

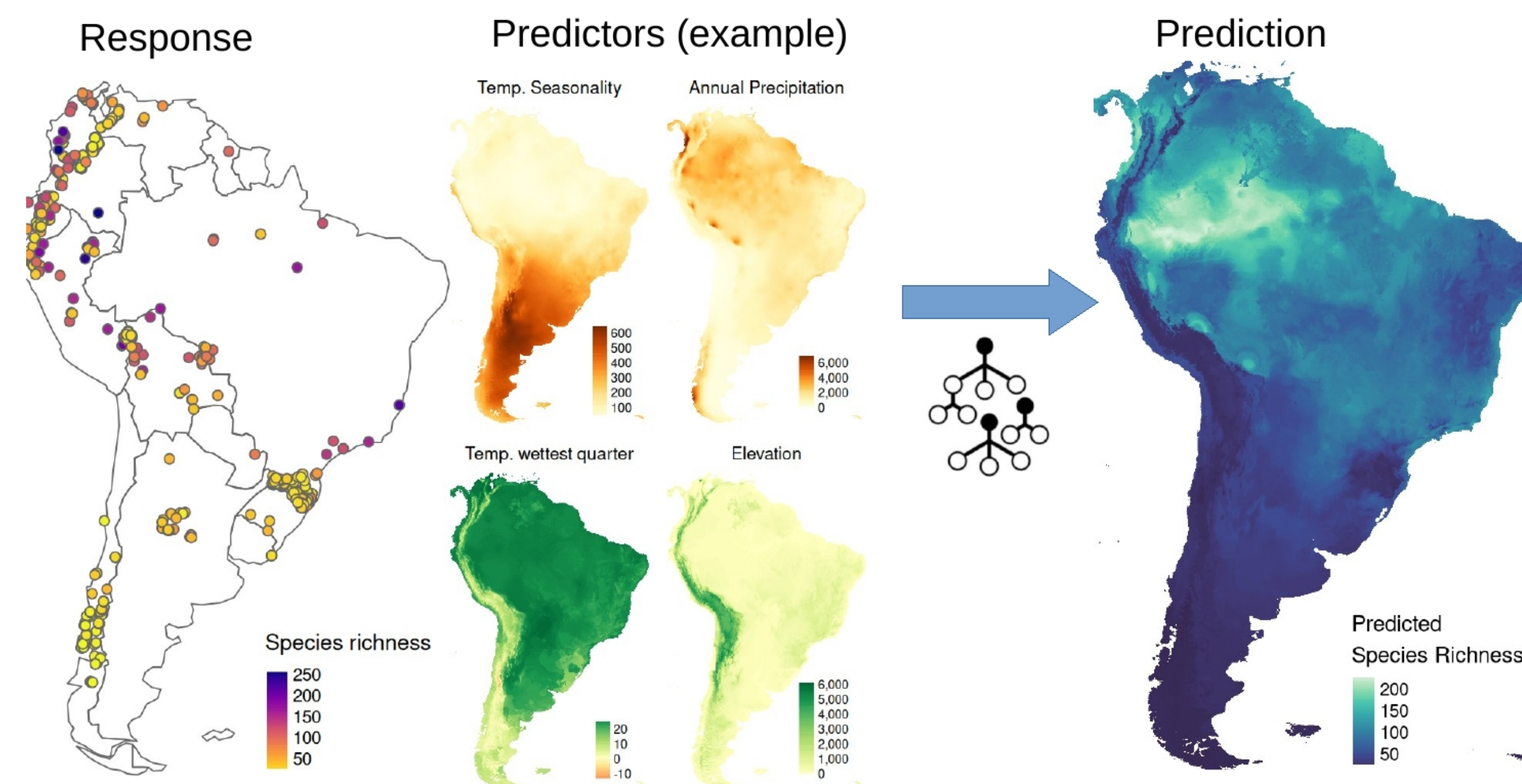


STeMP: A Spatio-Temporal Modelling Protocol

Jan Linnenbrink, Jakub Nowosad, Marvin Ludwig, Anna Frederike Jablotschkin, Hanna Meyer

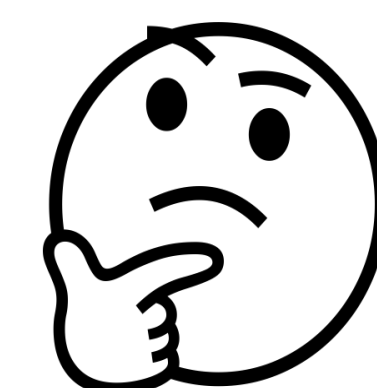
Background: Spatio-temporal machine learning is increasingly used in geosciences to derive spatial continuous data ("maps") of the environment based on limited field observations and remote sensing data.

Problem: Models and their predictions often lack transparency and reproducibility due to inadequate reporting of assumptions, modelling decisions, uncertainties and applicability, which calls for a standardized documentation. Existing model documentation standards, however, are either generic (e.g., for general ML) or tailored to very specific applications. A protocol for spatio-temporal ML is missing.



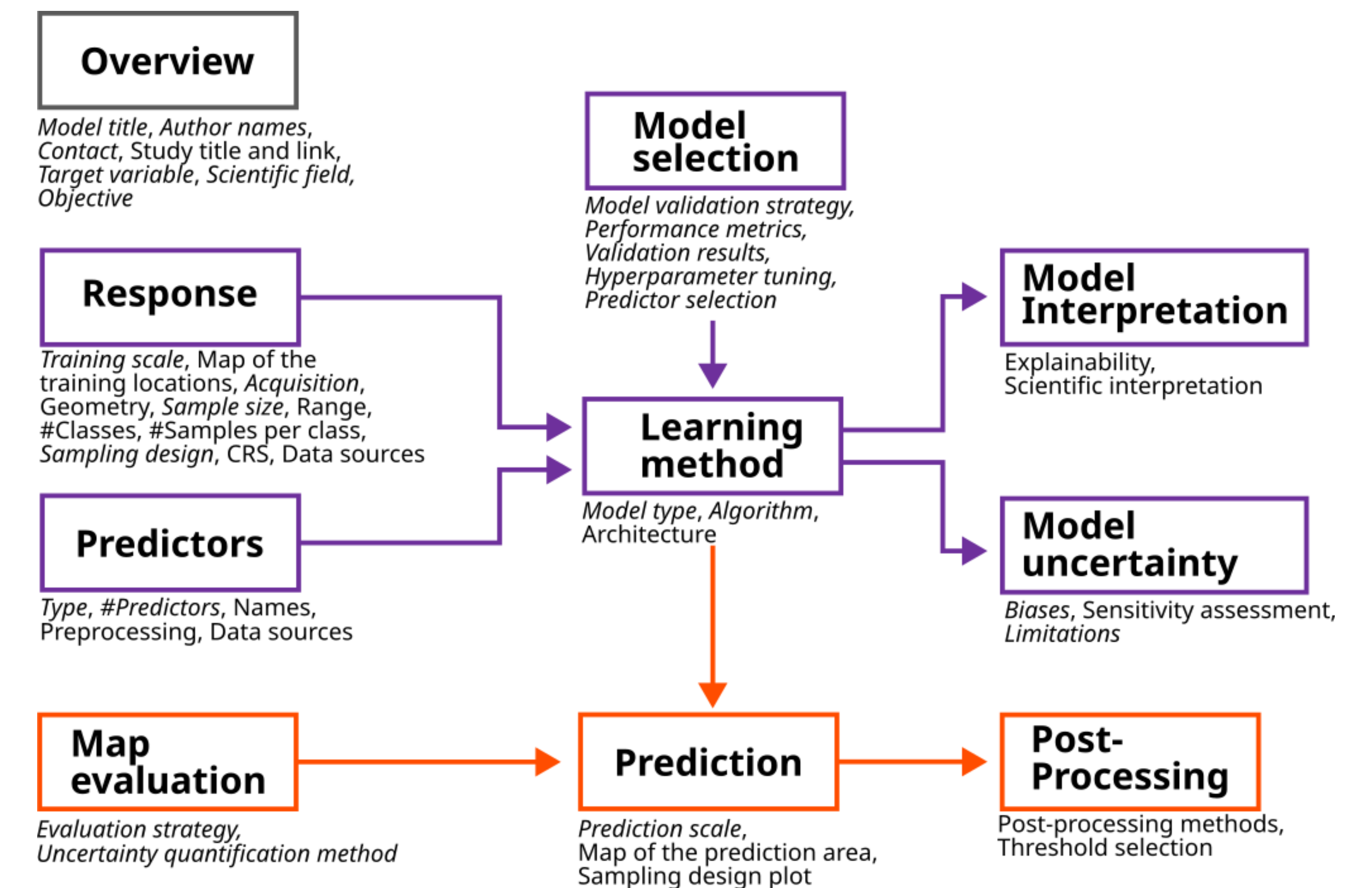
Example application: Modelling plant species richness of South America

- Which data have been used for model training?
- What is the source and resolution of the predictors?
- How was the model tuned?
- How was the performance measured?
- Where do the reported performance statistics hold?
- What is the expected uncertainty in a specific area?
- Can the model be applied beyond South America?
- Is it still valid when the climate changes?
- Under which assumptions does the model hold?
- ...



Aim of STeMP: To guide and standardize model development and reporting in spatio-temporal ML applications.

The elements of STeMP: The protocol is organized in three sections: Overview, Model and Prediction. The Model and Prediction sections are further divided into modules, which are aligned with the standard modelling workflow.



Implementation as a semi-automated WebApp: Users can upload models and geospatial objects from which key protocol fields are automatically filled and diagnostic figures generated. Warnings are raised when potential pitfalls arise.

Community contribution via GitHub: To enable continuous and open contribution by the community, the WebApp is hosted on Github:

