25th International Congress on Modelling and Simulation, Darwin, NT, Australia, 9 to 14 July 2023
mssanz.org.au/modsim2023

Multiverse experiments for enhancing reproducibility and resilience of statistical modelling via case studies in family wellbeing and biodiversity conservation

S. Low-Choy [a,b,c], K. Modecki [d,e], D. Vasco [b,d], I. Alkhairy [f,g] and W. Hallgren [c]

[a] Arts, Education and Law Group, Griffith University; [b] Griffith Institute of Educational Research, Griffith University; [c] Centre for Planetary Health and Food Security, Griffith University; [d] School of Applied Psychology, Menzies Health Institute Queensland, Griffith University; [e] Berkman Klein Centre for Internet and Society, Harvard University; [f] School of Information and Computing Technology, Griffith University; [g] Mathematics Department, Umm Al-Qura University, Saudi Arabia
Email: s.low-choy@griffith.edu.au

Abstract: A multiverse may be conceived of as a *hypothetical* set of universes, where the future unfolds along alternate pathways or realities, or as a *real* set of universes that co-exist simultaneously. Similarly, probabilities may be interpreted as the proportion of universes where an event occurs, especially for risk assessment of novel events. By extension, the multiverse idea also applies to modelling via computer experiments that comprise many model runs (cf universes), each with their own inputs—data, parameters, settings, etc (pathways or realities). Hence, a modelling multiverse is a collection of model runs that experiment with inputs, here illustrated by three case studies.

Smartphone case study. A multiverse of ordinal regressions (Modecki et al. 2020) each examined the effect of smartphone usage on parenting quality, as moderated by technoference (the degree to which technology interferes with parenting). Previous studies had typically contemplated only a single measure for each variable. Our multiverse of regressions varied the choice of measures for the outcome, explanatory, and moderator variables. Altogether 84 variable-sets were chosen using a fully factorial design. We used effect sizes and model fit diagnostics to understand how results varied across choice of measures.

Species distribution model (SDM) case study. In practice, many studies report a single model. However, statistical modelling guidelines now advocate sensitivity analysis, of several models. In these contexts, multiverse analyses may serve as "meta-research", that experiments with multiple model choices, at the same time. Assessing model uncertainty in this way is comprehensive but computationally intensive. Yet, multiverses are still finite. Hallgren et al. (2019) applied the One-Factor-At-a-Time experimental design to construct a SDM multiverse, by manipulating algorithms, inputs and several model settings. Implementation was streamlined through a Virtual Laboratory for Biodiversity and Climate Change.

Model choices guided by risk analysis case study. In parallel to the SDM case study, Alkhairy (2019) asked an expert to methodically assess carefully selected sets of SDM settings and options. These were analysed via Bayesian error-in-variables models to identify the best model configuration (most likely occurrence, least uncertainty) and settings (main effects). Statistical performance was compared for several experimental designs, contrasting traditional One-Factor-At-a-Time with Taguchi and fractional factorial designs.

Driven by limitations on resources, experimental design has traditionally helped researchers make good use of physical resources. As described above, experimental design for multiverses may help modellers exploit virtual resources. Importantly, the multiverse enhances reproducibility, avoiding the "cherry picking" effect of a single model run. Conclusions for all case studies would be very different had we chosen a single model specification, such as the typical single variable set, with default model settings. Concluding remarks refer to emerging literature on how multiverses afford studies some "resilience" to criticisms regarding model choice.

REFERENCES

Alkhairy, I. 2019, Designing and Encoding Scenario-based Expert Elicitation for Large Conditional Probability Tables, Ph.D. thesis, School of Information, Computing and Technology, Griffith University.

Hallgren, W., Santana, F., Low-Choy, S., Zhao, Y., Mackey, B. 2019. Species distribution models can be highly sensitive to algorithm configuration. Ecological Modelling, 408, 108719.

Modecki, K.L., Low-Choy, S., Uink, B.N., Vernon, L., Correia, H., Andrews, K. 2020. Tuning into the real effect of smartphone use on parenting: a multiverse analysis. Journal Of Child Psychology and Psychiatry, and Allied Disciplines, 61(8), 855–865.

Keywords: *Experimental design, computer experiments, reproducibility*