

Replication assessment for SEAcross v1.0

Christine Hertler^{1,2}, Jan-Olaf Reschke^{1,2,3}, Ericson Hölzchen^{4,5}, Iwan Pramesti Anwar⁶, Mika Puspaningrum⁶, Nicolas Büscher³, Emmanuel Kipruto Ngetich^{1,2,3}

¹ ROCEEH Research Center, Senckenberg Research Institute, Frankfurt am Main, Germany

² ROCEEH Research Center, Heidelberg Academy of Sciences and Humanities, Heidelberg, Germany

³ Paleobiology and Environment Group, Institute of Ecology, Evolution, and Diversity, Dept. of Biosciences, Goethe University, Frankfurt am Main, Germany

⁴ Trier Lab for Social Simulation (TRILAB), Trier University, Trier, Germany

⁵ German Center for Artificial Intelligence (DFKI), Cognitive Social Simulation, Trier University, Trier, Germany

⁶ Bandung Institute of Technology, Faculty of Earth Sciences and Technology, Bandung, Indonesia

Because the SEAcross ABM contains several sources of stochasticity, e.g. the initial positions of the agents in the source area, resource distribution, the probability by which the agent is allowed to deviate from heading towards the most attractive patch (“decision randomness”) and the probability by which agents avoid entering the sea strait (“hydrophobicity coefficient”), consecutive model runs will not yield identical results. A certain number of model runs is required to produce statistically robust values for the output variables. Because the impact of the factors varies with the output selected, several output variables need to be monitored.

In order to identify the number of runs required we performed a replication assessment (Hoad et al. 2010, Lorig 2018). In order to assess the required number of replications, we consider the values of each output variable as normally distributed and calculate the cumulative standard deviation from the cumulative statistical mean. We choose a 95 % confidence interval for the output variables displayed in the model.

	Factor	Factor level
Agent	Decision-randomness	0.1
Agent	Hydrophobia-coefficient	0.25
Agent	Bodymass	3,680
Agent	Head-body-length	6
Agent	Number-founder-agents	20
Environment	Water-barrier-size	70
Environment	Current-direction	90
Environment	current-speed	0.1

Tab. 1: Factor specifications applied in the replication assessment

Results and discussion

1. Output variables 'Crossing success' (CSR)

The CSRs are assessed for both, individuals and founder populations. For individuals a precision of the model lower than the 0.05 desired precision is reached after 13 replications, and for founder populations after 14 replications (Fig. 1).

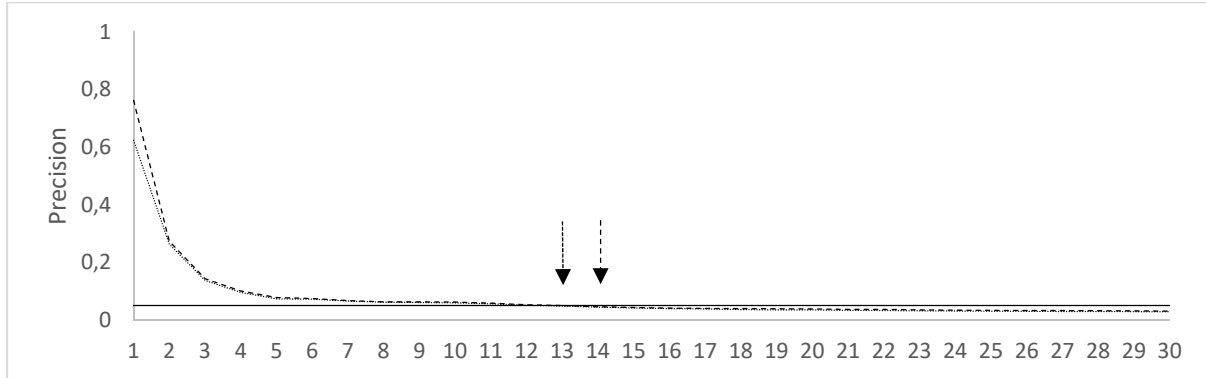


Fig. 1: Precision for both of the CSR output variables, individual (dotted), and founder populations (dashed); the confidence interval is displayed as solid line. Desired precisions for both of the variables are indicated by arrows.

2. Output variables 'Causes of disappearance' (COD)

We monitor four different causes of disappearance in the SEAcross model. These are exhaustion and the various ways to leave the map, i.e. over land in the source area, over land in the target area and being washed off the map while crossing the sea strait. Tab. 2 lists the minimum number of replications required to reach the 0.05 criterion of desired precision.

Cause of disappearance (COD)	Minimum number of replications
Exhaustion	16
Leave map water	4
Leave map source area	not reached after 50 replications
Leave map target area	not reached after 50 replications

Tab. 2: Replications after which precision falls below the confidence interval of 0.05

With respect to both sets of output factors monitored in the SEAcross model, the required minimum number of replications does not exceed 20. We opt for 25 replications. The number of replications is rather low for simulation-based models.

This does, however, not ensure robusticity for terrestrial movements, in which the agents leave the border of the map on land, either in the source or the target area. We are only monitoring both of the output factors for completeness and do not intend a detailed discussion. If this is intended at any stage, the replication assessment should be repeated with a sufficient number of model runs.

References

Hoad K, Robinson S, Davies R. 2010: Automated selection of the number of replications for a discrete-event simulation. *Journal of the Operational Research Society* 61: 1632–1644.

Lorig F. 2018. Hypothesis-Driven Simulation Studies. Assistance for the Systematic Design and Conducting of Computer Simulation Experiments. Dissertation, Trier University.