Lett.* **14**, 124007 (2019).

4. Chantry, M., Christensen, H., Dueben, P. & Palmer, T. Opportunities and challenges for machine learning in weather and climate modelling: hard, medium and soft AI. *Philos. Trans. Royal Soc. A: Mathe. Phys. Eng. Sci.* **379**, 20200083 (2021).
5. Kaack, L., Donti, P., Strubell, E. & Rolnick, D. Artificial intelligence and climate change: Opportunities, considerations, and policy levers to align AI with climate change goals (2021). <https://eu.boell.org/en/2020/12/03/artificial-intelligence-and-climate-change>.
6. Rolnick, D. et al. Tackling Climate Change with Machine Learning. *ACM Comp. Surv.* **55**, 42:1–42:96 (2022).
7. Adger, W. N., Brown, I. & Surminski, S. Advances in risk assessment for climate change adaptation policy. *Philos. Trans. Royal Soc. A Mathe. Phys. Eng. Sci.* **376**, 20180106 (2018).
8. Zscheischler, J. et al. Future climate risk from compound events. *Nat. Clim. Chan.* **8**, 469–477 (2018).
9. Peng, B., Guan, K., Pan, M. & Li, Y. Benefits of seasonal climate prediction and satellite data for forecasting u.s. maize yield. *Geophys. Res. Lett.* **45**, 9662–9671 (2018).
10. Jones, A. E. & Morse, A. P. Application and Validation of a Seasonal Ensemble Prediction System Using a Dynamic Malaria Model. *J. Clim.* **23**, 4202–4215 (2010).
11. Wu, W. et al. Ensemble flood forecasting: Current status and future opportunities. *WIREs Water* **7**, e1432 (2020).
12. Caminade, C. et al. Impact of climate change on global malaria distribution. *Proc. Natl. Acad. Sci.* **111**, 3286–3291 (2014).
13. Colón-González, F. J. et al. Probabilistic seasonal dengue forecasting in Vietnam: A modelling study using superensembles. *PLOS Med.* **18**, e1003542 (2021).
14. Wilby, R. L. & Dessai, S. Robust adaptation to climate change. *Weather* **65**, 180–185 (2010).
15. Eyring, V. et al. Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization. *Geosc. Model Dev.* **9**, 1937–1958 (2016).
16. Sutton, R. T. Climate Science Needs to Take Risk Assessment Much More Seriously. *Bull. Am. Meteorol. Soc.* **100** (2019).
17. Monier, E. et al. Toward a consistent modeling framework to assess multi-sectoral climate impacts. *Nat. Commun.* **9**, 660 (2018).
18. Zebisch, M. et al. The vulnerability sourcebook and climate impact chains - a standardised framework for a climate vulnerability and risk assessment. *Int. J. Clim. Chan. Strat. Manage.* **13**, 35–59 (2021).
19. Tonmoy, F. N., Rissik, D. & Palutikof, J. P. A three-tier risk assessment process for climate change adaptation at a local scale. *Clim. Chan.* **153**, 539–557 (2019).
20. Palmer, T. & Stevens, B. The scientific challenge of understanding and estimating climate change. *Proc. Natl. Acad. Sci.* **116**, 24390–24395 (2019).
21. Slingo, J. et al. Ambitious partnership needed for reliable climate prediction. *Nat. Clim. Change* **12**, 499–503 (2022).
22. Etritch, G. et al. Enhancing digital elevation models for hydraulic modelling using flood frequency detection. *Remote Sensing. Env.* **217**, 506–522 (2018).
23. Casali, Y., Aydin, N. Y. & Comes, T. Machine learning for spatial analyses in urban areas: a scoping review. *Sustainable Cities Soc.* **85**, 104050 (2022).
24. Tramblay, Y. et al. Challenges for drought assessment in the mediterranean region under future climate scenarios. *Earth-Sci. Rev.* **210**, 103348 (2020).
25. Scheidegger, S. & Bilonis, I. Machine learning for high-dimensional dynamic stochastic economies. *J. Comput. Sci.* **33**, 68–82 (2019).
26. Bentivoglio, R., Isufi, E., Jonkman, S. N. & Taormina, R. Deep learning methods for flood mapping: a review of existing applications and future research directions. *Hydrol. Earth Sys. Sci.* **26**, 4345–4378 (2022). Publisher: Copernicus GmbH.
27. Sanderson, H. et al. How do climate risks affect corporations and how could they address these risks? *SN Appl. Sci.* **1**, 1720 (2019).
28. Nativi, S., Mazzetti, P. & Craglia, M. Digital ecosystems for developing digital twins of the earth: The destination earth case. *Remote Sensing.* **13** (2021). <https://www.mdpi.com/2072-4292/13/11/2119>.
29. Overpeck, J. T., Meehl, G. A., Bony, S. & Easterling, D. R. Climate data challenges in the 21st century. *Science* **331**, 700–702 (2011).
30. Jain, P. et al. A review of machine learning applications in wildfire science and management. *Env. Rev.* **28**, 478–505 (2020).
31. Mamlakis, A., Ebert-Uphoff, I. & Barnes, E. Explainable Artificial Intelligence in Meteorology and Climate Science: Model Fine-Tuning, Calibrating Trust and Learning New Science, 315–339 (Springer International Publishing, Cham, 2022). https://doi.org/10.1007/978-3-031-04083-2_16.
32. Chakraborty, D. et al. Scenario-based prediction of climate change impacts on building cooling energy consumption with explainable artificial intelligence. *Appl. Energy* **291**, 116807 (2021).
33. Cheong, S.-M., Sankaran, K. & Bastani, H. Artificial intelligence for climate change adaptation. *WIREs Data Mining and Knowledge Discovery* **14**59 (2022). <https://wires.onlinelibrary.wiley.com/doi/abs/10.1002/widm.1459>.
34. Bommasani, R. et al. On the opportunities and risks of foundation models. *CoRR* **abs/2108.07258** (2021). Preprint at <https://arxiv.org/abs/2108.07258>, 2108.07258.
35. Stewart, A. J. et al. Torchgeo: deep learning with geospatial data. In *Proceedings of the 30th International Conference on Advances in Geographic Information Systems, SIGSPATIAL '22* (Association for Computing Machinery, New York, NY, USA, 2022). <https://doi.org/10.1145/3557915.3560953>.
36. Wang, Y., Fang, Z., Hong, H. & Peng, L. Flood susceptibility mapping using convolutional neural network frameworks. *J. Hydrol.* **582**, 124482 (2020).
37. Bjånes, A., De La Fuente, R. & Mena, P. A deep learning ensemble model for wildfire susceptibility mapping. *Ecol. Inform.* **65**, 101397 (2021).
38. Meuriot, O. & Jones, A. Scalable coastal inundation mapping using machine learning. In *NeurIPS 2021 Workshop on Tackling Climate Change with Machine Learning* (2021). <https://www.climatechange.ai/papers/neurips2021/61>.
39. Reichstein, M. et al. Deep learning and process understanding for data-driven Earth system science. *Nature* **566**, 195–204 (2019).
40. Tran, H. et al. Development of a Deep Learning Emulator for a Distributed Groundwater-Surface Water Model: ParFlow-ML. *Water* **13**, 3393 (2021).
41. Giladi, N., Ben-Haim, Z., Nevo, S., Matias, Y. & Soudry, D. Physics-aware down-sampling with deep learning for scalable flood modeling. In Ranzato, M., Beygelzimer, A., Dauphin, Y., Liang, P. & Vaughan, J. W. (eds.) *Advances in Neural Information Processing Systems*, vol. 34, 1378–1389 (2021). <https://proceedings.neurips.cc/paper/2021/file/0a3b5a7a477d359746061d41c3a04fd6-Paper.pdf>.
42. Yang, J., Jakeman, A., Fang, G. & Chen, X. Uncertainty analysis of a semi-distributed hydrologic model based on a Gaussian Process emulator. *Env. Model. Software* **101**, 289–300 (2018).
43. Wagener, T., Reinecke, R. & Pianosi, F. On the evaluation of climate change impact models. *WIREs Climate Change* **7**72 (2022). <https://onlinelibrary.wiley.com/doi/abs/10.1002/wcc.772>.
44. Singhal, S., Aneja, S., Liu, F., Real, L. V. & George, T. IFM: a scalable high resolution flood modeling framework. In *European Conference on Parallel Processing*, 692–703 (Springer, 2014).
45. Vitart, F. Evolution of ecmwf sub-seasonal forecast skill scores. *Quarterly J. Royal Meteorol. Soc.* **140**, 1889–1899 (2014).
46. The GPyOpt authors. GPyOpt: A bayesian optimization framework in python. <http://github.com/SheffieldML/GPyOpt> (2016).
47. Snoek, J., Larochelle, H. & Adams, R. P. Practical Bayesian Optimization of Machine Learning Algorithms. In Pereira, F., Burges, C. J., Bottou, L. & Weinberger, K. Q. (eds.) *Advances in Neural Information Processing Systems*, vol. 25 (2012). <https://proceedings.neurips.cc/paper/2012/file/05311655a15b75fab86956663e1819cd-Paper.pdf>.
48. Hoese, T. & Kuenzer, C. Object detection and image segmentation with deep learning on earth observation data: A review-part i: Evolution and recent trends. *Remote Sens.* **12**, 1667 (2020).
49. Hoese, T., Bachofer, F. & Kuenzer, C. Object detection and image segmentation with deep learning on earth observation data: A review-part ii: Applications. *Remote Sens.* **12**, 3053 (2020).
50. Kurth, T. et al. Exascale deep learning for climate analytics. In *SC18: International Conference for High Performance Computing, Networking, Storage and Analysis*, 649–660 (2018).
51. Vinuesa, R. et al. The role of artificial intelligence in achieving the Sustainable Development Goals. *Nat. Commun.* **11**, 233 (2020).
52. Fraccaro, P. et al. Deploying an artificial intelligence application to detect flood from sentinel 1 data. *Proceedings of AAAI22* (2022).
53. Matgen, P. et al. Feasibility assessment of an automated, global, satellite-based flood-monitoring product for the copernicus emergency management service. *Eur. Commission, Ispra, JRC119812, EUR* **30073**, 1–47 (2019).
54. Bonafilia, D., Tellman, B., Anderson, T. & Issenberg, E. Sen1floods11: a georeferenced dataset to train and test deep learning flood algorithms for sentinel-1. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops* **210–211** (2020).
55. Nemni, E., Bullock, J., Belabbes, S. & Bromley, L. Fully convolutional neural network for rapid flood segmentation in synthetic aperture radar imagery. *Remote Sens.* **12**, 2532 (2020).
56. Directorate Space, Security and Migration, European Commission Joint Research Centre (EC JRC). Copernicus emergency management service. <https://emergency.copernicus.eu/>. Accessed: 2022-07-27.
57. United Nations Institute for Training and Research. Unosat rapid mapping service. <https://www.unitar.org/maps/unosat-rapid-mapping-service>. Accessed: 2022-07-27.
58. Edwards, B. et al. CIMF: Climate impact modelling framework. In *ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (2022). <https://ai4good.org/wp-content/uploads/2022/08/FE2022-CIMF-Climate-impact-modelling-framework.pdf>.
59. Callaghan, M. et al. Machine-learning-based evidence and attribution mapping of 100,000 climate impact studies. *Nat. Clim. Chan.* **11**, 966–972 (2021).

60. USGS. Global Multi-resolution Terrain Elevation Data 2010 (GMTED2010). <https://doi.org/10.5066/F7J38R2N>.
61. Buchhorn, M. et al. Copernicus Global Land Service: Land Cover 100m: collection 3: epoch 2019: Globe (2020). <https://doi.org/10.5281/zenodo.3939050>.
62. Met Office Hadley Centre. UKCP18 Short Event Case Studies of Historical and Future Sea Surface Elevation around the UK. <https://catalogue.ceda.ac.uk/uuid/58c393f773504caaad48cdb6310e17b2> (2018).
63. USGS. Global land cover characteristics data base version 2.0. <https://www.usgs.gov/media/files/global-land-cover-characteristics-data-base-readme-version2>.
64. Hersbach, H. et al. The ERA5 global reanalysis. *Quarterly J. Royal Meteorol. Soc.* **146**, 1999–2049 (2020).

ACKNOWLEDGEMENTS

This work was supported by the Hartree National Centre for Digital Innovation, a collaboration between STFC and IBM. ECMWF data products ©2022 European Centre for Medium-range Weather Forecasts (ECMWF). Contains OS data ©Crown copyright and database right 2022. Contains public sector information licensed under the Open Government Licence v3.0.

AUTHOR CONTRIBUTIONS

A.J. wrote the manuscript, coordinated the illustrative experiments, and co-designed the impact modelling workflows, along with B.E. N.S. and B.E. created the underlying Climate Impact Modelling Framework computing infrastructure used to run the experiments, with guidance from K.W. O.M. developed the ML model for coastal flood and carried out the experiments for the Hull case study. T.I. wrote the dynamic inundation model used in the case study. J.K. developing the AI techniques for uncertainty quantification along with S.E., and wrote the section “AI surrogates of simulations”. P.F. developed the AI flood detection model and wrote the section “AI to leverage geospatial big data”. A.J., K.W. and S.A. conceived the study, with input from the other authors. All authors reviewed and approved the manuscript.

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/s41612-023-00388-1>.

Correspondence and requests for materials should be addressed to Anne Jones.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023