Ceylon Journal of Science 54 (4) 2025: 959-961
DOI: https://doi.org/10.4038/cjs.v54i4.9207

EDITORIAL

The power of AI in ecological and environmental sciences

Fifteen years ago, I, together with five co-authors, published a paper that suggested some fundamental patterns of how ecosystems respond to experiments that simulate future environmental conditions (Leuzinger et al., 2011). The underlying idea dates back even further, it sprung into my mind during a talk by Claus Beier, who at the time was one of the academic leaders of CLIMAITE, a Danish multifactor global change experiment (Mikkelsen et al., 2008). Based on an early version of a Belgian data base (see Dieleman et al., 2012), I argued that the longer a drought, warming, N addition, or elevated CO_2 treatment lasts, the smaller the ecosystem response, e.g. in terms of carbon fluxes and pools, water relations, but also biodiversity shifts, etc. (Hypothesis 1, H1). Similarly, I found that adding more than one global change driver in an experiment again tended to dampen the overall ecosystem response (Hypothesis 2, H2). To date, the paper has been cited about 430 times, and a lot of additional evidence has appeared since it was published, both in terms of combining global change drivers (e.g., Radolinski et al., 2025), and the evolution of ecosystem responses over time (e.g., Högberg et al., 2024 - possibly the longest running global change experiment in history). Fundamentally, responses can go in three different ways, both with experimental factor combination, and over time. They can yield smaller than additive (antagonistic, Albert et al., 2013), additive (Zavaleta et al., 2003), or larger than additive (synergistic, Reich et al., 2020) responses, and those can decrease (Norby et al., 2010), remain unchanged (Springer et al., 2007), or increase over time (Komatsu et al., 2019). Here, I briefly revisit this topic, one that remains fascinating to me, informally testing and discussing the use of AI to help analyse large amounts of text and data bases for ecological pattern finding.

Rather than providing an exhaustive or systematic analysis, my goal is to use spark ideas for further research, particularly using the available literature, existing data bases (Van Sundert et al., 2023), and AI assisted methods (Scotti et al., 2025). To this end, I first used the paid version of Grok (grok.com) and asked it to analyse the abstracts of all publications that cite my 2011 paper, and categorise them into those that directly support one of the two hypotheses (H1 or H2), and those which do not (suggest the opposite, i.e. static or amplified ecosystem responses). Importantly, the definitions of 'synergistic' and 'antagonistic' can vary (Orr et al., 2020). I stick to the terms 'dampened' (smaller than the reference) and 'amplified' (larger than the reference) to simply refer to an increasing vs. a decreasing (absolute) response. Grok found about half (H1) and a fifth (H2) of the abstracts relevant for such a classification. In regard to H1, 45 were found to support response dampening, 35 amplification, and 40 supported a stagnant or mixed response over time. For H2, 22 were found to support effect dampening, 13 amplification, and 27 an unchanged or mixed response when multiple global change drivers were combined in experiments. I randomly examined three abstracts from each category and did not find a case where my assessment contradicted Grok's classification. This provides some support at least for the general trends described in our 2011 paper. Of course, the caveats are many, and a much more thorough verification of Grok's job would be required. However, the point I want to make here is on the respective time commitment: Grok screened 430 abstracts in 2 minutes and 37 seconds (add to this my time needed to export the abstracts and prompt Grok, say 10 minutes), which stands in stark contrast to the approximately 5 minutes I needed to assess a single abstract manually. At 430 abstracts, this is close to an entire working week. However, the contrast becomes vastly more impressive once the task is to assess 1,000 or 10,000 abstracts, or even entire publications. Grok's time commitment would increase too, but far from proportionally, as would mine.

As a second example, I used the paid version of Julius.ai, a dedicated AI bot for statistical analysis, together with the latest version of the MESI data base (Van Sundert et al., 2023), again to test H1 and H2. The data set contains 57,046 observations and 61 variables, and meta data that are required to understand how it is constructed. A single prompt was enough to get a detailed analysis (a weighted regression approach that accounts for data scarcity of longer running experiments): On average, the absolute effect size decreases by 2.2% per year of experimental duration ($P<0.01$, $R^2_{adj} = 0.006$), and the provided plot I did not ask for looks convincing. To test H2, I asked Julius.ai to run a weighted regression model of absolute effect size predicted by the number of experimental global change drivers. Julius tested a few different regression models, and favoured a quadradic model that shows a reverse U-shaped pattern: an increase in effect size from one to two, but then a decrease from two to three drivers. Again, without claiming that these superficial analyses can be fully trusted, there is some support for H1 and H2. How long did it take Julius to do the analysis (c. 180 lines of code), and how long would it take to do them manually? The former question is easy to answer – under 2 minutes, the latter is a bit more difficult, but I don't think I know anyone who could do this, from familiarisation with the data set to the final plots and model selection, in less than two, much more realistically though, in less than four hours – 60-120 times longer than Julius.

Finally, let me more systematically look at time investment,

**Corresponding Author's Email: sebastian.leuzinger@aut.ac.nz*
https://orcid.org/0000-0001-9306-5281

error frequency, and the explanatory power in context of the above little trial. Earlier studies have estimated the error rates of AI abstract screening at 10%, and it is clear that such methods can only get better (Li et al., 2024). Errors will continue to happen, both in manual and automated data extraction from text. While human error is biased, accidental, and highly dependent on the individual doing the work, the type of errors introduced by AI will be more systematic, easier to remediate, and not researcher dependent. It becomes clear that with the cost (in terms of time, sample size etc.) for data extraction from text becoming negligible, *even* at a relatively high error rate, the opportunity for meaningful pattern finding increases dramatically. To illustrate this – a classification model with 70% accuracy and a sample size of 10 has a statistical power of about 15%, but the same model reaches a power of over 99% at a sample size of 100. Power here indicates the chance of detecting the true accuracy as significantly different from random guessing, which is not quite what we might be after. However, this still illustrates how a large sample size (or, again, minimal cost or time) rapidly and potently outweighs error. In the case of the data analysis example using Julius.ai, my assessment becomes very anecdotal, and nothing more than a personal opinion and outlook on the future: going forward, if the time my students and I spend on data analysis is reduced to, conservatively speaking, a fiftieth of what it used to be, then ecological analysis and pattern finding will, undoubtedly, become much more efficient and powerful.

My demonstration of how two fundamental ecological patterns (effect dampening in relation to H1 and H2) proposed in a 15 year old publication (Leuzinger et al., 2011) could be confirmed in an AI assisted exercise that took barely an hour of my time, yet drew upon 430 abstracts, and a data set with 3.5 million entries, unequivocally shows the power of AI. While it is likely that thorough scrutiny of the presented exercise would prompt further questions, and likely require additional analyses to be conducted, it is also highly unlikely that those would overthrow the found evidence completely. The presented tools are only two of a myriad of powerful AI assistants now available. While some disciplines have picked up the use of AI early on (Koçak & Akçalı, 2024) the environmental and ecological sciences seem to lag behind, and there is definitely a lot of underused potential, particularly in the analysis of global data bases and ecological pattern finding, as well as in conservation ecology (Berger-Tal et al., 2024). Moving forward, I encourage researchers, particularly in my field, to unleash this enormous potential in our research. Importantly, I used paid services, and my demonstration would not have worked using the free base versions of Grok and Julius.ai. Universities should therefore provide licenses for both students and lecturers, as there is an equally urgent need to include these tools in our teaching. Without a positive and forward-looking attitude towards AI tools, we are unable to prepare a future generation of ecologists, who need to provide solutions to rapidly emerging environmental problems.

REFERENCES

Albert, K. R., Boesgaard, K., Ro-Poulsen, H., Mikkelsen, T. N., Andersen, S., & Pilegaard, K. (2013). Antagonism between elevated CO2, nighttime warming, and summer drought reduces the robustness of PSII performance to freezing events. *Environmental and Experimental Botany*, *93*, 1–12. https://doi.org/10.1016/J.ENVEXPBOT.2013.03.008

Berger-Tal, O., Wong, B. B. M., Adams, C. A., Blumstein, D. T., Candolin, U., Gibson, M. J., Greggor, A. L., Lagisz, M., Macura, B., Price, C. J., Putman, B. J., Snijders, L., & Nakagawa, S. (2024). Leveraging AI to improve evidence synthesis in conservation. *Trends in Ecology & Evolution*, *39*(6), 548–557. https://doi.org/10.1016/J.TREE.2024.04.007

Dieleman, W. I. J., Vicca, S., Dijkstra, F. A., Hagedorn, F., Hovenden, M. J., Larsen, K. S., Morgan, J. A., Volder, A., Beier, C., Dukes, J. S., King, J., Leuzinger, S., Linder, S., Luo, Y., Oren, R., De Angelis, P., Tingey, D., Hoosbeek, M. R., & Janssens, I. A. (2012). Simple additive effects are rare: A quantitative review of plant biomass and soil process responses to combined manipulations of CO <inf>2</inf> and temperature. *Global Change Biology*, *18*(9). https://doi.org/10.1111/j.1365-2486.2012.02745.x

Högberg, P., Lucas, R. W., Högberg, M. N., Skyllberg, U., Egnell, G., Larson, J., & Binkley, D. (2024). What happens to trees and soils during five decades of experimental nitrogen loading? *Forest Ecology and Management*, *553*. https://doi.org/10.1016/j.foreco.2023.121644

Koçak, M., & Akçalı, Z. (2024). Development trends and knowledge framework of artificial intelligence (AI) applications in oncology by years: a bibliometric analysis from 1992 to 2022. *Discover Oncology*, *15*(1), 1–24. https://doi.org/10.1007/S12672-024-01415-0/FIGURES/15

Komatsu, K. J., Avolio, M. L., Lemoine, N. P., Isbell, F., Grman, E., Houseman, G. R., Koerner, S. E., Johnson, D. S., Wilcox, K. R., Alatalo, J. M., Anderson, J. P., Aerts, R., Baer, S. G., Baldwin, A. H., Bates, J., Beierkuhnlein, C., Belote, R. T., Blair, J., Bloor, J. M. G., … Zhang, Y. (2019). Global change effects on plant communities are magnified by time and the number of global change factors imposed. *Proceedings of the National Academy of Sciences of the United States of America*, *116*(36), 17867–17873. https://doi.org/10.1073/PNAS.1819027116/SUPPL_FILE/PNAS.1819027116.SD01.CSV

Leuzinger, S., Luo, Y., Beier, C., Dieleman, W., Vicca, S., & Körner, C. (2011). Do global change experiments overestimate impacts on terrestrial ecosystems? *Trends in Ecology and Evolution*, *26*(5). https://doi.org/10.1016/j.tree.2011.02.011

Li, M., Sun, J., & Tan, X. (2024). Evaluating the effectiveness of large language models in abstract screening: a comparative analysis. *Systematic Reviews*, *13*(1), 1–17. https://doi.org/10.1186/S13643-024-02609-X/TABLES/9

Mikkelsen, T. N., Beier, C., Jonasson, S., Holmstrup, M., Schmidt, I. K., Ambus, P., Pilegaard, K., Michelsen, A., Albert, K., Andresen, L. C., Arndal, M. F., Bruun, N., Christensen, S., Danbæk, S., Gundersen, P., Jørgensen, P., Linden, L. G., Kongstad, J., Maraldo, K., … Sverdrup, H. (2008). Experimental design of multifactor climate change experiments with elevated CO2, warming and drought: the CLIMAITE project. *Functional Ecology*, *22*(1), 185–195. https://doi.org/10.1111/J.1365-2435.2007.01362.X

Norby, R. J., Warren, J. M., Iversen, C. M., Medlyn, B. E., & McMurtrie, R. E. (2010). CO2enhancement of forest productivity constrained by limited nitrogen availability. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(45), 19368–19373. https://doi.org/10.1073/PNAS.1006463107/SUPPL_FILE/PNAS.201006463SI.PDF

Orr, J. A., Vinebrooke, R. D., Jackson, M. C., Kroeker, K. J., Kordas, R. L., Mantyka-Pringle, C., van den Brink, P. J., de Laender, F., Stoks, R., Holmstrup, M., Matthaei, C. D., Monk, W. A., Penk, M. R., Leuzinger, S., Schäfer, R. B., & Piggott, J. J. (2020). Towards a unified study of multiple stressors: Divisions and common goals across research disciplines. In *Proceedings of the Royal Society B: Biological Sciences* (Vol. 287, Issue 1926). https://doi.org/10.1098/rspb.2020.0421

Radolinski, J., Vremec, M., Wachter, H., Birk, S., Brüggemann, N., Herndl, M., Kahmen, A., Nelson, D. B., Kübert, A., Schaumberger, A., Stumpp, C., Tissink, M., Werner, C., & Bahn, M. (2025). Drought in a warmer, CO2-rich climate restricts grassland water use and soil water mixing. *Science*, *387*(6731), 290–296. https://doi.org/10.1126/SCIENCE.ADO0734/SUPPL_FILE/SCIENCE.ADO0734_SM.PDF

Reich, P. B., Hobbie, S. E., Lee, T. D., Rich, R., Pastore, M. A., & Worm, K. (2020). Synergistic effects of four climate change drivers on terrestrial carbon cycling. *Nature Geoscience 2020 13:12*, *13*(12), 787–793. https://doi.org/10.1038/s41561-020-00657-1

Scotti, K. L., Young, S., Gainey, M. A., & Lan, H. (2025). Artificial Intelligence and Automation in Evidence Synthesis: An Investigation of Methods Employed in Cochrane, Campbell Collaboration, and Environmental Evidence Reviews. *Cochrane Evidence Synthesis and Methods*, *3*(5), e70046. https://doi.org/10.1002/CESM.70046

Springer, C., Physiology, R. T.-T., & 2007, undefined. (2007). Photosynthetic responses of forest understory tree species to long-term exposure to elevated carbon dioxide concentration at the Duke Forest FACE experiment. *Academic.Oup.ComCJ Springer, RB ThomasTree Physiology, 2007•academic.Oup.Com*. https://academic.oup.com/treephys/article-abstract/27/1/25/1674516

Van Sundert, K., Leuzinger, S., Bader, M. K. F., Chang, S. X., De Kauwe, M. G., Dukes, J. S., Langley, J. A., Ma, Z., Mariën, B., Reynaert, S., Ru, J., Song, J., Stocker, B., Terrer, C., Thoresen, J., Vanuytrecht, E., Wan, S., Yue, K., & Vicca, S. (2023). When things get MESI: The Manipulation Experiments Synthesis Initiative—A coordinated effort to synthesize terrestrial global change experiments. *Global Change Biology*, *29*(7). https://doi.org/10.1111/gcb.16585

Zavaleta, E. S., Shaw, M. R., Chiariello, N. R., Mooney, H. A., & Field, C. B. (2003). Additive effects of simulated climate changes, elevated CO2, and nitrogen deposition on grassland diversity. *Proceedings of the National Academy of Sciences of the United States of America*, *100*(13), 7650–7654.

Sebastian Leuzinger
School of Science,
Auckland University of Technology,
34 St Paul St, Auckland 1010,
New Zealand