This is a video transcript of Module 10 of the Online Open Course in Species Distribution Modelling. To access the full suite of videos click here.

Online Open Course - Species Distribution Modelling

Powered by the Biodiversity and Climate Change Virtual Laboratory

---

Module 10 Case studies in the BCCVL

Welcome to the last module of the online open course in species distribution modelling. In this module, I am going to show you 4 different case studies that highlight the variety of research questions and applications that can be addressed with species distribution models. And I will show you how you can run these models in the BCCVL, the Biodiversity and Climate Change Virtual Laboratory. This is an online tool that let's you run species distribution models in a few easy steps.

In this module, I will first explain what the BCCVL is, and how you can use the BCCVL to run your species distribution models and climate change projections. And I will take you through 4 different case studies on how species distribution models and climate change projections have helped advance our knowledge in a number of different areas.

The Biodiversity and Climate Change Virtual Laboratory is a decision support and scenario planning tool that simplifies the process of biodiversity and climate change modelling. It brings together multiple biological, environmental and climate change datasets, algorithms and experiments in one easy-to-use, convenient, online location.

The BCCVL offers species distribution models with 17 different algorithms, across the 4 categories that we have spoken about in this course. One of the standout features of this tool is the ability to run all 17 algorithms at the same time, and compare the outputs of the algorithms at the click of a button.

Once you have run your species distribution model you can then use your results to run a climate change experiment across a number of different global warming scenarios, general circulation models, and years.

Case study 1: Corroboree frog

Let's start with our first case study. This is the Corroboree frog, one of Australia's most threatened vertebrates. They are only found in a 400 square kilometer patch in the subalpine regions of southern New South Wales and Victoria. The species is listed on the IUCN Red List as a critically endangered species, with fewer than 100 individuals observed in the wild in recent years. The decline of the Corroboree frogs is caused by several threats such as habitat loss, infections, predation by feral animals, and climate change.

And this is Peter, our ecologist, and he is determined to prevent the extinction of the Corroboree frog. So he wants to run a species distribution model in combination with a climate change projection to model the current and future distribution of this species, so he can evaluate which areas he needs to protect to maintain the population of this species.

As we have learned in previous modules of this course, it is advised to do a bit of research on your species, so you can make an appropriate selection of environmental variables to include in the model. Peter has searched the scientific literature, and found the following facts:
This critically endangered species needs shallow pools in wet grasslands and heath to breed, and is therefore affected by a changing climate. Especially long dry periods that occur frequently in Australia affect these breeding breeding pools, resulting in egg and tadpole deaths. The ongoing decline and the small population of this Corroboree frog remaining make it likely that the species will become extinct in the very near future.

Keeping this information in mind, Peter starts his search for occurrence data and environmental data that he needs to run a species distribution model. He finds a dataset with only 24 unique occurrence points in the Atlas of Living Australia. Although his dataset is not very large, the species is restricted to a small area, and these records are likely still a good representation of the environmental conditions that are suitable for the survival of the species. Because Peter doesn't have any absence data, he will generate pseudo-absence points in the BCCVL. For the environmental data, Peter is using the Australian Current Climate dataset and he has selected the bioclimatic variables that would best dictate a location suitable for the Corroboree frog.

I will now show you how Peter ran this model in the BCCVL.

After logging into the BCCVL, Peter first has a look at the available datasets. For species data, he can import data from ALA or upload his own dataset. And then there is a variety of data collections available for climate data and other environmental variables such as water availability, land cover and soil type.

Peter starts with selecting his species data by importing a dataset from the Atlas of Living Australia. He enters the name of the species and can then either look at the occurrence points at the ALA website or import them directly into the BCCVL. After the import is complete, he can view the map with the occurrence points. To check whether all data points are accurate, he downloads the occurrence dataset, and then cleans the data as we explained in module 3. He then uploads the cleaned dataset into the BCCVL, and the map now shows the dataset with only one occurrence point per location.

The next step is to design the experiment. In the BCCVL, you can run five different experiments divided in primary and secondary experiments. Peter selects the Species Distribution Modelling experiment, which brings him to a page with 7 different tabs where he can select each of the aspects of his model.

In the first tab he can give his experiment a title and a description.

In the next tab, he needs to select his occurrence dataset. He can search through the datasets with keywords on the left side. Peter selects the clean dataset of the corroboree frog, and can view the occurrence points again on the map.

The next step is to select absence data. Peter does not have any true absence data, so selects the option to generate pseudo-absence points.

The following tab is where Peter can select the environmental variables for his model. In the menu on the left he can select the type of data, the resolution, collection and source. The datasets that match the criteria then show up on the right, and Peter selects the current Australian climate dataset. He can view each of the layers that are available in this dataset on the map, and then select which of the environmental variables he wants to include in his model. From his research Peter knows that the Corroboree frog is affected by rainfall and temperature and he chooses 6 different variables to include in his model.

Under the constraints tab, Peter has the option to constrain his model to a particular geographic region, in this case the entire Australian continent.

The last step is to select the algorithms for the model. There are currently 17 different algorithms available in the BCCVL across the four different categories that we discussed in modules 5, 6 and 7. For each algorithm that you select, you can set the configuration options. There is information available for each of these options to guide you in your selection. If you don’t change the values of the options, the model uses the default configuration options of the algorithm package.

Under the last tab you can start your experiment. You can see that there are now 6 models running, one for each algorithm that Peter selected. You can turn off your computer, go and grab a coffee, and you will receive an e-mail once the experiment is completed, so you can then go back and view the results.

So Peter ran six different algorithms in the BCCVL, and he is curious to see what the maps look like with the predicted current distribution of the Corroboree frog. And these are the results. As you can see, there are slight differences between the different algorithms, but the core area of the Corroboree frog distribution is the same. Peter checks the performance metrics of all models, and finds that the Artificial Neural Network performed best, and he selects this algorithm for the climate change projection that he wants to run.

He also has a look at the response curves to see the effect of the environmental variables on the probability of occurrence for the frogs. I highlight two examples here: the left curve in this plot is the response for annual mean temperature, and shows that high temperatures lead to a zero probability of occurrence for the frog. The right plot shows the response for annual rainfall, and indeed shows that the probability of occurrence is high when rainfall is high.

Peter does not only want to know what the current distribution of the Corroboree frog is, and which areas are suitable for this species to survive now, but he also wants to look at their future distribution, to see how climate change might affect this species, and which areas will still be suitable in the future, so he can prioritize these areas for conservation. So, his next step is to project the predicted distribution of one of the algorithms, in this case, Artificial Neural Networks, into the future. He uses the Australia Future Climate dataset, and the general circulation model that was developed by CSIRO in Australia. And he selects two different global warming scenarios: an intermediate scenario, RCP 4.5, and the business-as usual scenario, RCP 8.5, and projects the distribution for 2045 and 2085.

Let's go back to the BCCVL, and I'll show you how to run a climate change experiment.

Back in the BCCVL, Peter goes to the experiments page again, and now selects to run a Climate Change experiment.

This brings him again to a page with different tabs to select each of the aspects of the model. He first gives the experiment a title and a description.

In the second tab, he can select the species distribution model that he wants to use for his projection. After selecting an experiment, all the different algorithms that he ran are listed. Because ANN performed the best, he selects this algorithm for the projection.

The next tab is where he selects the future climate data. On the left side, Peter can select which emission scenario he wants to use, as well as the future year of projection, the general circulation model and the resolution of the dataset. Peter selects two different scenarios for two different years using the CSIRO MK3 circulation model. The datasets that match these criteria are shown on the right, and he selects these four to run his model.

After reviewing all aspects of the model, Peter starts the experiment in the last tab, and we can see that now 4 different models are running. Let's have a look at the results.

Let's have a look what the future holds for the Corroboree frog. On the left is the map with the current distribution, predicted with the Artificial Neural Networks algorithm that Peter used for his climate change projections. He ran two different global warming scenarios, RCP 4.5 and RCP 8.5. RCP 4.5 represents an intermediate scenario, in which we use some technologies to limit carbon emission, but the earth is still predicted to warm with about 2.4 degrees Celsius by the year 2100. Under this scenario, we see that the suitable area for Corroboree frogs to survive is already declining, with still quite some areas left in 2045, but a significant reduction by 2085. The RCP 8.5 scenario is the ‘business as usual scenario’ in which we don’t limit our carbon emissions and keep going as we do right now. This scenario shows a faster reduction in suitable areas for the Corroboree frog, with only a small portion of the currently suitable area left by 2085.

So it is clear that the future doesn’t look good for the Corroboree frog. But with the help of species distribution models in combination with future climate change projections, ecologists like Peter, and managers can make decisions about which areas they should prioritize for conservation. Even under the most pessimistic, but unfortunately quite likely global warming scenario, there was a small area that would still be suitable for this species to survive and we could make sure to minimize other impacts such as habitat fragmentation and destruction and that could be a good start to protect this vulnerable species from going extinct.

Case study 2: Lantana

Our second case study is about Lantana, a plant that is regarded as one of the worst weeds in Australia. It was brought to Australia as an ornamental garden plant in 1841, but quickly established itself in the wild and is now spread along most of the eastern Australian coastline. It forms a threat to native plant species, such as eucalyptus trees as it competes with seedlings for nutrients and light. It also reduces pasture productivity, and it has been estimated that this weed incurs an economical cost of more than 20 million dollar per year for control management.

Hazel, our invasive species specialist wants to find out how lantana will be affected by future climates to understand whether the impact of this species will get worse or not. Like Peter, she takes the approach of running a species distribution model with a climate change projection in the BCCVL.

Let’s first again check the facts that Hazel found about Lantana. This toxic species is mostly found in coastal and sub-coastal areas of Australia, and thrives under warm climates with high rainfall, so it is negatively impacted by cooler temperatures and reduced precipitation. Hazel will take these facts into account when she selects the environmental data for her model.

Hazel uses the BCCVL to import the occurrence data for Lantana from the Atlas of Living Australia. After cleaning the data, she has an occurrence dataset of over 10,000 records. She selects the option in the BCCVL to randomly generate 10,000 pseudo-absence points throughout Australia. For her environmental dataset, she chooses a suite of environmental variables from the Australian Current Climate dataset at a 1 km resolution.

Let’s have a look at the results that Hazel got. All six algorithms that Hazel ran in the BCCVL show a similar current distribution of Lantana, widely spread along the eastern coast of Australia. The ROC curve and evaluation statistics that are displayed here are the results of the

Generalized Additive Model. This is the algorithm that Hazel chooses to use for her Climate Change projection.

These are three of the response curves for some of the temperature variables that Hazel put in her model. The two bottom graphs show that Lantana is negatively affected by cold temperatures, but the top graph shows that when it comes to maximum temperature, there might be an optimum, and thus a reduced probability of occurrence in temperatures about 30 degrees Celsius.

Hazel runs a Climate Change Experiment in the BCCVL as her next step, with the results from the Generalized Additive Model, and for her climate change projection, she selects the Australian General Circulation Model, CSIRO Mark 3. She wants to see the projection for the best possible scenario, represented by RCP 2.6, which predicts a warming of the earth of approximately 1.5 degrees Celsius by the year 2100. She also selects the worst possible case, RCP 8.5, which is the business-as-usual scenario. Hazel projects the results of the Generalized Additive Model with these two scenarios for 2 different years in the future, 2045 and 2085.

And these are the results that she got. On the left, we have again displayed the distribution map under current climate conditions. Let's first look at the RCP 2.6 scenario. Even under this relatively optimistic, but very ambitious global warming scenario, we still see a likely increase in the spread of this weed. If we look at the business as usual scenario, we see an enormous spread of the weed by 2085 in northern Australia, and also a movement further land inwards on the eastern side of the country.

So, unfortunately for Hazel, the future does not look very good in her efforts to eliminate the Lantana weed. It is clear that this weed is able to survive in a range of climatic conditions, shown by its widespread distribution throughout Australia. All future climate change scenarios show an increase in the spread of the weed. But remember that in this instance, we have only modelled the species distribution in relation to climate variables. To continue her research, Hazel can investigate which non-climatic variables influence Lantana, such as soil conditions. It is however clear that management actions need to be put in place across the country to control the spread of this weed in the future.

Case study 3: Hendra virus

Our third case study is about flying foxes, and more so about a disease that they carry: the Hendra virus. While the flying foxes themselves don't show any signs of illness when they carry the disease, they can transfer the disease to horses and humans for whom an infection can be life-threatening.

Let me introduce you to Linus. He is a horse stud owner near Byron Bay who lost two of his horses last year to the Hendra virus. He is looking to relocate his stable to land in areas where flying foxes are absent. And he is looking for a long term solution. He wants to move to an area where flying foxes are absent, now but also in the future.

Linus asked his data analyst to look up some information about the grey-headed flying fox and provide him with some maps of the current and potential future distribution of this species. And this is what his data analyst found: the grey headed flying fox is the largest species of bat in Australia. They live in a variety of habitats from rainforests to woodlands and swamps. They roost during the day and head off late afternoon to forage for food. While they are not a migratory species, populations of the grey headed flying fox will move from place to place as various plant species blossom. These bats are known to carry Hendra virus and were considered responsible for the many horse deaths in 2010. Research shows that flying foxes are negatively impacted by extreme temperature events and reduced rainfall.

The next step is to import the occurrence data that is available on the Atlas of Living Australia into the BCCVL. After cleaning up the data, Linus's data analyst has an occurrence dataset with over 5,000 points, and he is going to generate 10,000 pseudo-absence points in the BCCVL. He selects 8 climate variables from the Australian Current Climate dataset to run his species distribution models. And he decides to run 6 different algorithms, including both machine learning algorithms as well as statistical regression algorithms.

And these are the distribution maps from each of these six algorithms. They all generally follow the pattern of the occurrence data. The data analyst checks the performance metrics and decides to use the results of the Generalized Linear Model for his climate change projection.

He also checks the response curves, and these three show the probability of occurrence related to temperature and rainfall. We saw in the facts that flying foxes are susceptible to extreme heat events, which is reflected in the response curve for maximum temperature of the warmest month. They seem to be less affected by minimum temperatures, as the response curve for the relation with minimum temperature of the coldest month shows a flat line. The impact of rainfall is displayed in the bottom graph, that shows a negative impact of reduced precipitation.

Like Hazel and Peter, he also decides to use the general circulation model that was developed by the CSIRO in Australia, and he selects two different Representative Concentration Pathways, and two different years into the future.

The results of the climate change projection show that an intermediate RCP, 4.5, doesn't show that much of a change in comparison to the current distribution of the grey-headed flying fox. This means that if we undertake actions to stop increasing carbon emissions and limit global warming, the distribution of this species will remain similar to its current distribution. But if we look at the distribution predicted based on a RCP 8.5 scenario, we can see a shift towards more southern areas by 2085.

So what is the data analyst going to advise Linus? It is clear that his current location in Byron Bay will remain a suitable area for grey-headed flying foxes for many years to come under each of the climate change scenarios. Only the most severe 'business as usual' global warming scenario showed a southward movement of grey headed flying foxes in the future. So if Linus wants to move his stables, going south is probably not such a good idea, but he might decide to go further landwards and move to more western areas in Queensland or New South Wales.

Case study 4: Wine industry

Our last case study highlights how also the agricultural industry can benefit from species distribution models. In this case, we are looking at the wine industry, and in particular at the environmental conditions suitable for grape vines to grow.

Verona and Ben run a family-owned winery and cellar door in the Granite Belt, a wine region in Queensland in Australia. Verona and Ben want to prepare themselves for the future, and get a better understanding of how climate change is going to impact the growth of grapevines.

Let's first check the facts again. Climate has a major influence on the growth of the grapevines, and prolonged high temperatures can have a negative impact on the quality of the grapes, but also affect the colour and taste of the wine. But this is not the only environmental factor that is of importance. It appears that the interaction between climate and soil conditions is also very important. Verona and Ben will thus check whether there is an environmental dataset available with soil conditions that they can use to run their species distribution model.

They first have a look at the data that is available in the Atlas of Living Australia, but come to the conclusion that there are less than 200 records available, not very representative of each of the wine regions in Australia. Therefore they run a search in Google Earth for Australian vineyards, and use the coordinates of these vineyards as additional records, which gives them a total of almost 700 records. They check carefully whether these occurrence points represent the major wine areas in the country, and generate pseudo-absence points in the BCCVL.

Because they have learnt that both climate and soil are important factors that can limit the suitability of areas for grapevines to grow, they make a selection of environmental data that include both climatic variables from the Australian Current Climate dataset, as well as four different soil variables from the National Soil Grids dataset.

These are the distribution maps with the predictions of six different algorithms. It is clear that algorithms with more complex interactions between variables and the response, such as Random Forest and Multivariate Adaptive Regression Splines predict slightly lower probability of occurrences, but generally they predict these probabilities in the same areas as the other algorithms. Verona and Ben select Maxent, the maximum entropy model as the algorithm to use for climate change projections based on the performance metrics.

As in the other case studies, Verona and Ben select a general circulation model, two different global warming scenarios or RCPs and two different future periods to project the results of their species distribution model into the future.

The results show that for both the intermediate RCP 6 as well as the business as usual, RCP 8.5 scenario, the predicted distribution won't be severely affected in the first 30 years. For 2085, both scenarios show a slight decrease in probability of occurrence around the northernmost boundaries of the grapevine and a reduced suitable area in western Australia.

So how is it looking for Verona and Ben? It looks like they can safely continue their grapes where they are, but under a severe global warming scenario they might have to look for mitigation strategies. As they are located on the northern range of the suitable area for this species, they could investigate the possibility to change to grape varieties that are able to withstand warmer temperatures.

We have come to the end of this online open course in species distribution modelling. I would like to thank you very much for watching, and more so hope that this course has learnt you the essentials about species distribution models. Don't forget that these are complex models that need and deserve time and effort to fully comprehend and understand. I would like to encourage you to go and explore the species data, the environmental data and the different algorithms that are available and that you can use to predict the distribution of your species of interest now and in the future. And don't forget, that you can run these models in the Biodiversity and Climate Change Virtual Laboratory!

Attribution

Please cite this video as follows:

Huijbers CM, Richmond SJ, Low-Choy SJ, Laffan SW, Hallgren W, Holewa H (2016) SDM Online Open Course, module 10: Case studies in the BCCVL. Biodiversity and Climate Change Virtual Laboratory, http://www.bccvl.org.au/training/. DDMMYY of access.

Acknowledgements

Lead partners: Griffith University, James Cook University

Thanks to: University of New South Wales, Macquarie University, University of Canberra

Funded by: National eResearch Tools and Resources Project (NeCTAR)

The BCCVL is supported by the National eResearch Tools and Resources Project (NeCTAR), an initiative of the Commonwealth being conducted as part of the Super Science Initiative and financed from the Education Investment Fund, Department of Education. The University of Melbourne is the lead agent for the delivery of the NeCTAR project and Griffith University is the sub-contractor.