**Table 1.** Multi-assessment framework providing a list of useful metrics to assess, interpret or compare jSDMs across different ecological facets (rows) at the species, community or overall level. Underlined metrics are used in this study. The R code used to compute the metrics is provided in GitHub repository.

|  | Model outputs | Example of derived-metrics for model interpretation | Example of derived-metrics for model evaluation | | Example of performance measures to assess the explanatory/predictive power of models**[[1]](#footnote-1)** | |
| --- | --- | --- | --- | --- | --- | --- |
|  |  |  | Presence/Absence | Abundance | Presence/Absence | Abundance |
| Species level | Abundance, occurrence probability, environmental coefficients | Variable importance (e.g. LIME, SHAP[[2]](#footnote-2)), Response curve[[3]](#footnote-3), Heatmap of environmental coefficients, Variance partitioning | Number of Presence/Absence, Proportion of occupied sites | Total abundance, site-specific abundance | AUC, Kappa, F1-Score | RMSE, MAE, R2, Correlation between predicted and observed values |
| Alpha diversity | Site-specific community composition | Diversity index (e.g. Shannon entropy, Simpson-Gini index), Total abundance, Total richness, Proportion of rare/abundant species | | | Differences between predicted and observed values, RMSE, MAE, R2, Correlations (e.g. Kendall, Pearson) between observed and predicted alpha or beta diversity indices | |
| Beta diversity | Pairwise dissimilarity (e.g. Jaccard/Bray-Curtis)[[4]](#footnote-4),[[5]](#footnote-5), Total Beta diversity, Turnover, Nestedness, Local Contribution to Beta Diversity (LCBD), Species Contribution to Beta Diversity (SCBD) | | |
| Overall assessment  (all sites) | Regional community composition | Diversity index (e.g. Shannon entropy, Simpson-Gini index), Total abundance, Total richness, Proportion of rare/abundant species | | | Average over all species: AUC, Kappa, F1-Score | Average over all species: RMSE, MAE, R2, Correlation between predicted and observed values |
| Residual correlation matrix | Co-occurrence network analysis (e.g centrality, number of degrees) | Comparison with observed or reconstructed networks (expert-based or estimated e.g. based on trophic analyses)[[6]](#footnote-6), using e.g. correlations. residual correlation index (δ), | | | |
| Trait-based regressioncoefficients | Traits-environment response curves, Heatmap of traits-environment coefficients | Qualitatively, based on literature and/or expert knowledge[[7]](#footnote-7) | | | |

1. All performance measures can theoretically be compared between models. For instance, we here measured differences between models using a measure of relative change in RMSE or AUC relative to the Bench model using Eq. 1. Other measures could be correlations between model predictions. [↑](#footnote-ref-1)
2. See Ryo *et al.* ([2021](#ref-Ryo_2021)). [↑](#footnote-ref-2)
3. To ease model comparison and interpretation, we propose to summarize the information contained in species response curves using the framework initially proposed by Rigal *et al.* (2020) for classifying species temporal trajectories based on their trend, acceleration, direction and velocity. Applied to regression coefficients, it allows to classify the response of species to each environmental variable into several shapes that are easy to interpret, to link with ecological theory, and to compare across models. [↑](#footnote-ref-3)
4. For jSDM assessment, pairwise dissimilarities can be computed on the observed site-by-species matrix and on the predicted one. Comparing these values (e.g. through correlation analysis or simply through differences) will inform on how well the model reproduces/predict beta diversity patterns. Alternatively, pairwise dissimilarities can be computed between the observed taxa composition of a sample and its predicted one. These dissimilarities then become a metric to assess model performance based on species-composition predictions. [↑](#footnote-ref-4)
5. For jSDMs comparisons, pairwise dissimilarities computed between the observed taxa composition of a sample and its predicted one can be compared across models (e.g. through correlations) to assess to what extent differences between predicted and observed taxa composition are congruent across different models. Alternatively, comparing correlations between pairwise dissimilarities computed on the observed site-by-species matrix and on the predicted one will inform on which model best predict beta diversity patterns. [↑](#footnote-ref-5)
6. Species interaction networks can be reconstructed under certain conditions using the residual correlation matrices estimated by jSDM (see Momal *et al.* ([2020](#ref-Momal_2020))). The comparison between these reconstructed interaction networks and already known interaction networks (based on trophic data, experimental data, expert knowledge or qualitative information on species interactions) can serve as a means of model validation. [↑](#footnote-ref-6)
7. Comparing modelled species trait-environment responses (e.g., signs, shape of response curves) with expected responses (e.g. from theory, experiments or expert knowledge) can also serve to validate qualitatively the models. [↑](#footnote-ref-7)