ODMAP Protocol

Obligatory
Objective: Explanation/Inference
Objective: Mapping/Interpolation
Objective: Transfer/Forecast
Optional

ODMAP section	ODMAP subsection	ODMAP element
Overview	Model objective / Model purpose	SDM objective:
		o ecological inference/explanation
		o mapping
		o forecast/transfer
		Main target output: e.g., suitable vs. unsuitable habitat, continuous habitat suitability index, abundance
	Taxon & ecological scale	Taxon names: e.g., names of subspecies, species, genus, families
		Ecological level: e.g., operational taxonomic units, individuals, populations, species, communities
	Location	Location of study area
	Biodiversity data overview	Specify data source: e.g., own field data or data from external provider
		Specify observation type: e.g., standardised monitoring data, expert knowledge, citizen science, heterogenous types
		Specify data type: e.g., presence-only, presence/absence, counts, GPS locations (from individual tracking data)
		Specify spatial sampling design, if applicable: e.g., random, uniform, environmentally stratified, opportunistic
		Time period of data collection
		State (range of) sample size (incl. prevalence)
	Spatial and temporal scale	Spatial resolution and extent, type of extent boundary (e.g., natural or political)
		Temporal resolution and extent
	Conceptual model	Hypotheses about species-environment relationships
		Response variable: e.g. presence/absence, abundance, species richness
		Justification of considered predictor variables and their scales
	Assumptions	State critical model assumptions (cf. Table 2)
	SDM algorithms	State modelling and ensemble techniques used (justified vis. objectives and assumptions)
		Model complexity
	Model workflow	Conceptual description of modelling steps including model fitting, assessment and prediction
	Software	Specify modelling platform incl. version, key packages used, availability of source codes and data
Data	Biodiversity data	Details on external biodiversity data source: e.g., URL/DOI, accession date, database version
		Details on taxonomic reference system
		Details on observation type, if applicable: e.g., standardised monitoring data, expert knowledge, citizen science, heterogenous types
		Details on spatial and temporal sampling design, temporal replications, nestedness
		Details on sample size per taxon: e.g., number of observations/counts, prevalence
		Details on absence data collection, if applicable
		Details on background data derivation, if applicable: e.g., spatial and temporal extent, spatial and temporal buffer, bias correction (e.g. target group sampling)
		Details on potential errors and biases in data, if applicable: e.g., detection probability, misidentification potential, geo-referencing errors, sampling bias
		Details on data cleaning/filtering steps, if applicable: e.g., taxonomically, spatially, temporally, outlier presence/treatment
		Details on scaling, if applicable: e.g., rasterisation of polygon maps, spatial and temporal thinning, measures to address spatial uncertainties
	Data partitioning	Selection of training data (for model fitting)
		Selection of validation data (withheld from model fitting, used for estimating prediction error for model selection, model averaging or ensemble): e.g., cross-validation method
		Selection of test (truly independent) data, *sensu* Hastie, et al. (2009)
	Environmental data/predictor variables	Details on data sources: e.g., URL/DOI, accession date, database version
		Details on measurements errors and bias, when known
		Spatial and temporal resolution and extent
		Details on data processing and on spatial, temporal and thematic scaling: e.g. upscaling/downscaling, transformations, normalisations, thematic aggregations (e.g. of land cover classes), measures to address spatial uncertainties
		Details on dimension reduction of variable set, if applicable – if model-based, this should be contained in **M**odel section (element: Details on pre-selection of variables)
	Transfer data for projection	Details on data sources: e.g., URL/DOI, accession date, database version
		Models and scenarios used
		Spatial and temporal resolution and extent
		Details on data processing and scaling (see above)
		Quantification of novel environmental conditions and novel environmental combinations: e.g., distance to training data
Model	Multicollinearity	Methods for identifying and dealing with multicollinearity (Dormann, et al. 2013) or justification if multicollinearity is not explicitly dealt with
	Variable pre-selection	Details on pre-selection of variables, if applicable
	Model settings / model complexity	Details on model complexity and models settings for all selected algorithms (including default settings of specific platforms/packages)
		Weighting of data
		Details on relevant model settings for extrapolation beyond sample range, if applicable: e.g., clamping
	Model estimates	Assessment of model coefficients
		Details on quantification of uncertainty in model coefficients, e.g. resampling
		Variable importance
	Model selection / Model averaging / Ensembles	Details on model selection strategy: e.g. information-theoretic approach for variable selection, shrinkage and regularization
		Details on model averaging: e.g. derivation of weights
		Details on ensemble method: e.g. initial conditions (input data)
	Non-independence correction/analyses	Method for addressing spatial autocorrelation in residuals
		Method for addressing temporal autocorrelation in residuals
		Method to account for nested data: e.g., fixed and random effects
	Threshold selection	Details on threshold selection, if applicable: transforming continuous predictions into binary predictions
Assessment	Performance statistics	Performance statistics estimated on training data
		Performance statistics estimated on validation data (from data partitioning)
		Performance statistics estimated on test (truly independent) data, if applicable
	Response shapes	Plausibility check: e.g., partial response plots, evaluation strips, inflated response plots
Prediction	Prediction output	Prediction unit
		Post-processing, e.g. clipping, reprojection
	Uncertainty quantification	Algorithmic uncertainity, if applicable
		Uncertainty in input data, if applicable
		Error propagation in Hierarchical/Bayesian models, if applicable
		Uncertainty in scenarios (e.g. climate models, land use models, storylines)
		Visualisation/treatment of novel environments: e.g., masking
	Map display	Plausibility check