

# MASTER THESIS

Course code: BIO5010

Name / candidate.:

Oddbjørn Larsen / 5

---

Modelling the distribution for *Bohemanella frigida* at the northern range boundary: Habitat preferences and morphological features

---

Date: 29. May 2020

Total pages: 75

## Contents

Contents.....	i
Abstract .....	1
1 Introduction .....	2
2 Method .....	5
2.1 Study area.....	6
2.2 The first species distribution model.....	7
2.3 The second and evolving species distribution model.....	12
2.4 Finding grasshoppers at high probability locations.....	14
2.5 Estimation of plant species composition.....	14
2.6 Environmental sensors .....	15
2.7 Measuring morphological features.....	16
2.8 Statistical analysis .....	18
3 Results .....	20
3.1 The first species distribution model .....	20
3.2 The second and evolving species distribution model.....	23
3.3 Plant species composition .....	28
3.4 Explanatory variables impact on the occurrence of the grasshoppers .....	31
3.5 Environmental data .....	32
3.6 Morphological features .....	38
4 Discussion .....	40
4.1 Distribution models .....	40
4.2 Plant communities .....	42
4.3 Explanatory variables impact on the occurrence of the grasshoppers .....	42
4.4 Environmental data .....	43
4.5 Morphological features .....	44
4.6 Future prospects .....	45
4.7 Conclusion.....	47
5 Acknowledgements .....	49
References .....	50
Appendix .....	55

## Abstract

The grasshopper species *Bohemanella frigida* is widespread in mountainous regions in Norway. Although grasshoppers have not been known to prefer particular host plants, a strong correlation between vegetation dominated by *Vaccinium myrtillus* and *Calluna vulgaris* had been established by previous research. At the northern edge of the range at a *B. frigida* location in Vesterålen, another type of vegetation pattern was observed, where herbs seemed to dominate the habitat.

To test if this was a repeating pattern in the area, several species distribution models were created to test if new *B. frigida* locations could be found with a species distribution model. Where these models predicted a high probability of *B. frigida* occurrence, the vegetation was sampled, and sensors were deployed to further investigate if vegetation, microclimate or other factors had an effect on habitat selection. These explanatory variables were analysed to see which variable effected *B. frigida* habitat selection. Morphological features were measured to see if they differentiate at the northern edge boundary.

With one of the species distribution models, four new *B. frigida* habitats were located on Hinnøya. A similar trend was found in these locations, where *Vaccinium myrtillus* together with where herbs and ferns were among the most abundant plants, these habitats were warmer and drier than comparable locations indicating a correlation between *B. frigida* and microclimatic variations. Slope was found to be the only significant explanatory variable in *B. frigida* occurrence. Larger individuals than described in the literature was found, and different theories are discussed, which could explain these differences.

We conclude that species distribution models can be used to find *B. frigida* occurrences, slope is the most significant explanatory variable to explain the occurrences. Vegetation is different in *B. frigida* habitats at the northern range boundary and together with microclimatic conditions might explain their habitat selection, to defiantly prove this, further research would be required.

## 1 Introduction

The simplest ecological question one can ask is, why are organisms of a particular species present in some places and absent from others? (Krebs, 2009). There are three main conditions which need to be met for a species to occupy a place and maintain a population, the species has to reach the place and disperse there, the abiotic environmental conditions must be ecophysiologicaly suitable for the species and the biotic environment must be suitable for the species (Guisan, Thuiller, & Zimmermann, 2017).

Abiotic conditions can interact with biotic processes such as predation and competition to limit a species range (Fryxell, Sinclair, & Caughley, 2014). In the centre of a species range, there is likely a broader range of locations and suitable habitats to meet the species requirements. At the opposite end of a species range, specifically the high-latitude leading-edge boundaries, populations may become restricted to a particular habitat. This may be due to these habitats supporting a high population growth because the resources needed for survival and reproduction are available, or because other abiotic factors such as unusually warm microclimates can enable populations to persist despite increased climate-related mortality at range margins (Oliver, Hill, Thomas, Brereton, & Roy, 2009).

The fundamental niche for a species is defined by the set of resources and environmental conditions that allow a single species to persist in a particular region (Schoener, 1989). However, the fundamental niche is rarely, if ever seen in nature, because competing species restrict a given species to a narrower range of conditions. This niche is the observed or realized niche of the species in the community (Fryxell et al., 2014). It can be difficult or impossible to find all factors that can affect the niche because we can never be sure that we have measured all relevant factors which describe the fundamental niche of a species (Fryxell et al., 2014). Habitat suitability modelling largely depends on prior ecological and ecophysical knowledge about factors that physiologically and ecologically determine species distributions (Guisan et al., 2017).

Many insects that are found on mountains are either accidental or temporary visitors from lower altitudes (Mani, 1968). Large numbers of insects are lifted with warm updraft winds to high altitudes only to be chilled to death by the colder climate. Then there are hypsobiont or high-altitude insects which are residents in mountainous habitats. These insects are evolved and adapted to conditions found in the high altitudes, they complete their lifecycle in the mountains (Mani, 1968).

The northern migratory grasshopper (*Bohemanella frigida*) is an hypsobiont insect in the order orthoptera which includes grasshoppers, locusts, katydids, and crickets (Holst, 1986). The order consists of almost 24.000 described species in up to 40 families, they have a hemimetabolous development where eggs are laid singly or in batches into plants or soil, and the nymphs resemble small adults except in the lack of development of wings and genitalia (Gullan, Cranston, & McInnes, 2014, p. 502). Orthopterans are typically elongate cylindrical, up to 12 cm long and with enlarged hind legs for jumping (Gullan et al., 2014, p. 502).

*B. frigida* is also known as *Melanoplus frigida* as it belongs to the grasshopper subfamily Melanoplinae, and used to belong to the nearctic-restricted tribe Melanoplini, but is now placed in the holarctically distributed Podismini tribe (Litzenberger & Chapco, 2001). The current view is that it belongs to the Podismini tribe and that North American populations were established by the Bering land bridge, but research into mitochondrial genes of this species has support for placing the grasshopper back into the Melanoplini tribe (Litzenberger & Chapco, 2001).

*B. frigida* is typically found in the boreo-alpine regions, within Europe it is found in the alps at elevations of 2100-2600 m ASL and in Fennoscandia as low as 500 m ASL (Harz, 1975). Outside Europe it can be found in northern Russia, Siberia, the Altai mountains, northern Mongolia, Manchuria, Kamchatka, Alaska and northern Canada (Harz, 1975; Mani, 1968).

Body size ranges from 16-21 mm for males and 16.6-26 mm for females (Harz, 1975, p. 248; Holst, 1986), the pronotum ranges from 3.5-4.5 mm in males and 4.1-5.7 mm for females (Harz, 1975), while the tegmen ranges from 5.2 – 7 mm in males and 5.2 – 8.0 mm for females, hence most individuals are brachypterous and macropterous individuals seldom occurs (Finch, Löffler, & Pape, 2008). The widespread occurrence of wing reduction or absence of wings in alpine grasshoppers strongly suggests that this is an essential factor of adaption to high altitudes (Sømme, 1989). *B. frigida* is not known to stridulate as the sound known from other grasshopper species are caused by friction between the tegmen and a series of tubercles, or articulated pegs on the legs (Uvarov, 1977).

It prefers southern faced slopes dominated by *Vaccinium myrtillus* and *Juncus trifidus* (Hansen, 2009; Ottesen, 1992). Newer research also found that *B. frigida* prefers southern facing slopes but dominated by *Vaccinium myrtillus* and at lower altitudes also dominated by *Calluna vulgaris* and it was also observed at southern exposed foot slopes dominated by mats of the grass *Nardus stricta* (Finch et al., 2008).

Internal temperature in grasshoppers is positively affected by the temperature of the air, by radiant heat from the sun, the ground and objects on it, and by the metabolic production of heat; while losses of heat may be due to convection and long-wave radiation from the body, as well as to the evaporation of water (Uvarov, 1977, p. 207). In order to reach their optimum temperature, *B. frigida* is known to climb up on low vegetation and expose themselves to direct sunlight until they are sufficiently warmed up (Finch et al., 2008). Another factor that can affect how fast grasshoppers reach their optimum internal temperature is air humidity; studies on *Schistocerca* showed that the temperature equilibration is reached faster in humid air than in dry air (Bodenheimer, 1929, pp. 439-443; Uvarov, 1977, p. 208).

*B. frigida* had only one publicly recorded observation in the Vesterålen and Lofoten archipelago (including Hinnøya) before the researcher reported a new location in 2017. This new location had been known by the researcher, and among local inhabitants for several decades, but was never recorded in any public database (Global Biodiversity Information Facility, 2019). This location had different vegetation than surrounding areas, and was dominated by herbs and grasses and seemed to differ from previous research that found *Vaccinium myrtillus* and *Calluna vulgaris* to be the dominant vegetation in *B. frigida* habitations (Finch et al., 2008).

Given the infrequent reporting of *B. frigida*, their status as a hypsobiont species and the importance of species distribution data in nature management, several research questions about predicting and locating *B. frigida* in Vesterålen and Lofoten were developed. These were:

### **Research questions:**

1. Can new *B. frigida* locations be found with a species distribution model?
2. Is *B. frigida* restricted to a particular habitat at its northern range margin compared with habitats at the range centre?
3. In empty locations where the species distribution model predicts a high probability for *B. frigida*, can microclimate explain the lack of grasshoppers?
4. Are there any morphological differences in *B. frigida* at the northern edge?

## 2 Method

To find new grasshopper locations within the study area and compare these new sites with the known grasshopper location on Hinnøya, species distribution models (SDM) were created. These models predict areas of high probability for grasshopper habitats. Two different models were created. The first model used reported *B. frigida* location from all of Norway, where the location precision was less than or equal to 100 meters. The second model used grasshopper locations from within the study area where the precision of all locations was less than five meters. This model was evolving, as new locations were confirmed as presence or absence locations, they were added to the model and the model was calibrated.

Areas with a high probability of *B. frigida* occurrence were searched and plant communities were recorded in these high probability areas. If grasshoppers were found, they would be photographed so that morphological features could be measured later.

Environmental sensors were placed in six different locations without grasshoppers, but with a high probability of grasshoppers from the distribution models and sensors were placed in five different confirmed grasshopper locations. These sensors measured soil moisture (percent), temperature (Celsius) and precipitation (scaled) to see if some of these variables would explain the lack of grasshoppers in high probability locations.

## 2.1 Study area

The study area was located northernmost in Nordland county and some parts east on Hinnøya are located in Troms and Finnmark county. Species distribution models were created for the entire study area marked in Figure 1 and high probability areas on Hinnøya, Langøya, Hadseløya and Austvågøya were explored.

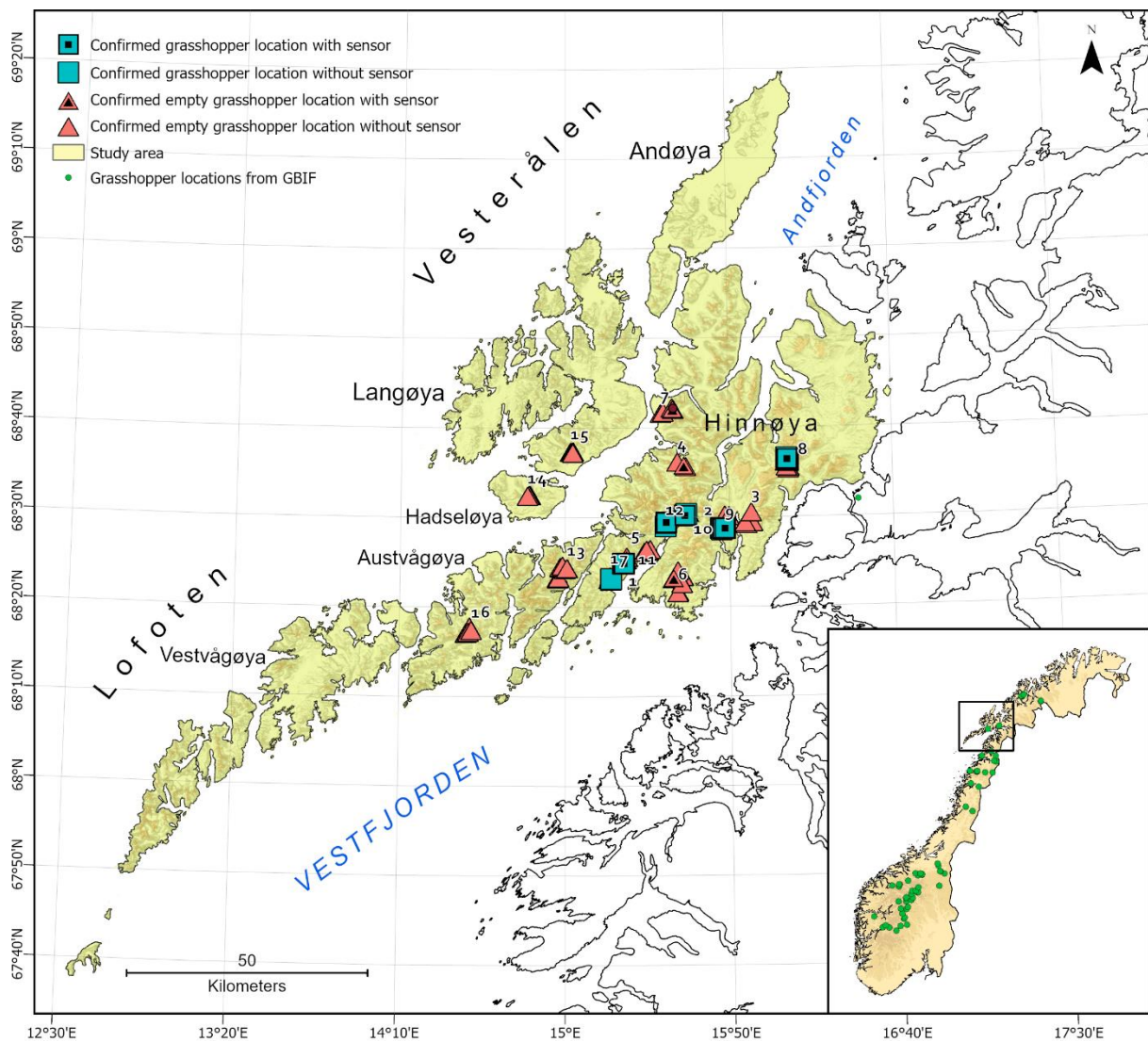


Figure 1 The study area marked with yellow colour is situated north in Norway in Nordland county. Some parts of Hinnøya are within the borders of Troms and Finnmark county. An overview map of Norway with the study area marked with a black outline is situated in the lower right corner. *B. frigida* from the Global Biodiversity Information Facility are marked with green circles (Global Biodiversity Information Facility, 2019).



## 2.2 *The first species distribution model*

### **Variables used**

To model the distribution of a species, it is vital to know the species ecological niche (Elith & Graham, 2009; Elith & Leathwick, 2009). Previous research has found that *Bohemanella frigida* prefers southern facing slopes and a habitat that consists of shrub-rich heath dominated by *Vaccinium myrtillus* and *Calluna vulgaris* in mountainous habitats as high as the low alpine belt, and they can also occur in sunny clearings of birch forest. Soil surface temperature and soil moisture conditions in these areas are crucial for the colonization of cold area by grasshoppers (Finch et al., 2008).

Several predictor rasters were used as input to map possible locations in Vesterålen and Lofoten where the grasshopper could be found. All the predictor rasters had a resolution of 30 x 30 m. To account for their distribution in the mountainous areas, a digital elevation model for Norway with a resolution of 10 x 10 m was resampled to 30 x 30 m using a bilinear resampling technique (McCoy, Johnston, & Kopp, 2002, p. 80). A slope and an aspect raster were created from the resampled digital elevation model and also used as input predictors.

To include areas that could account for a potential vegetation preferences, a normalized difference vegetation index (NDVI) was used together with a classified vegetation raster (Satveg). NDVI is a commonly used vegetation index and is based on the visible and near-infrared spectra of light, and is often used to detect proportions of live green vegetation (Lopez & Frohn, 2017). Satveg is a classified vegetation map for Norway based on satellite data, thematic topographic maps and digital elevation models, it is divided into 25 classes (Appendix A) and has a spatial resolution of 30 meters (B. E. Johansen, 2009). To account for moisture conditions a normalized difference moisture index (NDMI) was included in the model, NDMI is based on the near-infrared and shortwave infrared spectra and is used to determine vegetation water content (Jin & Sader, 2005). A soil composition index (SCI), also based on the near-infrared and the shortwave infrared spectra was included to detect the chemical soil composition (Al-Khaier, 2003).

The three indices were created in the google earth engine (Gorelick et al., 2017) with satellite data from the sentinel-2 satellites. These satellites have 13 different bands and a spatial resolution down to 10 meters in four of the bands, making them useful for land monitoring (Drusch et al., 2012). NDVI was calculated by the formula  $(NIR - RED) / (NIR + RED)$ ,

NDMI with the formula  $(NIR - SWIR) / (NIR + SWIR)$ , and SCI with the formula  $(SWIR - NIR) / (SWIR + NIR)$ . Data were filtered for the period for the months of July and August of 2018, with a maximum cloud cover percentage of 5 percent to create complete rasters of the entire study area.

To account for abiotic variables that could affect the distribution of the grasshoppers, bioclimatic variables BIO1-BIO19 (Table 1) from Worldclim were used (Fick & Hijmans, 2017).

*Table 1 Bioclimatic variables used with the first species distribution model with their code, meaning and unit of measurements. These bioclimatic variables represent annual trends.*

Code	Variable	Units
BIO1	Annual Mean Temperature	°C *10
BIO2	Mean Diurnal Range	Mean of monthly (max temp - min temp)
BIO3	Isothermality	BIO2/BIO7 *100
BIO4	Temperature Seasonality	Standard deviation *100
BIO5	Max Temperature of Warmest Month	°C *10
BIO6	Min Temperature of Coldest Month	°C *10
BIO7	Temperature Annual Range	BIO5-BIO6
BIO8	Mean Temperature of Wettest Quarter	°C *10
BIO9	Mean Temperature of Driest Quarter	°C *10
BIO10	Mean Temperature of Warmest Quarter	°C *10
BIO11	Mean Temperature of Coldest Quarter	°C *10
BIO12	Annual Precipitation	mm
BIO13	Precipitation of Wettest Month	mm
BIO14	Precipitation of Driest Month	mm
BIO15	Precipitation Seasonality	Coefficient of Variation
BIO16	Precipitation of Wettest Quarter	mm
BIO17	Precipitation of Driest Quarter	mm
BIO18	Precipitation of Warmest Quarter	mm
BIO19	Precipitation of Coldest Quarter	mm

### **Multicollinearity in the predictor variables**

Multicollinearity is used to describe if two or more predictor variables in a statistical model are linearly related. Many statistical routines in ecology are sensitive to collinearity and can lead to unstable parameter estimates, inflated standard errors on estimates and consequently biased inference statistics (Dormann et al., 2013).

To check the predictor variables for multicollinearity a variance inflation factor (VIF) test was applied to the predictor variables with a threshold of 0.7, recommended by Green (1979). If

one of the variables has a strong linear relationship with at least one other variable, the correlation coefficient would be close to one, and the VIF value for that variable would be large (Fox, 2015, pp. 341-348). A VIF value greater than ten is a signal that the model has a collinearity problem (Chatterjee & Hadi, 2015; Guisan et al., 2017; Naimi, Hamm, Groen, Skidmore, & Toxopeus, 2014). The VIF-test used, first find a pair of variables which have the maximum linear correlation (higher than the set threshold) and excludes the one with the higher VIF value and the procedure is repeated until no variable with a high correlation coefficient (greater than the set threshold) with other variables remains (Naimi et al., 2014). Only these remaining variables were used in the species distribution modelling.

### **Response variable: known locations vs. pseudo-absence points**

Reported locations of *Bohemanella frigida* from the Norwegian Biodiversity Information Centre (NBIC) that were located within Norway together with pseudo-absence points were used as response variables (Global Biodiversity Information Facility, 2019). Data collected from data warehouses such as NBIC, can contain duplicate records or data with faulty coordinates and it is vital to clean the data (Hijmans & Elith, 2013). All the reported locations of *Bohemanella frigida* were checked, and if an observation was found to be faulty it was removed before further processing and not used in the species distribution modelling. Each reported location has a precision value associated with the location. This precision value is set by the user when an observation is sent to the Norwegian Biodiversity Information Centre (NBIC). If a data point is imprecise, the information from the explanatory variables cannot be retrieved accurately from the data point coordinate, and this can have an effect on model fit and accuracy (Guisan et al., 2017, pp. 136-140). All non-faulty observations with a precision less than or equal to 100 meters were kept.

Observations from the NBIC do not contain observed absences, but only observed presences. Making probabilistic species distribution models from observed presence only can be complicated because the data for a given species may originate from an extensive collection of partly designed, usually biased observations of presence (Guisan et al., 2017, pp. 130-131). This can cause problems when inferring the habitat suitability of widely distributed species because nothing is known about their distribution in under-sampled areas, there are however different statistical approaches which can be used to predict species distributions from presence-only data. One method is to use information from the entire study area, either called

pseudo-absence or background data and combine that data with the existing presence data (Guisan et al., 2017, p. 131).

How the pseudo-absence data is collected, how many data points one should collect, how the absence and presence data should be weighted, and which strategy to use when collecting these points have been discussed thoroughly (Chefaoui & Lobo, 2008; VanDerWal, Shoo, Graham, & Williams, 2009; Warton & Shepherd, 2010). In this study, 10.000 pseudo-absence points were created with a random distribution, divided into ten sets of replicates, where each replicate had 1000 pseudo-absence points. The presence and absence points were equally weighted. These settings were suggested to be used with the generalized linear model (GLM), generalized additive model (GAM), classification tree analysis (CTA) and random forest (RF) modelling techniques (VanDerWal et al., 2009).

### **Habitat suitability modelling**

Biomod2 is the updated object-orientated version of the biomod package for R (Thuiller, Georges, Engler, & Breiner, 2013). This package, as its predecessor (biomod), can be used to ensemble multiple models of species distribution, forecast these predictions into future or past climate or land-change scenarios. Models can be tested with a wide range of approaches and enable the examination of species environment relationships (Thuiller, Lafourcade, Engler, & Araújo, 2009).

The biomod2 package includes ten different modelling techniques to describe and model the relationship between a given species and its environment (Thuiller et al., 2016). All ten modelling techniques were used, namely: GLM, generalized boosting model (GBM), GAM, CTA, artificial neural network (ANN), flexible discriminant analysis (FDA), multiple adaptive regression splines (MARS), RF, surface range envelop (SRA) and maximum entropy (MAXENT).

The original dataset was split in two, 80 percent of the data was used to calibrate the models, while the remaining 20 percent was used to evaluate the models, this split follows the Pareto principle, where 80 percent of the effects can be explained by 20 percent of the causes (Sanders, 1987). This process of calibration and evaluation was repeated ten times. This sort of cross-validation allows for a quite robust test of the models when independent data are not available (Thuiller et al., 2016).

Each model that is produced can be evaluated using different evaluation methods. For models which generate presence-absence predictions, the kappa statistic is the most used measure to test the performance the model, but it has been criticized for being inherently dependent on prevalence (the proportion of locations that are occupied) and that this dependency introduces statistical artefacts to estimates of predictive accuracy. The true skill statistics (TSS) corrects for the dependence kappa has on prevalence, while still keeping all the advantages of kappa (Allouche, Tsoar, & Kadmon, 2006). Because of this, TSS was used as the evaluation statistic for model assembly.

### **Model assembling**

The models built with the biomod2 modelling tool were assembled with the ensemble modelling function. This function use projections of individual models and ensemble them to build an ensemble of species projections over space and time (Thuiller et al., 2016). TSS was used as the evaluation metric when assembling the different models, where the TSS threshold was set to be higher or equal to the top ten performing models measured by TSS.

Explanatory variables importance was calculated with the variables\_importance function from the biomod2 package, version 3.3-7.1 (Thuiller et al., 2013). This function is based on Breiman's random forest algorithm (Breiman, 2001). A single response variable is 'shuffled' around and model predictions are made with this single variable, a correlation is computed between references predictions and the 'shuffled' one, and a return score of 0-1 is returned, where the higher value has more influence on the model (Thuiller et al., 2016).

The assembled model was evaluated with the evaluate function from biomod2, where binary transformed model predictions are compared to species occurrences, and mean TSS score, mean sensitivity and mean specificity were calculated for the assembled model (Thuiller et al., 2013). TSS score ranges from -1 to +1, where values close to zero and lower indicate a performance no better than random chance and a value of one, indicates perfect agreement (Fletcher & Fortin, 2018). Sensitivity is the true positive rate of predictions (percentage of presences correctly predicted), while specificity is the true negative rate of predictions (percentage of absences correctly predicted), both metrics ranges from 0-1 (Guisan et al., 2017, pp. 253-254). After the fieldwork was complete, all plots where data was sampled were used to measure the relative accuracy of each model.

### ***2.3 The second and evolving species distribution model***

After searching for the grasshopper in six different areas where the first SDM predicted high habitat suitability and not finding them in any of these areas, a new SDM was created. This SDM was created after re-finding a grasshopper location on Hinnøya reported in 1946 and seeing similarities between this site and the other known location within the study area. The similarities were the angle of the slope, a relatively small habitat, proximity to water and similar vegetation. The cell size used in the first SDM (30 x 30 m) could have been too large to detect these smaller suitable habitats accurately, therefore the cell size was decreased for the second SDM.

#### **Explanatory variables used**

A new set of six response variables were created for the second SDM, four topographic variables and one climatic variable were created from lidar data, and an NDVI variable was created from Sentinel-2 satellite data. The four topographic variables had a resolution of 1 x 1 meter, while the climatic and the NDVI variable had a resolution of 10 x 10 meters and was resampled to 1 x 1 meter.

The four topographic variables were all generated from a digital terrain model: slope (degrees), aspect (unfolded or 'linearized') after (Evans, Oakleaf, Cushman, & Theobald, 2014) and distance to water. The distance to water variable was created by first creating a filled digital terrain model using the fill tool in ArcGIS Pro (ESRI, 2020). This tool locates sinks and peaks in a digital terrain model and fills the sinks and removes the peaks. The filled digital terrain model was used as input in a flow direction raster; this raster is used to determine the direction of flow from every cell in the raster using the D8 method (O'Callaghan & Mark, 1984). The flow direction raster was then used as input in a flow accumulation raster where the accumulated flow is the accumulated weight of all cells flowing into each downslope cell in the output raster.

The next step used in creating a distance to water variable was to identify the stream networks, this was done by using a conditional statement, by using the flow accumulation raster created in the previous step as input, and identifying all cells that had > 10000 other cells flowing into them, a new stream network raster is created. This raster now has cells with the value 1 or NA, where each cell that has the value 1 is defined as a river. The output from

the conditional statement was converted into a feature using the stream to feature tool and this feature was used as input in the Euclidean distance tool in ArcGIS Pro (ESRI, 2020).

A climatic variable was created in the form of a solar radiation raster where solar radiation is measured in  $\text{WH/m}^2$ . The digital terrain model over the study area was split into 50 x 50 km blocks and solar radiation was calculated for each block. The area solar radiation tool in ArcGIS Pro (ESRI, 2020) uses the mean latitude of the input together with selected days to generate the output. Solar radiation for (day number 182 to 243) July and August 2019 were used as input days; day interval was set to 14 and hour interval was set to 0.5. The results were merged using the mosaic to new raster tool in ArcGIS Pro (ESRI, 2020) and resampled to 1 x 1 meter.

A normalized difference vegetation index (NDVI) with a resolution of 10 x 10 meters was created with the `sen2r` package, version 1.3.2 in R (Ranghetti & Busetto, 2020). The time window to gather data was set between the 1st of July 2019 to 31st of July 2019 with a maximum cloud coverage of 5 percent. The resulting NDVI raster was resampled to 1 x 1 meter.

## **Response variables**

Thirteen points from two confirmed grasshopper sites and 36 points from six unoccupied grasshopper locations were used in the second SDM. These points were split in 70 percent for prediction and 30 percent were used in evaluation. As new locations were examined, they were added either to the confirmed *B. frigida* locations or added as true absence locations.

When insufficient true absence data are available, adding pseudo-absence data may be useful (Georges & Thuiller, 2013). It is also suggested to make several pseudo-absence selections when true absences are not numerous enough, using this method, the influence of the pseudo-absence data could be estimated later on in the modelling (Thuiller et al., 2016).

From here onwards, the model was assembled in the same manner as the first SDM, including removal of variables with collinearity problems, generation of pseudo-absence points, and model assembling and evaluation.

#### ***2.4 Finding grasshoppers at high probability locations***

The grasshoppers were found by looking down at the ground while walking in mountainous areas where the SDM's had predicted a high probability of *B. frigida* habitats. From previous experience, they seemed to have a tendency to jump away when they notice a threat approaching, making them easy to notice as they jump away. In high probability areas, both with and without grasshoppers present, vegetation was sampled and the type of weather was registered. A finer search was conducted in the absent areas to lower the risk of defining an actual *B. frigida* location as unoccupied.

#### ***2.5 Estimation of plant species composition***

To see if the grasshoppers within the study area preferred another type of plant species composition than other research had found elsewhere in Norway (Finch et al., 2008), estimation of plant cover were conducted in locations where they were present and in areas where the species distribution models predicted a high probability, but no grasshoppers were found.

The visual estimation of species cover method was used, where the percent of each plant is estimated visually. With this method, the vertical projection of each species is projected onto an imagined two-dimensional plane that is parallel to the ground, taking into account all the above-ground plant parts. Because of this, plant cover can be over 100%, allowing for overlapping plants within the canopy and the estimates can include both species rooted inside and outside the plot boundary (Halbritter et al., 2020). The plot size was set to 0.5 x 0.5 m.



## 2.6 Environmental sensors

Environmental sensors were used to measure air temperature, moisture in the soil and precipitation. These sensors were built around the (UNO R3 ATmega328P) Arduino system of microcontrollers (Banzi & Shiloh, 2014). A data logger shield (DLS 2.0 Xtreme low power data logger shield for Arduino) was connected to the Arduino microcontroller, which controlled the Arduino and turned it on at a set interval of 59 minutes. Each time the Arduino was turned on, it measured the temperature from a DS18B20 Temperature Sensor, moisture in the ground from three LM393 comparator soil moisture sensors and precipitation from an LM393 comparator raindrops detection sensor module. This data was then saved to an SD-card in the Arduino data logger.

The Arduino, the data logger and a 9V battery were contained in a waterproof box (Figure 2), and the sensors were connected by wires and placed outside of the box. Each moisture sensor had a wire length of 2 meters and was placed in the ground 2 meters left, right and downhill from the waterproof box. The temperature sensors wire had a length of 1 meter and were placed above the ground, but within the vegetation. The sensor that measured precipitation was screwed on top of the box in such a way that rain could fall directly on it.



Figure 2 The sensor box, on the left side is a sensor box placed in a habitat without grasshoppers, the rain sensor can be seen on top, and wires going to the soil moisture sensors can be seen in the bottom right corner. On the right is a photo of the inside of a sensor box, where the Arduino low power shield can be seen together with the battery and miscellaneous wires.

## ***2.7 Measuring morphological features***

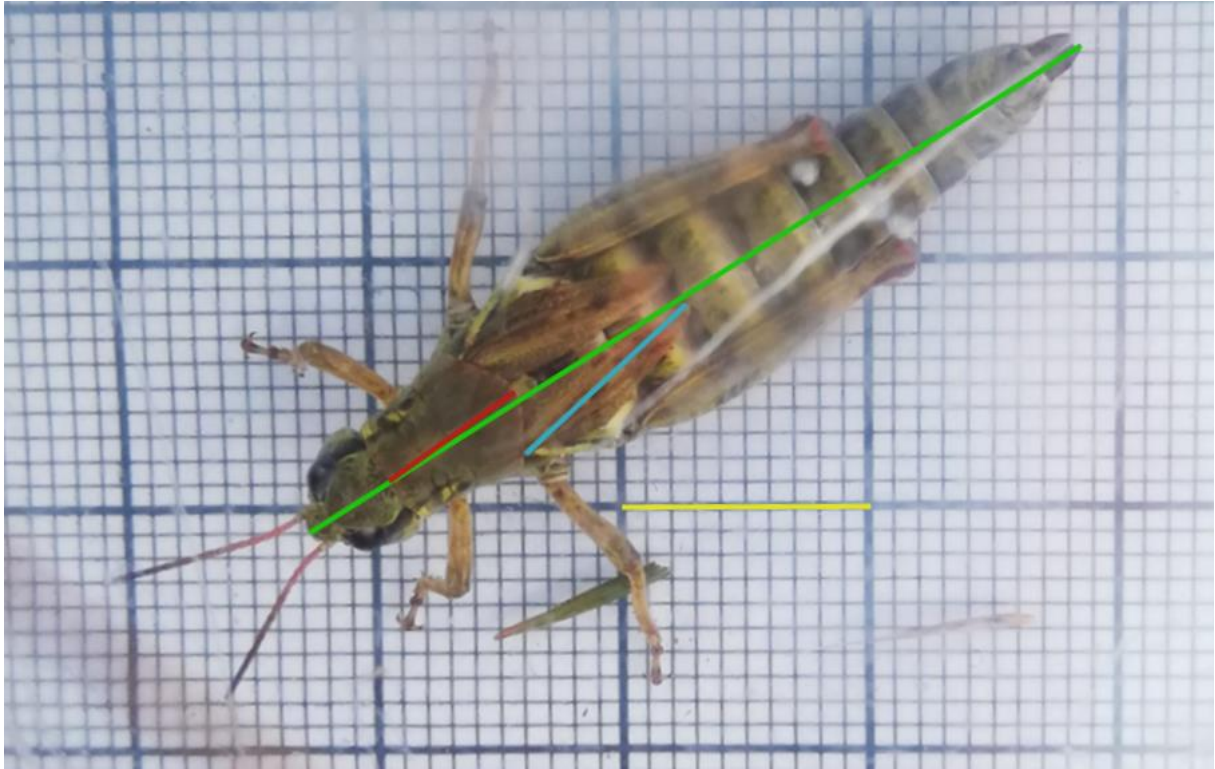
Body length is usually measured from the front of the head to the tip of the hind femora in grasshopper species which have wings that do not extend beyond the tip of the abdomen; this measurement usually corresponds with the tip of the abdomen in fresh specimens (Capinera, Scott, & Walker, 2004). The body length, the length of the pronotum and the length of the tegmen were measured in accordance with Carbonell (2002, pp. 12-15).

The grasshoppers were captured in a plastic box that had millimetre paper glued to the bottom. They were left in the box until they settled down and did not jump around inside the box, then a picture was taken from above the box, at as close to 90 degrees perpendicular to the box as possible.

The images were processed in Digimizer (MedCalc Software Ltd, 2020), where a line was drawn between two centimetre endpoints of the millimetre paper to calibrate the image (Figure 3). A second line was drawn along the length of the tip of the head to the tip of the abdomen to measure the specimen's body length (Figure 3), a third line was drawn along the length of the pronotum (Figure 3) and a fourth line along the length of the tegmen (Figure 3).

To check if the measurements in Digimizer were correct, a tape measurement band was placed next to the box used to measure the grasshoppers. The tape measurement band was placed one centimetre above the millimetre paper to represent the height of a grasshopper in the box. The millimetre paper was used to calibrate the image in the same manner as the photos with grasshoppers were, and ten measurements on the tape measurement band between two centimetre marks were used to calculate the standard deviation of the measurements.

The differences in morphological features for females and males did not meet parametric assumptions, hence a Wilcoxon signed-rank test was used to test differences between the two sexes. An alpha level of 0.05 was considered as statistically significant.



*Figure 3 Measurement of a grasshopper from Digimizer software. The yellow line is a calibration line between two centimetre marks. The red line is the length of the pronotum. The green line is the length from the fastigium to the end of the abdomen (body length) and the blue line is the length of the tegmen.*

## 2.8 Statistical analysis

### Vegetation

The vegetation data were analysed with the `labdsv` package, version 2.0-1 (D. W. Roberts, 2019). A vegetation table was created, and the constancy of the vegetation table was summarized where the presence or absence of grasshoppers was used as an effect on the relative frequencies. The constancy index is the proportion of plots of that the species is in, and is a number between one and zero, where species with a value of one is found in all the plots (D. Roberts, n.d.). Rare species with a constancy less than 0.1 (found in less than 10% of the plots) were not included in the constancy analysis to reduce their influence on the result as rare species are more likely to be in the plots by chance (Lawesson, 2000).

Detrended correspondence analysis of the vegetation data was created with the `vegan` package, version 2.5-6 (Oksanen et al., 2019) to show the distribution of all species and plots along the main gradients to explain the variation in the floristic data. Rare species were down-weighted with the 'iweigh' option. Environmental variables collected from each plot were passively fitted with the `envfit` function from the `vegan` package (Oksanen et al., 2019).

### Explanatory variables impact on the occurrence of the grasshoppers

Vegetation data in the form of DCA axis 1, type of weather when searching a location and the six variables used in the second species distribution model were examined to find which of the variables had an effect on *B. frigida* habitat presence, the response variables were presence or absence of *B. frigida* and the binomial family was used. Generalized linear mixed-effects models were fitted in R (R Core Team, 2019) using the `lm4` package, version 1.1-21 (Bates, Maechler, Bolker, & Walker, 2015). Location was treated as a random effect, and all the numeric response variables were standardized using the `scale` function. A model selection table of models with combinations (subsets) of fixed effect was created with the `dredge` function in the `MuMIn` R-package, version 1.43.15 (Barton, 2019), with the `bobyqa` optimizer and a maximum of 100.000 iterations. All models were ranked with the Akaike Information Criterion (AIC), where models with lower AIC scores and higher weights are assumed to be better (Burnham & Anderson, 2004). Overdispersion was checked with the `dispersion_glm` function in the R-package `blemco`, version 1.4 (Korner-Nievergelt et al., 2015)

## Sensor data

All data entries from the sensors were collected and merged in R, version 3.6.2 (R Core Team, 2019). Outliers in the form of faulty readings were removed and the data were split into two categories depending on if the sensor was placed in a habitat with grasshoppers or a habitat without grasshoppers. The three moisture sensors were combined into one variable, where the average of the three sensors was calculated. The two data sets were then analysed with the `ddply` function in the `dplyr` R-package, version 0.8.4 (Wickham, François, Henry, & Müller, 2020), where mean, median, maximum and minimum values were calculated for each day the sensors were turned on. The soil moisture data were converted into percentages.

The differences in temperature for locations with and without *B. frigida* did not meet parametric assumptions. Hence a Wilcoxon signed-rank test was used to test differences between the two types of location. An alpha level of 0.05 was considered as statistically significant.

### 3 Results

#### 3.1 *The first species distribution model*

Of the original 26 explanatory variables, twelve of the variables remained after testing for multicollinearity, where these twelve variables had a VIF score less than ten. Max correlation was found between bio8 ~ bio5: 0.65, where bio8 had a VIF score of 2.64 and bio5 had a VIF score of 3.65.

Two hundred fifty-six observations of *B. frigida* were initially gathered from the Norwegian Biodiversity Information Centre through the Global Biodiversity Information Facility (Global Biodiversity Information Facility, 2019). Eight of these observations were removed by manually checking for outliers and faulty coordinates. Sixty-two observations were removed as they had the same X and Y coordinates as another observation and 115 of the observations were removed as they had an accuracy greater than 100 meters. The final selection had 71 observations with a mean precision accuracy of 58.6 m spread among 56 different municipalities. These points were used as presence points in the biomod2 modelling.

As no true absence points were available when this model was created, only 10.000 pseudo-absence points were used. These were split into ten different datasets, with 1000 pseudo-absence points in each dataset. Of the ten different modelling techniques, nine were used as the SRE modelling did not work with factorial categories from satveg. The observations were split into 80 percent for training and 20 percent for validation.

The ten best performing models according to their TSS score were chosen to be used in model assembly and the TSS threshold was set to 0.448 to select these ten best performing models.

The mean TSS score from the assembled model was 0.64, indicating a good functioning model, which can predict *B. frigida* potential habitats better than random chance. Sensitivity of the assembled model was 0.76, which means that the model predicted correctly in 76 percent of the presence data used for validation. The specificity was 0.85, which means that the model predicted correctly in 85 percent of the absence data used for validation. The map created from the assembled model can be seen in Figure 4.

Relative probability values from 137 locations where the vegetation was sampled was used to calculate mean predicted values in locations with and without grasshoppers. This first species distribution model had a mean relative probability score in plots without grasshoppers of 0.57 (n = 92) and in plots with grasshoppers 0.61 (n = 45), meaning that this model was not very successful in distinguishing *B. frigida* habitats from other areas.

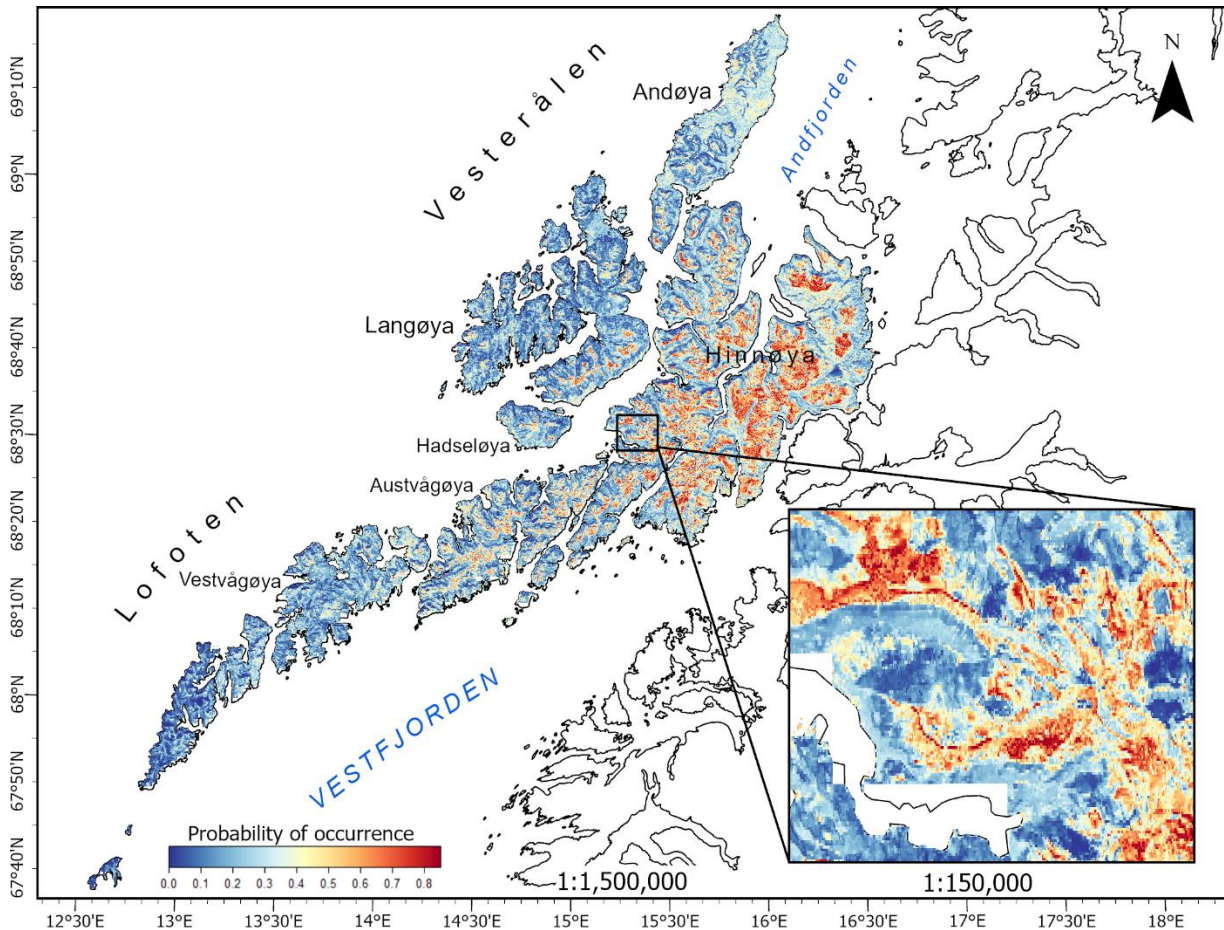


Figure 4 A map created with the first species distribution model, where a high probability of *B. frigida* habitats are marked with red colour, and low probability are blue (range 0.005-0.849). A sample area where the scale is ten times smaller to visualize details is overlaid in the right bottom corner.

Variables importance analysis showed that BIO5 (maximum temperature of the warmest month) was the most important variable, where temperatures around 12-15°C had the highest probability of occurrence (Table 2). The second most important variable in this SDM was the classified vegetation variable ‘satveg’ which had the highest probability of occurrence in class number seven (Bilberry- low fern birch forest) (Table 2). Probability of occurrences were higher in areas with an elevation over 500 meters, in areas with healthy vegetation and in southerly faced slopes around 20 degrees (Figure 5).

Table 2 Explanatory variables used in the first species distribution model ranked by importance. The values are a representation of the ranges of the explanatory variables used with this SDM and originates from all the presence and pseudo-absence points used in the model. BIO5 is the maximum temperature of the warmest month, BIO18 is the precipitation of the warmest quarter, BIO15 is precipitation seasonality, BIO3 is isothermality, BIO2 is mean diurnal range and BIO8 is the mean temperature of the wettest quarter.

Rank	Variable	Min	1st Qtl	Mean	3rd Qtl	Max	Units
1	BIO5	72.00	136.00	152.69	167.00	217.00	°C *10
2	Satveg	1.00	6.00	12.69	18.00	26.00	-
3	DEM	-0.99	241.75	557.62	822.12	2288.77	m
4	BIO18	140.00	224.00	274.87	311.00	625.00	mm
5	BIO15	14.00	24.00	27.95	31.00	51.00	Coefficient of Variation
6	BIO3	19.00	24.00	25.70	27.00	31.00	BIO2/BIO7 *100
7	SCI	-0.97	-0.29	-0.14	0.00	0.47	-
8	BIO2	38.00	59.00	67.11	76.00	95.00	Mean of monthly (max temp - min temp)
9	Aspect	0.02	86.99	179.38	267.91	359.99	degree
10	NDVI	-0.49	0.40	0.49	0.67	0.86	-
11	BIO8	-107.00	24.00	55.38	91.00	145.00	°C *10
12	Slope	0.00	3.41	11.02	15.17	72.50	Degree

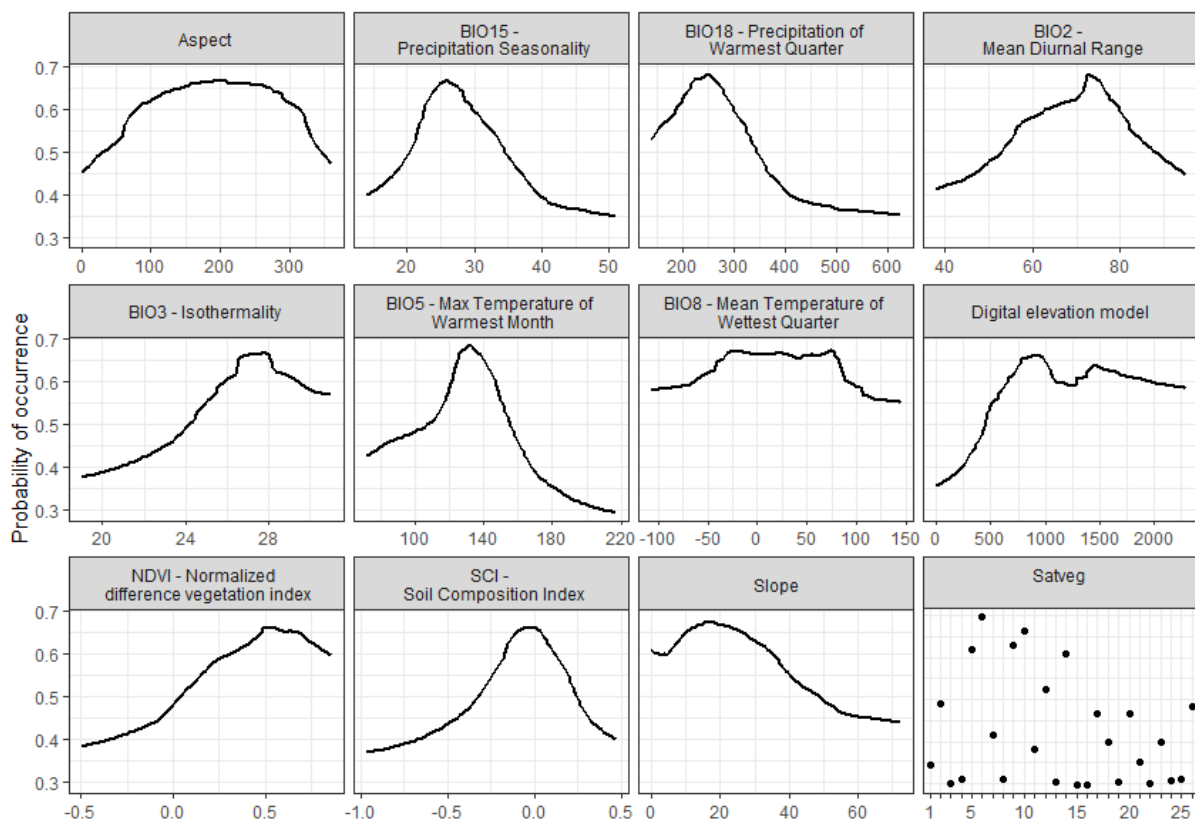


Figure 5 Response curves of the explanatory variables for the ensemble model of the first species distribution model. The highest scoring categorical variable from satveg was class number six, bilberry- low fern birch forest, followed by class number ten, tall-grown lawn vegetation (B. Johansen, Aarrestad, & Øien, 2009).



### 3.2 The second and evolving species distribution model

The six explanatory variables created for the second SDM had no collinearity problems, max correlation was found between slope ~ dtm: 0.47, where slope had a VIF score of 1.51 and the digital elevation model had a VIF score of 1.54.

13 observations from two confirmed grasshopper locations (Appendix C and Appendix J) and 36 observations from seven confirmed unoccupied locations (Appendix D – Appendix J) were used as response variables the first time this model was assembled. The observations had been collected with a GPS with an accuracy of less than five meters. 10.000 pseudo-absence points were created and split into ten different sets with 1000 pseudo-absence points in each set, and all ten different modelling techniques were used. The map created with this model is presented in Figure 6, high probability areas of occurrence was less frequent than in the first SDM.

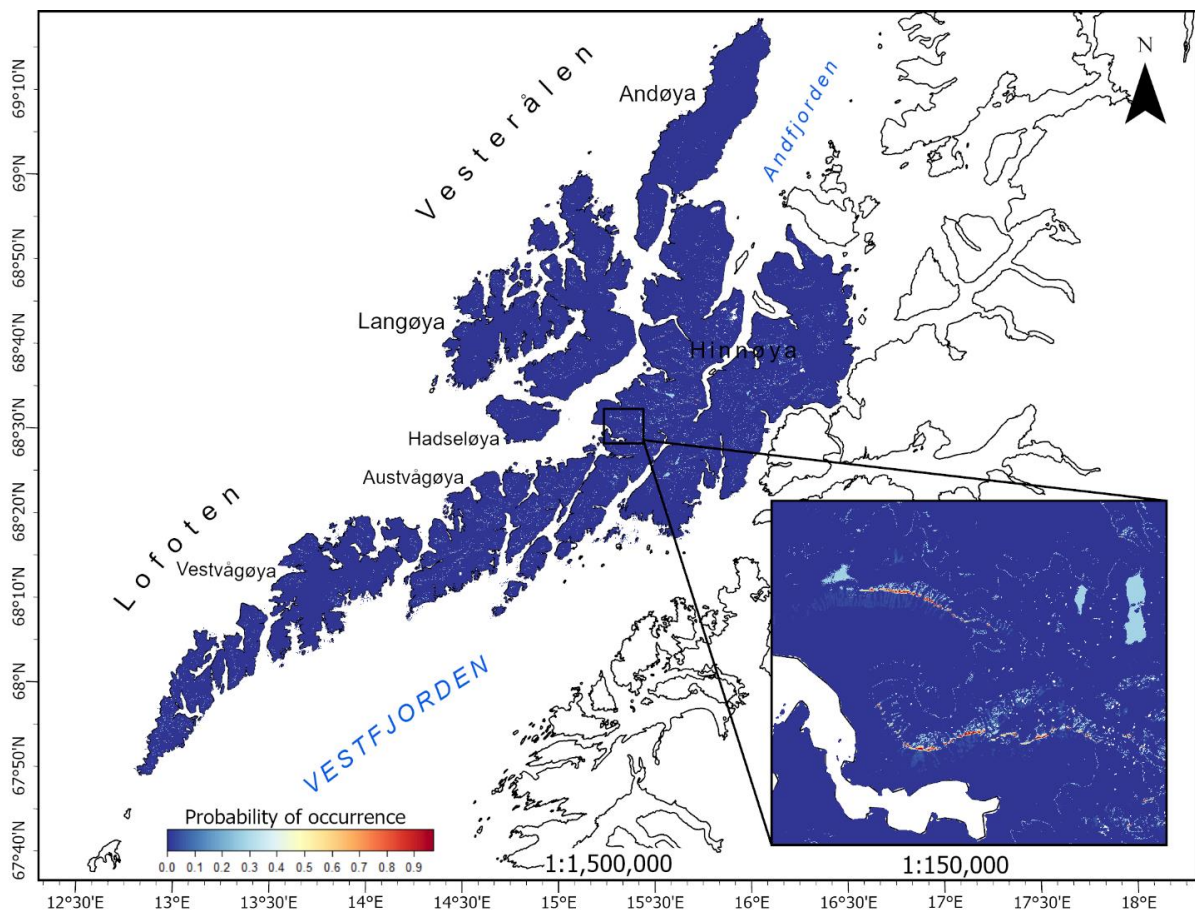


Figure 6 The map created with the second species distribution model, where a high probability of *B. frugida* habitats are marked with red colour, and low probability are blue (range: 0-0.972). A sample area where the scale is ten times smaller to visualize details is overlaid in the right bottom corner. The model used to create this map used 13 observations of *B. frugida* from two different locations and 36 observations from seven confirmed unoccupied locations.

The mean TSS score for the assembled model was 0.93 indicating a near perfect fit for the assembled model, sensitivity was 0.96 which means that the model predicted correctly in 96 percent of the presence data used for validation. Specificity was 0.79 which means that the model predicted correctly in 79 percent of the absence data used for validation. Relative probability values from 137 locations where the vegetation was sampled was used to calculate mean predicted values in locations with and without grasshoppers. This second model with 13 presence and 36 absence points had an average relative probability score in plots without grasshoppers of 0.35 (n = 92) and in plots with grasshoppers 0.72 (n = 45), which means this model was more accurate in distinguishing *B. frigida* habitats from other areas than the first species distribution model.

Elevation was the most important explanatory variable in this model (Table 3). This can also be seen in Figure 7, where there was a spike in probability of occurrence around 430 meters. Solar radiation and distance to water was similar in the two locations where the presence data came from, the grasshoppers were in areas with a high amount of solar radiation (approximately 240000 WHm<sup>2</sup>) and close to water, less than 50 meters. The probability of occurrence with the slope and aspect variable was more spread out, but higher in southern faced slopes around 20-50 degrees.

*Table 3 Response variables used in the second species distribution model ranked by importance. The values are an representation of the ranges of the explanatory variables used with this SDM and originates from all the presence, absence and pseudo-absence points used in the model.*

Rank	Variable	Min	1st Qtl	Mean	3rd Qtl	Max	Units
1	Digital elevation model	-0.04	34.94	212.15	339.99	1205.99	m
2	Solar radiation	27893	162299	177682	200929	260207	WH/m2
3	Slope	0.00	6.66	20.63	31.87	86.66	degree
4	Distance to water	0.00	14.00	51.59	73.54	952.05	m
5	Aspect	0.01	96.75	183.24	272.58	359.98	degree
6	Normalized difference vegetation index	-0.96	0.62	0.69	0.83	0.92	-

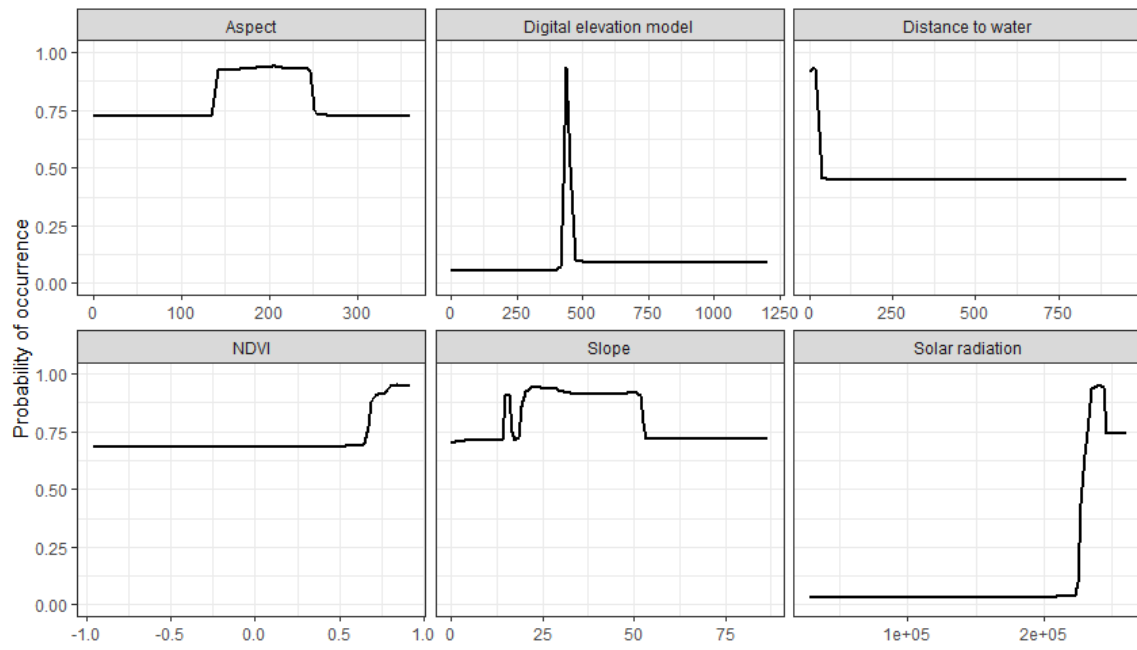


Figure 7 Response curves for the explanatory variables from the second SDM, built with 13 presence observations from two different locations and 36 absence observations from seven different locations.

## The second species distribution model with all observations

When the fieldwork was complete the final model consisted of 45 confirmed observations in six different locations and 92 empty observations from eleven locations. Mean TSS score for the assembled model with all observations was 0.75, indicating a very good model which can predict *B. frigida* habitats with confidence. Sensitivity was 0.94 which means that the model predicted correctly in 94 percent of the presence data used for validation. Specificity was 0.80 which means that the model predicted correctly in 80 percent of the absence data used for validation. Compared with the first version of this model, the final model has similar sensitivity and specificity but a lower TSS score, which led to a less certain model, but with a lesser change of overpredicting *B. frigida* habitat probability.

Relative probability values from 137 locations where the vegetation was sampled was used to calculate mean predicted values in locations with and without grasshoppers. This final species distribution model had an average score in plots without grasshoppers of 0.25 (n = 92) and in plots with grasshoppers 0.72 (n = 45), which shows that the final model was more accurate in predicting unoccupied areas than the first version of this model (Figure 8).

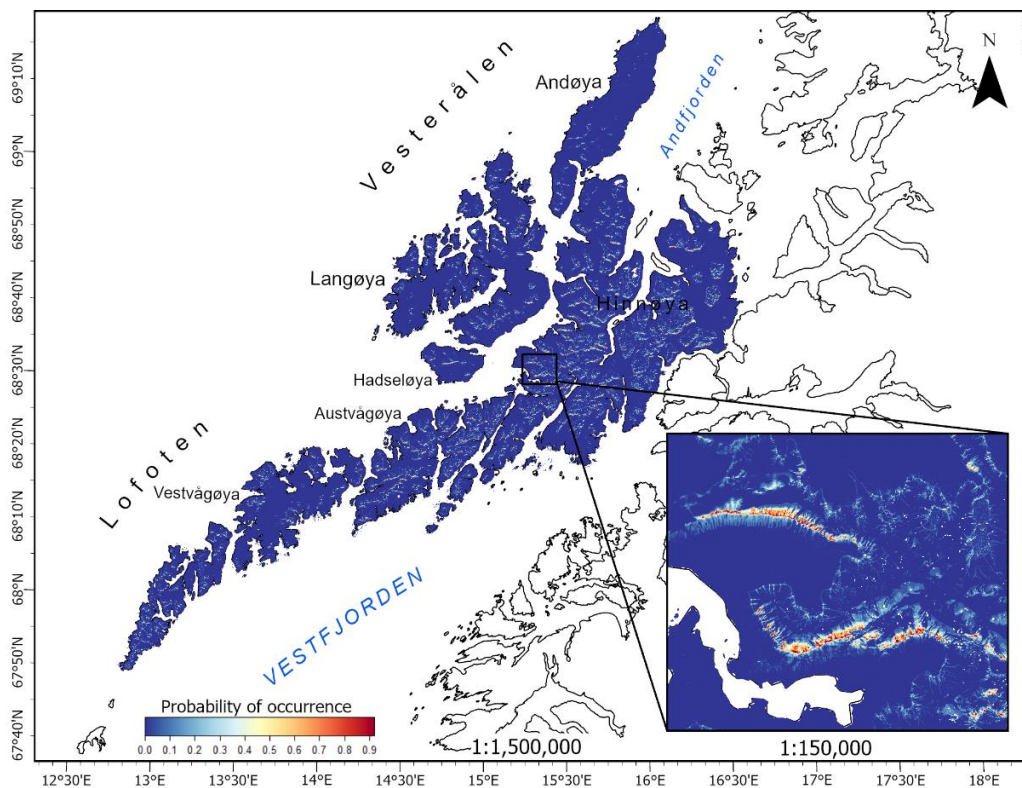


Figure 8 The final map created with the second species distribution model, where high probability of *B. frigida* habitats are marked with red colour, and low probability are blue. A sample area where the scale is ten times smaller to visualize details is overlaid in the right bottom corner. The model used to create this map used 45 observations of *B. frigida* from six different locations.

The importance of the explanatory variables was similar as for the first version of this SDM, aspect and slope had changed places, indicating that aspect was more important than slope (Table 4). The response curves widened in the final model, where *B. frigida* probability of occurrence now had a more extensive range of potential habitats (Figure 9). Elevation now had a high probability between 350-550 meters, but with a peak around 420 meters (Figure 9). The model predicted high probability in southern faced slopes where the angle was greater than 25 degrees.

Table 4 Explanatory variables used in the final species distribution model, ranked by importance.

Rank	Variable	Min	1 <sup>st</sup> Qtl	Mean	3 <sup>rd</sup> Qtl	Max	Units
1	Digital elevation model	-0.19	36.26	213.79	343.52	1080.97	m
2	Solar radiation	27252	163750	178394	201379	264819	WH/m <sup>2</sup>
3	Aspect	0.00	96.43	182.41	271.98	359.96	degree
4	Distance to water	0.00	13.67	51.33	72.09	944.54	m
5	Slope	0.00	6.50	20.41	31.61	83.21	degree
6	Normalized difference vegetation index	-0.62	0.62	0.69	0.84	0.92	-

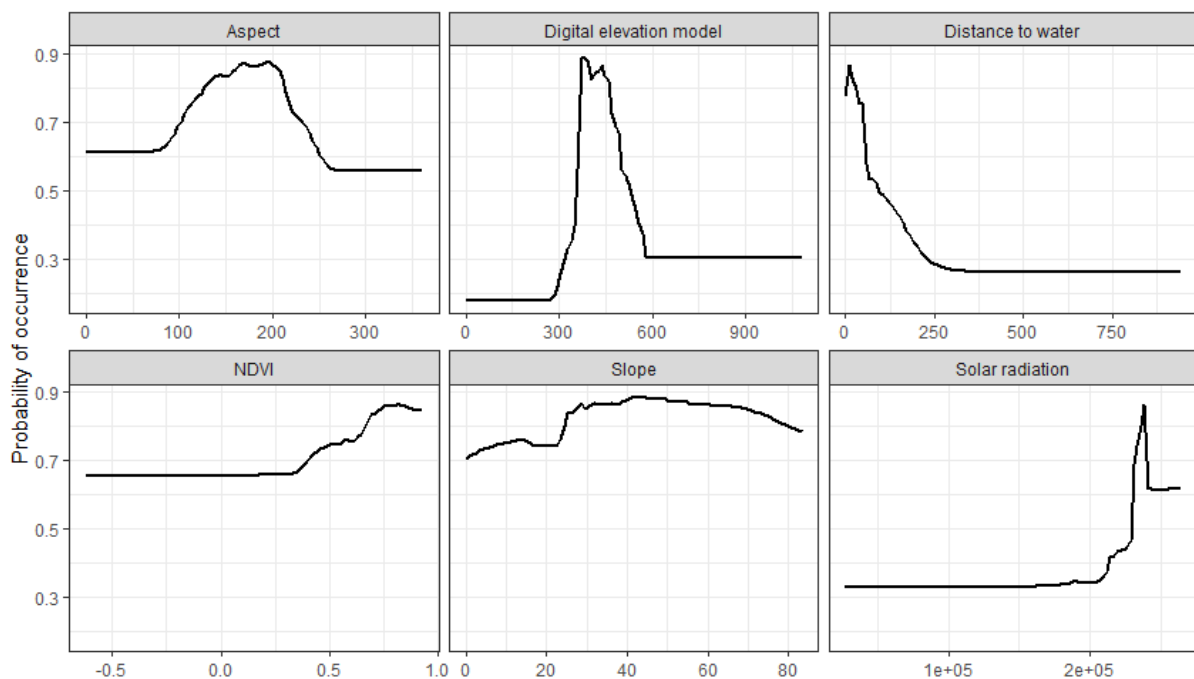


Figure 9 Response curves for the explanatory variables from the second SDM, where all observations of *B. frigida* and absences were included.

### 3.3 Plant species composition

Rare species were suppressed from the constancy analysis, where species that had a constancy score less than 0.1 were not included. *Vaccinium myrtillus* had the highest score both in plots with grasshoppers and in plots where grasshoppers were not present (Figure 10). In plots with grasshoppers, the two grasses *Avenella flexuosa* and *Anthoxanthum odoratum* were ranked as number two and three followed by the herb *Rumex acetosa* as the fourth species with the highest constancy score (Figure 10). Of the 39 species with a constancy value greater than 0.1 in plots where grasshoppers were present, 19 of the species were herbs and ferns, 7 were woody shrubs, 6 were grasses and 7 were lichens, mosses and liverworts (Figure 10).

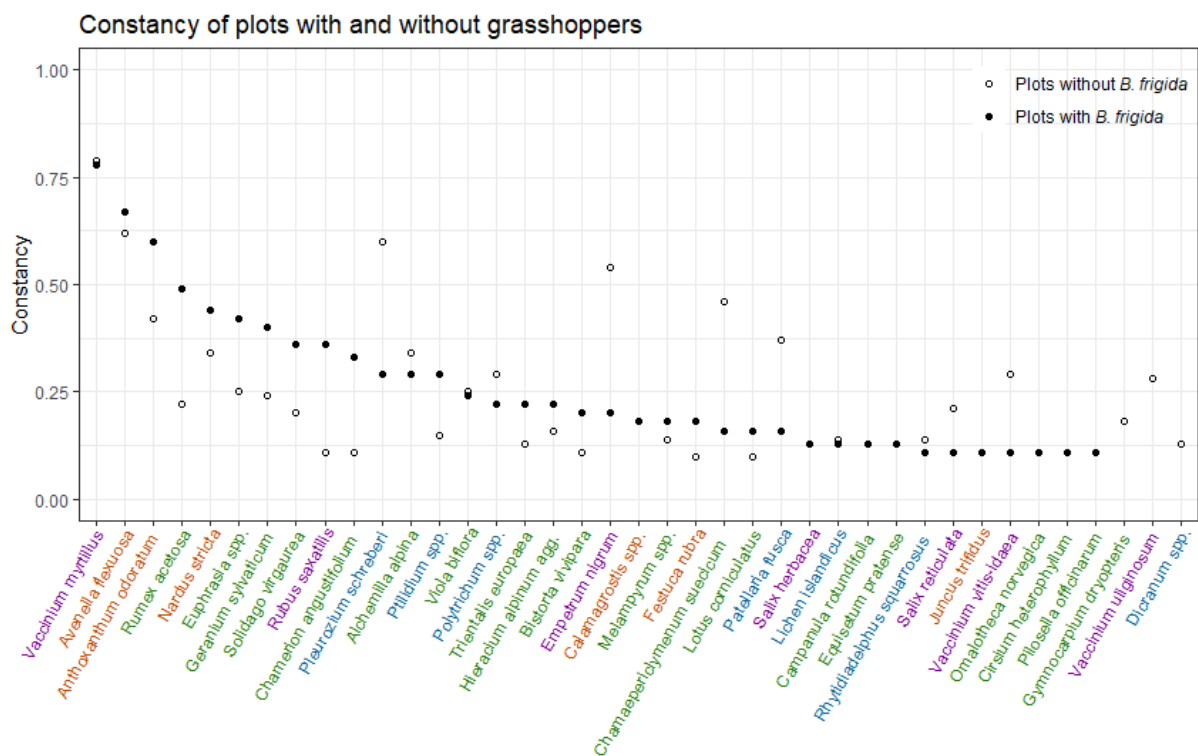


Figure 10 Constancy values for plants sampled in plots with and without *B. Frigida* present. Purple species are woody shrubs, brown species are grasses, green species are herbs and ferns, and blue species are lichens, mosses and liverworts.

The detrended correspondence analysis shows the distribution of the floristic data along the two DCA axis. Eigenvalues were 0.623 for axis one and 0.316 for axis two. Environmental variables were fitted to the plot and can be seen as arrows in Figure 11, these explanatory variables were added to the DCA plot, and were not included in the DCA analysis, but it is possible to see how they relate to the DCA analysis. The longest arrow of the explanatory variable was slope, followed by solar radiation and height above sea level (DTM). Most of the plots with *B. frigida* are located on the left side, which could imply an association with slope.

On the left side of axis one in the DCA analysis, in Figure 12, the floristic data are grouped into four groups and most of the grasses together with the herbs and ferns can be seen on the left side of axis one together with most of the *B. frigida* observations which could imply an association between herbs and ferns, and *B. frigida*. Most of the woody shrubs and mosses, lichens and liverworts are on the right side of axis one with some of the *B. frigida* observations, indicating that *B. frigida* sometimes are found in this type of vegetation.

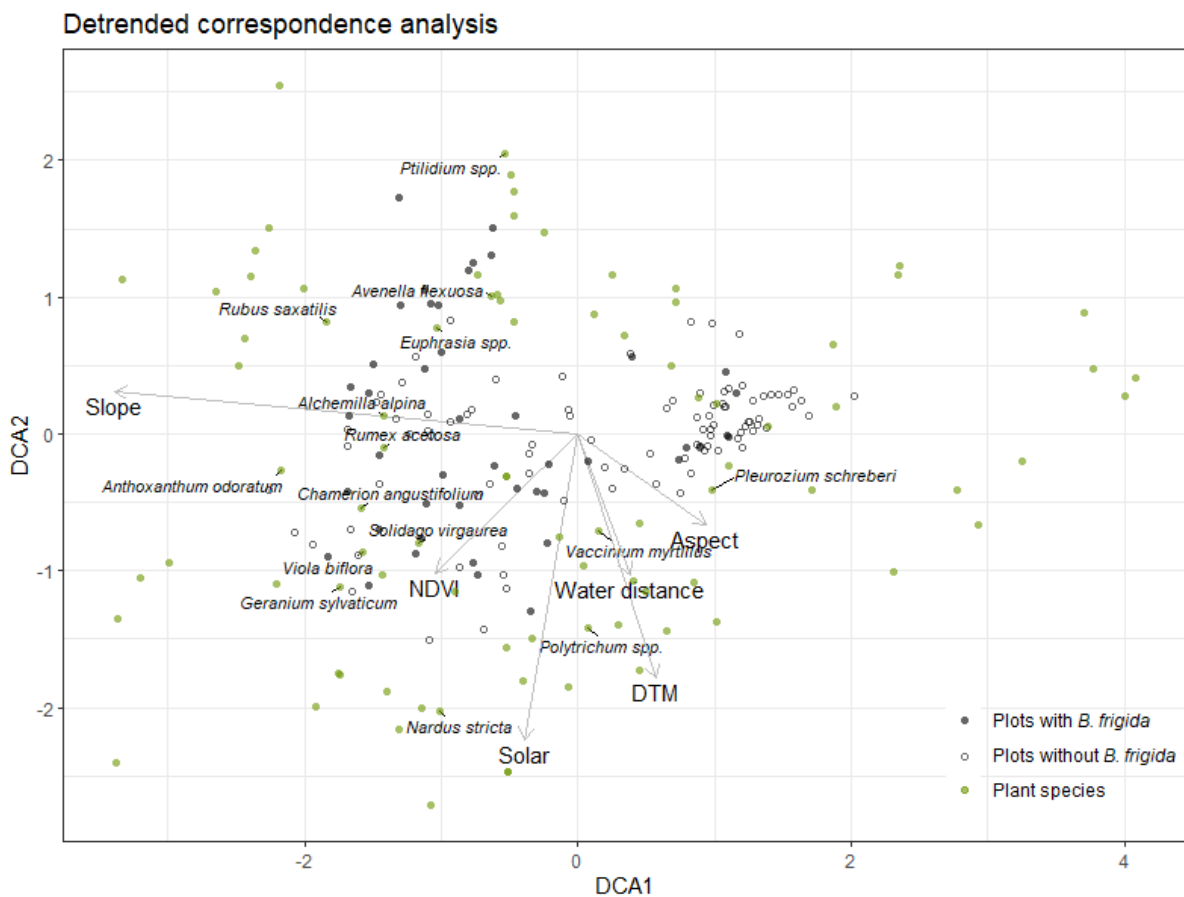


Figure 11 Detrended correspondence analysis (DCA) of the floristic data sampled from 137 plots within the study area. The 15 plant species with the highest constancy value according to the constancy analysis have their Latin name added to the plot. Explanatory variables are fitted to this plot in the form of arrows are not a part of the DCA analysis.

## Detrended correspondence analysis

Plotted by plant species

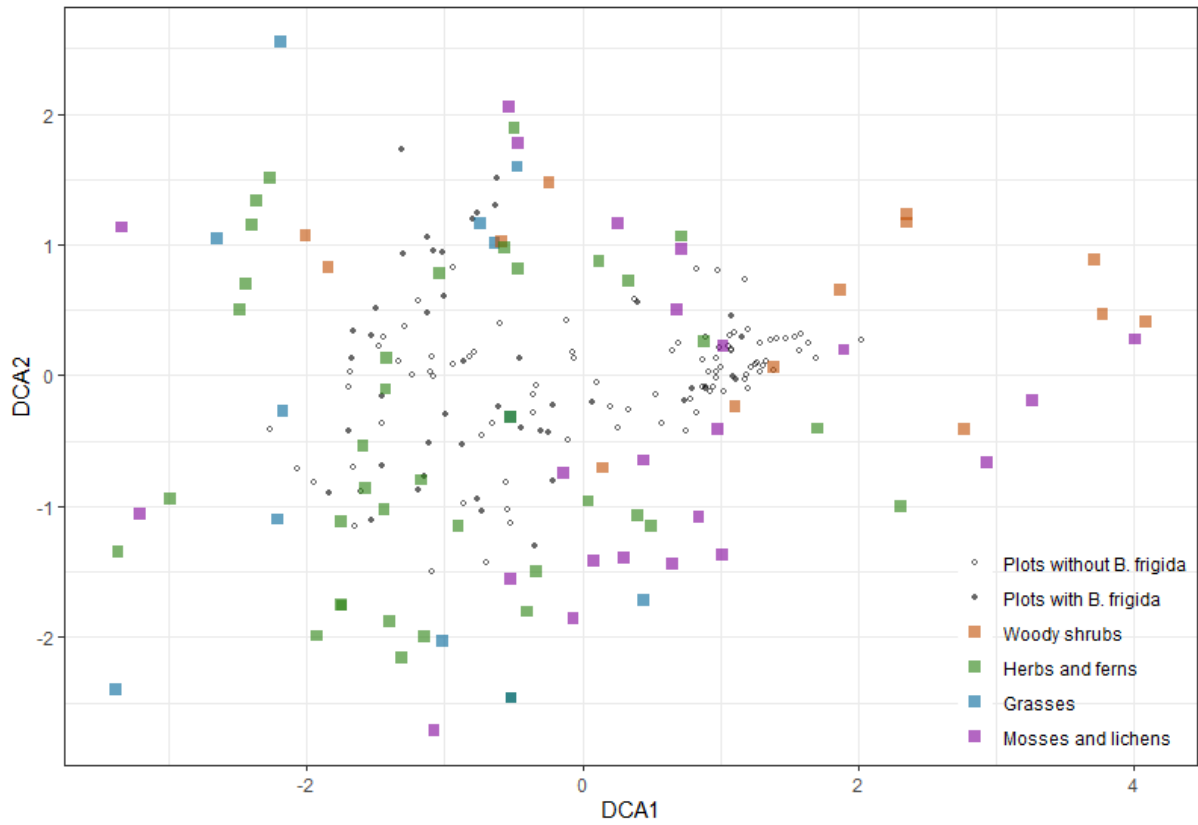


Figure 12 Detrended correspondence analysis (DCA) of the floristic data sampled from 137 plots within the study area. All the floristic data are grouped into four groups: woody shrubs which include species like *Calluna vulgaris* and *Phyllodoce caerulea*, herbs and ferns, grasses and mosses and lichens which includes liverworts.



### 3.4 Explanatory variables impact on the occurrence of the grasshoppers

Analysing the explanatory variables, vegetation in form of DCA axis one, NDVI, elevation, aspect, distance to water, solar radiation, slope and type of weather in a generalized linear mixed-effects model showed that slope had the highest effect on *B. frigida* occurrence (Table 5). The best model according to AICc score and weight, was the model that only included aspect and slope (

Table 6). In this model, slope had a p-value of 0.005, while aspect had a p-value of 0.1577, the dispersion for the model was 0.52. Diagnostic plots for the best model is presented in appendix B.

Table 5 The six top models ranked by AICc from the dredge function. The initial model was build with the following line: `glmer(Gresshopper ~ DCA + ndvi+ dtm + aspect + water_dist + Weather + solar + slope +(1|Location), data = dat.scaled , family = binomial(link = "logit"),na.action = na.fail, control=glmerControl(optimizer="bobyqa", optCtrl = list(maxfun = 100000)))`

Intercept	Aspect	DCA	Slope	Water_dist	Weather	df	logLik	AICc	delta	weight
-7.933	-1.225	.	3.134	.	+	6	-30.786	74.3	0	0.210
-6.118	-1.324	-0.9436	2.725	.	+	7	-29.674	74.4	0.04	0.206
-7.008	.	.	2.597	.	+	5	-32.021	74.6	0.25	0.186
-8.397	-1.394	.	3.441	-0.9622	+	7	-30.041	75.1	0.77	0.143
-5.414	.	-0.7957	2.275	.	+	6	-31.203	75.2	0.84	0.139
-6.511	.	.	2.637	-0.7926	+	6	-31.380	75.5	1.19	0.116

Table 6 The best performing model according to AICc.

<i>Dependent variable:</i>	
Grasshopper	
Aspect	-1.117 (0.839)
Slope	2.912*** (1.047)
Weatherclear	5.196* (2.744)
Weatherrain	-0.074 (1.989)
Constant	-10.459** (4.314)
Observations	137
Log Likelihood	-32.060
Akaike Inf. Crit.	76.119
Bayesian Inf. Crit.	93.639

Note: \*p<0.1; \*\*p<0.05; \*\*\*p<0.01

### 3.5 Environmental data

#### Temperature

The data collected from the sensors consisted of 25,935 observations, where each observation includes one temperature reading, rain sensor values and three moisture sensor values. Every sensor system recorded its values every 59th minute. The temperature was measured in Celsius and the rain and the soil moisture sensors measured wetness on an index from 0-1024, where 0 is wet and 1024 is dry. The measurements for the soil moisture were converted into percentages. In total, 25 sensor boxes were placed in six locations without grasshoppers and five locations with grasshoppers (Figure 1). One of the sensors was destroyed by an animal, and another stopped working during the summer, data from these were included in the analysis until they stopped working.

Mean temperature for plots with *B. frigida* in the time period 3rd of July 2019 to 20th of September 2019 was 12.49°C (Figure 13), in sites without *B. frigida* the mean temperature was 11.45°C. The difference was significantly different between the two groups during the period (p-value < 0.001, V = 2511).

Median temperature for plots with *B. frigida* in the time period 3rd of July 2019 to 20th of September 2019 was 11.12°C (Figure 13), in sites without *B. frigida* the mean temperature was 10.47°C. The difference was significantly different between the two groups during the period (p-value < 0.001, V = 2356).

The maximum temperature was 57.94°C (Figure 13) in plots with *B. frigida*, and 48.19°C in plots without *B. frigida*. The difference was significantly different between the two groups during the period (p-value < 0.001, V = 2399).

The minimum temperature was -0.25°C (Figure 13) in plots with *B. frigida*, and -1.19°C in plots without *B. frigida*. The difference was not significantly different between the two groups during the period (p-value = 0.1297, V = 1584).

The differences in temperature implies a preference for warmer habitats for *B. frigida*, both the mean and maximum temperatures were significantly higher. The minimum temperature was not significantly different between the plots with and without *B. frigida*, which could be explained by lower and similar temperatures in the night and higher temperatures during the daytime.

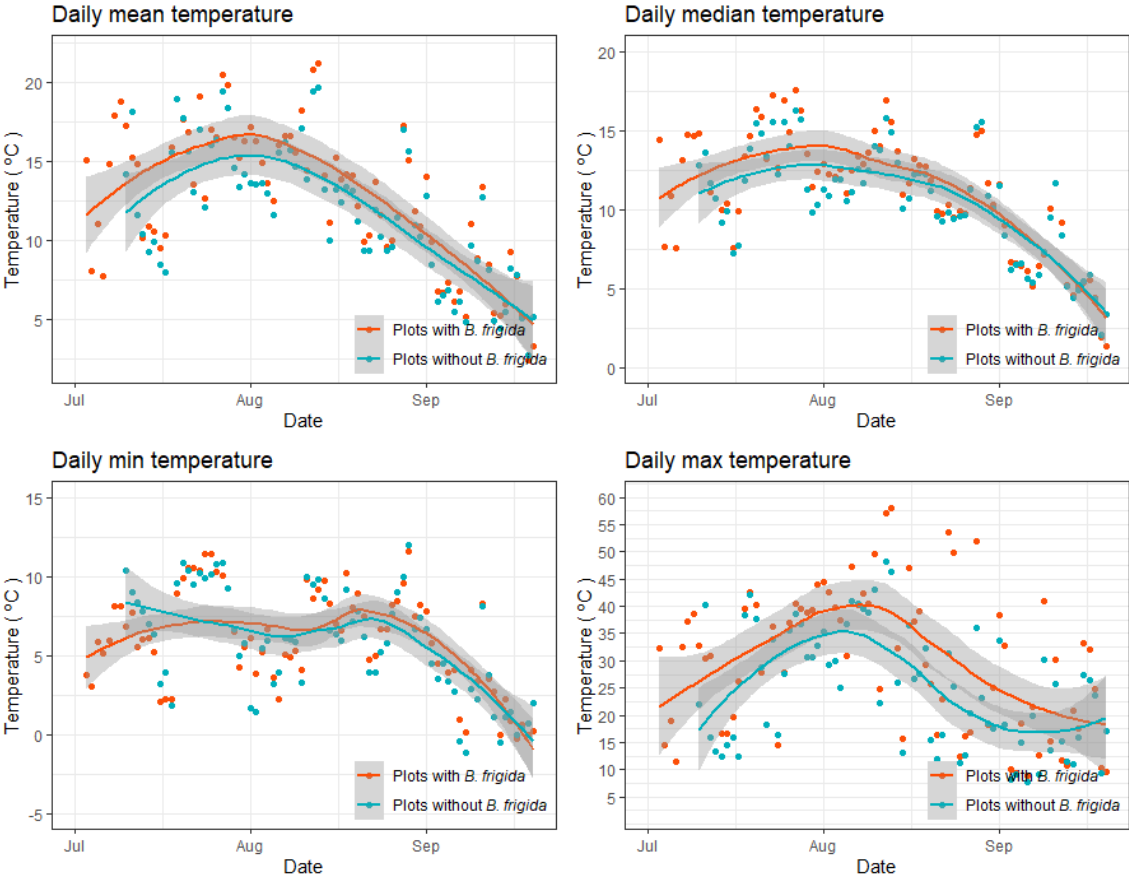


Figure 13 Daily mean, median, minimum and maximum temperature readings from 25 sensors stations in six confirmed *B. frigida* locations and five empty locations from 3rd of July 2019 to 20th of September 2019 in Vesterålen and Lofoten.

## Soil moisture

As for the temperature data, the soil moisture data consisted of 25,935 separate observations of moisture values, three moisture readings were collected from each sensor box and an average of these three readings were used in the analysis.

Mean soil moisture for plots with *B. frigida* in the time period 3rd of July 2019 to 20th of September 2019 was 23.07%, in sites without *B. frigida* the mean soil moisture was 28.38% (Figure 14). The difference was significantly different between the two groups during the period (p-value < 0.001, V = 116).

Median soil moisture for plots with *B. frigida* in the time period 3rd of July 2019 to 20th of September 2019 was 21.26%, in sites without *B. frigida*, the mean soil moisture was 28.99% (Figure 14). The difference was significantly different between the two groups during the period (p-value < 0.001, V = 43.5).

The maximum soil moisture was 47.36% in plots with *B. frigida*, and 64.23% in plots without *B. frigida* (Figure 14). The difference was not significantly different between the two groups during the period (p-value = 0.291, V = 1543).

The minimum soil moisture was 4.99% in plots with *B. frigida*, and 10.77% in plots without *B. frigida* (Figure 14). The difference was significantly different between the two groups during the period (p-value = 0.003, V = 822).

Soil moisture decreased in all plots during the summer until mid-august, when it started increasing again. The mean soil moisture was significantly lower in plots with *B. frigida* as was the minimum soil moisture. The soil moisture seemed to follow the trend of the temperatures, where the temperature increased until mid-august, the soil moisture decreased until mid-august. *B. frigida* seems to have a correlation with drier habitats compared to unoccupied high probability locations.

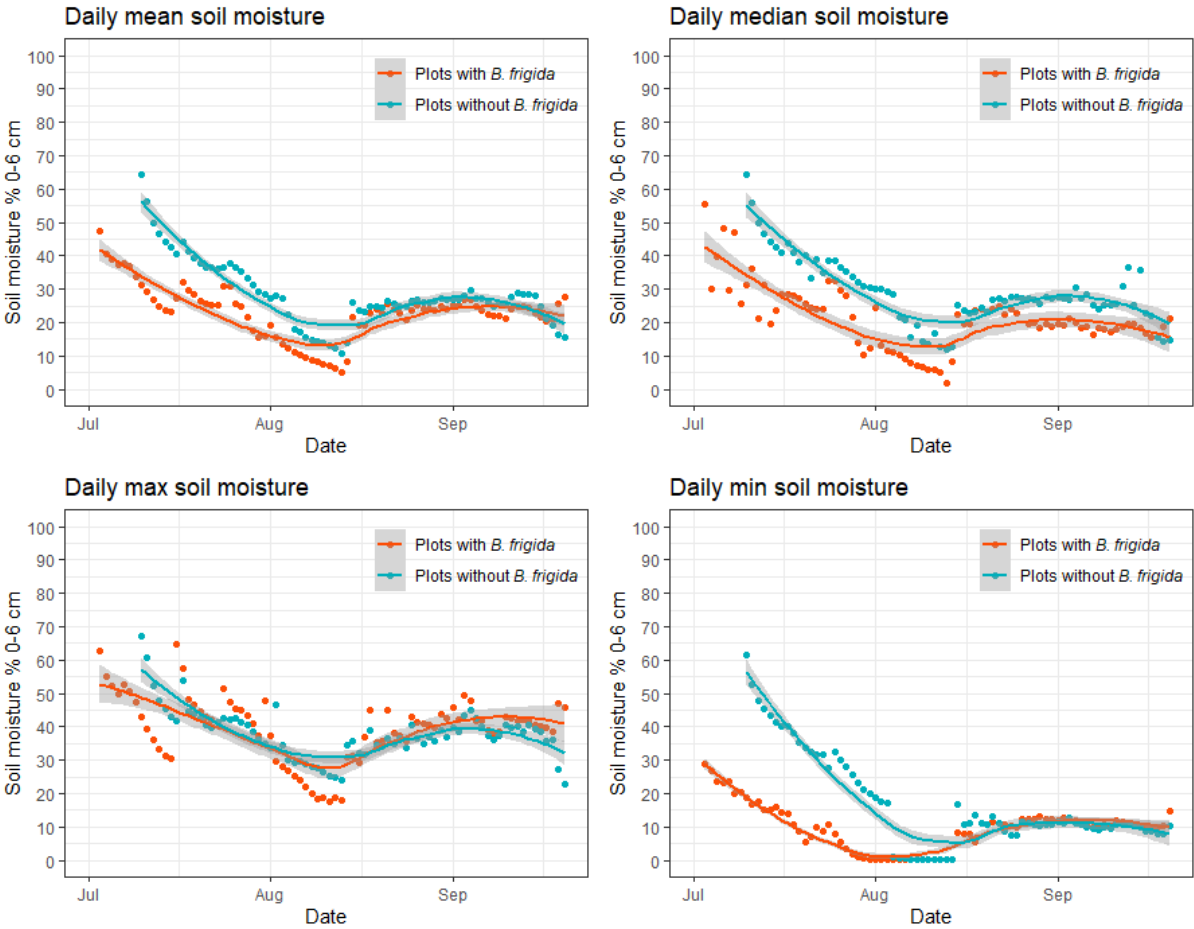


Figure 14 Soil moisture percentages for the period 3<sup>rd</sup> of July 2019 to 20<sup>th</sup> of September 2019 in locations with and without *B. frigida* located in Vesterålen and Lofoten.

## Precipitation

The rain sensors also collected 25.935 data points, of these 14.580 were collected in sites with grasshoppers and 11.355 were collected in sites without grasshoppers.

Mean rain values for plots with *B. frigida* in the time period 3<sup>rd</sup> of July 2019 to 20<sup>th</sup> of September 2019 were 973.71, in sites without *B. frigida* the mean rain values was 973.92 (Figure 15). The difference was not significantly different between the two groups during the period (p-value = 0.578, V = 1249).

Median rain values for plots with *B. frigida* in the time period 3<sup>rd</sup> of July 2019 to 20<sup>th</sup> of September 2019 were 993.93, in sites without *B. frigida* the mean rain values was 999.08 (Figure 15). The difference was not significantly different between the two groups during the period (p-value = 0.345, V = 1328).

The maximum rain values were both 1023 in plots with *B. frigida*, and in plots without *B. frigida* (Figure 15). The difference was not significantly different between the two groups during the period (p-value = 0.116, V = 88).

The minimum rain values were 11.00 in plots with *B. frigida*, and 347.15 in plots without *B. frigida* (Figure 15). The difference was significantly different between the two groups during the period (p-value = 0.027, V = 948.5).

The rain sensors measured raindrops and not the volume of rain and lower values from the rain sensor is an indication of more precipitation. Only the minimum rain values were significantly different in sites with and without *B. frigida*, which means locations with *B. frigida* received more precipitation than locations without *B. frigida* at the same time period as temperatures were rising, which implies a wetter habitat even with warmer temperatures in *B. frigida* habitats. The daily max rain values were all in the high range of values, this is probably caused by the sensors taking a measurement sometime during a day when there was no precipitation.

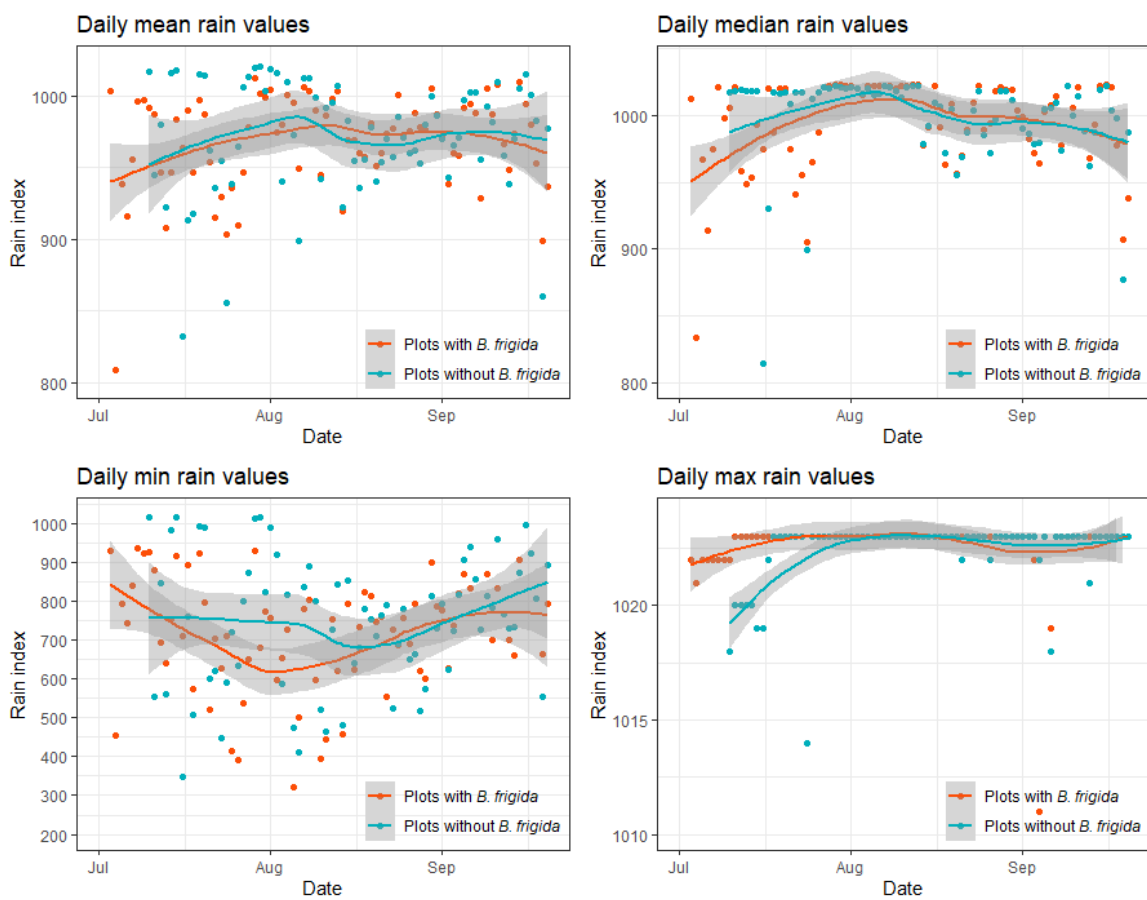


Figure 15 Daily rain values collected from 25 sensors in the study area from 3rd of July 2019 to 20th of September 2019. A higher value on the rain index indicates a drier rain sensor, and lower values indicates more rain on the rain sensor. The daily minimum rain values show that sites with *B. frigida* received more rain than sites without *B. frigida*.

### 3.6 Morphological features

Pictures were taken of 80 grasshoppers in six different locations between 3rd of July and 17th of September, of these 50 were identified as male and 30 as female. Body length was measured for all 80 individuals; Maximum body lengths were 3.836 cm for females and 2.427 cm for males, the females were significantly larger than the males (p-value < 0.001, W = 1232). The minimum body length for females was 1.654 cm and 0.887 for males, with a mean of 2.885 cm for females and 2.009 cm for males.

The tegmen were measured on 26 of the females and 24 of the males and had a range of 0.351-1.047 cm for females and 0.277-0.819 cm for males (Table 7), the difference between females and males was significant (p-value < 0.005, W = 1232). No macropterous grasshoppers were found. Protonum was measured on 29 of the females and 29 of the males and had a length of 0.367-0.820 cm for females and 0.305-0.529 for males (Table 7), the difference between females and males were significant (p-value < 0.001, W = 743). The measurements, with maximum values from the literature is presented in Figure 16.

With a separate test using a tape measurement band and the same method which was applied to the grasshoppers, the standard deviation was measured to be 0.0166 cm (n=10).

Table 7 Mean, minimum, standard deviation and maximum values for three morphological features measured on 80 *B. frigida* grasshoppers within the study area in Vesterålen and Lofoten. The values are in centimetres.

Sex	n	Body length				Tegmen				Protonum			
		Min	Mean	Sd	Max	Min	Mean	Sd	Max	Min	Mean	Sd	Max
F	30	1.654	2.885	0.675	3.836	0.351	0.770	0.239	1.047	0.367	0.609	0.108	0.820
M	50	0.887	2.009	0.353	2.427	0.277	0.607	0.143	0.819	0.305	0.451	0.063	0.529
All	80	0.887	2.338	0.653	3.836	0.277	0.692	0.213	1.047	0.305	0.530	0.119	0.820



### Morphological features measured on *B. frigida*

With expected maximum values from literature

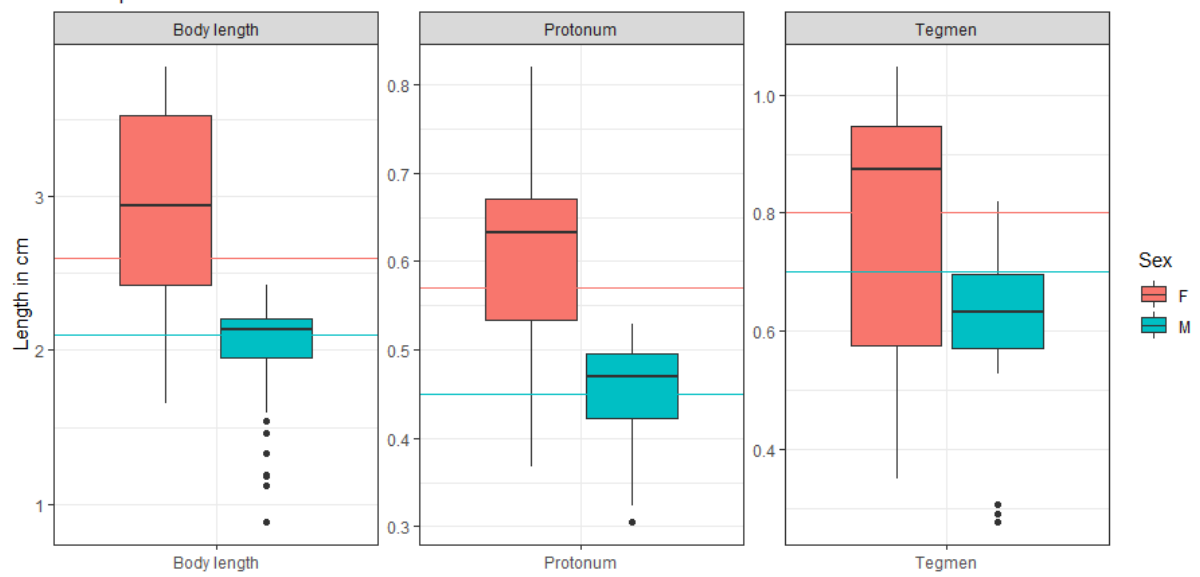


Figure 16 Mean, minimum and maximum values for three morphological features (body length, protonum and tegmen) measured on 30 *B. frigida* females and 50 *B. frigida* males. The red and blue horizontal lines are maximum values from the literature (Harz, 1975; Holst, 1986).

## 4 Discussion

The aim of the research was to explore the habitats of *B. frigida* at its northern edge range in Vesterålen and Lofoten. A different group of plants in the only known grasshopper site within the study area seemed to be of importance compared to previous research. Another aspect of the research was to see if potential grasshopper habitats could be predicted using species distribution modelling where slope, aspect and elevation were among the predictor variables. In trying to quantify the realized niche for the *B. frigida* at the boundary of their northern range, all collected variables were tested to see which one was most significant. Microclimate was measured to see if the lack of grasshoppers in unoccupied high probability areas could explain their absence. Three morphological features were measured to see if there were any differences between *B. frigida* populations in the study area and the literature.

### 4.1 Distribution models

The first species distribution model overpredicted grasshopper probability of occurrence, where larger areas in mountainous terrain had a high probability. When arriving in one of these areas, it was clear that it would be challenging to find the exact location of a grasshopper habitat, like searching for the famous needle in the haystack. Six locations (Appendix D through I) where this model predicted a high probability were visited before concluding another approach or model was needed.

The only location known from publicly available records within the study area (Appendix J), reported in 1946, was visited and *B. frigida* were found after a day of searching. This place was very similar to the other known site within the study area (Appendix C) with regards to a very steep slope, the vegetation, that water was close by and that the grasshoppers seemed to be restricted to a small area. This area was too small to be correctly detected with an SDM with a spatial resolution of 30 x 30 meters. For example, the slope in degrees would not reflect the real steepness of the area but average out over 30 x 30 meters. In location number 8 (Figure 1), where one of the sensors was placed, the slope value with a spatial resolution of 30 x 30 meters was 17.5 degrees, while with a resolution of 1 x 1 meter, the slope was 47.7 degrees. To incorporate these smaller habitats, a new model with a smaller spatial resolution was required.

The second model used a spatial resolution of 1 x 1 meters for the topographical variables, and resampled solar radiation and NDVI variables down to 1 x 1 meters. The NDVI variable for the second SDM was generated with another program than the NDVI variable for the first SDM, this was due to size restrictions when using google earth engine. This SDM used only the grasshopper locations found within the study area, which had a precision better than 5 meters. To begin with, the model only had 13 presence locations and 36 true absence locations. This led to a model which picked out tightly limited areas with a narrow band of environmental conditions and the model evaluation had an unusually high TSS and sensitivity score. The data in the second SDM used 70 percent of the data for prediction and 30 percent for validation, while the first SDM used an 80/20 split of the data, this different split was due to few observations, which caused the model to fail with an 80/20 split. Despite the unusual high TSS and sensitivity score, using the map (Figure 6) this model made, five new locations (Appendix K, Appendix L, Appendix M, Appendix N and Appendix S) on Hinnøya were searched and four of these had grasshoppers. As more observations were added to the model, the narrow band of environmental conditions widened and the TSS, sensitivity and specificity score was lowered.

The model predicted high probability on other islands within the study area and two locations on Austvågøya (Appendix N and Appendix Q), one on Hadseløya (Appendix O), and one location on Langøya (Appendix P) were searched, but no grasshoppers were found on these islands. With island biogeography in mind, *B. frigida* habitats could perhaps be considered as an island, where they are separated by valleys, rivers and lakes, which could be difficult for a grasshopper to traverse. However, MacArthur and Wilson (1967, pp. 114-115) argue in their book that mountain tops cannot be considered as an island in the sense of island biogeography because the space separating them is not barren of competitors. This could however be the case for why grasshoppers were found on only one of the four islands which were examined, were immigrants must come to a new island often enough to maintain a population (MacArthur & Wilson, 1967, p. 115).

No new grasshopper locations were found while using the first SDM as a reference to where to search; the input was *B. frigida* location from all of Norway with a precision less than 100 meters. This, together with a spatial resolution of 30 meters might have introduced a bias since coarser resolutions blur the environmental variability contained in higher resolution variables (Guisan et al., 2017, pp. 136-140). The map created from this SDM overpredicted

possible grasshopper habitats where high probability areas were several square kilometres, making it close to impossible to find the assumed small grasshopper habitats.

Another approach that was not used would be to use the *B. frigida* locations from all of Norway with the smaller 1 x 1 meter resolution, but with the method used, this would require several weeks of processing before a prediction map was created since information from the entire study area in form of pseudo-absence data would be required (Guisan et al., 2017, p. 131).

The second and evolving SDM had no variables to differentiate between different climate conditions, especially westward into Lofoten, which if included could potentially have lowered the predicted habitability in these areas compared to Hinnøya.

#### **4.2 Plant communities**

48.7 percent of the species from the constancy analysis (Figure 10) where herbs and ferns, woody shrubs and lichens, mosses and liverworts accounted for 17.9 percent of the species, and 15.3 percent of the species was grasses. This indicated that herbs and ferns are abundant in *B. frigida* within the study area. *Vaccinium myrtillus* had the highest constancy score in plots where grasshoppers were both present and absent, meaning they are the most dominant plant in all plots regardless if grasshoppers are there or not. Other research found that *Vaccinium myrtillus* and *Calluna vulgaris* was the dominating plants in *B. frigida* habitats and only four different herbs were included in their plant species composition analysis (Finch et al., 2008). In this study *Calluna vulgaris* was not a dominating species in either type of habitat and herbs and ferns were the most abundant plant group in grasshopper habitats.

The detrended correspondence analysis, showed that slope followed by solar radiation is the most essential gradients to explain structure in the plant species community.

#### **4.3 Explanatory variables impact on the occurrence of the grasshoppers**

Possible explanatory variables were tested to check the importance of these with regards to *B. frigida* habitat preferences. The best model included slope and aspect as the most influential variables, were only slope had a p-value lower than 0.05. Two of the best model by AICc

score included DCA (vegetation) and distance to water, which could imply that these variables has an impact in *B. frigida* habitat selection.

#### **4.4 Environmental data**

The temperature in sites with grasshoppers were, on average one degree warmer than sites without grasshoppers. The average minimum temperature was quite close in the different sites, this was properly due to the temperature during nights were similar, whereas in the day the sun was more prevalent in sites where grasshoppers were present. This can also be seen in the daily maximum temperatures, where temperatures in plots with grasshoppers were consistently higher compared with plots without grasshoppers. The highest temperature measured was 57.94°C, this was in location 8 (Figure 1), these unusually high measurements are thought to be caused by direct sunlight onto the temperature sensor. At the same day and time as this maximum temperature was recorded, another location (location 12, figure 1) also measured temperatures above 50°C.

The sites with grasshoppers were drier than sites without grasshoppers, this difference became smaller at the end of the season, when the daily soil moisture in sites with grasshoppers surpassed the daily mean soil moisture in sites without grasshoppers. The reason for the grasshopper's sites being drier, is probably an effect of these sites being warmer and getting more solar radiation, which led to the ground drying up. Rain values were similar in sites with and without grasshoppers except for the daily minimum rain values which was significantly different, which could imply more precipitation in *B. frigida* habitats. Another explanation could be that these locations have more moisture in the air which condensed onto the rain sensors.

The solar radiation that reaches the earth (insolation) and the thinner air at higher altitudes in combinations with diverse local conditions differentiates microclimatic conditions in alpine regions from the macroclimatic conditions found over larger mountainous areas. The mean atmospheric temperatures can be low during the summer at high altitudes, but rapid and intense insolation accounts for higher temperatures in most microclimatic niches of the high altitude insects (Mani, 1968, pp. 32-33). The higher temperatures found in locations with *B. frigida* could be an indication of a microclimate, where high temperature together with lower soil moisture and more precipitation make these locations suitable for them. The right temperature and soil moisture content is proven to be essential for egg development in other

species in the genus *Melanoplus* (Mukerji & Gage, 1978). Previous research found soil moisture values in *B. frigida* habitats to be relatively low (20%-40%) in both spring and autumn (Finch et al., 2008). These values were similar to soil moisture values measured in *B. frigida* habitats within the study area, although no direct egg depositions were found, this is an indication that 20%-40% soil moisture values is the favourable soil moisture content for *B. frigida*. In locations without grasshoppers, the soil moisture was higher in the early summer and together with lower temperatures could be an explanation of why the species distribution models predicted high probability, but no grasshoppers were present: there were microclimatic conditions affecting *B. frigida* occurrence that were not included in the species distribution models.

The temperature probes used in this study were laying in between the vegetation, whereas Finch et al. (2008) had probes at +15, -1 and -15 centimetres above and below the ground. As the temperature difference can be extremely different in the sun and in the shade in high altitudes (Mani, 1968, pp. 20-21), a uniform location of a temperature probe should be in place, where the sun cannot directly affect the readings as seen with temperatures reaching nearly 60°C in this study.

#### ***4.5 Morphological features***

Three different parts of the grasshoppers were measured, the body length, the tegmen and the pronotum, the reason for measuring these parts were that they could easily be measured from photos after the field work was complete. Body length in females (mean 2.8 cm, maximum 3.8 cm) was 47.5 percent larger than what (Harz, 1975; Holst, 1986) has reported (1.6-2.6 cm). For males (Harz, 1975; Holst, 1986) reported 1.6-2.1 cm while the measured maximum body length was 2.427 cm, 15 percent longer than reported. The minimum body length measured was 0.887 cm, which could indicate that some individuals were still in their nymph stage of life when measured.

The other measurements were consistently larger than those reported in the literature. Female tegmen were 30 percent larger, and male tegmen 17 percent larger, while female pronotum were 44 percent larger and male pronotum 17 percent larger, compared with the values from Harz (1975). To ensure that the measurements were correct, the same method was used on a tape measurement band, where the standard deviation was found to less than one millimetre.

Interestingly Harz (1975, p. 248) said that the individuals with the smallest size he measured came from Norway at around 500 meters, which is quite opposite to what was found in Vesterålen and Lofoten at roughly the same elevation. An explanation for this increase in size is presented in three different theories:

Theory 1: *B. frigida* habitats near their northern range boundary have microclimatic conditions which make them warmer than other locations:

*B. frigida* habitats in the study area were at a lower altitude (average 420 meters), than most other known *B. frigida* locations in Europe (Harz, 1975). Reduction in body size as altitude increases has been seen in other high altitude insects such as the *Amara*, *Nebria* and *Trechus* beetles (Mani, 1968, p. 58). The special microclimatic conditions where *B. frigida* has its habitats within the study area are warmer which leads to a larger body size.

Theory 2: Vesterålen and Lofoten is further north than other *B. frigida* locations, so it's colder:

The increase in size in *B. frigida* could perhaps be explained by Bergmann's rule, which states that within a broadly distributed taxonomic clade, a species size will increase at higher latitudes or altitudes (Bergmann, 1848). Although this rule is mainly applied to species which are endothermic, some researchers have found evidence for this rule in ectothermic species (Olalla-Tárraga, Rodríguez, & Hawkins, 2006; Shelomi, 2012). More research would be needed in order to confirm if Bergmann's rule applies to *B. frigida*, where temperatures would be measured with the same method over several years within different latitudes.

Theory 3: The *B. frigida* populations on Hinnøya are separated from other populations.

This suggested separation restricts immigration from other populations following the principals of island biogeography by (MacArthur & Wilson, 1967), in which *B. frigida* on Hinnøya has been isolated from populations on the main land and created a metapopulation.

#### **4.6 Future prospects**

To further improve the species distribution models and discover new *B. frigida* locations in within the study area and other places, climatic variables such as the bioclim variables at a lower spatial resolution than currently available and explanatory variables which include soil moisture together with different vegetation indices with a better spatial resolution than data from the Sentinel-2 satellite provides, could improve the predictions in future species

distribution models. More measurements of the morphological features of *B. frigida* over several years in Vesterålen and Lofoten to see if the increase in size is consistent and not just a one-year fluke because of an unusually warm period. To further quantify the niche of *B. frigida*, measurements of the size of their habitat and the density in these habitats could help to model their niche.

*B. frigida* eggs are laid below the surface in the autumn and hibernate through the winter until the snowmelt (Finch et al., 2008). Although the temperature may drop significantly below zero degrees Celsius, the soil beneath the snow do not freeze, the presence of heavy snow-cover ensures optimal conditions for hibernation for high altitude insects (Mani, 1968) p.71. With climate change predicting warmer temperatures, more precipitation in form of rain and fewer days with snow cover (Hanssen-Bauer et al., 2009), this could potentially be a threat for the survival of *B. frigida*.



## 4.7 Conclusion

### 1. Can new *B. frigida* locations be found with a species distribution model?

Species distribution models can be suitable to map possible *B. frigida* habitats, the habitats within the study area are quite small and models which use a low spatial resolution are needed to detect them accurately. *B. frigida* has been known to be a widespread species in mountainous areas (Holst, 1986; Nilssen & Lundmo, 2013), but they do seem to be more restricted at the northern edge of their range boundary, which makes species distribution models with the right spatial resolution and the right explanatory variables suitable to map potential *B. frigida* habitats.

### 2. Is *B. frigida* restricted to a particular habitat at its northern range margin compared with habitats at the range centre?

*B. frigida* at the boundary of its northern range is found at lower altitudes than most other places in Norway (Finch et al., 2008; Global Biodiversity Information Facility, 2019; Harz, 1975). Elevation was the third most important explanatory variable in the first species distribution model and the most important explanatory variable in the final species distribution model. Vegetation was not found to be significant for the distribution of *B. frigida* in any of the analyses, *Vaccinium myrtillus* was the dominating plant within the study area both in locations with and without grasshoppers, with a slightly larger constancy score in sites without grasshoppers. *Calluna vulgaris* was not an important plant in *B. frigida* habitats as reported by Finch et al. (2008) in the continental region of Norway. Herbs and ferns were abundant in *B. frigida* habitats compared to habitats at its range centre (Finch et al., 2008).

In the generalized linear mixed-effects model, only slope was found to be of significance, as such *B. frigida* cannot be said to be restricted to a particular habitat at its northern range margin compared with habitats at the range centre. Nevertheless, they are more often found in steep southerly faced slopes around 40 degrees with an elevation around 420 meters and a vegetation pattern dominated by *Vaccinium myrtillus* and herbs and ferns.

**3. In empty locations where the species distribution model predicts a high probability for *B. frigida*, can microclimate explain the lack of grasshoppers?**

Microclimatic conditions were warmer, and the soil was drier in locations with high probability and confirmed presence of *B. frigida* compared with high probability locations without *B. frigida*. A connection between *B. frigida* and microclimate was discussed by (Finch et al., 2008) where the link between *B. frigida* habitat preference and microclimatic conditions such as soil moisture and temperature could not be satisfactorily answered. The daily mean temperatures and daily mean soil moisture were quite different in locations with and without grasshoppers and these results imply that temperature and soil moisture could explain the lack of grasshoppers in high probability areas, but more research would be needed to establish this for certain.

**4. Are there any morphological differences in *B. frigida* at the northern edge?**

Morphological features were measured, and unexpected results in the form of larger body size, larger pronotum and larger tegmen than previously described was discovered (Harz, 1975; Holst, 1986). The reason for these morphological differences are not certain, but different theories are discussed in chapter 4.5.

## 5 Acknowledgements

First of all, I want to thank my supervisor Amy Eycott for the support during this endeavor, also thanks to co-supervisors Sam Stayert and Joseph Chipperfield.

Thanks to my late grandfather, without him, this thesis would be about something entirely else. In my younger years he showed me how to use *B. frigida* as fishing bait. Although I doubt their effectiveness as bait, a spark was ignited, which lasted some 20 years before I again could explore the world of the grasshoppers.

A shout out to Kristoffer Myrland for help with soldering, cutting a lot of wires and testing the sensors into late nights. Thank you, Christopher Reppe, for feedback and tips and thank you Knut Botten for support and feedback during these last years.

Thanks to the landowners for letting me park on their land and leaving out sensors during the 2019 season.

Finally, a huge thanks to my parents for support through my time as a student and during the fieldwork.

## References

- Al-Khaier, F. (2003). *Soil salinity detection using satellite remote sensing*. (Master's thesis). ITC,
- Allouche, O., Tsoar, A., & Kadmon, R. (2006). Assessing the accuracy of species distribution models: prevalence, kappa and the true skill statistic (TSS). *Journal of applied ecology*, *43*(6), 1223-1232.
- Banzi, M., & Shiloh, M. (2014). *Getting started with Arduino: the open source electronics prototyping platform*. Sebastopol: Maker Media, Inc.
- Barton, K. (2019). MuMIn: Multi-Model Inference. Retrieved from <https://CRAN.R-project.org/package=MuMIn>
- Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, *67*(1), 1-48.
- Bergmann, C. (1848). *Über die Verhältnisse der Wärmeökonomie der Thiere zu ihrer Grösse*. Gottinger studies.
- Bodenheimer, F. S. (1929). Studien zur Epidemiologie. Ökologie und Physiologie der afrikanischen Wanderheuschrecke: (*Schistocerca gregaria* Forsk.). *Zeitschrift für Angewandte Entomologie*, *15*(3), 435-557.
- Breiman, L. (2001). Random forests. *Machine learning*, *45*(1), 5-32.
- Burnham, K. P., & Anderson, D. R. (2004). Multimodel inference: understanding AIC and BIC in model selection. *Sociological methods & research*, *33*(2), 261-304.
- Capinera, J. L., Scott, R. D., & Walker, T. J. (2004). *Field guide to grasshoppers, crickets, and katydids of the United States*. Ithaca: Cornell University Press.
- Carbonell, C. S. (2002). *The grasshopper tribe Phaeopariini (Acridoidea: Romaleidae)*. Philadelphia: Orthopterists' Society, Academy of Natural Sciences, Dept. of Entomology.
- Chatterjee, S., & Hadi, A. S. (2015). *Regression analysis by example*: John Wiley & Sons.
- Chefaoui, R. M., & Lobo, J. M. (2008). Assessing the effects of pseudo-absences on predictive distribution model performance. *Ecological modelling*, *210*(4), 478-486.
- Dormann, C. F., Elith, J., Bacher, S., Buchmann, C., Carl, G., Carré, G., . . . Leitão, P. J. (2013). Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. *Ecography*, *36*(1), 27-46.
- Drusch, M., Del Bello, U., Carlier, S., Colin, O., Fernandez, V., Gascon, F., . . . Martimort, P. (2012). Sentinel-2: ESA's optical high-resolution mission for GMES operational services. *Remote sensing of environment*, *120*, 25-36.

- Elith, J., & Graham, C. H. (2009). Do they? How do they? WHY do they differ? On finding reasons for differing performances of species distribution models. *Ecography*, 32(1), 66-77.
- Elith, J., & Leathwick, J. R. (2009). Species distribution models: ecological explanation and prediction across space and time. *Annual review of ecology, evolution, and systematics*, 40, 677-697.
- ESRI. (2020). ArcGIS Pro 2.5.0. Redlands, CA: Environmental Systems Research Institute. Retrieved from <https://www.esri.com/en-us/arcgis/products/arcgis-pro/resources>
- Evans, J., Oakleaf, J., Cushman, S., & Theobald, D. (2014). An ArcGIS toolbox for surface gradient and geomorphometric modeling, version 2.0-0.
- Fick, S. E., & Hijmans, R. J. (2017). WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. *International journal of climatology*, 37(12), 4302-4315.
- Finch, O. D., Löffler, J., & Pape, R. (2008). Assessing the sensitivity of *Melanoplus frigidus* (Orthoptera: Acrididae) to different weather conditions: a modeling approach focussing on development times. *Insect science*, 15(2), 167-178.
- Fletcher, R., & Fortin, M.-J. (2018). *Spatial ecology and conservation modeling : applications with R*. Cham: Springer.
- Fox, J. (2015). *Applied regression analysis and generalized linear models* (3rd ed. ed.). Los Angeles: Sage.
- Fryxell, J. M., Sinclair, A. R., & Caughley, G. (2014). *Wildlife ecology, conservation, and management, third edition*. Oxford: John Wiley & Sons.
- Georges, D., & Thuiller, W. (2013). *An example of species distribution modeling with biomod2*. Retrieved from [http://search.r-project.org/library/biomod2/doc/Simple\\_species\\_modelling.pdf](http://search.r-project.org/library/biomod2/doc/Simple_species_modelling.pdf)
- Global Biodiversity Information Facility. (2019). *Bohemanella frigida* (Boheman, 1846). Retrieved from <https://www.gbif.org/species/5098111>
- Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., & Moore, R. (2017). Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote sensing of environment*, 202, 18-27.
- Green, R. H. (1979). *Sampling design and statistical methods for environmental biologists*: John Wiley & Sons.
- Guisan, A., Thuiller, W., & Zimmermann, N. E. (2017). *Habitat suitability and distribution models: with applications in R*. Cambridge: Cambridge University Press.
- Gullan, P. J., Cranston, P. S., & McInnes, K. H. (2014). *The insects: an outline of entomology* (5th ed.). Chichester: Wiley-Blackwell.

- Halbritter, A. H., De Boeck, H. J., Eycott, A. E., Reinsch, S., Robinson, D. A., Vicca, S., . . . Grünzweig, J. M. (2020). The handbook for standardized field and laboratory measurements in terrestrial climate change experiments and observational studies (ClimEx). *Methods in Ecology and Evolution*, 11(1), 22-37.
- Hansen, L. O. (2009). Norges gresshopper og gresshoppelignende insekter. Retrieved from <https://www.nhm.uio.no/fakta/zoologi/insekter/norort/>
- Hanssen-Bauer, I., Drange, H., Førland, E., Roald, L., Børsheim, K., Hisdal, H., . . . Sorteberg, A. (2009). *Climate in Norway 2100*. Retrieved from <https://www.miljodirektoratet.no/globalassets/publikasjoner/M741/M741.pdf>
- Harz, A. (1975). *The Orthoptera of Europe II*. The Hague: W. Junk Publishers.
- Hijmans, R. J., & Elith, J. (2013). *Species distribution modeling with R*. Retrieved from <https://cran.r-project.org/web/packages/dismo/vignettes/sdm.pdf>
- Holst, K. T. (1986). *The Saltatoria (Bush-Crickets, crickets and grasshoppers) Of Northern Europe* (Vol. 16). Vinderup: Brill.
- Jin, S., & Sader, S. A. (2005). Comparison of time series tasseled cap wetness and the normalized difference moisture index in detecting forest disturbances. *Remote sensing of environment*, 94(3), 364-372.
- Johansen, B., Aarrestad, P. A., & Øien, D. I. (2009). *Vegetasjonskart for Norge basert på satellittdata*. Retrieved from
- Johansen, B. E. (2009). *Vegetasjonskart for Norge basert på Landsat TM/ETM+ data*. Retrieved from [ftp://ftp.miljodirektoratet.no/miljodata/satveg/Satveg\\_deling\\_nd\\_partnere\\_09\\_12\\_2009/dokumenter/Hovedrapport.pdf](ftp://ftp.miljodirektoratet.no/miljodata/satveg/Satveg_deling_nd_partnere_09_12_2009/dokumenter/Hovedrapport.pdf)
- Korner-Nievergelt, F., Roth, T., von Felten, S., Guelat, J., Almasi, B., & Korner-Nievergelt, P. (2015). *Bayesian Data Analysis in Ecology using Linear Models with R, BUGS and Stan*: Elsevier.
- Krebs, C. J. (2009). *Ecology : the experimental analysis of distribution and abundance* (6th ed. ed.). San Francisco: Benjamin Cummings.
- Lawesson, J. E. (2000). *A Concept for vegetation studies and monitoring in the Nordic countries* (Vol. 2000:517). København: Nordisk ministerråd.
- Litzenberger, G., & Chapco, W. (2001). A molecular phylogeographic perspective on a fifty-year-old taxonomic issue in grasshopper systematics. *Heredity*, 86(1), 54-59.
- Lopez, R. D., & Frohn, R. C. (2017). *Remote Sensing for Landscape Ecology: Monitoring, Modeling, and Assessment of Ecosystems* (2 ed. Vol. 1). Boca Raton: Routledge.
- MacArthur, R. H., & Wilson, E. O. (1967). *The theory of island biogeography* (Vol. 1). Princeton: Princeton university press.

- Mani, M. S. (1968). *Ecology and biogeography of high altitude insects*. The Hague: W. Junk Publishers.
- McCoy, J., Johnston, K., & Kopp, S. (2002). *Using ArcGIS Spatial Analyst : GIS by ESRI*. Redlands: ESRI.
- MedCalc Software Ltd. (2020). Digimizer 5.4.4. Gent, Belgium.
- Mukerji, M., & Gage, S. (1978). A model for estimating hatch and mortality of grasshopper egg populations based on soil moisture and heat. *Annals of the Entomological Society of America*, 71(2), 183-190.
- Naimi, B., Hamm, N. A., Groen, T. A., Skidmore, A. K., & Toxopeus, A. G. (2014). Where is positional uncertainty a problem for species distribution modelling? *Ecography*, 37(2), 191-203.
- Nilssen, A. C., & Lundmo, S. (2013). Gresshopper i Nord-Norge. *Insekter og andre småkryp*, 9-22.
- O'Callaghan, J. F., & Mark, D. M. (1984). The extraction of drainage networks from digital elevation data. *Computer vision, graphics, and image processing*, 28(3), 323-344.
- Oksanen, J., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., . . . Wagner, H. (2019). vegan: Community Ecology Package. Retrieved from <https://CRAN.R-project.org/package=vegan>
- Olalla-Tárraga, M. Á., Rodríguez, M. Á., & Hawkins, B. A. (2006). Broad-scale patterns of body size in squamate reptiles of Europe and North America. *Journal of Biogeography*, 33(5), 781-793.
- Oliver, T., Hill, J. K., Thomas, C. D., Brereton, T., & Roy, D. B. (2009). Changes in habitat specificity of species at their climatic range boundaries. *Ecology letters*, 12(10), 1091-1102.
- Ottesen, P. (1992). *Norges gresshopper* (Vol. 13). Ås: Norsk entomologisk forening.
- R Core Team. (2019). R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from <https://www.R-project.org/>
- Ranghetti, L., & Busetto, L. (2020). sen2r: Find, Download and Process Sentinel-2 Data. Retrieved from <http://sen2r.ranghetti.info>
- Roberts, D. (n.d.). R Labs for Vegetation Ecologists. Retrieved from <http://ecology.msu.montana.edu/labdsv/R/labs/>
- Roberts, D. W. (2019). labdsv: Ordination and multivariate analysis for ecology.
- Sanders, R. (1987). The Pareto principle: its use and abuse. *Journal of Services Marketing*, 1(2), 37-40.

- Schoener, T. W. (1989). *The ecological niche. In: Ecological Concepts: the Contribution of Ecology to an Understanding of the Natural World (ed. Cherrett, J.M.)*. Oxford: Blackwell Scientific.
- Shelomi, M. (2012). Where are we now? Bergmann's rule sensu lato in insects. *The American Naturalist*, 180(4), 511-519.
- Sømme, L. (1989). Adaptations of terrestrial arthropods to the alpine environment. *Biological Reviews*, 64(4), 367-407.
- Thuiller, W., Georges, D., Engler, R., & Breiner, F. (2013). biomod2: Ensemble platform for species distribution modeling. *R package version*, 2(7), r560.
- Thuiller, W., Georges, D., Engler, R., Breiner, F., Georges, M. D., & Thuiller, C. W. (2016). Package 'biomod2'.
- Thuiller, W., Lafourcade, B., Engler, R., & Araújo, M. B. (2009). BIOMOD—a platform for ensemble forecasting of species distributions. *Ecography*, 32(3), 369-373.
- Uvarov, B. (1977). *Grasshoppers and locusts. A handbook of general acridology Vol. 2. Behaviour, ecology, biogeography, population dynamics*. London: Centre for Overseas Pest Research.
- VanDerWal, J., Shoo, L. P., Graham, C., & Williams, S. E. (2009). Selecting pseudo-absence data for presence-only distribution modeling: how far should you stray from what you know? *Ecological modelling*, 220(4), 589-594.
- Warton, D. I., & Shepherd, L. C. (2010). Poisson point process models solve the “pseudo-absence problem” for presence-only data in ecology. *The Annals of Applied Statistics*, 4(3), 1383-1402.
- Wickham, H., François, R., Henry, L., & Müller, K. (2020). dplyr: A Grammar of Data Manipulation. Retrieved from <https://CRAN.R-project.org/package=dplyr>



## Appendix

### Appendix A Satveg classes

Table 8 Class number and type of vegetation for the satellite based vegetation map for Norway (Satveg) (B. Johansen et al., 2009).

Class number	Forest vegetation
1	Coniferous Forest – dense canopy layer
2	Coniferous forest and mixed forest - open canopy
3	Lichen rich pine forest
4	Low herb forest and broad leaved deciduous forest
5	Tall herb - tall fern deciduous forest
6	Bilberry- low fern birch forest
7	Crowberry birch forest
8	Lichen-rich birch forest
	Mire and open swamp vegetation
9	Ombrotrophic bog and low-grown lawn vegetation
10	Tall-grown lawn vegetation
11	Wet mires, sedge swamps and reed beds
	Mountainous vegetation
12	Exposed alpine ridges, scree and rock complex
13	Graminoid alpine ridge vegetation
14	Heather-rich alpine ridge vegetation
15	Lichen-rich heathland
16	Heather- and grass-rich early snow patch com.
17	Fresh heather and dwarf-shrub communities (u/l)
18	Herb-rich meadows (up-/lowland)
19	Grass and dwarf willow snow-patch vegetation
20	Bryophyte late snow patch vegetation
21	Glacier, snow and wet snow-patch vegetation
	Other
22	Water
23	Agricultural areas
24	Cities and built-up areas
25	Unclassified and shadow affected areas,

**Appendix B Diagnostic plots**

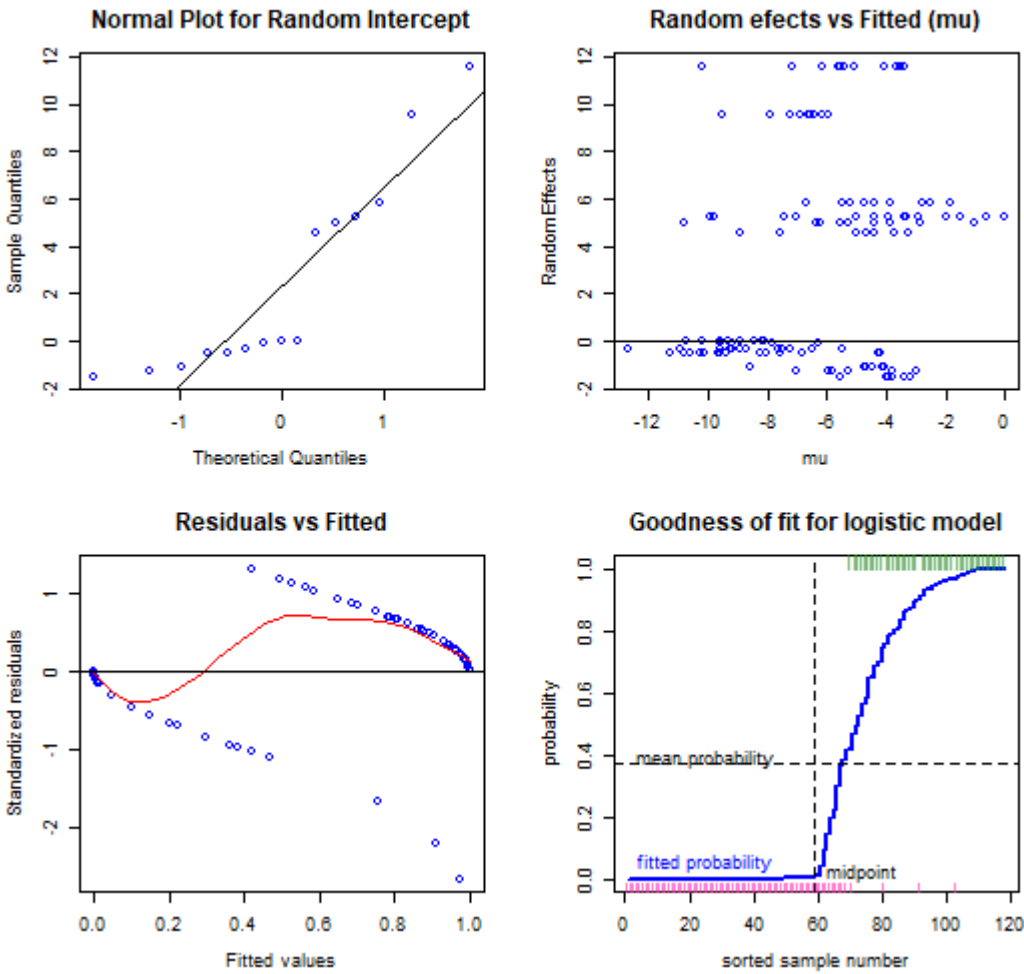


Figure 17 Diagnostic plots for the highest-ranking model where explanatory variables impact on the occurrence of the grasshoppers were modelled.

## Appendix C The different SDM's predictions at location number 1, Lappebu

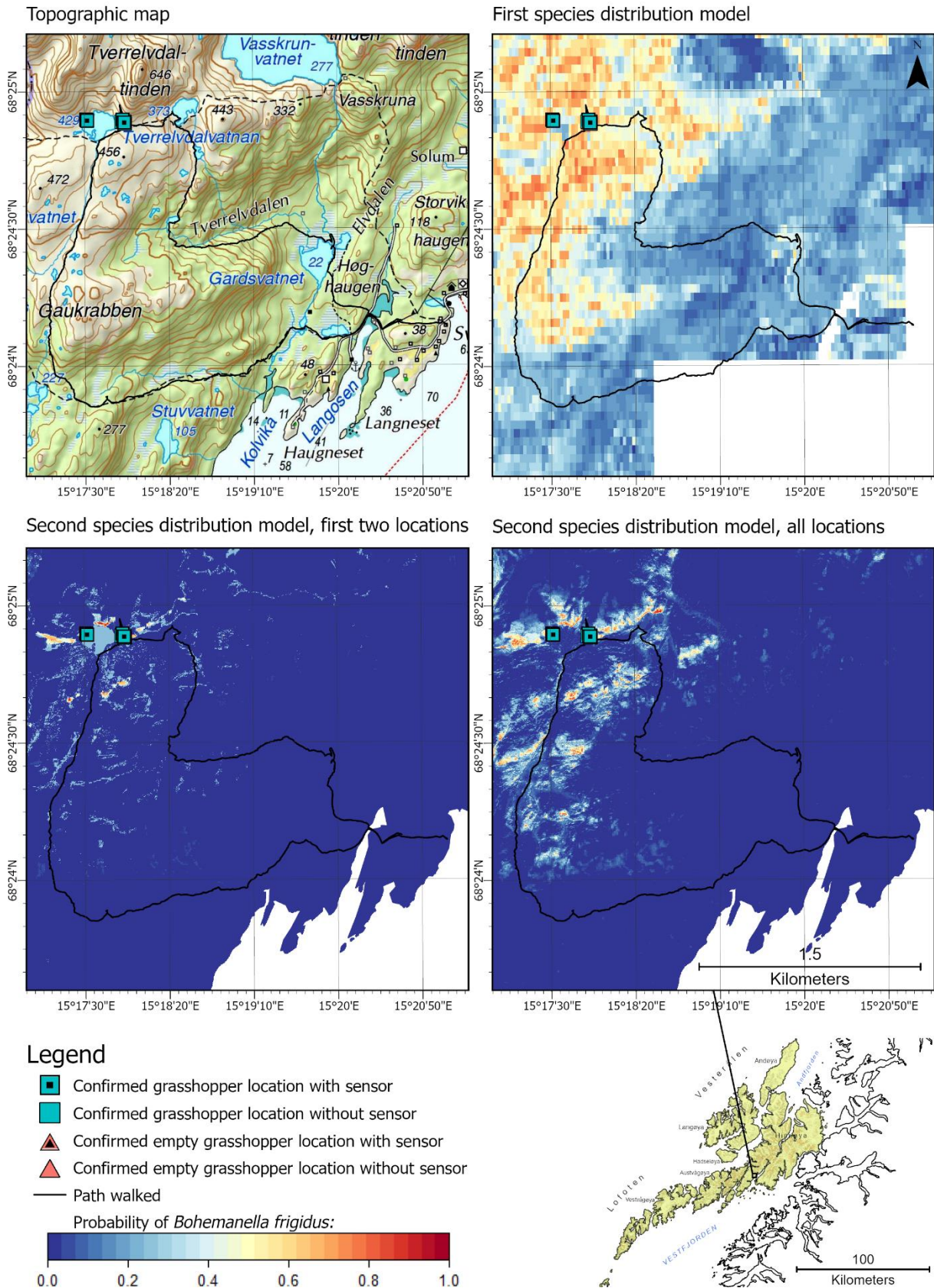


Figure 18 The two SDM's predictions of *B. frigidus* probability of occurrence in location 1 in Vesterålen and Lofoten

## Appendix D The different SDM's predictions at location number 2, Stordalsaksla

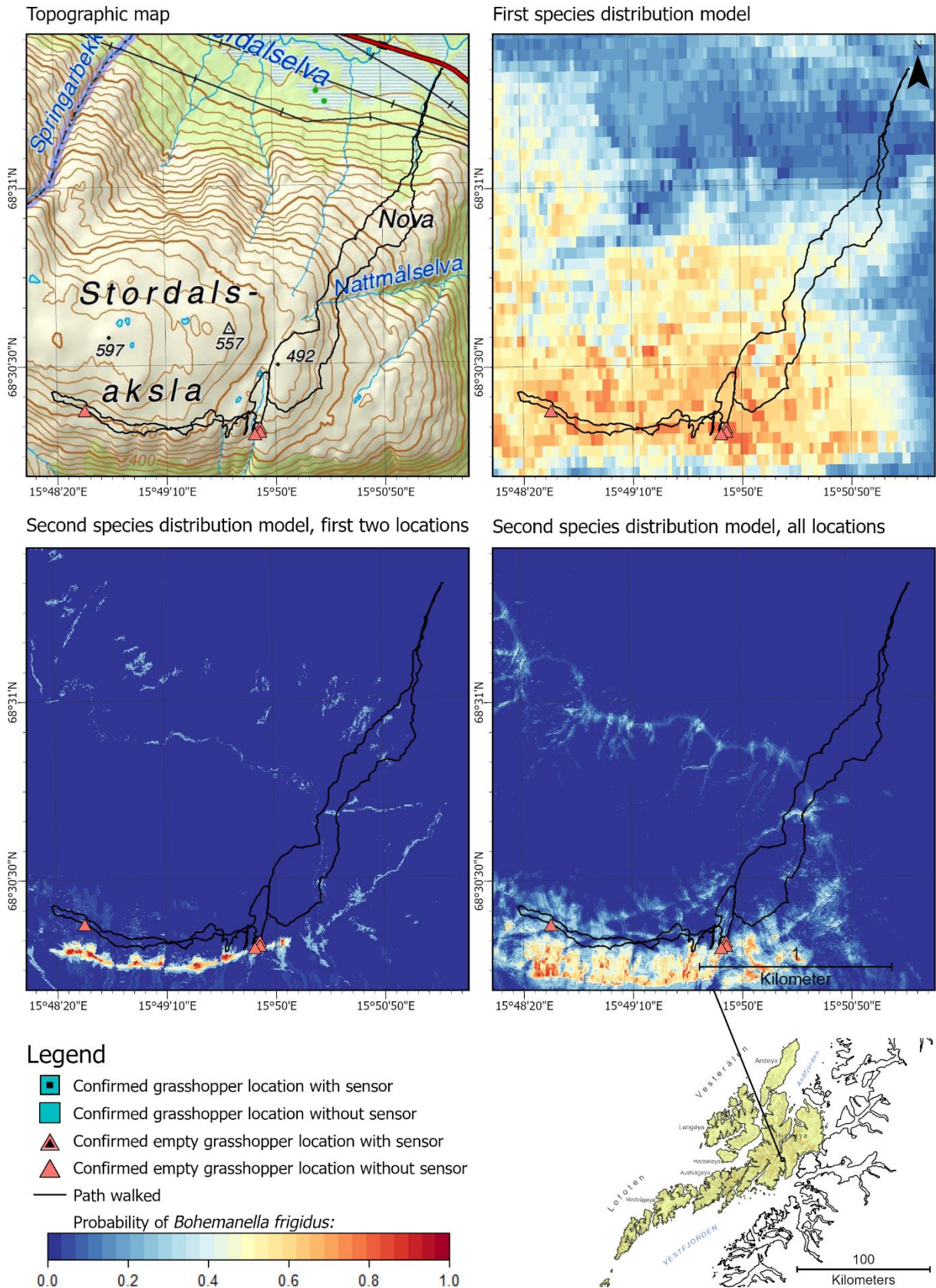


Figure 19 The two SDM's predictions of *B. frigidus* probability of occurrence in location 2 in Vesterålen and Lofoten

## Appendix E The different SDM's predictions at location number 3, Kanstadjfjellet

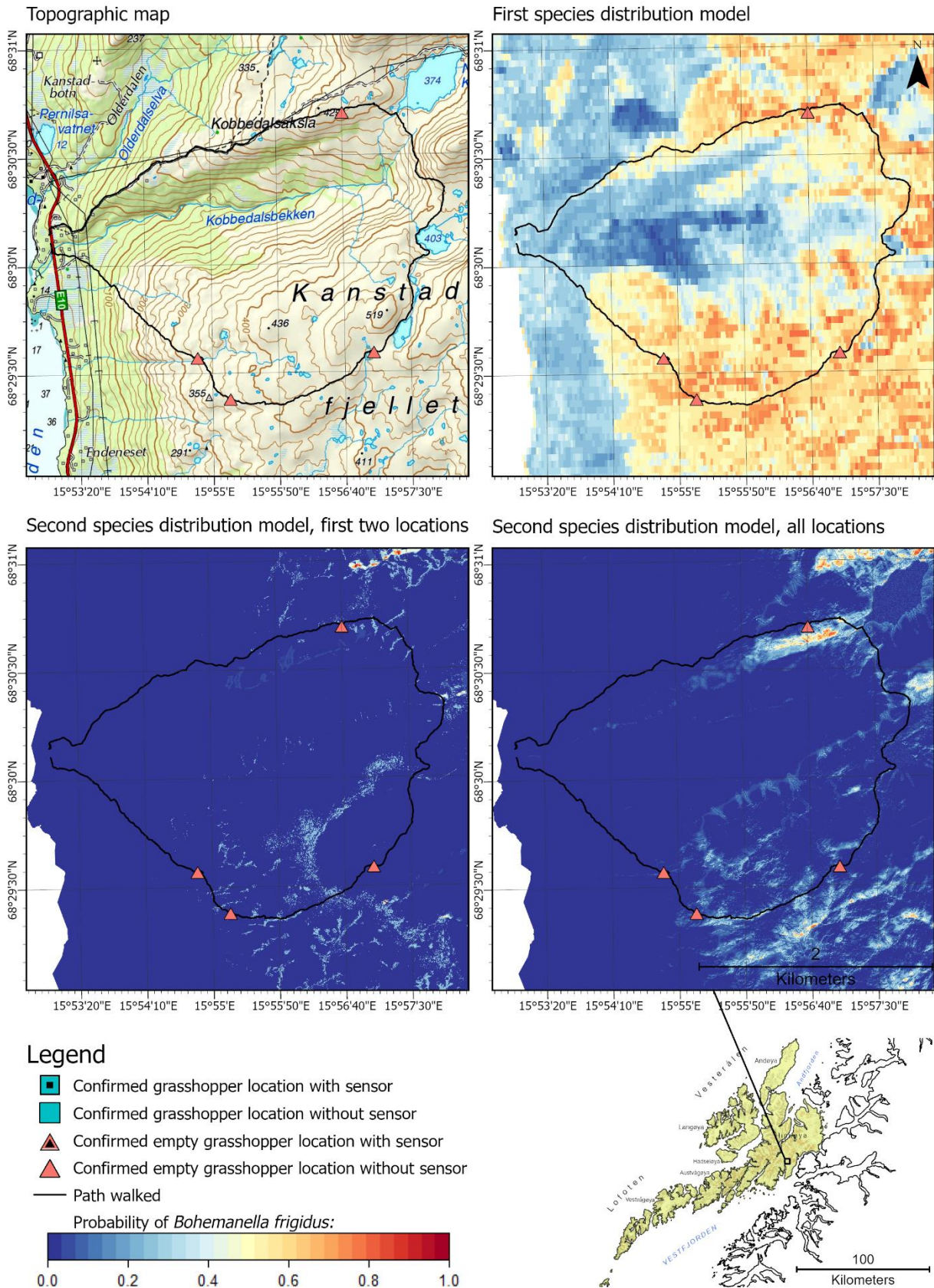


Figure 20 The two SDM's predictions of *B. frigidus* probability of occurrence in location 3 in Vesterålen and Lofoten

## Appendix F The different SDM's predictions at location number 4, Vangpollen

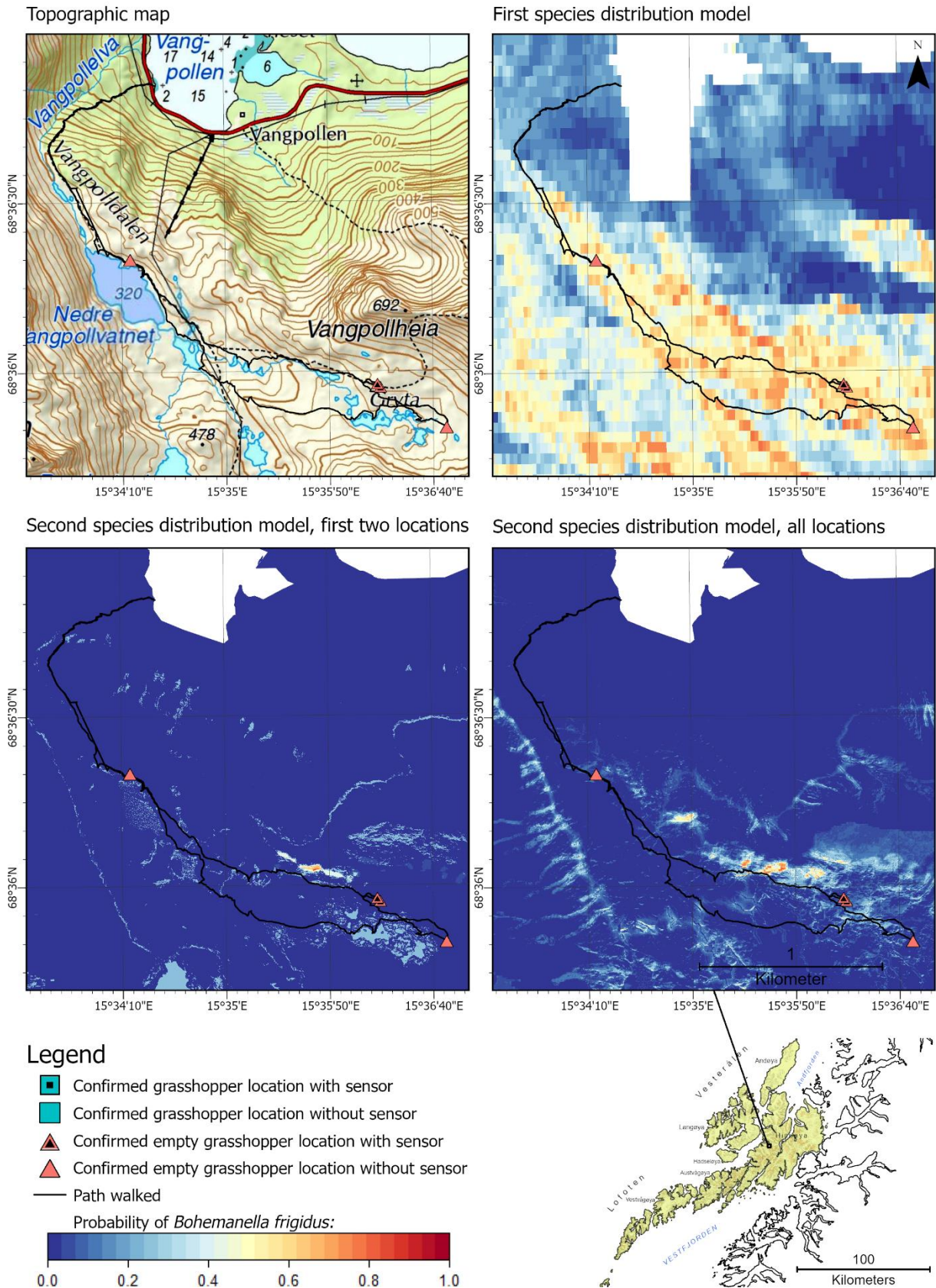


Figure 21 The two SDM's predictions of *B. frigidus* probability of occurrence in location 4 in Vesterålen and Lofoten

## Appendix G The different SDM's predictions at location number 5, Vikfjellet

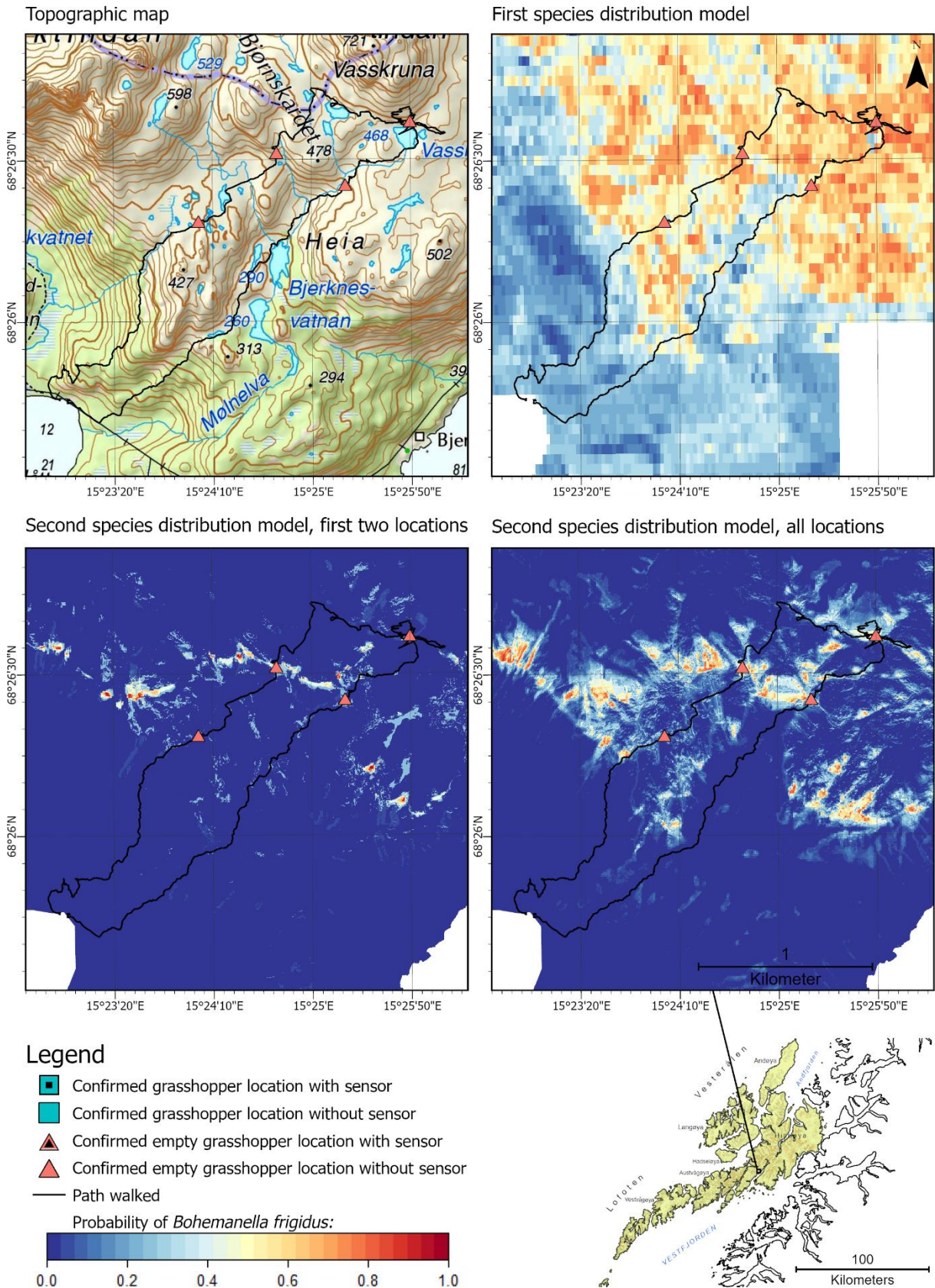


Figure 22 The two SDM's predictions of *B. frigidus* probability of occurrence in location 5 in Vesterålen and Lofoten

## Appendix H The different SDM's predictions at location number 6, Vestbygda

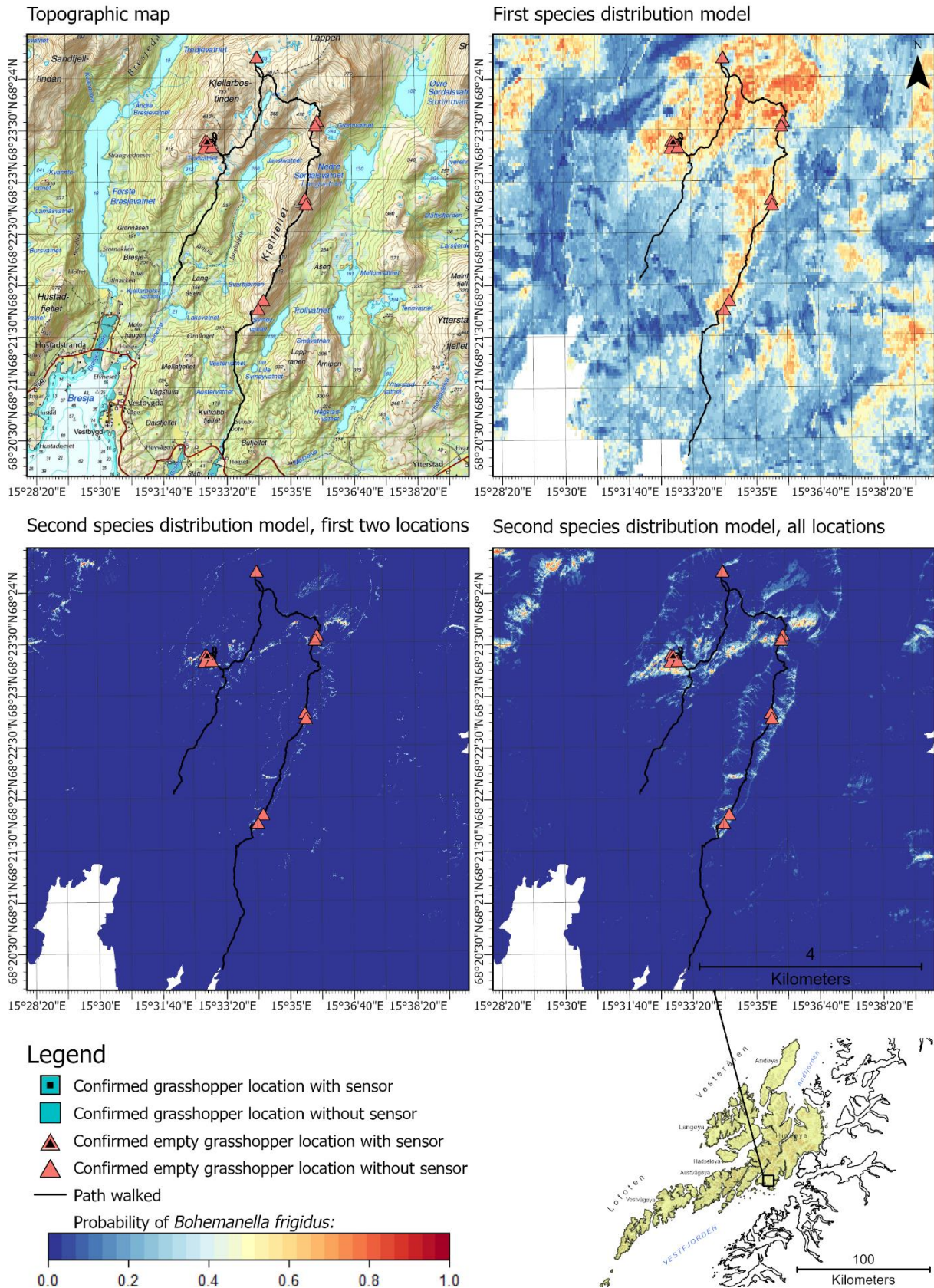


Figure 23 The two SDM's predictions of *B. frigidus* probability of occurrence in location 6 in Vesterålen and Lofoten



## Appendix I The different SDM's predictions at location number 7, Strandheia

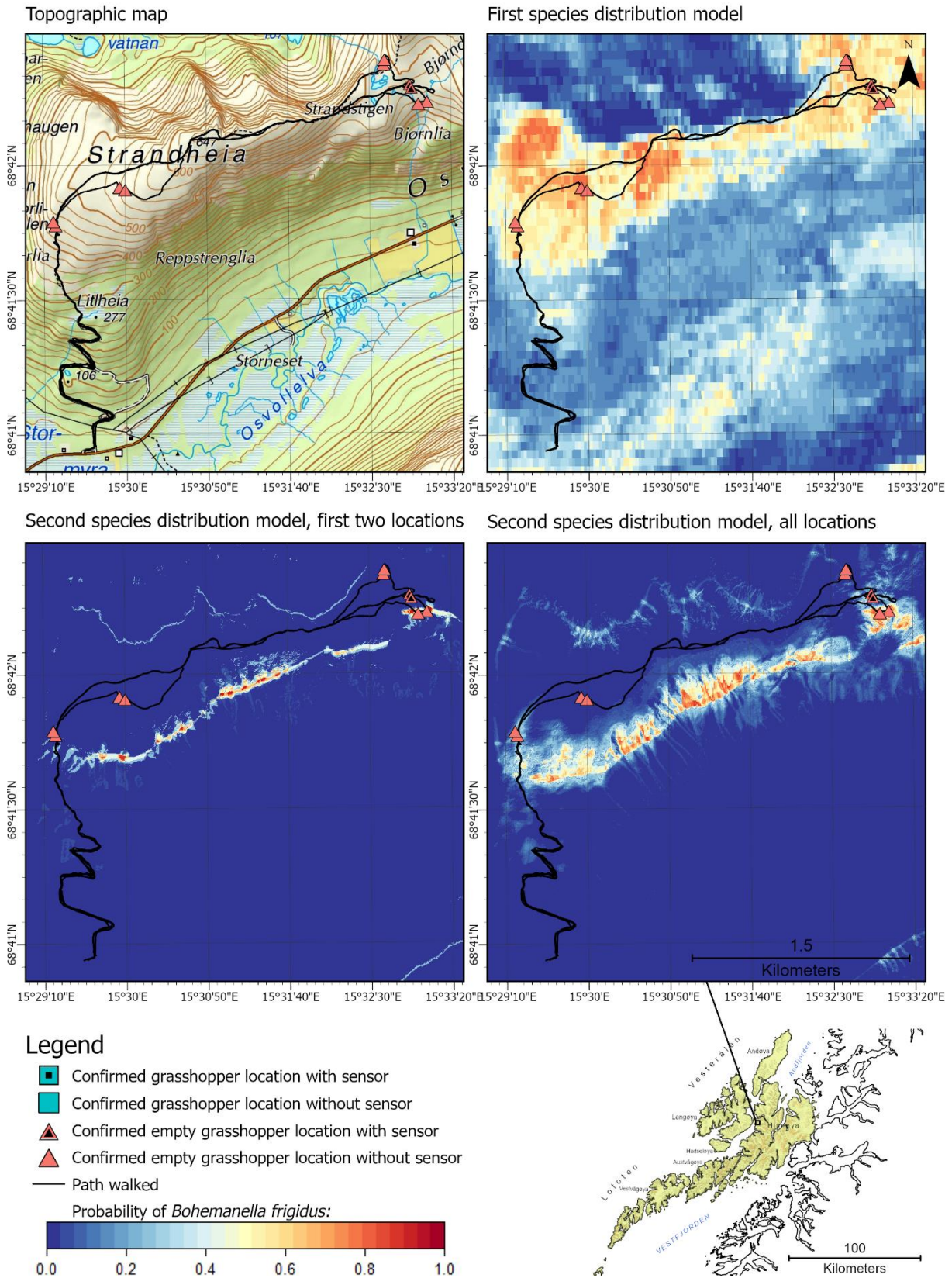


Figure 24 The two SDM's predictions of *B. frigidus* probability of occurrence in location 7 in Vesterålen and Lofoten

## Appendix J The different SDM's predictions at location number 8, Kongsvika

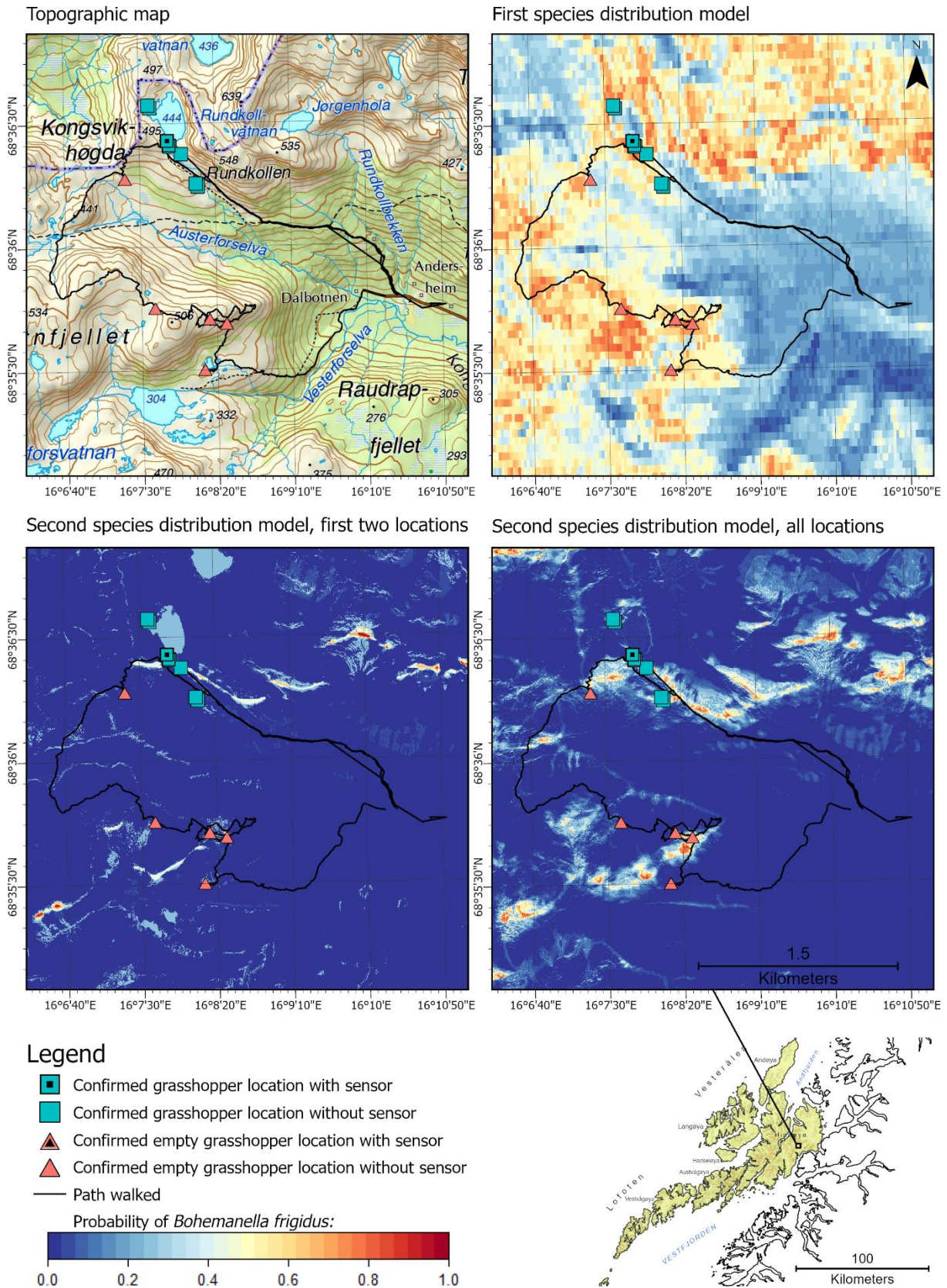


Figure 25 The two SDM's predictions of *B. frigidus* probability of occurrence in location 8 in Vesterålen and Lofoten

## Appendix K The different SDM's predictions at location number 9, Erikstadjordfjellet

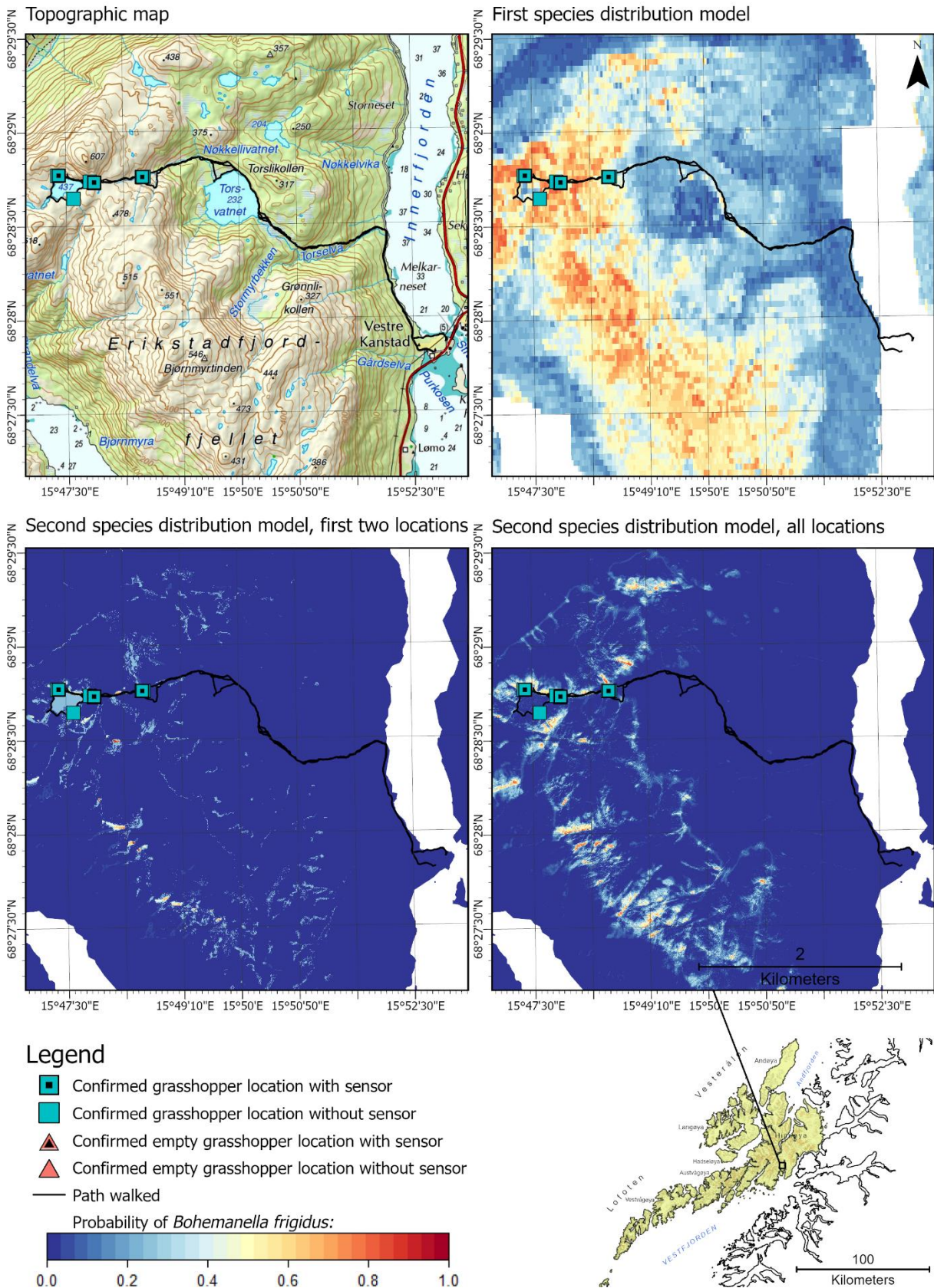


Figure 26 The two SDM's predictions of *B. frigidus* probability of occurrence in location 9 in Vesterålen and Lofoten

## Appendix L The different SDM's predictions at location number 10, Malmaksla

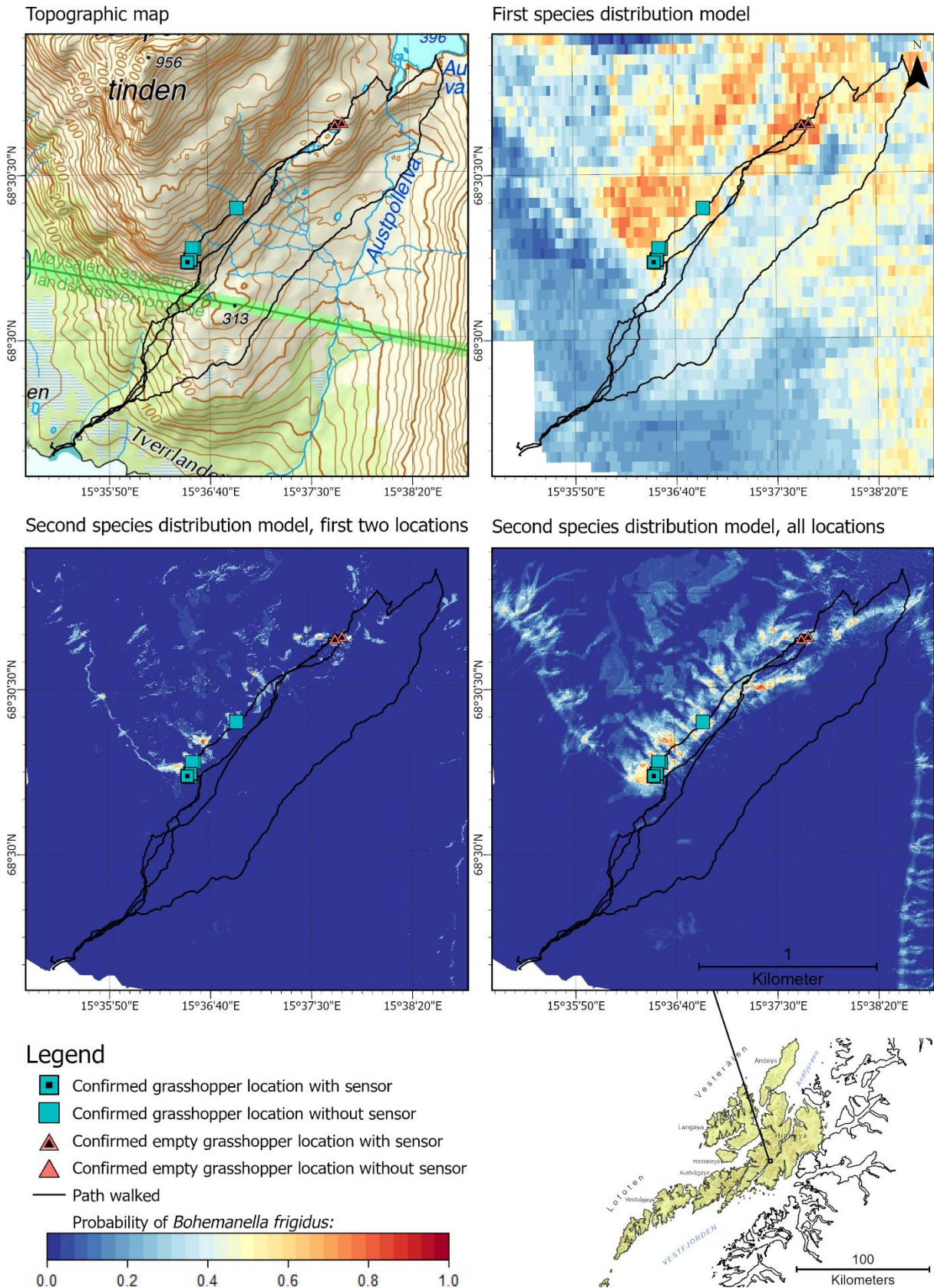


Figure 27 The two SDM's predictions of *B. frigidus* probability of occurrence in location 10 in Vesterålen and Lofoten

## Appendix M The different SDM's predictions at location number 11, Svartskardtinden

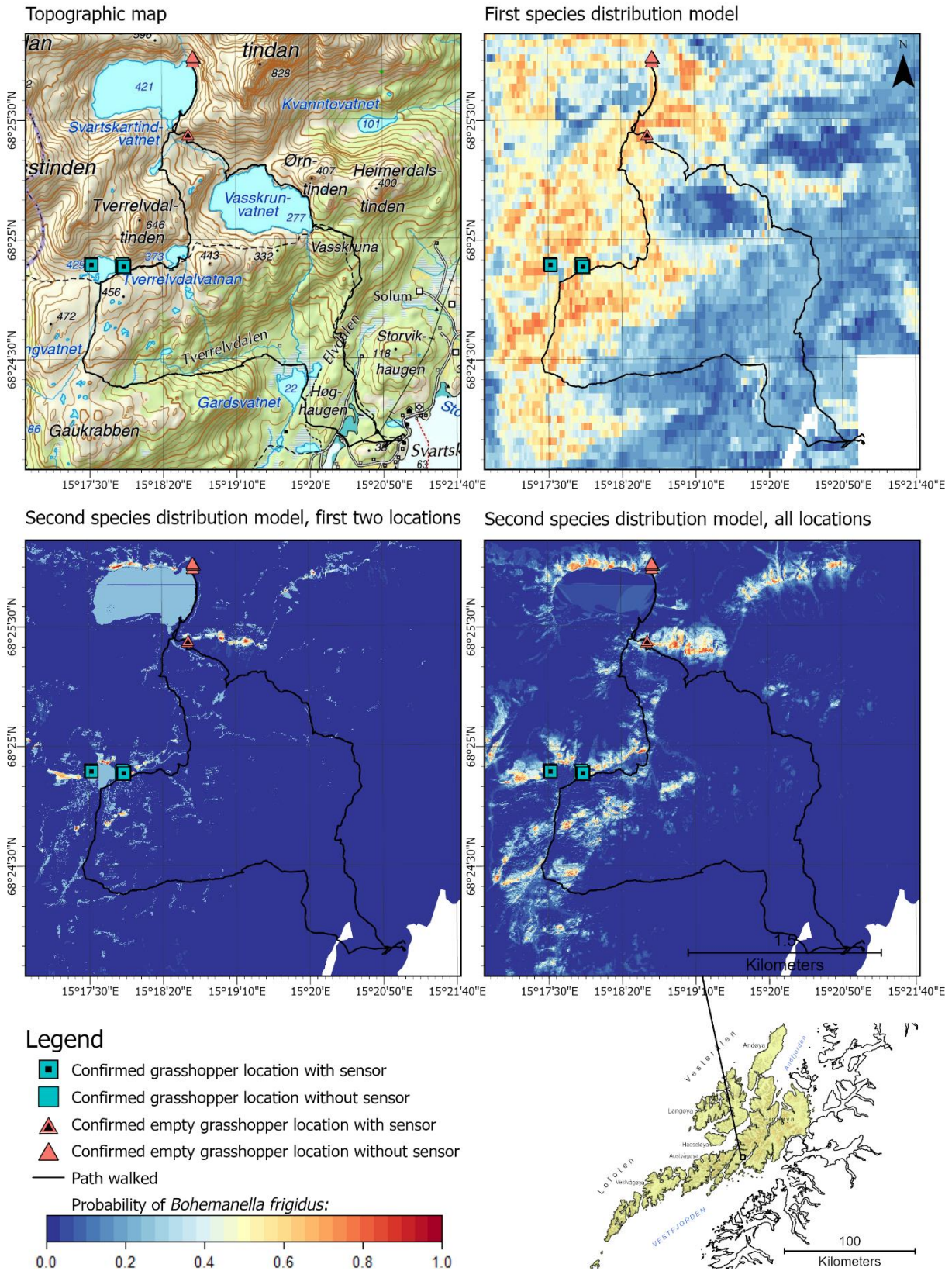


Figure 28 The two SDM's predictions of *B. frigidus* probability of occurrence in location 11 in Vesterålen and Lofoten

## Appendix N The different SDM's predictions at location number 12, Lakselvfjellet

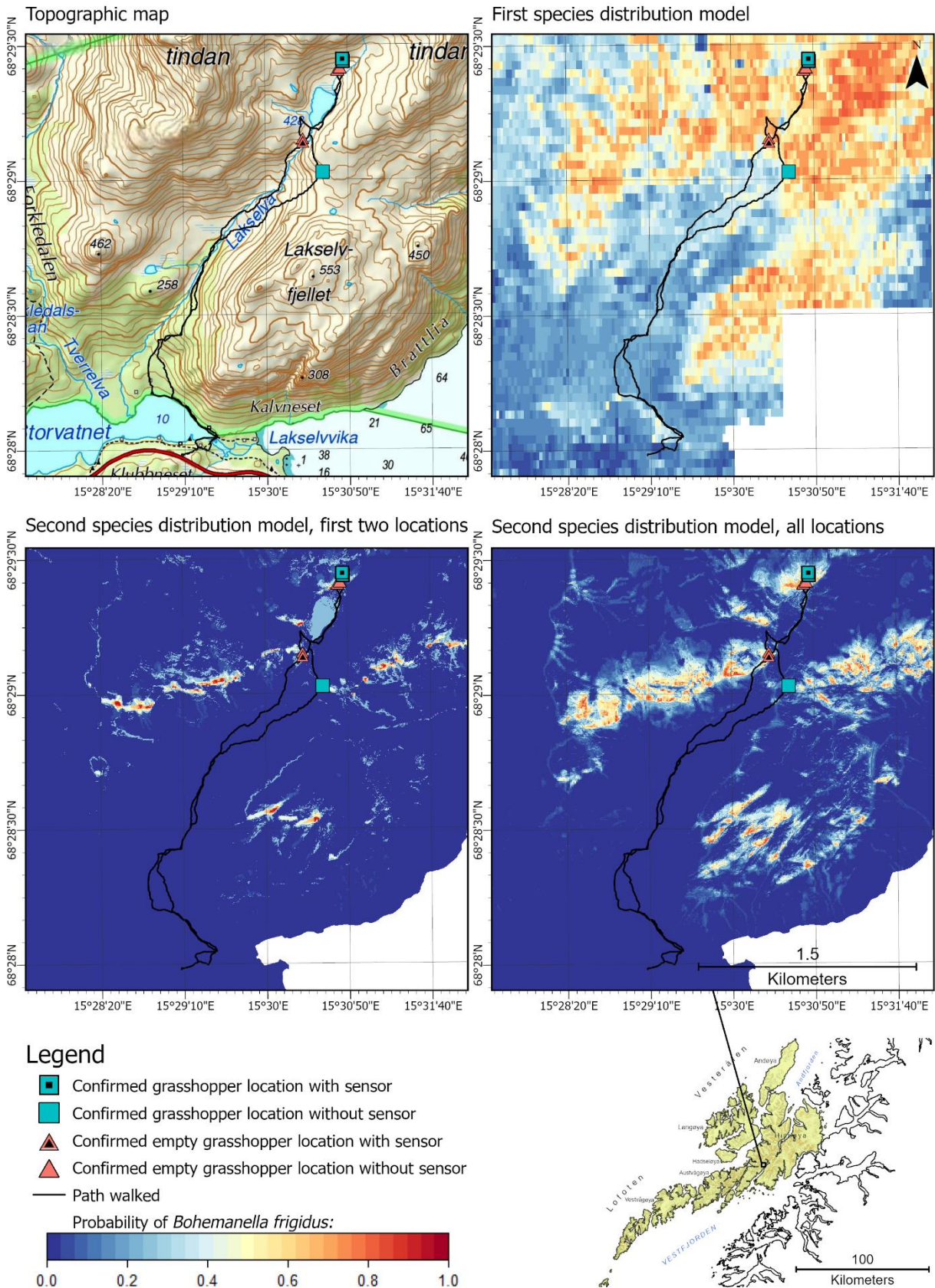


Figure 29 The two SDM's predictions of *B. frigidus* probability of occurrence in location 12 in Vesterålen and Lofoten

## Appendix O The different SDM's predictions at location number 13, Grunnfjorddalen

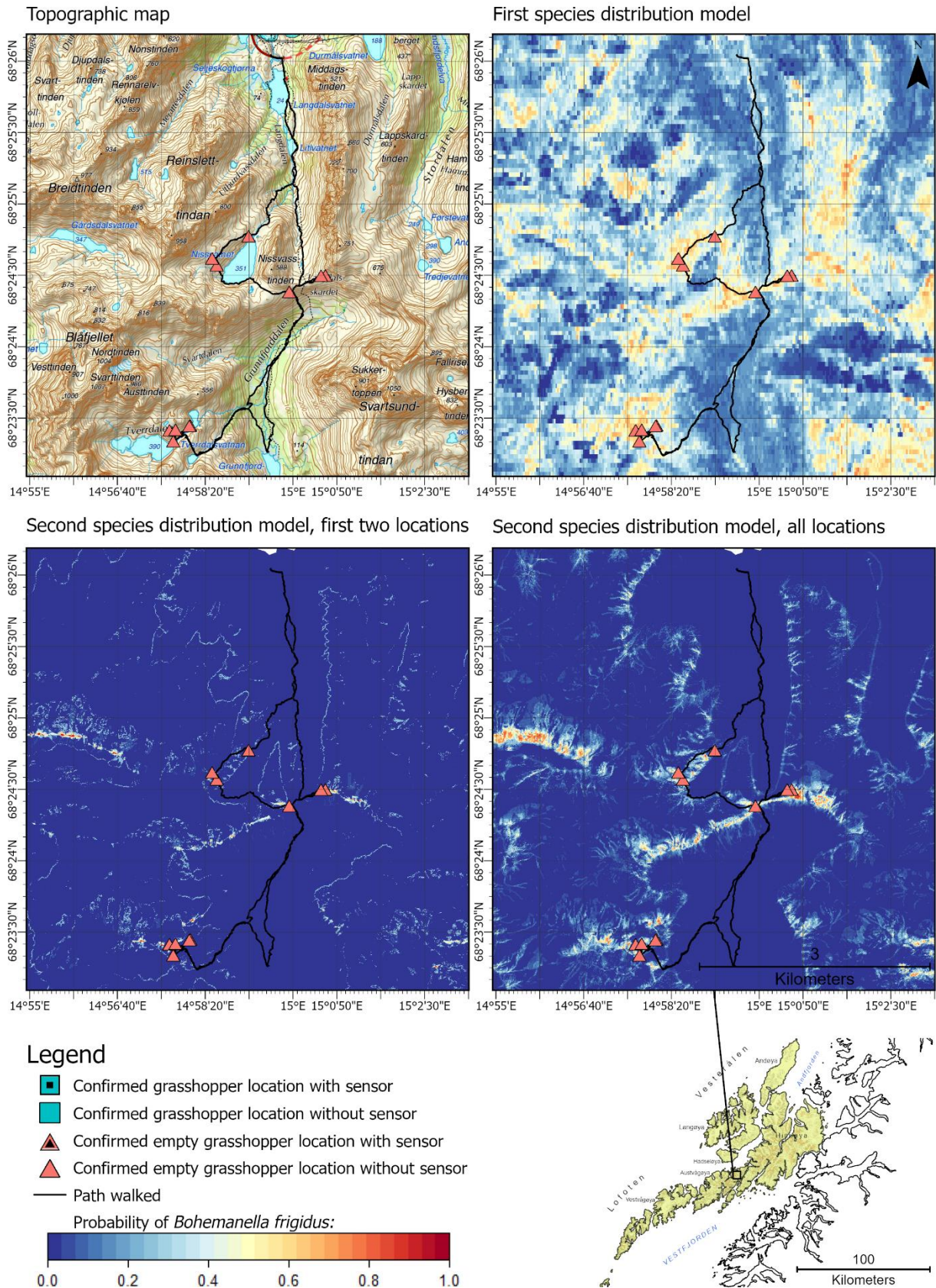


Figure 30 The two SDM's predictions of *B. frigidus* probability of occurrence in location 13 in Vesterålen and Lofoten

## Appendix P The different SDM's predictions at location number 14, Hadseløya

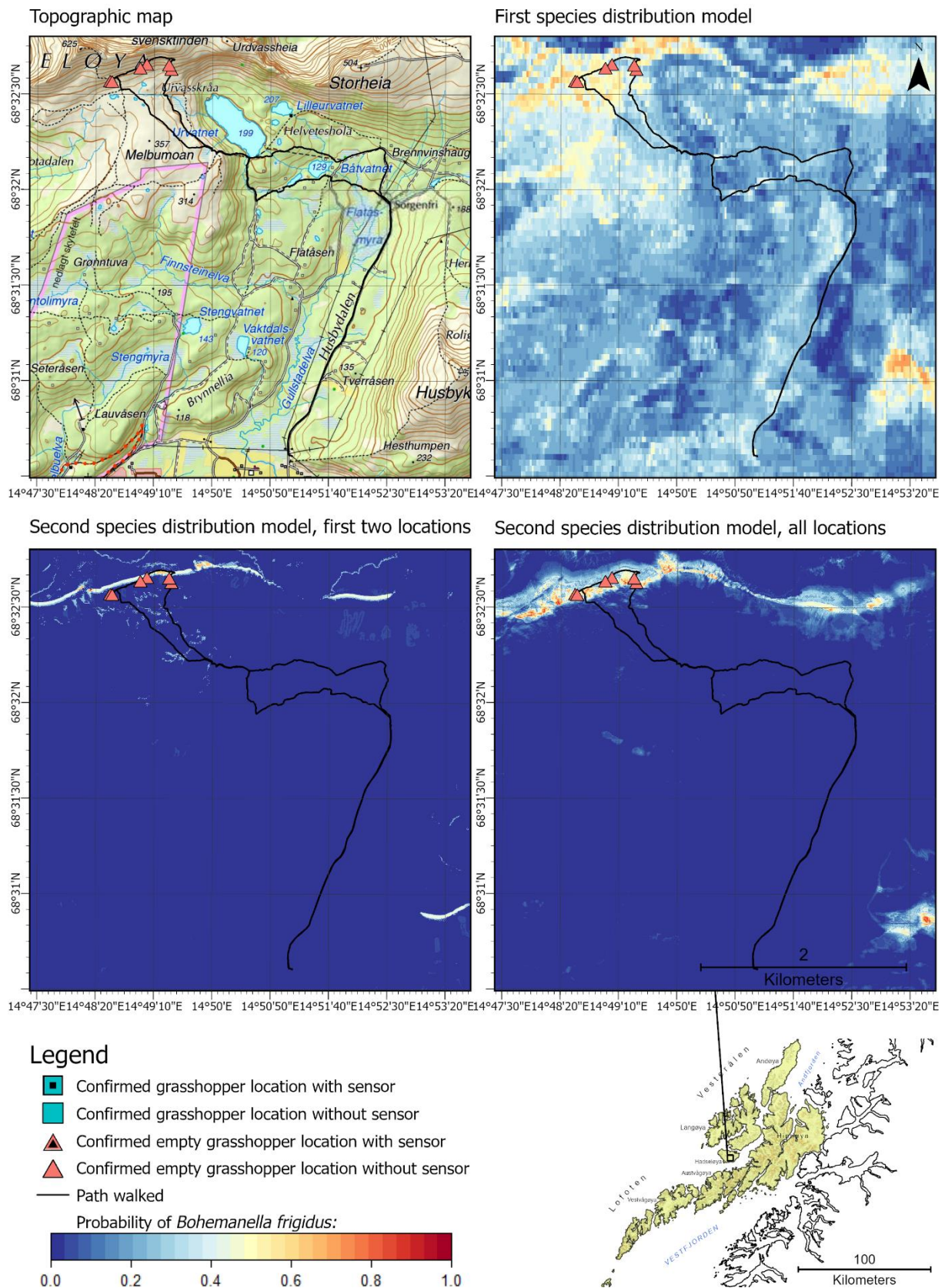


Figure 31 The two SDM's predictions of *B. frigidus* probability of occurrence in location 14 in Vesterålen and Lofoten



## Appendix Q The different SDM's predictions at location number 15, Langøya

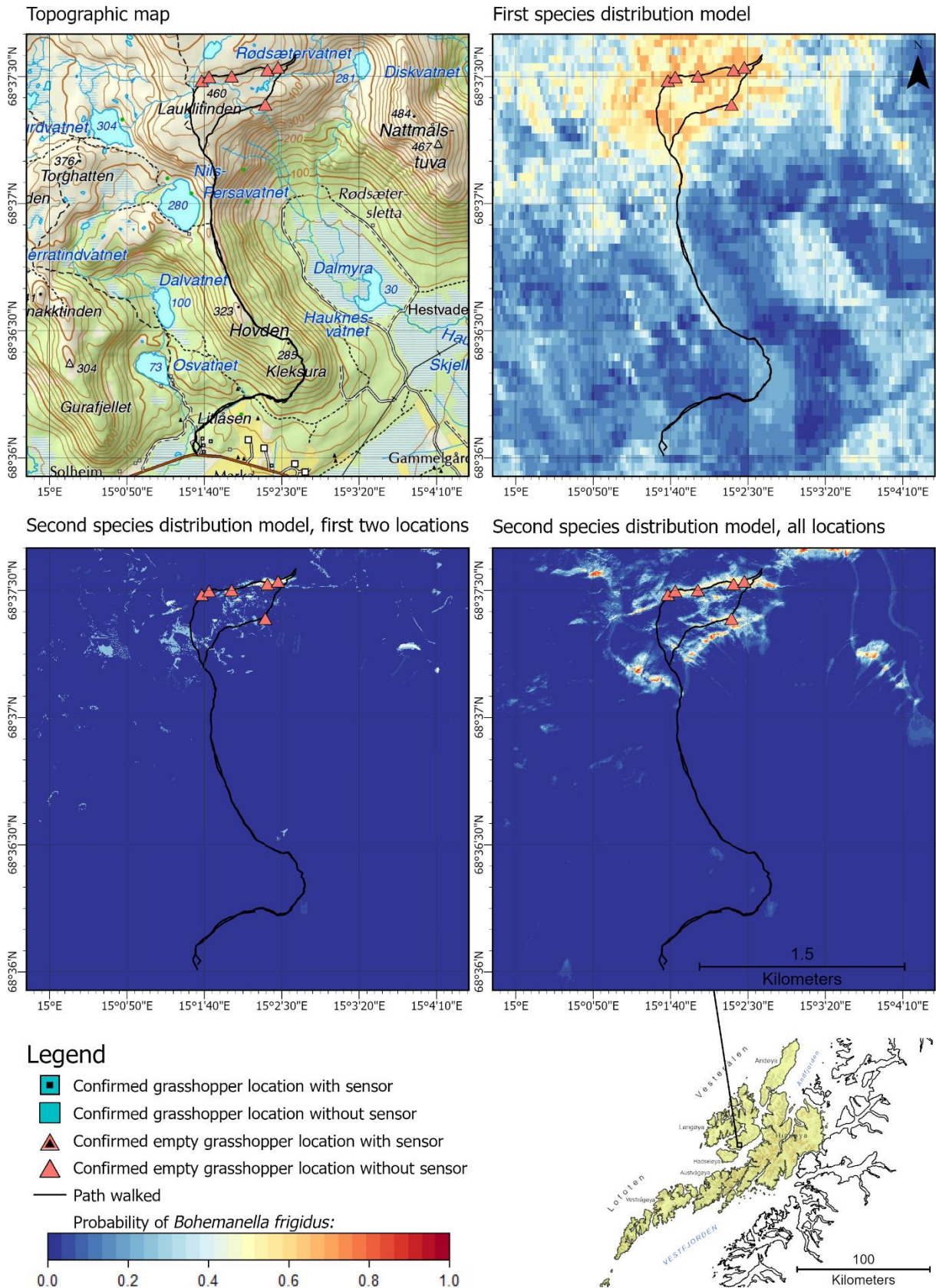


Figure 32 The two SDM's predictions of *B. frigidus* probability of occurrence in location 15 in Vesterålen and Lofoten

## Appendix R The different SDM's predictions at location number 16, Rundfjellet

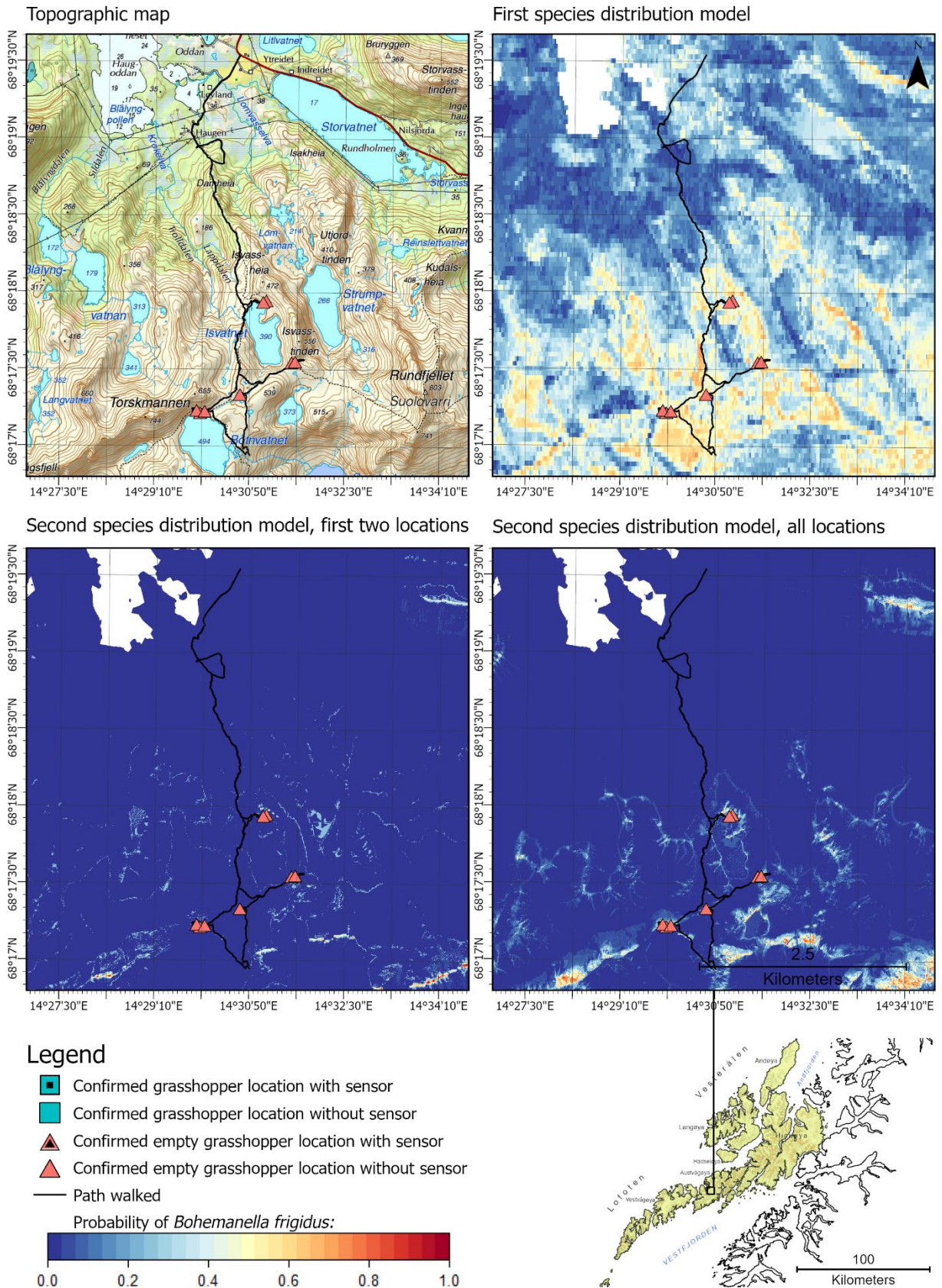


Figure 33 The two SDM's predictions of *B. frigidus* probability of occurrence in location 16 in Vesterålen and Lofoten

## Appendix S The different SDM's predictions at location number 17, Kvannkjosfjellet

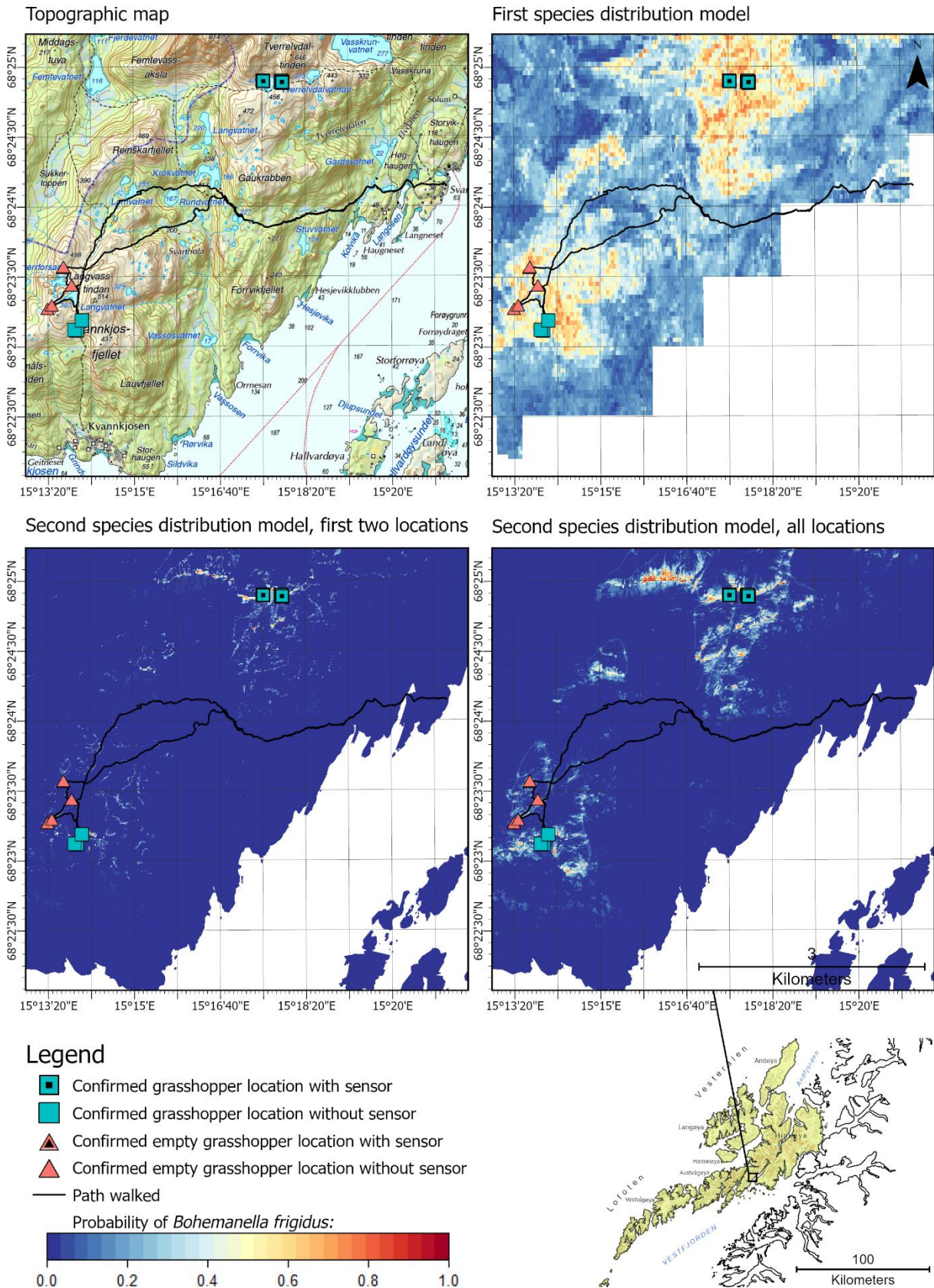


Figure 34 The two SDM's predictions of *B. frigidus* probability of occurrence in location 17 in Vesterålen and Lofoten