



Species Distribution Models: Models and Application for Sustainable Species Conservation

Koissi Savi (Ph.D.)

Founder and CEO of HealthDataInsights

April 25, 2024



- 1 Species Distribution Models (SDMs)
- 2 Types of SDMs and some Algorithms
- 3 Correlative SDMs
- 4 Mechanistic SDMs

Concepts

- **Species Distribution models also known as Ecological Niche Modeling** is a tool to study species geographic distributions or favorable habitats of a given species. This type of model relates presence-absence or presence-only occurrence data to explanatory landscape factors, producing estimates of suitable habitat



Concepts

- **Species Distribution models also known as Ecological Niche Modeling** is a tool to study species geographic distributions or favorable habitats of a given species. This type of model relates presence-absence or presence-only occurrence data to explanatory landscape factors, producing estimates of suitable habitat
- SDMs are widely used modeling frameworks in global change impact assessments for projecting potential future range shifts of species (IPBES 2016).



Concepts

- **Species Distribution models also known as Ecological Niche Modeling** is a tool to study species geographic distributions or favorable habitats of a given species. This type of model relates presence-absence or presence-only occurrence data to explanatory landscape factors, producing estimates of suitable habitat
- SDMs are widely used modeling frameworks in global change impact assessments for projecting potential future range shifts of species (IPBES 2016).
- The ease with which SDMs can gather model input data contributes to their widespread adoption. For example, online repositories include data on species distributions (e.g., GBIF and OBIS), individual animal locations (e.g., Movebank), climate (e.g., WorldClim and CHELSA), land cover, and other remote sensing products (e.g., Copernicus).



Concepts

- **Species Distribution models also known as Ecological Niche Modeling** is a tool to study species geographic distributions or favorable habitats of a given species. This type of model relates presence-absence or presence-only occurrence data to explanatory landscape factors, producing estimates of suitable habitat
- SDMs are widely used modeling frameworks in global change impact assessments for projecting potential future range shifts of species (IPBES 2016).
- The ease with which SDMs can gather model input data contributes to their widespread adoption. For example, online repositories include data on species distributions (e.g., GBIF and OBIS), individual animal locations (e.g., Movebank), climate (e.g., WorldClim and CHELSA), land cover, and other remote sensing products (e.g., Copernicus).
- SDMs are applied to determine which predictors limit a species' distribution, often referred to as variable importance (Bardie and Leung, 2017)



Concepts

- **Species Distribution models also known as Ecological Niche Modeling** is a tool to study species geographic distributions or favorable habitats of a given species. This type of model relates presence-absence or presence-only occurrence data to explanatory landscape factors, producing estimates of suitable habitat
- SDMs are widely used modeling frameworks in global change impact assessments for projecting potential future range shifts of species (IPBES 2016).
- The ease with which SDMs can gather model input data contributes to their widespread adoption. For example, online repositories include data on species distributions (e.g., GBIF and OBIS), individual animal locations (e.g., Movebank), climate (e.g., WorldClim and CHELSA), land cover, and other remote sensing products (e.g., Copernicus).
- SDMs are applied to determine which predictors limit a species' distribution, often referred to as variable importance (Bardie and Leung, 2017)



Concept of Niches

- Hutchinsons (1957) definition of the fundamental niche of a species: a **hypervolume** of environmental variables, "every point of which corresponds to a state of the environment which would permit the species to exist indefinitely."



Concept of Niches

- Hutchinsons (1957) definition of the fundamental niche of a species: a **hypervolume** of environmental variables, "every point of which corresponds to a state of the environment which would permit the species to exist indefinitely."
- From the standpoint of population ecology A species may exist for a period of time $t_0 < t < t_1$ in a place g of the world if its total instantaneous growth rate $\frac{dx_g}{dt} > 0$ is on the average non-negative during $t_0 < t < t_1$. It is clear to see that this definition disregards stochastic factors and evolutionary change.



Concept of Niches

- Hutchinsons (1957) definition of the fundamental niche of a species: a **hypervolume** of environmental variables, "every point of which corresponds to a state of the environment which would permit the species to exist indefinitely."
- From the standpoint of population ecology A species may exist for a period of time $t_0 < t < t_1$ in a place g of the world if its total instantaneous growth rate $\frac{dx_g}{dt} > 0$ is on the average non-negative during $t_0 < t < t_1$. It is clear to see that this definition disregards stochastic factors and evolutionary change.
- instantaneous growth rate =
intrinsic growth rate + density-dependent factors and the results of interactions with other species + population structure that determines arrival of dispersing

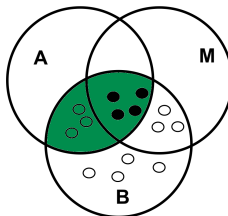


Concept of Niches

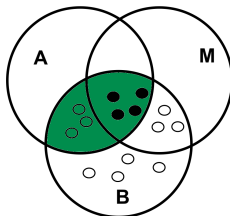
- Hutchinsons (1957) definition of the fundamental niche of a species: a **hypervolume** of environmental variables, "every point of which corresponds to a state of the environment which would permit the species to exist indefinitely."
- From the standpoint of population ecology A species may exist for a period of time $t_0 < t < t_1$ in a place g of the world if its total instantaneous growth rate $\frac{dx_g}{dt} > 0$ is on the average non-negative during $t_0 < t < t_1$. It is clear to see that this definition disregards stochastic factors and evolutionary change.
- instantaneous growth rate =
intrinsic growth rate + density-dependent factors and the results of interactions with other species + population structure that determines arrival of dispersing



Niche Theory explained with a diagram

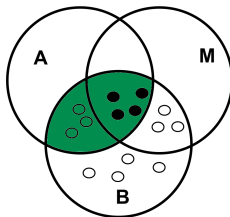


Niche Theory explained with a diagram



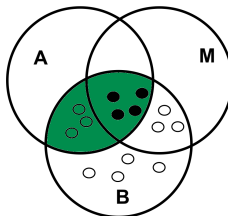
- The BAM diagram represents the factors that limit species distributions (Soberon 2007).

Niche Theory explained with a diagram



- The BAM diagram represents the factors that limit species distributions (Soberon 2007).
- A species can only survive in geographic areas where both the abiotic environmental conditions (A) and the prevailing biotic interactions (B) allow positive population growth.

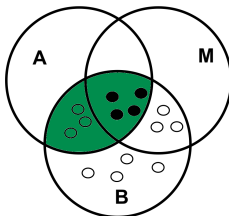
Niche Theory explained with a diagram



- The BAM diagram represents the factors that limit species distributions (Soberon 2007).
- A species can only survive in geographic areas where both the abiotic environmental conditions (A) and the prevailing biotic interactions (B) allow positive population growth.
- The intersection of A and B thus represents the potential distribution of the species, or the realized niche (the green ellipse).

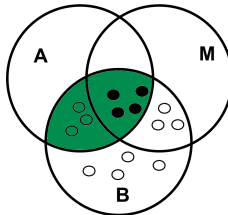


Niche Theory explained with a diagram



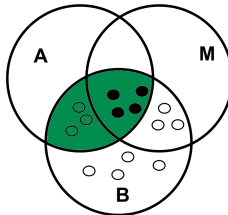
- The BAM diagram represents the factors that limit species distributions (Soberon 2007).
- A species can only survive in geographic areas where both the abiotic environmental conditions (A) and the prevailing biotic interactions (B) allow positive population growth.
- The intersection of A and B thus represents the potential distribution of the species, or the realized niche (the green ellipse).
- The movement capacity (M) of a species will determine which geographic area is accessible within the time period of interest.

Niche Theory explained with a diagram



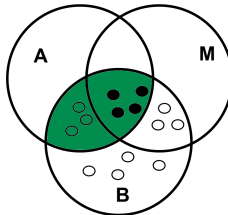
- The BAM diagram represents the factors that limit species distributions (Soberon 2007).
- A species can only survive in geographic areas where both the abiotic environmental conditions (A) and the prevailing biotic interactions (B) allow positive population growth.
- The intersection of A and B thus represents the potential distribution of the species, or the realized niche (the green ellipse).
- The movement capacity (M) of a species will determine which geographic area is accessible within the time period of interest.
- The intersection of A, B, and M represents the geographic area that is actually occupied by the species and where we can find source populations (filled circles).

Niche Theory explained with a diagram



- The BAM diagram represents the factors that limit species distributions (Soberon 2007).
- A species can only survive in geographic areas where both the abiotic environmental conditions (A) and the prevailing biotic interactions (B) allow positive population growth.
- The intersection of A and B thus represents the potential distribution of the species, or the realized niche (the green ellipse).
- The movement capacity (M) of a species will determine which geographic area is accessible within the time period of interest.
- The intersection of A, B, and M represents the geographic area that is actually occupied by the species and where we can find source populations (filled circles).
- Open circles represent sink populations with negative population growth.

Niche Theory explained with a diagram



- The BAM diagram represents the factors that limit species distributions (Soberon 2007).
- A species can only survive in geographic areas where both the abiotic environmental conditions (A) and the prevailing biotic interactions (B) allow positive population growth.
- The intersection of A and B thus represents the potential distribution of the species, or the realized niche (the green ellipse).
- The movement capacity (M) of a species will determine which geographic area is accessible within the time period of interest.
- The intersection of A, B, and M represents the geographic area that is actually occupied by the species and where we can find source populations (filled circles).
- Open circles represent sink populations with negative population growth.



SDM position vis-a-vis Niche Theory

- Although SDMs are clearly positioned within the niche theory, it is an ongoing debate whether the estimated species-environment relationship approximates the fundamental niche (area A in the BAM diagram), the realized niche (green area in BAM diagram), or the occupied niche (intersection of A, B, and M in BAM diagram).



SDM position vis-a-vis Niche Theory

- Although SDMs are clearly positioned within the niche theory, it is an ongoing debate whether the estimated species-environment relationship approximates the fundamental niche (area A in the BAM diagram), the realized niche (green area in BAM diagram), or the occupied niche (intersection of A, B, and M in BAM diagram).
- **Fundamental niche** was the set of all the states of the environment which would permit the species S to exist (Hutchinson 1957)



Type of models

What we term SDMs have also been called (sometimes with different emphases and meanings): bioclimatic models, climate envelopes, ecological niche models (ENMs), habitat models, resource selection functions (RSFs), range maps, and more loosely correlative models or spatial models.



Type of models

What we term SDMs have also been called (sometimes with different emphases and meanings): bioclimatic models, climate envelopes, ecological niche models (ENMs), habitat models, resource selection functions (RSFs), range maps, and more loosely correlative models or spatial models.

- **Models and usefulness**

- 1 **Linear multiple regression and discriminant function analyses** (Capen 1981, Stauffer 2002) = studies of species-habitat associations



Type of models

What we term SDMs have also been called (sometimes with different emphases and meanings): bioclimatic models, climate envelopes, ecological niche models (ENMs), habitat models, resource selection functions (RSFs), range maps, and more loosely correlative models or spatial models.

■ Models and usefulness

- 1 **Linear multiple regression and discriminant function analyses** (Capen 1981, Stauffer 2002) = studies of species-habitat associations
- 2 **Generalized linear models (GLMs)** = enabled pioneering regression-based SDMs that had much more sophistication and realism than was possible earlier (e.g., see Austin's work in 1970s and 1980s, cited in Austin 1985). The key structural features of GLMs (non-normal error distributions, additive terms, nonlinear fitted functions) continue to be useful and are part of many current methods including RSFs (Manly et al. 2002) and maximum entropy models (MaxEnt; Phillips et al. 2006).



Type of models

What we term SDMs have also been called (sometimes with different emphases and meanings): bioclimatic models, climate envelopes, ecological niche models (ENMs), habitat models, resource selection functions (RSFs), range maps, and more loosely correlative models or spatial models.

■ Models and usefulness

- 1 **Linear multiple regression and discriminant function analyses** (Capen 1981, Stauffer 2002) = studies of species-habitat associations
- 2 **Generalized linear models (GLMs)** = enabled pioneering regression-based SDMs that had much more sophistication and realism than was possible earlier (e.g., see Austin's work in 1970s and 1980s, cited in Austin 1985). The key structural features of GLMs (non-normal error distributions, additive terms, nonlinear fitted functions) continue to be useful and are part of many current methods including RSFs (Manly et al. 2002) and maximum entropy models (MaxEnt; Phillips et al. 2006).



Explanation vs Prediction

- **Explanation:** These models are used to seek insight, even if indirectly, into the causal drivers of species distributions (Mac Nally 2000). SDMs are still regularly used for such purposes, particularly in quantitative ecological studies (Leathwick & Austin 2001) and evolutionary biology (Graham et al. 2004).

Explanation vs Prediction

- **Explanation:** These models are used to seek insight, even if indirectly, into the causal drivers of species distributions (Mac Nally 2000). SDMs are still regularly used for such purposes, particularly in quantitative ecological studies (Leathwick & Austin 2001) and evolutionary biology (Graham et al. 2004).
- **Prediction:** This type of model is used in two main ways

Explanation vs Prediction

- **Explanation:** These models are used to seek insight, even if indirectly, into the causal drivers of species distributions (Mac Nally 2000). SDMs are still regularly used for such purposes, particularly in quantitative ecological studies (Leathwick & Austin 2001) and evolutionary biology (Graham et al. 2004).
- **Prediction:** This type of model is used in two main ways
 - 1 **The model-based interpolation to unsampled sites:** predictions are made to new sites within the range of environments sampled by the training data and within the same general time frame as that in which the sampling occurred. Typical applications include global analyses of species distributions, mapping within a region for conservation planning or resource management, and identifying suitable habitats for rare species (Guisan & Thuiller 2005)

Explanation vs Prediction

- **Explanation:** These models are used to seek insight, even if indirectly, into the causal drivers of species distributions (Mac Nally 2000). SDMs are still regularly used for such purposes, particularly in quantitative ecological studies (Leathwick & Austin 2001) and evolutionary biology (Graham et al. 2004).
- **Prediction:** This type of model is used in two main ways
 - 1 **The model-based interpolation to unsampled sites:** predictions are made to new sites within the range of environments sampled by the training data and within the same general time frame as that in which the sampling occurred. Typical applications include global analyses of species distributions, mapping within a region for conservation planning or resource management, and identifying suitable habitats for rare species (Guisan & Thuiller 2005)
 - 2 **The model-based extrapolation to unsampled sites:** The environments in these new times and places need to be carefully assessed, particularly for new combinations of predictor values or for predictor values outside their original ranges in the training data. Prediction of new geographic regions is a special case and has been termed transferability, but often without clear information on the environmental similarities and differences between the model fitting and prediction regions.

Explanation vs Prediction

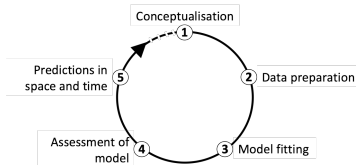
- **Explanation:** These models are used to seek insight, even if indirectly, into the causal drivers of species distributions (Mac Nally 2000). SDMs are still regularly used for such purposes, particularly in quantitative ecological studies (Leathwick & Austin 2001) and evolutionary biology (Graham et al. 2004).
- **Prediction:** This type of model is used in two main ways
 - 1 **The model-based interpolation to unsampled sites:** predictions are made to new sites within the range of environments sampled by the training data and within the same general time frame as that in which the sampling occurred. Typical applications include global analyses of species distributions, mapping within a region for conservation planning or resource management, and identifying suitable habitats for rare species (Guisan & Thuiller 2005)
 - 2 **The model-based extrapolation to unsampled sites:** The environments in these new times and places need to be carefully assessed, particularly for new combinations of predictor values or for predictor values outside their original ranges in the training data. Prediction of new geographic regions is a special case and has been termed transferability, but often without clear information on the environmental similarities and differences between the model fitting and prediction regions.

Scale of SDMs

Scale is relevant to the distributions of both species and environments and comprises both grain and extent.

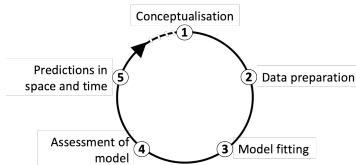
- **The extent (or domain)** usually reflects the purpose of the analysis. For instance, macroecological and global change studies tend to be continental to global in scope (e.g., Araujo & New 2007), whereas studies targeting detailed ecological understanding or conservation planning tend toward local to regional extents (Fleishman et al. 2001, Ferrier et al. 2002).
- **The Grain** usually describes properties of the data or analysis often the predictor variables and their grid cell size or polygon size, but also the spatial accuracy and precision of the species records (Dungan et al. 2002, Tobalske 2002)

Different phases of SDM building



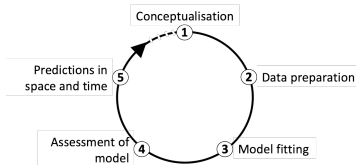
- **Conceptualisation:** main research objectives, select the model, main underlying assumptions of SDMs, e.g. pseudo-equilibrium with the environment; bias in the data (Guisan, Thuiller, and Zimmermann 2017), selection of predictors.

Different phases of SDM building



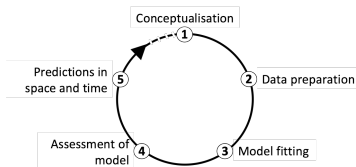
- **Conceptualisation:** main research objectives, select the model, main underlying assumptions of SDMs, e.g. pseudo-equilibrium with the environment; bias in the data (Guisan, Thuiller, and Zimmermann 2017), selection of predictors.
- **Data preparation:** Quality of data and decision about the train and test sample.

Different phases of SDM building



- **Conceptualisation:** main research objectives, select the model, main underlying assumptions of SDMs, e.g. pseudo-equilibrium with the environment; bias in the data (Guisan, Thuiller, and Zimmermann 2017), selection of predictors.
- **Data preparation:** Quality of data and decision about the train and test sample.
- **Model fitting:** Pivotal questions need to be asked during both the conception and this phase and some of them are

Different phases of SDM building



- How to deal with multicollinearity in the environmental data?
- How many variables should be included in the model (without overfitting) and how should we select these?
- How to select the final model or average the models?
- ... Guisan, Thuiller, and Zimmermann (2017)
- **Model assessment:** Strictly, checking the plausibility of the fitted species-environment relationship by visual inspection of response curves, and by assessing model coefficients and variable importance would also be part of the model assessment.
- **Predictions:** Inspected model and extrapolation behavior, and assessed predictive performance, it is finally time to make predictions in space and time.

Potential purposes of SDM from from Zurell et al. 2020

- Araújo et al. (2019) suggested three distinct SDM purposes

1 Explanation: (also termed inference) regards detailed analyses of species-environment relationships and aims to provide or test specific hypotheses about the main factors driving the species distributions.



Potential purposes of SDM from from Zurell et al. 2020

- Araújo et al. (2019) suggested three distinct SDM purposes

- 1 Explanation:** (also termed inference) regards detailed analyses of species-environment relationships and aims to provide or test specific hypotheses about the main factors driving the species distributions.
- 2 Mapping:** (also termed interpolation) means that the estimated species-environment relationships are used to map (or interpolate) the species distributions in the same geographic area and time period in which the model was calibrated.



Potential purposes of SDM from Zurell et al. 2020

- Araújo et al. (2019) suggested three distinct SDM purposes

- 1 Explanation:** (also termed inference) regards detailed analyses of species-environment relationships and aims to provide or test specific hypotheses about the main factors driving the species distributions.
- 2 Mapping:** (also termed interpolation) means that the estimated species-environment relationships are used to map (or interpolate) the species distributions in the same geographic area and time period in which the model was calibrated.
- 3 Transfer:** (also termed forecast or projection, but these terms are less precise) means that the estimated species-environment relationships are transferred to a different geographic region or time period future or past (Yates et al. 2018).



Potential purposes of SDM from Zurell et al. 2020

- Araújo et al. (2019) suggested three distinct SDM purposes

- 1 Explanation:** (also termed inference) regards detailed analyses of species-environment relationships and aims to provide or test specific hypotheses about the main factors driving the species distributions.
- 2 Mapping:** (also termed interpolation) means that the estimated species-environment relationships are used to map (or interpolate) the species distributions in the same geographic area and time period in which the model was calibrated.
- 3 Transfer:** (also termed forecast or projection, but these terms are less precise) means that the estimated species-environment relationships are transferred to a different geographic region or time period future or past (Yates et al. 2018).



Standard of SDM models

- **Gold standard:** Usually requires ideal data (seldom available) and next-generation modeling approaches that remain under development, as well as results obtained through multiple sources of evidence including manipulative experiments. Therefore, this level typically provides targets for excellence and directions for future research.



Best practice standard for SDM from Araujo et al. 2019

Standard of SDM models

- **Gold standard:** Usually requires ideal data (seldom available) and next-generation modeling approaches that remain under development, as well as results obtained through multiple sources of evidence including manipulative experiments. Therefore, this level typically provides targets for excellence and directions for future research.
- **(Silver standard):** Typically involving imperfect (but best available) data combined with analyses that allow uncertainty and bias to be reduced, accounted for, or at least estimated.



Standard of SDM models

- **Gold standard:** Usually requires ideal data (seldom available) and next-generation modeling approaches that remain under development, as well as results obtained through multiple sources of evidence including manipulative experiments. Therefore, this level typically provides targets for excellence and directions for future research.
- **(Silver standard):** Typically involving imperfect (but best available) data combined with analyses that allow uncertainty and bias to be reduced, accounted for, or at least estimated.
- **Bronze standard:** Encompasses data and procedures that represent the minimum currently acceptable practices for models to be included in SDM. It includes approaches to characterize and address limitations of data and models and to interpret their implications on the results.

Standard of SDM models

- **Gold standard:** Usually requires ideal data (seldom available) and next-generation modeling approaches that remain under development, as well as results obtained through multiple sources of evidence including manipulative experiments. Therefore, this level typically provides targets for excellence and directions for future research.
- **(Silver standard):** Typically involving imperfect (but best available) data combined with analyses that allow uncertainty and bias to be reduced, accounted for, or at least estimated.
- **Bronze standard:** Encompasses data and procedures that represent the minimum currently acceptable practices for models to be included in SDM. It includes approaches to characterize and address limitations of data and models and to interpret their implications on the results.
- **Deficient:** Involves the use of data and/or modeling practices that are considered unacceptable for models used in driving policy and practice.

Standard of SDM models

- **Gold standard:** Usually requires ideal data (seldom available) and next-generation modeling approaches that remain under development, as well as results obtained through multiple sources of evidence including manipulative experiments. Therefore, this level typically provides targets for excellence and directions for future research.
- **(Silver standard):** Typically involving imperfect (but best available) data combined with analyses that allow uncertainty and bias to be reduced, accounted for, or at least estimated.
- **Bronze standard:** Encompasses data and procedures that represent the minimum currently acceptable practices for models to be included in SDM. It includes approaches to characterize and address limitations of data and models and to interpret their implications on the results.
- **Deficient:** Involves the use of data and/or modeling practices that are considered unacceptable for models used in driving policy and practice.



Best practice standard for SDM from Araujo et al. 2019

Key Criteria affecting the quality of the SDM

- Quality of the response variable (usually species occurrence data)



Best practice standard for SDM from Araujo et al. 2019

Key Criteria affecting the quality of the SDM

- Quality of the response variable (usually species occurrence data)
- Quality of the predictor variables (usually environmental data)



Key Criteria affecting the quality of the SDM

- Quality of the response variable (usually species occurrence data)
- Quality of the predictor variables (usually environmental data)
- "Model building i.e., fitting a statistical relationship between species occurrence data and environmental data.



Key Criteria affecting the quality of the SDM

- Quality of the response variable (usually species occurrence data)
- Quality of the predictor variables (usually environmental data)
- "Model building i.e., fitting a statistical relationship between species occurrence data and environmental data.
- "Model evaluation aligns a balanced blend of realism, accuracy, and generality which can be assessed through these questions:



Best practice standard for SDM from Araujo et al. 2019

Key Criteria affecting the quality of the SDM

- Quality of the response variable (usually species occurrence data)
- Quality of the predictor variables (usually environmental data)
- "Model building i.e., fitting a statistical relationship between species occurrence data and environmental data.
- "Model evaluation aligns a balanced blend of realism, accuracy, and generality which can be assessed through these questions:
 - 1 How robust is the model to depart from the assumptions?



Key Criteria affecting the quality of the SDM

- Quality of the response variable (usually species occurrence data)
- Quality of the predictor variables (usually environmental data)
- "Model building i.e., fitting a statistical relationship between species occurrence data and environmental data.
- "Model evaluation aligns a balanced blend of realism, accuracy, and generality which can be assessed through these questions:
 - 1 How robust is the model to depart from the assumptions?
 - 2 How meaningful are the evaluation metrics used?



Key Criteria affecting the quality of the SDM

- Quality of the response variable (usually species occurrence data)
- Quality of the predictor variables (usually environmental data)
- "Model building i.e., fitting a statistical relationship between species occurrence data and environmental data.
- "Model evaluation aligns a balanced blend of realism, accuracy, and generality which can be assessed through these questions:
 - 1 How robust is the model to depart from the assumptions?
 - 2 How meaningful are the evaluation metrics used?
 - 3 How predictive is the model when tested against independent data?



Key Criteria affecting the quality of the SDM

- Quality of the response variable (usually species occurrence data)
- Quality of the predictor variables (usually environmental data)
- "Model building i.e., fitting a statistical relationship between species occurrence data and environmental data.
- "Model evaluation aligns a balanced blend of realism, accuracy, and generality which can be assessed through these questions:
 - 1 How robust is the model to depart from the assumptions?
 - 2 How meaningful are the evaluation metrics used?
 - 3 How predictive is the model when tested against independent data?



Type of SDMs

There are two main types of SDMs.

Type of SDMs

There are two main types of SDMs.

Correlative SDMs, also known as **climate envelope models**, **bioclimatic models**, or **resource selection function models**, model the observed distribution of a species as a function of environmental conditions.

Type of SDMs

There are two main types of SDMs.

Correlative SDMs, also known as **climate envelope models**, **bioclimatic models**, or **resource selection function models**, model the observed distribution of a species as a function of environmental conditions.

Mechanistic SDMs, also known as **process-based models** or **biophysical models**, use independently derived information about a species' physiology to develop a model of the environmental conditions under which the species can exist.

Type of SDMs

There are two main types of SDMs.

Correlative SDMs, also known as **climate envelope models**, **bioclimatic models**, or **resource selection function models**, model the observed distribution of a species as a function of environmental conditions.

Mechanistic SDMs, also known as **process-based models** or **biophysical models**, use independently derived information about a species' physiology to develop a model of the environmental conditions under which the species can exist.



Correlative SDMs

- Correlative SDMs model the observed distribution of a species as a function of geographically referenced climatic predictor variables using multiple regression approaches.



Correlative SDMs

- Correlative SDMs model the observed distribution of a species as a function of geographically referenced climatic predictor variables using multiple regression approaches.
- Correlative SDMs assume that species are at equilibrium with their environment and that the relevant environmental variables have been adequately sampled.

Correlative SDMs

- Correlative SDMs model the observed distribution of a species as a function of geographically referenced climatic predictor variables using multiple regression approaches.
- Correlative SDMs assume that species are at equilibrium with their environment and that the relevant environmental variables have been adequately sampled.
- Records of species absences are typically not as common as records of presences, thus often "random background" or "pseudo-absence" data are used to fit these models.

Correlative SDMs

- Correlative SDMs model the observed distribution of a species as a function of geographically referenced climatic predictor variables using multiple regression approaches.
- Correlative SDMs assume that species are at equilibrium with their environment and that the relevant environmental variables have been adequately sampled.
- Records of species absences are typically not as common as records of presences, thus often "random background" or "pseudo-absence" data are used to fit these models.
- Since correlative SDMs are models of a species' observed distribution, they are models of the **realized niche** (the environments where a species is found), as opposed to the **fundamental niche** (the environments where a species can be found, or where the abiotic environment is appropriate for the survival).

Correlative SDMs

- Correlative SDMs model the observed distribution of a species as a function of geographically referenced climatic predictor variables using multiple regression approaches.
- Correlative SDMs assume that species are at equilibrium with their environment and that the relevant environmental variables have been adequately sampled.
- Records of species absences are typically not as common as records of presences, thus often "random background" or "pseudo-absence" data are used to fit these models.
- Since correlative SDMs are models of a species' observed distribution, they are models of the **realized niche** (the environments where a species is found), as opposed to the **fundamental niche** (the environments where a species can be found, or where the abiotic environment is appropriate for the survival).



Family of correlative SDMs algorithms

- **Profile techniques:** Simple statistical techniques that use e.g. environmental distance to known sites of occurrence



Family of correlative SDMs algorithms

- **Profile techniques:** Simple statistical techniques that use e.g. environmental distance to known sites of occurrence

1 BIOCLIM



Family of correlative SDMs algorithms

- **Profile techniques:** Simple statistical techniques that use e.g. environmental distance to known sites of occurrence

- 1 BIOCLIM
- 2 DOMAIN

Family of correlative SDMs algorithms

- **Profile techniques:** Simple statistical techniques that use e.g. environmental distance to known sites of occurrence
 - 1 BIOCLIM
 - 2 DOMAIN
 - 3 Ecological Niche Factor Analysis (ENFA)

Family of correlative SDMs algorithms

- **Profile techniques:** Simple statistical techniques that use e.g. environmental distance to known sites of occurrence

- 1 BIOCLIM
- 2 DOMAIN
- 3 Ecological Niche Factor Analysis (ENFA)
- 4 Mahalanobis distance

Family of correlative SDMs algorithms

- **Profile techniques:** Simple statistical techniques that use e.g. environmental distance to known sites of occurrence

- 1 BIOCLIM
- 2 DOMAIN
- 3 Ecological Niche Factor Analysis (ENFA)
- 4 Mahalanobis distance
- 5 Isodar analysis

Family of correlative SDMs algorithms

- **Profile techniques:** Simple statistical techniques that use e.g. environmental distance to known sites of occurrence

- 1 BIOCLIM
- 2 DOMAIN
- 3 Ecological Niche Factor Analysis (ENFA)
- 4 Mahalanobis distance
- 5 Isodar analysis

- **Regression-based techniques**

Family of correlative SDMs algorithms

- **Profile techniques:** Simple statistical techniques that use e.g. environmental distance to known sites of

occurrence

- 1 BIOCLIM
- 2 DOMAIN
- 3 Ecological Niche Factor Analysis (ENFA)
- 4 Mahalanobis distance
- 5 Isodar analysis

- **Regression-based techniques**

- 1 Generalized linear model (GLM)

Family of correlative SDMs algorithms

- **Profile techniques:** Simple statistical techniques that use e.g. environmental distance to known sites of

occurrence

- 1 BIOCLIM
- 2 DOMAIN
- 3 Ecological Niche Factor Analysis (ENFA)
- 4 Mahalanobis distance
- 5 Isodar analysis

- **Regression-based techniques**

- 1 Generalized linear model (GLM)
- 2 Generalized additive model (GAM)

Family of correlative SDMs algorithms

- **Profile techniques:** Simple statistical techniques that use e.g. environmental distance to known sites of occurrence

- 1 BIOCLIM
- 2 DOMAIN
- 3 Ecological Niche Factor Analysis (ENFA)
- 4 Mahalanobis distance
- 5 Isodar analysis

- **Regression-based techniques**

- 1 Generalized linear model (GLM)
- 2 Generalized additive model (GAM)
- 3 Multivariate adaptive regression splines (MARS)

Family of correlative SDMs algorithms

- **Profile techniques:** Simple statistical techniques that use e.g. environmental distance to known sites of

occurrence

- 1 BIOCLIM
 - 2 DOMAIN
 - 3 Ecological Niche Factor Analysis (ENFA)
 - 4 Mahalanobis distance
 - 5 Isodar analysis
- **Regression-based techniques**
 - 1 Generalized linear model (GLM)
 - 2 Generalized additive model (GAM)
 - 3 Multivariate adaptive regression splines (MARS)
 - 4 Maxlike

Family of correlative SDMs algorithms

- **Profile techniques:** Simple statistical techniques that use e.g. environmental distance to known sites of occurrence

- 1 BIOCLIM
- 2 DOMAIN
- 3 Ecological Niche Factor Analysis (ENFA)
- 4 Mahalanobis distance
- 5 Isodar analysis

- **Regression-based techniques**

- 1 Generalized linear model (GLM)
- 2 Generalized additive model (GAM)
- 3 Multivariate adaptive regression splines (MARS)
- 4 Maxlike
- 5 Favourability Function (FF)

Family of correlative SDMs algorithms

- **Profile techniques:** Simple statistical techniques that use e.g. environmental distance to known sites of occurrence

- 1 BIOCLIM
- 2 DOMAIN
- 3 Ecological Niche Factor Analysis (ENFA)
- 4 Mahalanobis distance
- 5 Isodar analysis

- **Regression-based techniques**

- 1 Generalized linear model (GLM)
- 2 Generalized additive model (GAM)
- 3 Multivariate adaptive regression splines (MARS)
- 4 Maxlike
- 5 Favourability Function (FF)



Family of correlative SDMs algorithms (End)

- Machine learning techniques



Family of correlative SDMs algorithms (End)

- Machine learning techniques

- 1 MAXENT

Family of correlative SDMs algorithms (End)

- Machine learning techniques

- 1 MAXENT
- 2 Artificial neural networks (ANN)

Family of correlative SDMs algorithms (End)

■ Machine learning techniques

- 1 MAXENT
- 2 Artificial neural networks (ANN)
- 3 Genetic Algorithm for Rule Set Production (GARP)

Family of correlative SDMs algorithms (End)

■ Machine learning techniques

- 1 MAXENT
- 2 Artificial neural networks (ANN)
- 3 Genetic Algorithm for Rule Set Production (GARP)
- 4 Boosted regression trees (BRT)/gradient boosting machines (GBM)

Family of correlative SDMs algorithms (End)

■ Machine learning techniques

- 1 MAXENT
- 2 Artificial neural networks (ANN)
- 3 Genetic Algorithm for Rule Set Production (GARP)
- 4 Boosted regression trees (BRT)/gradient boosting machines (GBM)
- 5 Random forest (RF)

Family of correlative SDMs algorithms (End)

■ Machine learning techniques

- 1 MAXENT
- 2 Artificial neural networks (ANN)
- 3 Genetic Algorithm for Rule Set Production (GARP)
- 4 Boosted regression trees (BRT)/gradient boosting machines (GBM)
- 5 Random forest (RF)
- 6 Support vector machines (SVM)

Family of correlative SDMs algorithms (End)

■ Machine learning techniques

- 1 MAXENT
- 2 Artificial neural networks (ANN)
- 3 Genetic Algorithm for Rule Set Production (GARP)
- 4 Boosted regression trees (BRT)/gradient boosting machines (GBM)
- 5 Random forest (RF)
- 6 Support vector machines (SVM)
- 7 XGBoost (XGB)

Family of correlative SDMs algorithms (End)

■ Machine learning techniques

- 1 MAXENT
 - 2 Artificial neural networks (ANN)
 - 3 Genetic Algorithm for Rule Set Production (GARP)
 - 4 Boosted regression trees (BRT)/gradient boosting machines (GBM)
 - 5 Random forest (RF)
 - 6 Support vector machines (SVM)
 - 7 XGBoost (XGB)
- Furthermore, **ensemble models** can be created from several model outputs to create a model that captures components of each. Often the mean or median value across several models is used as an ensemble.
 - Similarly, **consensus models** are models that fall closest to some measure of the central tendency of all models; consensus models can be individual model runs or ensembles of several models.



Description

- Classification Tree Analysis (CTA), also known as Decision Tree Analysis, is a machine learning algorithm used for classification and regression tasks. In the context of SDM, CTA is applied to predict the presence or absence of a species based on environmental predictor variables. The algorithm recursively splits the dataset into subsets based on the predictor variables, creating a tree structure where each node represents a decision based on a specific variable. Here are some of the advantages of CTA:



Description

- Classification Tree Analysis (CTA), also known as Decision Tree Analysis, is a machine learning algorithm used for classification and regression tasks. In the context of SDM, CTA is applied to predict the presence or absence of a species based on environmental predictor variables. The algorithm recursively splits the dataset into subsets based on the predictor variables, creating a tree structure where each node represents a decision based on a specific variable. Here are some of the advantages of CTA:
- **Interpretability:** Decision trees are easy to interpret and visualize, making it simpler for researchers and stakeholders to understand the model's decision-making process.

Description

- Classification Tree Analysis (CTA), also known as Decision Tree Analysis, is a machine learning algorithm used for classification and regression tasks. In the context of SDM, CTA is applied to predict the presence or absence of a species based on environmental predictor variables. The algorithm recursively splits the dataset into subsets based on the predictor variables, creating a tree structure where each node represents a decision based on a specific variable. Here are some of the advantages of CTA:
- **Interpretability:** Decision trees are easy to interpret and visualize, making it simpler for researchers and stakeholders to understand the model's decision-making process.
- **Non-linearity Handling:** CTA can capture non-linear relationships between environmental variables and species occurrence, which might be missed by linear models.

Description

- Classification Tree Analysis (CTA), also known as Decision Tree Analysis, is a machine learning algorithm used for classification and regression tasks. In the context of SDM, CTA is applied to predict the presence or absence of a species based on environmental predictor variables. The algorithm recursively splits the dataset into subsets based on the predictor variables, creating a tree structure where each node represents a decision based on a specific variable. Here are some of the advantages of CTA:
- **Interpretability:** Decision trees are easy to interpret and visualize, making it simpler for researchers and stakeholders to understand the model's decision-making process.
- **Non-linearity Handling:** CTA can capture non-linear relationships between environmental variables and species occurrence, which might be missed by linear models.
- **No Assumption of Linearity:** Unlike some parametric models, CTA does not assume a linear relationship between predictors and the response, making it more flexible.

Description

- Classification Tree Analysis (CTA), also known as Decision Tree Analysis, is a machine learning algorithm used for classification and regression tasks. In the context of SDM, CTA is applied to predict the presence or absence of a species based on environmental predictor variables. The algorithm recursively splits the dataset into subsets based on the predictor variables, creating a tree structure where each node represents a decision based on a specific variable. Here are some of the advantages of CTA:
- **Interpretability:** Decision trees are easy to interpret and visualize, making it simpler for researchers and stakeholders to understand the model's decision-making process.
- **Non-linearity Handling:** CTA can capture non-linear relationships between environmental variables and species occurrence, which might be missed by linear models.
- **No Assumption of Linearity:** Unlike some parametric models, CTA does not assume a linear relationship between predictors and the response, making it more flexible.
- **Variable Importance:** CTA provides a measure of variable importance, helping identify which environmental factors are most influential in predicting species distribution.



Use cases and weaknesses of CTA

- Habitat Suitability Modeling; Predicting Range Shifts; Understanding the ecological requirements and preferences of a species.

Use cases and weaknesses of CTA

- Habitat Suitability Modeling; Predicting Range Shifts; Understanding the ecological requirements and preferences of a species.
- Decision trees can be prone to overfitting, capturing noise in the training data and leading to poor generalization on new data.



Use cases and weaknesses of CTA

- Habitat Suitability Modeling; Predicting Range Shifts; Understanding the ecological requirements and preferences of a species.
- Decision trees can be prone to overfitting, capturing noise in the training data and leading to poor generalization on new data.
- Small changes in the data can result in different tree structures, impacting the stability of the model.



Use cases and weaknesses of CTA

- Habitat Suitability Modeling; Predicting Range Shifts; Understanding the ecological requirements and preferences of a species.
- Decision trees can be prone to overfitting, capturing noise in the training data and leading to poor generalization on new data.
- Small changes in the data can result in different tree structures, impacting the stability of the model.
- CTA can be sensitive to outliers, potentially leading to biased predictions.



Use cases and weaknesses of CTA

- Habitat Suitability Modeling; Predicting Range Shifts; Understanding the ecological requirements and preferences of a species.
- Decision trees can be prone to overfitting, capturing noise in the training data and leading to poor generalization on new data.
- Small changes in the data can result in different tree structures, impacting the stability of the model.
- CTA can be sensitive to outliers, potentially leading to biased predictions.

CTA model building and evaluation

- **Accuracy:** Accuracy is the ratio of correctly predicted instances to the total instances. It measures the overall correctness of the model.

CTA model building and evaluation

- **Accuracy:** Accuracy is the ratio of correctly predicted instances to the total instances. It measures the overall correctness of the model.

$$Accuracy = \frac{TruePositives + TrueNegative}{Total} \quad (1)$$

- **Sensitivity:** Sensitivity is the ratio of correctly predicted positive observations to the total actual positives. It measures the ability of the model to capture all positive instances.

CTA model building and evaluation

- **Accuracy:** Accuracy is the ratio of correctly predicted instances to the total instances. It measures the overall correctness of the model.

$$Accuracy = \frac{TruePositives + TrueNegative}{Total} \quad (1)$$

- **Sensitivity:** Sensitivity is the ratio of correctly predicted positive observations to the total actual positives. It measures the ability of the model to capture all positive instances.

$$Sensitivity = \frac{TruePositives}{ActualPositive} \quad (2)$$

- **Specificity:** Specificity is the ratio of correctly predicted negative observations to the total actual negatives. It measures the ability of the model to avoid false positives.

CTA model building and evaluation

- **Accuracy:** Accuracy is the ratio of correctly predicted instances to the total instances. It measures the overall correctness of the model.

$$Accuracy = \frac{TruePositives + TrueNegative}{Total} \quad (1)$$

- **Sensitivity:** Sensitivity is the ratio of correctly predicted positive observations to the total actual positives. It measures the ability of the model to capture all positive instances.

$$Sensitivity = \frac{TruePositives}{ActualPositive} \quad (2)$$

- **Specificity:** Specificity is the ratio of correctly predicted negative observations to the total actual negatives. It measures the ability of the model to avoid false positives.

$$Specificity = \frac{TrueNegatives}{ActualNegative} \quad (3)$$

- **Precision:** Precision is the ratio of correctly predicted positive observations to the total predicted positives. It measures the accuracy of positive predictions made by the model.

CTA model building and evaluation

- **Accuracy:** Accuracy is the ratio of correctly predicted instances to the total instances. It measures the overall correctness of the model.

$$Accuracy = \frac{TruePositives + TrueNegative}{Total} \quad (1)$$

- **Sensitivity:** Sensitivity is the ratio of correctly predicted positive observations to the total actual positives. It measures the ability of the model to capture all positive instances.

$$Sensitivity = \frac{TruePositives}{ActualPositive} \quad (2)$$

- **Specificity:** Specificity is the ratio of correctly predicted negative observations to the total actual negatives. It measures the ability of the model to avoid false positives.

$$Specificity = \frac{TrueNegatives}{ActualNegative} \quad (3)$$

- **Precision:** Precision is the ratio of correctly predicted positive observations to the total predicted positives. It measures the accuracy of positive predictions made by the model.

$$Precision = \frac{TruePositive}{PredictedPositive} \quad (4)$$



Concept

- In the context of Species Distribution Modeling (SDM), an ensemble model refers to a modeling approach where multiple individual models are combined to produce a more robust and accurate prediction.



Concept

- In the context of Species Distribution Modeling (SDM), an ensemble model refers to a modeling approach where multiple individual models are combined to produce a more robust and accurate prediction.
- The idea is that by aggregating predictions from multiple models, the weaknesses of individual models can be mitigated, and the overall performance can be improved. Common techniques for creating ensemble models include averaging, stacking, and boosting.



Concept

- In the context of Species Distribution Modeling (SDM), an ensemble model refers to a modeling approach where multiple individual models are combined to produce a more robust and accurate prediction.
- The idea is that by aggregating predictions from multiple models, the weaknesses of individual models can be mitigated, and the overall performance can be improved. Common techniques for creating ensemble models include averaging, stacking, and boosting.
- Ensemble models often outperform individual models by reducing overfitting and capturing a broader range of patterns in the data.

Concept

- In the context of Species Distribution Modeling (SDM), an ensemble model refers to a modeling approach where multiple individual models are combined to produce a more robust and accurate prediction.
- The idea is that by aggregating predictions from multiple models, the weaknesses of individual models can be mitigated, and the overall performance can be improved. Common techniques for creating ensemble models include averaging, stacking, and boosting.
- Ensemble models often outperform individual models by reducing overfitting and capturing a broader range of patterns in the data.
- Ensembles are more robust to noise and variations in the data, as they aggregate information from multiple sources.

Concept

- In the context of Species Distribution Modeling (SDM), an ensemble model refers to a modeling approach where multiple individual models are combined to produce a more robust and accurate prediction.
- The idea is that by aggregating predictions from multiple models, the weaknesses of individual models can be mitigated, and the overall performance can be improved. Common techniques for creating ensemble models include averaging, stacking, and boosting.
- Ensemble models often outperform individual models by reducing overfitting and capturing a broader range of patterns in the data.
- Ensembles are more robust to noise and variations in the data, as they aggregate information from multiple sources.
- Ensemble models can be more complex and computationally intensive than individual models, especially when using a large number of base models.

Concept

- In the context of Species Distribution Modeling (SDM), an ensemble model refers to a modeling approach where multiple individual models are combined to produce a more robust and accurate prediction.
- The idea is that by aggregating predictions from multiple models, the weaknesses of individual models can be mitigated, and the overall performance can be improved. Common techniques for creating ensemble models include averaging, stacking, and boosting.
- Ensemble models often outperform individual models by reducing overfitting and capturing a broader range of patterns in the data.
- Ensembles are more robust to noise and variations in the data, as they aggregate information from multiple sources.
- Ensemble models can be more complex and computationally intensive than individual models, especially when using a large number of base models.
- The interpretability of ensemble models may be lower compared to individual models, as it involves combining predictions from multiple sources.

Concept

- In the context of Species Distribution Modeling (SDM), an ensemble model refers to a modeling approach where multiple individual models are combined to produce a more robust and accurate prediction.
- The idea is that by aggregating predictions from multiple models, the weaknesses of individual models can be mitigated, and the overall performance can be improved. Common techniques for creating ensemble models include averaging, stacking, and boosting.
- Ensemble models often outperform individual models by reducing overfitting and capturing a broader range of patterns in the data.
- Ensembles are more robust to noise and variations in the data, as they aggregate information from multiple sources.
- Ensemble models can be more complex and computationally intensive than individual models, especially when using a large number of base models.
- The interpretability of ensemble models may be lower compared to individual models, as it involves combining predictions from multiple sources.



Application

- Ensemble models are particularly effective when dealing with diverse and heterogeneous datasets.



Application

- Ensemble models are particularly effective when dealing with diverse and heterogeneous datasets.
- If there is uncertainty about which individual model is the best, ensemble models can help capture different perspectives.



Application

- Ensemble models are particularly effective when dealing with diverse and heterogeneous datasets.
- If there is uncertainty about which individual model is the best, ensemble models can help capture different perspectives.
- Ensembles can be useful in situations where overfitting is a concern with individual models.



Application

- Ensemble models are particularly effective when dealing with diverse and heterogeneous datasets.
- If there is uncertainty about which individual model is the best, ensemble models can help capture different perspectives.
- Ensembles can be useful in situations where overfitting is a concern with individual models.



Concept

- An envelope model is a type of ecological niche model that identifies the environmental conditions within which a species is likely to occur. It uses associations between aspects of climate and known occurrences of species across landscapes of interest to define sets of conditions under which species are likely to maintain viable populations.

Concept

- An envelope model is a type of ecological niche model that identifies the environmental conditions within which a species is likely to occur. It uses associations between aspects of climate and known occurrences of species across landscapes of interest to define sets of conditions under which species are likely to maintain viable populations.
- Envelope models are often simpler and computationally less demanding than some machine learning-based models.

Concept

- An envelope model is a type of ecological niche model that identifies the environmental conditions within which a species is likely to occur. It uses associations between aspects of climate and known occurrences of species across landscapes of interest to define sets of conditions under which species are likely to maintain viable populations.
- Envelope models are often simpler and computationally less demanding than some machine learning-based models.
- The results are often easy to interpret, making them accessible for practitioners and researchers with varying levels of statistical expertise.

Concept

- An envelope model is a type of ecological niche model that identifies the environmental conditions within which a species is likely to occur. It uses associations between aspects of climate and known occurrences of species across landscapes of interest to define sets of conditions under which species are likely to maintain viable populations.
- Envelope models are often simpler and computationally less demanding than some machine learning-based models.
- The results are often easy to interpret, making them accessible for practitioners and researchers with varying levels of statistical expertise.
- The performance of envelope models can be sensitive to the quality and representativeness of input data.

Concept

- An envelope model is a type of ecological niche model that identifies the environmental conditions within which a species is likely to occur. It uses associations between aspects of climate and known occurrences of species across landscapes of interest to define sets of conditions under which species are likely to maintain viable populations.
- Envelope models are often simpler and computationally less demanding than some machine learning-based models.
- The results are often easy to interpret, making them accessible for practitioners and researchers with varying levels of statistical expertise.
- The performance of envelope models can be sensitive to the quality and representativeness of input data.



Application

- Envelope models assume that the species' niche is in equilibrium with the environmental conditions. This may not hold true in cases of rapid environmental change.

Application

- Envelope models assume that the species' niche is in equilibrium with the environmental conditions. This may not hold true in cases of rapid environmental change.
- Non-ML envelope models may assume linear relationships between species occurrences and environmental variables, which may not capture complex ecological responses.

Application

- Envelope models assume that the species' niche is in equilibrium with the environmental conditions. This may not hold true in cases of rapid environmental change.
- Non-ML envelope models may assume linear relationships between species occurrences and environmental variables, which may not capture complex ecological responses.

Concept

- MaxEnt (Maximum Entropy) is a popular machine learning-based algorithm used for Species Distribution Modeling (SDM). It's widely used due to its ability to handle presence-only data and its capacity to model complex relationships between species occurrences and environmental variables.

Concept

- MaxEnt (Maximum Entropy) is a popular machine learning-based algorithm used for Species Distribution Modeling (SDM). It's widely used due to its ability to handle presence-only data and its capacity to model complex relationships between species occurrences and environmental variables.
- MaxEnt operates based on the principle of maximum entropy, seeking a probability distribution that is maximally uncertain (maximum entropy) while satisfying constraints imposed by the available information (species occurrences and environmental variables). In the context of SDM, MaxEnt aims to find the probability distribution of environmental variables given the observed species occurrences.

Concept

- MaxEnt (Maximum Entropy) is a popular machine learning-based algorithm used for Species Distribution Modeling (SDM). It's widely used due to its ability to handle presence-only data and its capacity to model complex relationships between species occurrences and environmental variables.
- MaxEnt operates based on the principle of maximum entropy, seeking a probability distribution that is maximally uncertain (maximum entropy) while satisfying constraints imposed by the available information (species occurrences and environmental variables). In the context of SDM, MaxEnt aims to find the probability distribution of environmental variables given the observed species occurrences.
- The objective function in MaxEnt is formulated as:

Concept

- MaxEnt (Maximum Entropy) is a popular machine learning-based algorithm used for Species Distribution Modeling (SDM). It's widely used due to its ability to handle presence-only data and its capacity to model complex relationships between species occurrences and environmental variables.
- MaxEnt operates based on the principle of maximum entropy, seeking a probability distribution that is maximally uncertain (maximum entropy) while satisfying constraints imposed by the available information (species occurrences and environmental variables). In the context of SDM, MaxEnt aims to find the probability distribution of environmental variables given the observed species occurrences.
- The objective function in MaxEnt is formulated as:

$$\text{Maximize } H(p) = - \sum_i p(x_i) \log(p(x_i)) + \lambda \sum_j w_j f_j(x) \quad (5)$$

where $H(p)$ is the entropy; $p(x_i)$ is the probability of the environmental variable x_i ; λ is a regularization parameter; w_j are weights assigned to environmental features $f_j(x)$



Mechanistic SDMs

- Mechanistic SDMs utilize physiological data from controlled field or laboratory studies to determine the environmental conditions necessary for species persistence.



Mechanistic SDMs

- Mechanistic SDMs utilize physiological data from controlled field or laboratory studies to determine the environmental conditions necessary for species persistence.
- Simple models may establish threshold values beyond which a species cannot survive.

Mechanistic SDMs

- Mechanistic SDMs utilize physiological data from controlled field or laboratory studies to determine the environmental conditions necessary for species persistence.
- Simple models may establish threshold values beyond which a species cannot survive.
- Complex models comprise multiple sub-models, such as micro-climate conditions derived from macro-climate data, body temperature from micro-climate conditions, and biological rates (e.g., survival, fecundity) based on body temperature (thermal performance curves), resource requirements, and population dynamics.

Mechanistic SDMs

- Mechanistic SDMs utilize physiological data from controlled field or laboratory studies to determine the environmental conditions necessary for species persistence.
- Simple models may establish threshold values beyond which a species cannot survive.
- Complex models comprise multiple sub-models, such as micro-climate conditions derived from macro-climate data, body temperature from micro-climate conditions, and biological rates (e.g., survival, fecundity) based on body temperature (thermal performance curves), resource requirements, and population dynamics.
- Mechanistic SDMs incorporate causal mechanisms and excel at extrapolation and non-equilibrium scenarios.

Mechanistic SDMs

- Mechanistic SDMs utilize physiological data from controlled field or laboratory studies to determine the environmental conditions necessary for species persistence.
- Simple models may establish threshold values beyond which a species cannot survive.
- Complex models comprise multiple sub-models, such as micro-climate conditions derived from macro-climate data, body temperature from micro-climate conditions, and biological rates (e.g., survival, fecundity) based on body temperature (thermal performance curves), resource requirements, and population dynamics.
- Mechanistic SDMs incorporate causal mechanisms and excel at extrapolation and non-equilibrium scenarios.
- However, they demand extensive physiological data collection and validation, which may be challenging to obtain.

Mechanistic SDMs

- Mechanistic SDMs utilize physiological data from controlled field or laboratory studies to determine the environmental conditions necessary for species persistence.
- Simple models may establish threshold values beyond which a species cannot survive.
- Complex models comprise multiple sub-models, such as micro-climate conditions derived from macro-climate data, body temperature from micro-climate conditions, and biological rates (e.g., survival, fecundity) based on body temperature (thermal performance curves), resource requirements, and population dynamics.
- Mechanistic SDMs incorporate causal mechanisms and excel at extrapolation and non-equilibrium scenarios.
- However, they demand extensive physiological data collection and validation, which may be challenging to obtain.
- These models rely on numerous assumptions and parameter estimates, potentially leading to complexity.

Mechanistic SDMs

- Mechanistic SDMs utilize physiological data from controlled field or laboratory studies to determine the environmental conditions necessary for species persistence.
- Simple models may establish threshold values beyond which a species cannot survive.
- Complex models comprise multiple sub-models, such as micro-climate conditions derived from macro-climate data, body temperature from micro-climate conditions, and biological rates (e.g., survival, fecundity) based on body temperature (thermal performance curves), resource requirements, and population dynamics.
- Mechanistic SDMs incorporate causal mechanisms and excel at extrapolation and non-equilibrium scenarios.
- However, they demand extensive physiological data collection and validation, which may be challenging to obtain.
- These models rely on numerous assumptions and parameter estimates, potentially leading to complexity.



Thank You