JUDITH A. VERSTEGEN
Institute for Geoinformatics
University of Münster
Email: j.a.verstegen@uni-muenster.de

Judith Verstegen obtained a Bachelor of Science (BSc) in Earth Sciences at Utrecht University in 2007, and a Master of Science (MSc) in GeoInformation Science at Wageningen University, finishing this Masters cum laude in 2010 with a thesis on agent based modelling of the spatial planning process of urban expansion. She continued as a PhD candidate at Utrecht University in the BE-Basic program, a public-private partnership developing industrial bio-based solutions for a sustainable society. She defended her PhD thesis with the title “Quantifying and reducing uncertainty in land use change model projections - Case studies in the implications of increased bioenergy demands” successfully in February 2016.

During her PhD, she also contributed to the educational programs of Earth Sciences and Geographical Information Management and Applications (GIMA) by supervising MSc students during their theses, supervising BSc students during fieldwork, giving lectures and computer practicals and preparing course materials for a variety of BSc and MSc courses, thus obtaining the qualifications for teaching higher education (BKO).

Verstegen has worked at the Institute for Geoinformatics, University of Münster as a junior professor (comparable to assistant/associate professor in other countries) since 2016. There, she leads the Geosimulation Modeling lab. She is interested in geosimulation models (agentbased models, cellular automata), spatial statistics, error propagation methods, and spatial optimization. Currently she works primarily with models of land use change and urban dynamics. The main aim of her research is to better quantify the predictive value of geosimulation models. She has published 18 peer-reviewed articles in scientific journals, and contributed to a range of conferences, such as Agile, EGU, AGU, CAMUSS, and CUPUM.

Verstegen currently teaches in two MSc programs: Geoinformatics, and Geospatial Technologies. She coordinates the courses Geosimulation Modeling, and Python in GIS, and contributes to a range of other courses. She has supervised over 20 BSc and MSc students on their thesis. Furthermore, she is the head of the Graduate School for Geoinformatics, and a member of the editorial board of *Computers & Geosciences* and *Environmental Modeling & Assessment*.

A Plea for Statistical Analyses of Geosimulation Model Projections

The model development cycle

Spatial data science is concerned with the development of methods to gain understanding of the processes in a certain system, and to project the potential future state of this system. The two aims of gaining process understanding and making projections are strongly related in an iterative manner (e.g., van Vliet *et al.,* 2016), as explained in the following. Process understanding, stemming from either existing theories or the analysis of historical data, allows for the formulation of a conceptual model of the system. The conceptual model can be translated into a computational model that simulates the system dynamics over space and time, a so-called geosimulation model. When this geosimulation model is run for a historical period, the simulated system state (model output) can be compared with the observed system state (empirical data). The interpretation of the inevitable differences between the two, may lead to enhanced process understanding, such that the conceptual model can be adapted, and the next iteration can start.

Uncertainty quantification—stochastic modelling vs. model ensembles

Irrespective of the number of iterations carried out, a geosimulation model always remains (and should remain - cf. Occam's Razor) a simplification and thereby an inexact representation of the corresponding real-world system. Broadly speaking, two approaches are in use[1] to account for the uncertainty resulting from this. The first approach is to define the processes probabilistically in the geosimulation model and to perform error propagation with the resulting stochastic model to quantify the uncertainty in the model outcomes (e.g., Verstegen *et al.,* 2016). The second approach is to pick a set of separate discrete models that represent a set of possible process descriptions and perform multi-model ensemble runs to build a set of outcomes (e.g. Alexander *et al.,* 2017). So, both approaches result in a distribution of model outcomes that represents, for one or multiple variables, a range of potential values and their probabilities given the model inputs and the current knowledge about the system.

Scenario-thinking

For many systems, especially the systems humans may influence or steer, one is interested in not one future system state, but a set of future states related to a set of story lines. The quantitative descriptions of these storylines are called scenarios. Scenario-thinking has become common in the last two decades, particularly through the climate change mitigation scenarios developed over several successive IPCC reports (e.g. IPCC, 2014). A scenario analysis with a stochastic model (and error propagation) or a model ensemble thus results in a set of model outcome distributions, i.e. one distribution for each scenario (Figure 1).

The problem

Unfortunately, model outcome distributions are not used to the fullest at the moment. At best, the median (or worst, the mean), and minimum and maximum (or x^{th} and y^{th} percentile values) are reported, as in Figure 1. Firstly, this does not give the full picture, and secondly, end users (e.g. policy makers) typically do not know what to do with three values per scenario, and are therefore likely to resolve to using the median c.q. mean only. Thereby valuable uncertainty information is discarded.

[1] Although uncertainty quantification methods are in frequent use in the physical modelling domains (climate science, hydrology, etc.), they are only sporadically used in social science domains (land use change, urban dynamics, etc.).

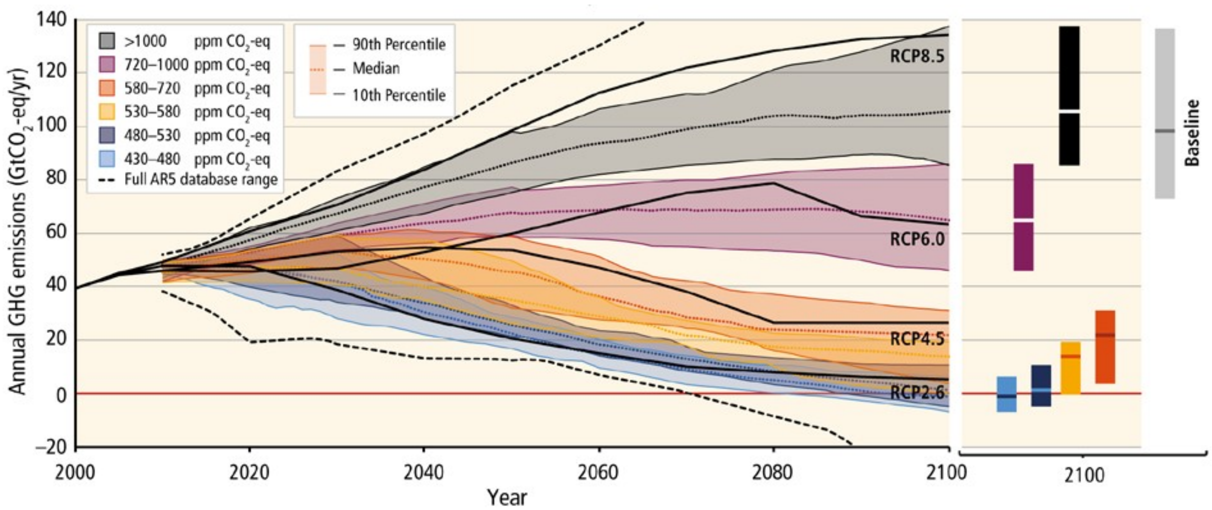

Figure 1: GHG emission pathways 2000-2100: All AR5 scenarios, taken from Figure SPM.11a (IPCC, 2014).

Towards a solution

Though error propagation modelling and model ensemble usage are relatively new approaches, that were boosted by the ever increasing computing capacities, the more general situation of having one or more outcome distributions is certainly not new. Scientific domains that deal with experiments, the collection of data samples, or user studies have a developed a wide range of statistical methods to derive information from one or more distributions of values of a certain variable. I want to argue that these methods offer the geosimulation modelling community an excellent opportunity to get more information and more usable answers out of stochastic models and model ensembles.
Here, I will give two examples of how questions that may arise from Figure 1 can be answered with statistical methods stemming that have their origin in the analysis of experimental data. One may ask: How certain can we be (or can we be at least X% certain) that we will reach annual GHG emissions of zero or less in 2100 in the light blue scenario? Furthermore, since the ranges of the light blue and dark blue scenario outputs seem to overlap a great deal, one may also ask: Does the light blue scenario (that requires more effort) result in mean emissions that are significantly lower than the mean emissions of the dark blue scenario in 2100? If the answer is no, we might not need to bother to spend the extra money that the light blue scenario requires.

Neither of these two questions can be answered with only mean, minimum and maximum of the output distribution from the model ensemble. The modellers, who have the full output distribution, can compute the exceedance probability with respect to 0 gigatonne CO_2-equivalent per year, and thereby answer the first question. The second question pertains to hypothesis testing. The question is analogous to a very common question in e.g. medical research: Does the treatment group outperform the control group? A wide range of methods is available to answer this type of question. A t-test may be used in case the output distributions match the conditions for a parametric test. Otherwise, non-parametric alternatives, such as the Kolmogorov-Smirnov test, may be applied. Furthermore, in case one wants to compare multiple scenarios among each other, an ANOVA with post-hoc test may be exploited (Basso *et al.,* 2009). To the best of my knowledge, none of this is currently being done with geosimulation model projections.

Therefore, I argue that the spatial data science community should look into the applicability of statistical methods to geosimulation model output distributions, especially in the context of scenario analysis. Some methods and tests may be directly applicable whereas others may need to be adapted. Some may be applicable to model ensembles (usually containing a low number of realizations, e.g. 7), but not to the outputs of error propagation modelling (usually yielding a very high number of realizations, e.g. 10 000), or the other way around. The ultimate goal would be a set of transferrable statistical methods to evaluate geosimulation model output distributions, accompanied by guidelines to pick the right method for the type of model, type of uncertainty quantification method, and type of scenario. A bonus of such statistical methods is that the added value of uncertainty analysis becomes more concrete, which may convince domains in the geosimulation community in which uncertainty analysis is not practiced at this moment, to start doing so.

References

Alexander, P., Prestele, R., Verburg, P.H., Arneth, A., Baranzelli, C., Batista e Silva, F., Brown, C., Butler, A., Calvin, K., Dendoncker, N., Doelman, J.C., Dunford, R., Engström, K., Eitelberg, D., Fujimori, S., Harrison, P.A., Hasegawa, T., Havlik, P., Holzhauer, S., Humpenöder, F., Jacobs-Crisioni, C., Jain, A.K., Krisztin, T., Kyle, P., Lavalle, C., Lenton, T., Liu, J., Meiyappan, P., Popp, A., Powell, T., Sands, R.D., Schaldach, R., Stehfest, E., Steinbuks, J., Tabeau, A., van Meijl, H., Wise, M.A. and Rounsevell, M.D.A. (2017) Assessing uncertainties in land cover projections, *Global Change Biology* 23(2):. 767–781. DOI: 10.1111/gcb.13447.

Basso, D., Pesarin, F., Salmaso, L. and Solari, A. (2009) *Permutation Tests for Stochastic Ordering and ANOVA: Theory and Applications with R.* Springer Science & Business Media. DOI: 10.1007/978-0387-85956-9.

IPCC (2014) *Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change.* Geneva, Switzerland: IPCC.

van Vliet, J., Bregt, A.K., Brown, D.G., van Delden, H., Heckbert, S. and Verburg, P.H. (2016) A review of current calibration and validation practices in land-change modeling, *Environmental Modelling & Software* 82: 174–182. DOI: 10.1016/j.envsoft.2016.04.017.

Verstegen, J.A., van der Hilst, F., Woltjer, G., Karssenberg, D., de Jong, S.M. and Faaij, A.P.C. (2016) What can and can't we say about indirect land-use change in Brazil using an integrated economic land-use change model? *GCB Bioenergy* 8: 561–578. DOI: 10.1111/gcbb.12270.