Foundational AI could usher in a new era for models of all life on Earth

Joseph Millard[1]*, Arman Pili[1], Katherine Berthon[1], David A. Coomes[2], Srinivasan Keshav[3], Anil Madhavapeddy[3], Robert J. Fletcher[1], Lynn V. Dicks[1]

[1]Department of Zoology, University of Cambridge, Cambridge, UK
[2]Department of Plant Sciences, Conservation Research Institute, University of Cambridge, Cambridge, UK
[3]Department of Computer Science and Technology, University of Cambridge, William Gates Building, 15 JJ Thomson Avenue, Cambridge, United Kingdom

*To whom correspondence should be addressed: jwm57@cam.ac.uk

Deep-learned foundational AI could usher in a new era for the simulation of whole ecosystems, argue Joseph Millard and colleagues.

"Time to model all life on Earth"[1]. It's been just over a decade now since Drew Purves and colleagues wrote this sentence, challenging ecologists to develop models for simulating whole ecosystems. The authors then introduced the Madingley model[1], a new form of rule-based general ecosystem model built for the simulation of both terrestrial and marine systems. Taking inspiration from general circulation models and the climate, Purves and colleagues argued that to fully understand how ecosystems will respond to human activities and environmental change, simulation of these ecosystems would be key. This is likely still true. Process-based models such as Madingley can theoretically predict events that haven't yet occurred, but might do under future conditions.

Developed bottom-up from the flow of carbon, plants in the Madingley model were represented via spatially explicit rules for primary production and biomass allocation, driven by environmental inputs such as temperature, precipitation, and remotely sensed net primary productivity. Higher trophic levels were grouped on the basis of traits such as diet, metabolism, and body size, with interactions parameterised on the basis of allometric and metabolic scaling laws[2]. Grouping organisms in this way allowed simulation all the way from plankton and aphids up to whales and elephants.

The Madingley model represented a significant step-forward for ecosystem modelling. But although highly innovative, the grand challenge of reasonably simulating whole ecosystems has never been fully resolved. Whilst general ecosystem models have advanced since the Madingley model[3], they still face fundamental limitations. Practically they are difficult to parameterise, apply, and validate at scales such that they are useful for ecology and conservation. Conceptually, it remains difficult to represent important ecological processes such as organismal intelligence, and to simulate across orders of magnitude in scales of time, space or organismal size[4]. We think that the simulation of ecosystems may be about to take another step forwards, but using a different approach. It's possible that foundational AI elevates our ability to simulate ecosystems.

A new philosophy on simulation

Since the development of Madingley in 2013, neural networks have emerged as a new paradigm in the prediction of natural phenomena. Deep-learning (i.e., a multilayered neural network) has shown that where clean training data and compute are easily available, these models can significantly outperform equivalent physical rule-based algorithms. This has been true for protein folding[5]. It has been true for complex video games[6]. It has been true for weather prediction[7]. In each of these domains, deep-learning has raised the level of what's tractable for simulation.

For ecology and conservation, deep-learning has revolutionised prediction in individual domains such as visual and audio recognition[8], with tagged datasets facilitating supervised learning. But for the simulation of ecosystems, it's currently not clear how a deep-learning model might be trained, or even what it might look like if we did. We think there might be a way. How? A solution might come from repurposing and combining discrete forms of foundational AI.

On the shoulders of scale

In 2020, neural network scaling laws arrived[9]. A team of researchers at OpenAI demonstrated that through scaling-up 3 parameters—neural network model size, dataset size, and the quantity of compute—large language models (LLM), a class of self-supervised learning models, get predictably better at predicting the next word[9].

ChatGPT was a bet on the philosophy of scale. Build a very large neural network for predicting the next word, train it on a large amount of text and compute, and hope that LLM scaling continues to hold.

The bet paid off. The eventual result was an LLM that could not only use language in a human-like manner, but also appeared to gain capabilities across many novel tasks[8]. We now call this form of model "foundational AI"—a pretrained, self-supervised general-purpose neural network, on which further fine-tune improvements can be made—and it is rapidly becoming ubiquitous in many parts of society, lifting the level on what an individual can do with a computer. Scaling has since been applied across multiple other domains, with new forms of foundational AI appearing across multiple modalities (e.g., speech, text, images). Foundational AI models require no expert labelling, and thus can exploit large, unlabeled, noisy observational data. Ecologists now use foundational AI models to accelerate their analytical workflows[11]. Applications for the development of ecological models themselves are growing[12].

Workshop a way forward

As part of an ARIA (Advanced Research + Invention Agency) funded pre-programme discovery project, over four days we invited world-leading experts in computer science and ecology to two workshops in Cambridge, UK. Our aim was to explore the potential of deep-learned foundational AI in ecology, particularly with respect to the forecasting of ecosystem resilience and the development of a new type of general ecosystem model. We focussed on two types of foundational AI: multi-modal Earth-observation models and generative agent-based simulation. In the context of ecosystem modelling, our reasoning was that Earth-observation models might help parameterize the dynamics of vegetation growth, whereas generative agent-based models might better enable the simulation of animal populations.

A multi-modal Earth-observation model uses self-supervised learning to compress large quantities of Earth observation data into vector embeddings. These embeddings embody significant information about the structure of the Earth's surface, often including multiple data modalities, in a form that's both repurposable and easy to handle. Two examples are TESSERA[13] and AlphaEarth[14], respectively developed by the University of Cambridge and Google Deepmind. Where some form of large-scale spatially explicit task is concerned, multi-modal Earth-observation models will likely come to form an initial building block across many analytical pipelines.

Generative agent-based simulation is a new form of agent-based model, in which behaviour of individual agents emerges from multiple generative AI models running simultaneously. Concordia is one example platform[15], inspired by the game Dungeons and Dragons. Developed by Google Deepmind, in Concordia each agent is an instantiation of an LLM, hosted such that it can interact with multiple other LLMs. One LLM agent is the 'Game Master', responsible for tracking the environment at any given time and retaining a history of all prior interactions. All other LLM agents are characters in the simulation, playing out their roles in real-time. Crucially, Concordia assumes that LLM agents, given their training, can act reasonably and apply common sense in any given circumstance.

Since LLMs are trained on human data, most work using generative agent-based models has concerned the simulation of human social behaviours. For example, generative agent-based simulation has been used to show that strongly-held views on COVID-19 were underpinned by an ideological landscape in existence before the pandemic[16]. But if LLMs contain significant ecological knowledge, and they can reason about how this knowledge relates to behaviour, then it's possible generative agent-based simulation will also be applicable in ecology. Promisingly, recent research has shown that the model weights alone (i.e., without consulting the web) of an LLM do indeed possess a significant degree of ecological knowledge[17].

Over the course of our programme discovery project and workshop, a broad idea emerged: If the embedding of a multi-modal Earth-observation model can be fed to an environment tracking LLM, and other LLMs can interpret that embedding reasonably, then perhaps there could be a way to simulate specific ecological interactions, for any specific location on Earth's surface. The benefit of a modelling platform such as Concordia is that the environment in which agents interact can be parameterised, held constant, and tracked by the Game Master. What this means in the context of an ecosystem model, is that the ecological phenomena of interest can be teased apart from the environment in which it exists. That might mean we could simulate the spread of the reintroduction of beavers, for example, without needing to explicitly parameterise their environment. In other words, the fabric and nuance of ecological and environmental reality would be learned into a foundational neural network, and then generated on-the-fly by the Game Master only as it's encountered in simulation.

The momentum is there

Generative agent-based general ecosystem simulation could be fruitful, but there is a problem with this type of approach. For any simulation, quantifying its validity is challenging. Without a held-out empirical truth with which to validate, it's not clear any prediction of a generative agent is reasonable. This will be true even where a model is parameterised via a multi-modal Earth-observation embedding.

Growing data sources might help solve this problem. Since the release of Madingley, total observations on iNaturalist have increased exponentially, now numbering 200 million unique observations[18]. Standardised and large databases of biodiversity data accumulated from the literature abound[19,20], further accelerated by the accumulation of evidence via automated approaches[22,23]. And momentum has grown for multi-scale standardised biodiversity monitoring[22]. Moving forward we must anticipate the ground-truth data we need, and then collect it deliberately. Growing biodiversity data should be held out from the training data underpinning foundational AI simulations, at least in part. Held-out data can then be used to post-train or validate the predictions of simulation, helping to ensure models are not just memorising, allowing a better test of generalisability. It's also possible that as data scales, we find new ways to directly train foundational AI ecological models, perhaps without needing to simulate generative agents at all[24].

The Madingley model represented an ambitious step forward, but ecosystem models have not yet gained traction with conservation decision makers. Conventional general ecosystem models are based on a set of bottom-up governing rules that are challenging to parameterize, contextualise precisely, and validate. We think this may now change. We think it's possible ecological simulation in the future might be learned rather than governed. Generative agent-based models and multi-modal Earth-observation models might be the way. Foundational AI could usher in a new era for the simulation of whole ecosystems.

Acknowledgements

Thanks to all ARIA workshop attendees across our two workshops held at Pembroke College, entitled "Foundational AI to forecast ecosystem resilience". Thanks especially to Mike Harfoot for providing comments before submission of this manuscript. This work was funded by ARIA under one of their Engineering Ecosystem Resilience pre-programme discovery projects. JM is funded by the Leverhulme Trust and the Isaac Newton Trust on an Early Career Fellowship.

References

1. Purves, D., Scharlemann, J. P. W., Harfoot, M., Newbold, T., Tittensor, D. P., Hutton, J., & Emmott, S. (2013). Time to model all life on Earth. Nature, 493(7432), 295–297. https://doi.org/10.1038/493295a
2. Brown, J. H., Gillooly, J. F., Allen, A. P., Savage, V. M., & West, G. B. (2004). Toward a metabolic theory of ecology. Ecology, 85(7), 1771–1789. https://doi.org/10.1890/03-9000
3. Töpper, J. P., Hoeks, S., Hartig, F., Kolstad, A., Tucker, M. A., Bartlett, J., Bærum, K. M., Chipperfield, J., Harfoot, M., Huijbregts, M. a. J., & Svenning, J. (2025). Advancing general ecosystem models (GEMs): Towards a mechanistic understanding of the biosphere in the light of the Anthropocene. Ecological Solutions and Evidence, 6(2). https://doi.org/10.1002/2688-8319.70053
4. Groner, V. P., Cook, J., Orme, C. D. L., Amarasekare, P., Comyn-Platt, E., Rallings, T., Joshi, J., & Ewers, R. M. (2025). Harmonizing nature's timescales in ecosystem models. Trends in Ecology & Evolution, 40(6), 575–585. https://doi.org/10.1016/j.tree.2025.03.011
5. Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., Bridgland, A., Meyer, C., Kohl, S. a. A., Ballard, A. J., Cowie, A., Romera-Paredes, B., Nikolov, S., Jain, R., Adler, J., ... Hassabis, D. (2021). Highly accurate protein structure prediction with AlphaFold. Nature, 596(7873), 583–589. https://doi.org/10.1038/s41586-021-03819-2
6. Vinyals, O., Babuschkin, I., Czarnecki, W. M., Mathieu, M., Dudzik, A., Chung, J., Choi, D. H., Powell, R., Ewalds, T., Georgiev, P., Oh, J., Horgan, D., Kroiss, M., Danihelka, I., Huang, A., Sifre, L., Cai, T., Agapiou, J. P., Jaderberg, M., ... Silver, D. (2019). Grandmaster level in StarCraft II using multi-agent reinforcement learning. Nature, 575(7782), 350–354. https://doi.org/10.1038/s41586-019-1724-z
7. Price, I., Sanchez-Gonzalez, A., Alet, F., Andersson, T. R., El-Kadi, A., Masters, D., Ewalds, T., Stott, J., Mohamed, S., Battaglia, P., Lam, R., & Willson, M. (2024). Probabilistic weather forecasting with machine learning. Nature, 637(8044), 84–90. https://doi.org/10.1038/s41586-024-08252-9

8. Borowiec, M. L., Dikow, R. B., Frandsen, P. B., McKeeken, A., Valentini, G., & White, A. E. (2022). Deep learning as a tool for ecology and evolution. Methods in Ecology and Evolution, 13(8), 1640–1660. https://doi.org/10.1111/2041-210x.13901
9. Kaplan, J., McCandlish, S., Henighan, T., Brown, T. B., Chess, B., Child, R., Gray, S., Radford, A., Wu, J., & Amodei, D. (2020). Scaling laws for neural language models. arXiv. https://doi.org/10.48550/arxiv.2001.08361
10. Bubeck, S., Chandrasekaran, V., Eldan, R., Gehrke, J., Horvitz, E., Kamar, E., Lee, P., Lee, Y. T., Li, Y., Lundberg, S., Nori, H., Palangi, H., Ribeiro, M. T., & Zhang, Y. (2023). Sparks of Artificial General Intelligence: Early experiments with GPT-4. arXiv. https://doi.org/10.48550/arxiv.2303.12712
11. Cooper, N., Clark, A. T., Lecomte, N., Qiao, H., & Ellison, A. M. (2024). Harnessing large language models for coding, teaching and inclusion to empower research in ecology and evolution. Methods in Ecology and Evolution, 15(10), 1757–1763. https://doi.org/10.1111/2041-210x.14325
12. Rafiq, K., Beery, S., Palmer, M. S., Harchaoui, Z., & Abrahms, B. (2025). Generative AI as a tool to accelerate the field of ecology. Nature Ecology & Evolution, 9(3), 378–385. https://doi.org/10.1038/s41559-024-02623-1
13. Feng, Z., Atzberger, C., Jaffer, S., Knezevic, J., Sormunen, S., Young, R., Lisaius, M.C., Immitzer, M., Jackson, T., Ball, J., Coomes, D.A., Madhavapeddy, A., Blake, A., & Keshav, S. (2025). TESSERA: Precomputed FAIR Global Pixel Embeddings for Earth Representation and Analysis. arXiv. https://arxiv.org/abs/2506.20380
14. Brown, C.F., Kazmierski, M.R., Pasquarella, V.J., Rucklidge, W.J., Samsikova, M., Zhang, C., Shelhamer, E., Lahera, E., Wiles, O., Ilyushchenko, S. and Gorelick, N., … Kohli, P. (2025). Alphaearth foundations: An embedding field model for accurate and efficient global mapping from sparse label data. arXiv. https://doi.org/10.48550/arXiv.2507.22291
15. Vezhnevets, A.S., Agapiou, J.P., Aharon, A., Ziv, R., Matyas, J., Duéñez-Guzmán, E.A., Cunningham, W.A., Osindero, S., Karmon, D. and Leibo, J.Z., (2023). Generative agent-based modeling with actions grounded in physical, social, or digital space using Concordia. arXiv. https://doi.org/10.48550/arXiv.2312.03664
16. Kozlowski, A. C., Kwon, H., & Evans, J. A. (2024). In Silico Sociology: Forecasting COVID-19 Polarization with Large Language Models. arXiv. https://arxiv.org/abs/2407.11190
17. Dorm, F., Millard, J., Purves, D., Harfoot, M. and Mac Aodha, O., (2026). Large language models possess some ecological knowledge, but how much? Ecological Informatics. 95(103699). https://doi.org/10.1016/j.ecoinf.2026.103699
18. Mason, B. M., Mesaglio, T., Heitmann, J. B., Chandler, M., Chowdhury, S., Gorta, S. B Z., Grattarola, F., Groom, Q., Hitchcock, C., Hoskins, L., Lowe, S. K., Marquis, M., Pernat, N., Shirey, V., Baasanmunkh, S., & Callaghan, C. T. (2025). iNaturalist accelerates biodiversity research. BioScience, 75(11), 953–965. https://doi.org/10.1093/biosci/biaf104
19. Hudson, L. N., Newbold, T., Contu, S., Hill, S. L. L., Lysenko, I., De Palma, A., Phillips, H. R. P., Alhusseini, T. I., Bedford, F. E., Bennett, D. J., Booth, H., Burton, V. J., Chng, C. W. T., Choimes, A., Correia, D. L. P., Day, J., Echeverría-Londoño, S., Emerson, S. R., Gao, D., . . . Purvis, A. (2016). The database of the PREDICTS (Projecting Responses of Ecological Diversity In Changing Terrestrial Systems) project. Ecology and Evolution, 7(1), 145–188. https://doi.org/10.1002/ece3.2579
20. Dornelas, M., Antão, L. H., Bates, A. E., Brambilla, V., Chase, J. M., Chow, C. F. Y., Fontrodona-Eslava, A., Magurran, A. E., Martins, I. S., Moyes, F., Sagouis, A., Adu-Acheampong, S., Acquah-Lamptey, D., Adam, D., Ajani, P. A., Albaina, A., Almaraz, P., An, J., Anderson, R. S., . . . Zvuloni, A. (2025). BioTIME 2.0: Expanding and improving a

database of biodiversity time series. Global Ecology and Biogeography, 34(5). https://doi.org/10.1111/geb.70003

21. Scheepens, D., Millard, J., Farrell, M., & Newbold, T. (2024). Large language models help facilitate the automated synthesis of information on potential pest controllers. Methods in Ecology and Evolution, 15(7), 1261–1273. https://doi.org/10.1111/2041-210x.14341
22. Gonzalez, A., Vihervaara, P., Balvanera, P., Bates, A. E., Bayraktarov, E., Bellingham, P. J., Bruder, A., Campbell, J., Catchen, M. D., Cavender-Bares, J., Chase, J., Coops, N., Costello, M. J., Czúcz, B., Delavaud, A., Dornelas, M., Dubois, G., Duffy, E. J., Eggermont, H., . . . Torrelio, C. Z. (2023). A global biodiversity observing system to unite monitoring and guide action. Nature Ecology & Evolution, 7(12), 1947–1952. https://doi.org/10.1038/s41559-023-02171-0
23. Besson, M., Alison, J., Bjerge, K., Gorochowski, T.E., Høye, T.T., Jucker, T., Mann, H.M. and Clements, C.F., 2022. Towards the fully automated monitoring of ecological communities. Ecology letters, 25(12), pp.2753-2775. https://doi.org/10.1111/ele.14123
24. Gardiner, R. et al. (2026). Towards ecologically meaningful foundation models. EcoEvoRxiv. https://doi.org/10.32942/X2VM1Z