Table S1.2. Classification of the purpose for which SDMs are used.

Classes	Definition
Explanation	Investigate a species' (causal) relationship with the environment.
Prediction	Map species' potential distributions within the same time period and geographic region as the data used to construct SDMs.
Projection	Project species distribution predictions into a different time period or location from the data used to construct SDMs. Also called 'transferring'.

Table S1.3. Classification of the conservation applications of SDMs.

Conservation application	
New species records	Spatially identifies areas in which the species is not currently recorded but might be found, and which should be surveyed.
Global change	Predicts spatial locations of areas that will change in suitability due to climate or land-use change.
Spatial prioritisation	Spatially identifies areas in which conservation would be valuable, or calculates a metric of conservation value of specific areas.
Habitat evaluation	Quantifies ability of landscape to support existing populations, or population decline in the landscape. Includes population viability analyses. Does not make spatially explicit recommendations about habitat management.
Biological invasions	Spatially identifies areas that could be threatened by, facilitate, or prevent biological invasions or the transmission of disease.
Translocation	Spatially identifies areas suitable for translocated populations, or calculates suitability of specific areas.
Restoration	Spatially identifies areas that would be appropriate for restoration, or measures the efficacy of restoring specific areas.

Table S2.1. Guidelines—Response variable.

Issue	Explanation	Prediction	Projection	Standard
1A. Sampling of response variables	Same as prediction or projection, depending on whether desired explanation is local or global respectively.	Sampled via systematically designed surveys demonstrated to encompass the major environmental gradients occupied by the taxon, and spatial extent of the taxon's occurrence, within the study area. Includes estimates of population demographic parameters (to identify self-sustaining populations), and taxon detectability. Information available on intensity of sampling at each site, and used to ensure sampling is unbiased.	Sampled via systematically designed surveys demonstrated to encompass the major environmental gradients occupied by the taxon, and spatial extent of the taxon's occurrence. Includes estimates of population demographic parameters (to identify self-sustaining populations) and taxon detectability. Information available on intensity of sampling at each site, and used to ensure sampling is unbiased.	Gold
	Same as prediction or projection, depending on whether desired explanation is local or global respectively.	Sampled via systematically designed surveys that encompass the major environmental gradients occupied by the taxon, and spatial extent of the taxon's occurrence, within the study area. Information on intensity of sampling at each site used to conduct post-hoc resampling/weighting to reduce bias; see box 3B.	Sampled via systematically designed surveys that encompass the major environmental gradients occupied by the taxon, and spatial extent of the taxon's occurrence. Information on intensity of sampling at each site used to conduct post-hoc resampling/weighting to reduce bias; see box 3B.	Silver
	Sampled via non-systematically designed surveys, with information on intensity of sampling at each site used to conduct post-hoc resampling/weighting to reduce bias; see box 3B). OR Sampled via non-systematically designed surveys, without information on intensity of sampling at each site, but post-hoc processing undertaken to reduce bias and yield geographically and environmentally representative samples; see box 3B).			Bronze
	Sampled via non-systematically designed surveys. No post-hoc resampling / weighting / processing to reduce bias or yield geographically and environmentally representative samples.			Deficient
1B. Identification of taxa (if species	ID provided by experts, based on multiple lines of evidence, which can be examined.			Gold
	ID provided by experts, based on a single line of evidence, which can be examined.			Silver
	ID based on heterogeneous sources. Records used without being checked by taxonomic experts but after being critically "cleaned" to			Bronze

Issue	Explanation	Prediction	Projection	Standard
occurrence used as response variable)	remove unreasonable records.			
	ID based on heterogeneous sources. Records used without being checked by taxonomic experts and without being critically “cleaned” by others to remove unreasonable records.			Deficient

Issue	Explanation	Prediction	Projection	Standard
1C. Spatial accuracy of response variable	Spatial accuracy of all records sufficiently high relative to spatial resolution of predictor variables such that all points are known to fall within the location to which they are assigned. The spatial accuracy matches the spatial resolution of predictor variables as defined in section 2B.			Gold
	Spatial accuracy of all records known and variable across records, so that some points might fall outside the location to which they are assigned. These potential locational errors integrated into formal uncertainty analysis (see 3D), and/or steps taken (and documented) to remove records with locational errors.			Silver
	Spatial accuracy not known or inconsistently quantified, but steps taken (and documented) to remove unreasonable records.			Bronze
	Spatial accuracy not known or inconsistently quantified, and no steps taken to remove unreasonable records.			Deficient
1D. Environmental extent across which response variable is sampled	Same as prediction or projection, depending on whether desired explanation is local or global respectively.	Multiple lines of evidence (in addition to occurrence data used to train SDMs) demonstrate that data cover the range of the taxon's environmental tolerances within the study area.	Multiple lines of evidence (in addition to occurrence data used to train SDMs) demonstrate that data cover the entire range of the taxon's environmental tolerances and that no evolutionary changes in species environmental tolerances have occurred in the projection space.	Gold
	Same as prediction or projection, depending on whether desired explanation is local or global respectively.	A single line of evidence (in addition to occurrence data used to train SDMs) demonstrates that data cover the entire environmental range of the study region.	A single line of evidence (in addition to occurrence data used to train SDMs) demonstrates that data cover the entire environmental extent of the known distribution of the taxon.	Silver
	Models fitted with the best available data on the known geographical extent of the taxon, but without evidence that the taxon's environmental tolerances are covered. Steps taken and documented to avoid impact of incomplete distribution data on results.	Steps are taken to avoid or flag extrapolation to conditions outside the environmental extent used to train the models.	Models fitted with the best available data on the known geographical extent of the taxon, but without evidence that the taxon's environmental tolerances are covered. Steps taken to avoid or flag extrapolation to conditions outside the extent of each predictor variable used to train the models.	Bronze
	No evidence provided that data cover the entire environmental range of the taxon.	No evidence provided that data cover the entire environmental range of the study region.	No evidence provided that data cover the entire environmental range of the study region. No steps taken to avoid or flag extrapolation to conditions outside the extent of each predictor used to train the models.	Deficient

Issue	Explanation	Prediction	Projection	Standard
1E. Geographic extent across which response variable is sampled (includes occurrence data and absence, pseudo-absence, or background data	Same as prediction or projection, depending on whether desired explanation is local or global respectively.	Samples restricted to and inclusive of all regions of the study area that are suitable for the taxa to establish populations, and that are accessible to the taxon (as demonstrated by multiple lines of evidence).	Samples restricted to and inclusive of all regions that are suitable for the taxa to establish populations and that are accessible to the taxon (as demonstrated by multiple lines of evidence).	Gold
	Same as prediction or projection, depending on whether desired explanation is local or global respectively.	Samples restricted to and inclusive of all regions of the study area that contain the full historical and current range of the focal taxon.	Samples restricted to and inclusive of all regions that contain the full historical and current range of the focal taxon (as demonstrated by a single line of evidence).	Silver
	Same as prediction or projection, depending on whether desired explanation is local or global respectively.	Samples derived only from regions within the study area reasonably justified to contain the full current range of the focal taxon.	Samples derived only from regions reasonably justified to contain the full current range of the focal taxon.	Bronze
	No justification of regions from which samples drawn, or samples derive from regions outside those reasonably deemed accessible to the taxon.			Deficient

Table S2.2. Guidelines—Predictor variables.

Issue	Explanation	Prediction	Projection	Standard
2A. Selection of candidate variables	Candidates include all proximal variables that multiple lines of evidence (in addition to occurrence data used to train SDMs) that can be shown to have a measurable effect on the taxon's distribution at the spatial scale examined. This must include, whenever relevant, a full range of environmental and biotic variables.			Gold
	Same as prediction or projection, depending on whether desired explanation is local or global respectively.	Candidates include proximal and/or distal variables that a single line of evidence (in addition to occurrence data used to train SDMs) shows to have a measurable association with the taxon at the spatial scale examined. This should include, whenever relevant, a range of environmental and biotic variables.	Candidates include proximal variables that a single line of evidence (in addition to occurrence data used to train SDMs) shows to have a measurable effect on the taxon's distribution at the spatial scale examined. This should include, whenever relevant, a range of environmental and biotic variables.	Silver
	Same as prediction or projection, depending on whether desired explanation is local or global respectively.	Candidates include proximal or distal variables theoretically justified as having an association with the taxon's distribution at the spatial scale examined.	Candidates include observationally, statistically or theoretically justified proximal and/or distal variables that have a measurable association with the taxon's distribution at the spatial scale examined. This should include, whenever possible, a range of environmental and biotic variables.	Bronze
	Same as prediction or projection, depending on whether desired explanation is local or global respectively.	No ecological justification of variable choice.	No ecological justification of variable choice and/or distal variables used without strong justification.	Deficient
2B. Spatial and temporal resolution of predictor variables	Variables directly measured at the temporal and spatial resolution at which multiple lines of evidence (in addition to occurrence data used to train SDMs) demonstrate that the taxon responds.			Gold
	Variables interpolated at the temporal and spatial resolution at which at least one line of evidence (in addition to occurrence data used to train SDMs) demonstrate the taxon responds.			Silver
	Variables interpolated at a resolution theoretically justified for the taxon.			Bronze
	Variables interpolated at a spatial and temporal resolution to which the taxon does not respond and/or without theoretical justification of resolution.			Deficient

Issue	Explanation	Prediction	Projection	Standard
2C. Uncertainty in predictor variables (both under current and projected conditions)	All sources of uncertainty in the predictors and their effects on model results quantified, mapped, and interpreted.			Gold
	Some of the perceived most important sources of uncertainty in the predictors (e.g. errors in geo-registration, measurement, interpolation) quantified and mapped.			Silver
	Possible sources of uncertainty in the predictors (e.g. errors in geo-registration, measurement, interpolation) and the effects these could have on the model acknowledged, and consequences for interpretation of the results discussed.			Bronze
	No consideration of uncertainty.			Deficient

Table S2.3. Guidelines—Model building.

Issue	Explanation	Prediction	Projection	Standard
3A. Model complexity	Same as prediction or projection, depending on whether desired explanation is local or global respectively.	Same as projection without the necessity of using independent data.	The optimal level of complexity is decided by constructing models using an appropriate method to deal with model complexity, performing comparison with multiple lines of independent data (see table 4B).	Gold
	Same as prediction or projection, depending on whether desired explanation is local or global respectively.	Same as projection without the necessity of using independent data.	The optimal level of complexity is decided by constructing models using an appropriate method to deal with model complexity, performing cross-validation or comparison with a single line of independent data.	Silver
	Broadly agreed rules of thumb are followed and/or the optimal level of complexity is decided using justified methods without independent data.			Bronze
	Complexity is not considered, or inappropriate methods are used to deal with it.			Deficient
3B. Treatment of bias and noise in response variables	Demonstrated that there are no geographical and environmental biases in response data. OR Model fully corrected for bias in response data, tested by performing comparison with independent data.			Gold
	Model corrected for major biases in response data, tested by performing internal cross-validation.			Silver
	Bias, and the effects these could have on the model and results, acknowledged and described.			Bronze
	No consideration of biases.			Deficient

Issue	Explanation	Prediction	Projection	Standard
3C. Treatment of collinearity	Demonstrated that there is no collinearity in data. OR Model construction is informed by a full mechanistic understanding of interactions among predictor variables so that the model is insensitive to collinearity.			Gold
	Fitting techniques used are known to be insensitive to collinearity. OR Demonstrated that the results are robust to changes in collinearity between predictor variables, including non-analogue combinations of predictor variables.			Silver
	Approximate methods are applied to deal with collinearity. OR Collinearity is acknowledged and described, as are the effects the known collinearity could have on the results.			Bronze
	Models use collinear variables and a fitting technique sensitive to collinearity without acknowledging the effects on the results,			Deficient
3D. Dealing with modelling and parameter uncertainty	Uncertainty arising from different modelling techniques, response data, and predictor variables is comprehensively characterized. Results are obtained from several SDM techniques that are representative of all appropriate current distribution modelling techniques in order to characterize model uncertainty (sometimes called ensemble modelling). Uncertainty is fully propagated through the modelling process in order to quantify, map, and interpret uncertainty in results. Biases arising from similarities among structures of model classes, and the effects these could have on results are discussed.			Gold
	Major suspected model and data uncertainties are characterized and: 1) known uncertainties are propagated through the model; or 2) the range of predictions built using different scenarios (including parameter and model technique) are quantified and mapped, and sensitivity analysis is conducted. Results are obtained from several SDM techniques that are representative of all appropriate current distribution modelling techniques in order to characterize model uncertainty. Biases arising from similarities among structures of model classes are quantified, accounted for, and discussed.			Silver
	Results are obtained from multiple SDM techniques but that are not representative of all appropriate current distribution modelling techniques. The effect of major suspected model uncertainties on the projections is quantified. Major suspected sources of data uncertainties are acknowledged, and their consequences for interpretation of the results are discussed.			Bronze
	Uncertainty is not dealt with (i.e., a single SDM technique with one set of parameters is used).			Deficient

Table S2.4. Guidelines—Model evaluation.

Issue	Explanation	Prediction	Projection	Standard
4A. Evaluation of model assumptions	Demonstrated lack of violation of, or robustness to, assumptions relevant for technique being used.			Gold
	Theoretically justified lack of violation of, or expected robustness to, assumptions of technique being used.			Silver
	Violation of major assumptions of technique being used characterized, and their consequences for interpretation of results discussed.			Bronze
	No check for violation of statistical assumptions.			Deficient
4B. Evaluation of model outputs	Same as prediction or projection, depending on whether desired explanation is local or global respectively.	Evaluated against multiple datasets that are statistically independent from the data used to train the models, but not necessarily from an independent location or time period.	Evaluated against multiple and diverse independent evaluation datasets, and/or corroboration with experimental testing.	Gold
	Same as prediction or projection, depending on whether desired explanation is local or global respectively.	Evaluated against data obtained by geographically structured sub-sampling of the training data.	Evaluated against at least one independent evaluation dataset.	Silver
	Same as prediction or projection, depending on whether desired explanation is local or global respectively.	Evaluated with non-independent data. Re-substitution used to estimate over-fitting. N – 1 Jackknife acceptable for very small sample sizes.	Evaluated with non-independent data obtained by sub-sampling the training data, with repetition. N – 1 Jackknife acceptable for very small sample sizes.	Bronze
	No evaluation at all or re-substitution alone.			Deficient

Issue	Explanation	Prediction	Projection	Standard
4C. Measures of model performance	Same as prediction.	Same as projection, but no characterization of temporal errors.	Measures of performance exhaustively summarize goodness of fit and discrimination. AND Patterns of uncertainty comprehensively characterized (spatial, temporal and environmental; see also 3E).	Gold
	Measures of performance summarize goodness of fit and calibration AND Major patterns of uncertainty comprehensively characterized (spatial, temporal and/or environmental; see also 3E)			Silver
	One or more major aspects of model performance measured and summarized.			Bronze
	No, cursory, or inappropriate measures of model performance.			Deficient