Vascular plant indicators for the Nature Index 2020

Olav Skarpaas

# 1. Introduction

This report describes the methods and results for estimation of indicator values for a set of vascular plant species indicators in the Nature Index (NI) for Norway 2020 (Table 1). These indicators are based on occurrence data available through the Global Biodiversity Information Facility (www.gbif.org; mirrored in the Norwegian Biodiversity Information Centre portal Artskart, artskart.artsdatabanken.no). For the last major update of NI in 2015, the states of these indicators were modelled across space and time using a hierarchical and dynamical spatiotemporal model (based on Wikle 2003; Wikle & Hooten 2010) This model had the advantage of accounting for noise in the observation process as well as capturing population developments through time by incorporating dependence on the previous states of the local and neighboring populations. These advantages came with a high computational cost. The analyses required methods that allow little experimentation (Bayesian MCMC) and low spatial resolution (100km by 100km). The coarse resolution potentially introduces prediction error (imprecision and bias) and also raises the issue of how well the observations of the indicators represent populations in the major ecosystem that they are supposed to represent in the Nature Index (woodland, mountain, open lowland and wetland) and how indicator observations and predictions relate to pressures. Therefore, the vascular plant indicators for NI 2020 were estimated with computationally less demanding methods, providing estimates with higher spatial resolution and higher precision. The main aim of this report is to document the estimation process, to make it fully reproducible, and to investigate indicator-ecosystem relations and effects of pressures.

The report is structured as follows. Section 2 describes the data and gives a brief assessment of the relationship between indicator species occurrences and land cover types (as a proxy for major ecosystems). Section 3 describes the methods for estimation of indicator states, including the reference state, and discusses some key aspects of these results. Section 4 gives an assessment of relationships between indicator trends and pressures. Finally, Section 5 presents some perspectives on methods and results, and summarizes the main conclusions. Additional information (R scripts and results) is provided in appendices.

Table 1. Indicator species for the Nature Index (NI) of Norway covered in this report, with indicator ID number in the NI database, indicator name in the NI database (Norwegian species name) and scientific name.

|  |  |  |
| --- | --- | --- |
| ID | NI indicator name | Scientific name |
| 5 | Alm | Ulmus glabra |
| 54 | Fjellvalmue | Papaver radicatum |
| 66 | Greplyng | Kalmia procumbens |
| 85 | Issoleie | Ranunculus glacialis |
| 100 | Kusymre | Primula vulgaris |
| 125 | Myrtelg | Thelypteris palustris |
| 130 | Olavsstake | Moneses uniflora |
| 151 | Prestekrage | Leucanthemum vulgare |
| 152 | Purpurlyng | Erica cinerea |
| 167 | Sennegras | Carex vesicaria |
| 180 | Solblom | Arnica montana |
| 196 | Sveltstarr | Carex pauciflora |
| 211 | Engmarihand | Dactylorhiza incarnata |
| 337 | Hvitmyrak | Rhynchospora alba |
| 338 | Brunmyrak | Rhynchospora fusca |
| 339 | Smalsoldogg | Drosera anglica |
| 340 | Dikesoldogg | Drosera intermedia |

# 2. Data

## Species occurrences and environmental covariates

Species occurrence data were downloaded from GBIF (see Appendix 1). Occurrences were counted within each 1km by 1km cell in five-year intervals. The total number of unique visits (i.e. the number of dates with records of one or more vascular plants within a cell over the five-year interval) was used as an offset to correct for sampling effort (see details on modelling below).

Environmental data were taken from Horvath et al. (2019), including landcover, climate, etc., aggregated to 1km resolution. These data were supplemented with nitrogen deposition data from Lund et al. (2012) at the municipal level. Note that none of these environmental data were used to estimate the state of indicators; they were only used to explore properties of the occurrence data in relation to land cover (next section) and to assess estimated trends in relation to environmental conditions and pressures (below).

## Land cover affinity

To get an indication of the “representativeness” of the indicators for their major ecosystems, species records (with coordinates) of each species were summarized by dominant land cover type in in 1km by 1km cells (Table 2). As expeced, the woodland indicators (*Ulmus glabra*, *Primula vulgaris*, *Moneses uniflora*) are frequently found in forest-dominated areas, and the mountain indicators (*Papaver radicatum*, *Kalmia procumbens*, *Ranunculus glacialis*) have high frequencies in areas dominated by barren land. *Ranunculus glacialis* and *Kalmia procumbens* are the only species with records in areas dominated by glaciers. Two of the indicators of open lowland, *Leucanthemum vulgare* and *Arnica montana*, are most frequently found in forest-dominated areas, while the third, *Erica cinerea*, is most often found in areas dominated by barren land. This reflects the ecology of the species: the two former species are associated with areas of extensive agriculture (pastures and hayfields), often found in areas dominated by forest and nowadays often in succession towards forest, while the latter is found in the coastal heathlands in between rocky areas with shallow soil. Finally, the wetland indicators (*Thelypteris palustris*, *Carex vesicaria*, *Carex pauciflora*, *Dactylorhiza incarnata*, *Rhynchospora alba*, *Rhynchospora fusca*, *Drosera anglica* and *Drosera intermedia*) have higher frequencies in mires than most other species, but are still most frequent in forest-dominated areas, where many small mires are found.

To summarize, the occurrences of the indicator species are more or less as expected with respect to dominant land cover at the scale of 1km by 1km. However, the dominant land cover at this spatial scale is not a good representation of the main ecosystem of open lowland and wetland species, as many habitats of these species are found in areas dominated by other land cover types (often forest).

Table 2. Percent of records in 1km by 1km cells with a given dominant land cover type for each of the indicator species.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Developed | Agricultural | Forest | Barren | Mire | Glacier | Freshwater | Marine |
| Ulmus glabra | 10.8 | 8.7 | 72.9 | 4.0 | 0.0 | 0.0 | 3.4 | 0.3 |
| Papaver radicatum | 0.0 | 0.0 | 45.5 | 54.5 | 0.0 | 0.0 | 0.0 | 0.0 |
| Kalmia procumbens | 0.0 | 0.1 | 19.2 | 78.3 | 0.5 | 0.1 | 1.8 | 0.0 |
| Ranunculus glacialis | 0.0 | 0.0 | 1.8 | 95.1 | 0.0 | 1.8 | 1.3 | 0.0 |
| Primula vulgaris | 1.4 | 10.6 | 65.3 | 19.5 | 0.2 | 0.0 | 2.7 | 0.2 |
| Thelypteris palustris | 6.9 | 19.0 | 70.7 | 0.0 | 1.7 | 0.0 | 1.7 | 0.0 |
| Moneses uniflora | 0.6 | 2.9 | 85.9 | 3.8 | 3.2 | 0.0 | 3.2 | 0.4 |
| Leucanthemum vulgare | 6.0 | 15.5 | 68.5 | 4.4 | 0.9 | 0.0 | 4.3 | 0.4 |
| Erica cinerea | 0.7 | 11.8 | 22.8 | 64.0 | 0.0 | 0.0 | 0.0 | 0.7 |
| Carex vesicaria | 5.4 | 12.0 | 65.7 | 4.1 | 0.5 | 0.0 | 11.3 | 1.1 |
| Arnica montana | 0.5 | 7.0 | 86.4 | 2.4 | 0.5 | 0.0 | 3.1 | 0.2 |
| Carex pauciflora | 0.1 | 1.3 | 75.9 | 15.7 | 5.2 | 0.0 | 1.9 | 0.0 |
| Dactylorhiza incarnata | 0.4 | 12.0 | 59.4 | 13.2 | 13.5 | 0.0 | 1.5 | 0.0 |
| Rhynchospora alba | 1.5 | 5.1 | 81.3 | 7.2 | 3.3 | 0.0 | 1.5 | 0.0 |
| Rhynchospora fusca | 1.7 | 4.3 | 80.0 | 8.7 | 1.7 | 0.0 | 3.5 | 0.0 |
| Drosera anglica | 0.1 | 1.4 | 75.1 | 15.8 | 5.7 | 0.0 | 1.7 | 0.2 |
| Drosera intermedia | 1.5 | 7.4 | 79.4 | 8.1 | 1.5 | 0.0 | 2.2 | 0.0 |

# 3. Modelling indicator values

## A generalized additive space-time model

To model the distribution of the indicator species in space and time we used Generalized Additive Models (GAM; Wood 2017). These models are flexible enough to capture complex non-linear patterns in space and time, while at the same time being computationally efficient and accessible through the extensive R package ‘mgcv’. This means (among others) that the analyses can be carried out with relatively high resolution without resorting to specialized software and supercomputing resources.

The response variable in the model is the count of observations of a species within 1km by 1km cells in a five-year interval, with a Poisson error distribution. To account for sampling effort, the total count of unique visits to each cell in the same time interval (as defined above) was included as an offset variable on the log scale. When the sampling effort is 1, the model is essentially a model of the rate of occurrence of the species in a sample from a cell. This is useful when making predictions from the model: we can then set the sampling effort to S = 1 (or log S = 0).

The only covariates in the models are spatial and temporal coordinates (x, y and year). Although other environmental variables clearly affect the species, the aim here is to estimate the state of the indicators in space and time - without the need for additional information (which may not be available at times and places where we might want to predict the state of the indicator). Assessment of relationships between estimated states of indicators and environmental pressures are carried out in a second step (below).

The model specification in ‘gam’ looks like this:

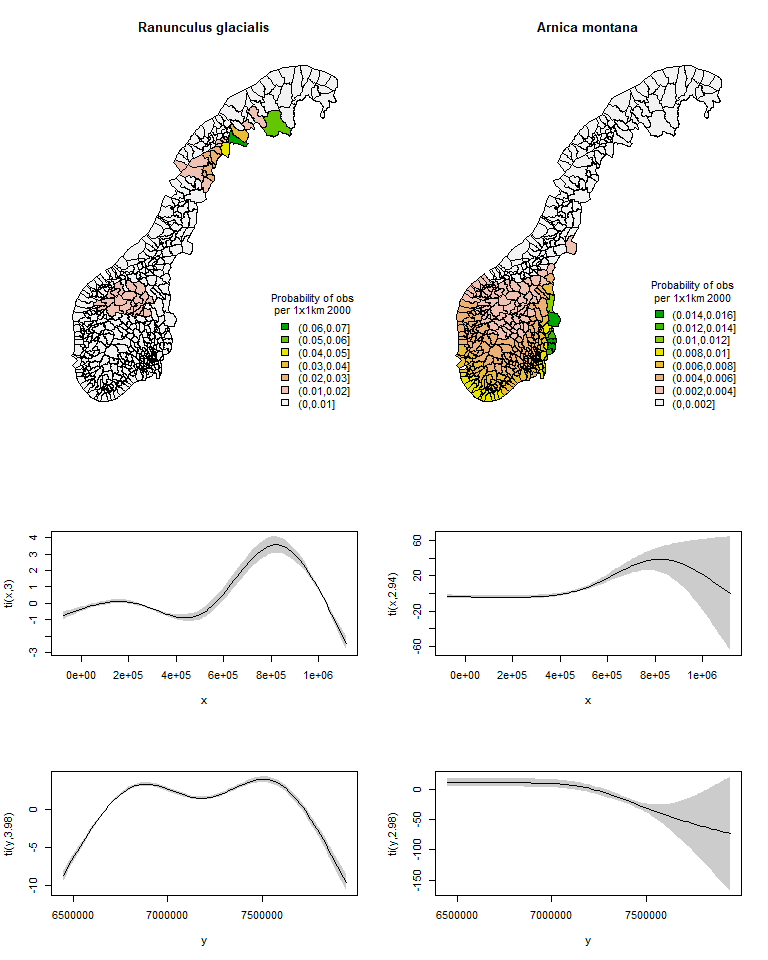
gam(Y ~ ti(x) + ti(y) + ti(year) + ti(x,year) + ti(y,year), family=poisson, offset=logS, data=d, gamma=3, select=TRUE)

where Y is the count of the species in five-year intervals, x and y are the spatial coordinates, time is represented by the first year of the time interval and logS is the natural logarithm of the sampling effort (unique visits) S in the time interval. These variables are all contained in the data set d (see Appendix 1). The model was fit to all available data for each species 1820-2019: to maximize data back in time no filtering on spatial precision was applied (in contrast to the assessment of land cover relations above). This is needed to estimate the reference state (initial attempts at filtering the data on precision lead to highly uncertainty backwards in time, and unrealistic trends in space and time). Low accuracy of observations may be less of a problem here than in many other applications: The resolution of NI is municipalities, which corresponds to the precision of some of the older occurrence data. Nevertheless, to reduce sensitivity to artificial patterns in the data, three-way interactions between spatial coordinates (x,y) and time (year) in the tensor smoothing functions ti() were omitted, and the smoothing parameter gamma was set to three (more smoothing than default).

The R code for estimation of the indicator values, including data download and export of estimates to the NI database, is given in Appendix 1.

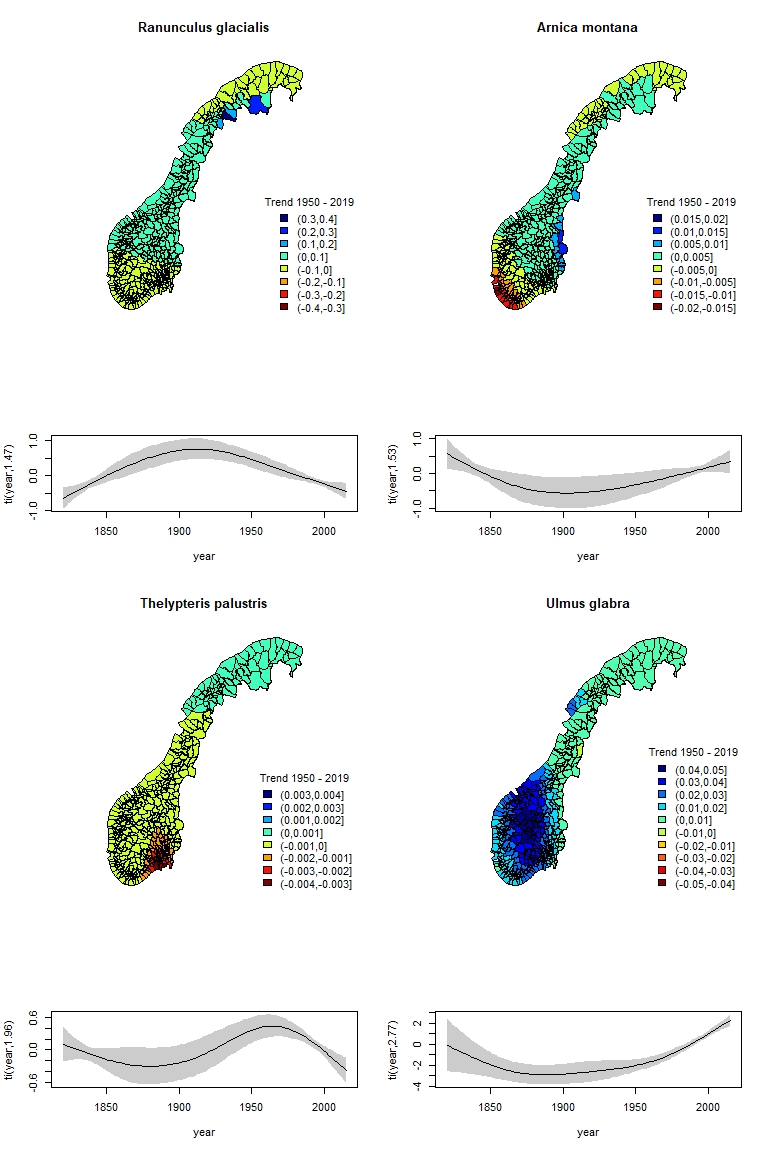
## Estimates of indicator values

The models suggest that there are spatiotemporal patterns in all species (see model summaries in Appendix 2). The models reproduced relatively complex well-known patterns in the distributions of the species. For instance, for the alpine herb *Ranunculus glacialis* the model predicts the bimodal distribution of the species in the southern and northern mountain massifs of Norway, as a result of bimodal response curves from west to east and south to north (Figure 1, left panels). For the lowland herb *Arnica montana*, the model predicts higher probability of occurrence towards the east and south, as a combined result of a unimodal response to the west-east gradient and a declining curve towards the north (Figure 1, right panels). Note that south-north and west-east patterns cannot be interpreted completely independently, as the northernmost areas of Norway are also furthest to the east. In the case of *Arnica montana* this results in higher prediction uncertainty towards the east and north, because the species is found in the southeast of the country, but not in the northeast. Unclear trends along individual spatial axes occur in several of the indicator species with more complex or limited distribution patterns, especially when moving (far) outside the known distribution of the species. However, in the Nature Index, these predictions are filtered away: only predictions within the known distribution (the definition area; see www.naturindeks.no) are used to calculate the index.



*Figure 1. Maps showing the estimated states of two indicator species in the year 2000. For each species the main effects of spatial coordinates west-east (x) and south-north (y) are plotted underneeth the map (tensor smooth, with estimated degrees of freedom) with confidence intervals (shaded areas).*

The models suggest that most species do not have strong overall temporal trends, but that there is spatial variability in the trends (Figure 2). Some of this is as expected; see the discussion of species distribution and trends in relation to climate and pressures below. However, there are also some results that contradict expectations, at least at first glance. For instance, for *Ulmus glabra*, the model estimates an overall positive trend with time. In the latest national Red list for species in Norway (Henriksen & Hilmo 2015), *Ulmus glabra* is listed as vulnerable (VU) because of elm disease, browsing and competition with introduced trees. This may have caused population declines, at least locally, and may cause larger declines in the future. However, this negative trend is not apparent in the analyses of this study. Regardless of choice of reference time within the modeled time span, *Ulmus glabra* has a positive development in recent times. This may be because the reported declines are only local.



*Figure 2. Maps showing the difference in predicted indicator values between 1950 and 2019 for four selected indicator species. For each species the main effect of time (year) across the entire time period of the data is plotted underneeth the map (tensor smooth, with estimated degrees of freedom). Note that in some cases the map predictions are (far) outside the range of the species (e.g. Arnica montana, Thelypteris palustris and Ulmus glabra in Northern Norway), and that such predictions are filtered away when calculating the Nature Index.*

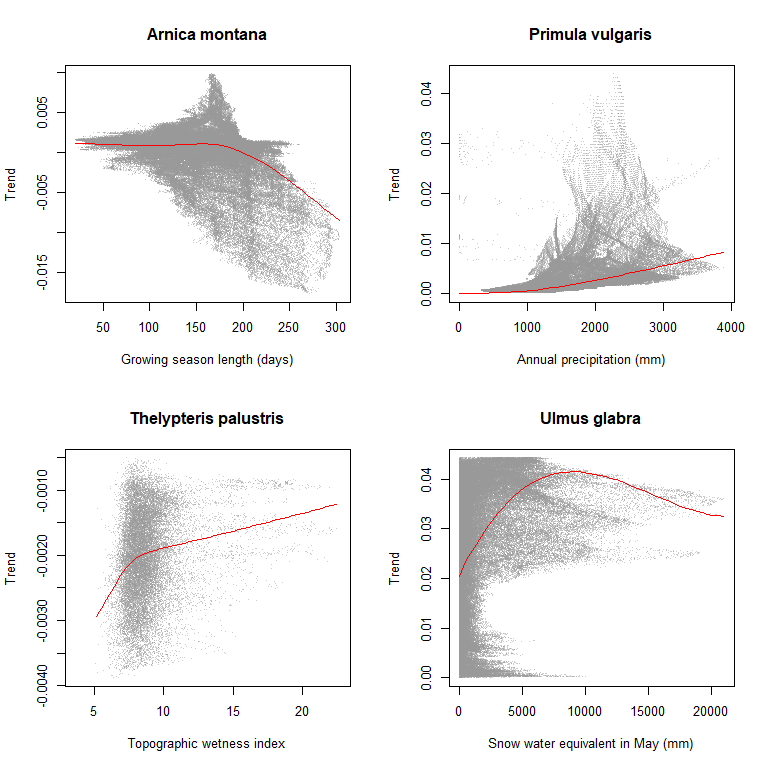
# 4. Assessing trends and pressures

This section considers a (limited) selection of pressures in relation to estimated trends (model predictions) in the state of the indicator species from 1950 to 2019. The results are based on comparisons of predicted trends across the entire grid within the convex polygon covering 99% of the observations of the species. The number of cells with predictions within these areas are still much larger than the number of cells with observational data. The relationships between trends and pressures are therefore not tested statistically (the degrees of freedom are strongly inflated).

There were no strong patterns in trends with dominant land cover (Table 3). A few possible patterns may still be worth a note. Four species had consistently positive mean trends for all land cover classes (*Ulmus glabra*, *Kalmia procumbens*, *Primula vulgaris* and *Drosera anglica*), and three species had consistently negative mean trends (*Thelypteris palustris*, *Carex vesicaria* and *Rhynchospora alba*). For the other ten species, the mean estimated trends varied across landcover types, but with relatively small amounts and high variability (not shown). Several species (e.g. *Moneses uniflora*, *Arnica montana*, *Carex pauciflora* and *Rhynchospora alba*) had low mean trends in developed and agricultural areas compared to areas dominated by other land cover types. This could be a random pattern (the variability is high), but may perhaps also reflect better performance of these species in areas with less intense human activity. There are also a few patterns that may reflect climate rather than land cover: *Ulmus glabra* has the strongest growth in areas dominated by glaciers and *Primula vulgaris* in marine areas (coast), which may be related to temperature and precipitation respectively. *Papaver radicatum* had the lowest growth rates in areas dominated by barren lands and glaciers. This may be related to better growing conditions for competitors in these areas due to increasing temperatures and/or less snow cover.

Table 3. Change from 1950 to 2019 in the predicted frequency (expressed as difference in percentage points) of the species in 1km by 1km cells with a given dominant land cover type within the minimum convex polygon covering 99% of the observations of the species (NaN means that the species is not observed in the land cover class).

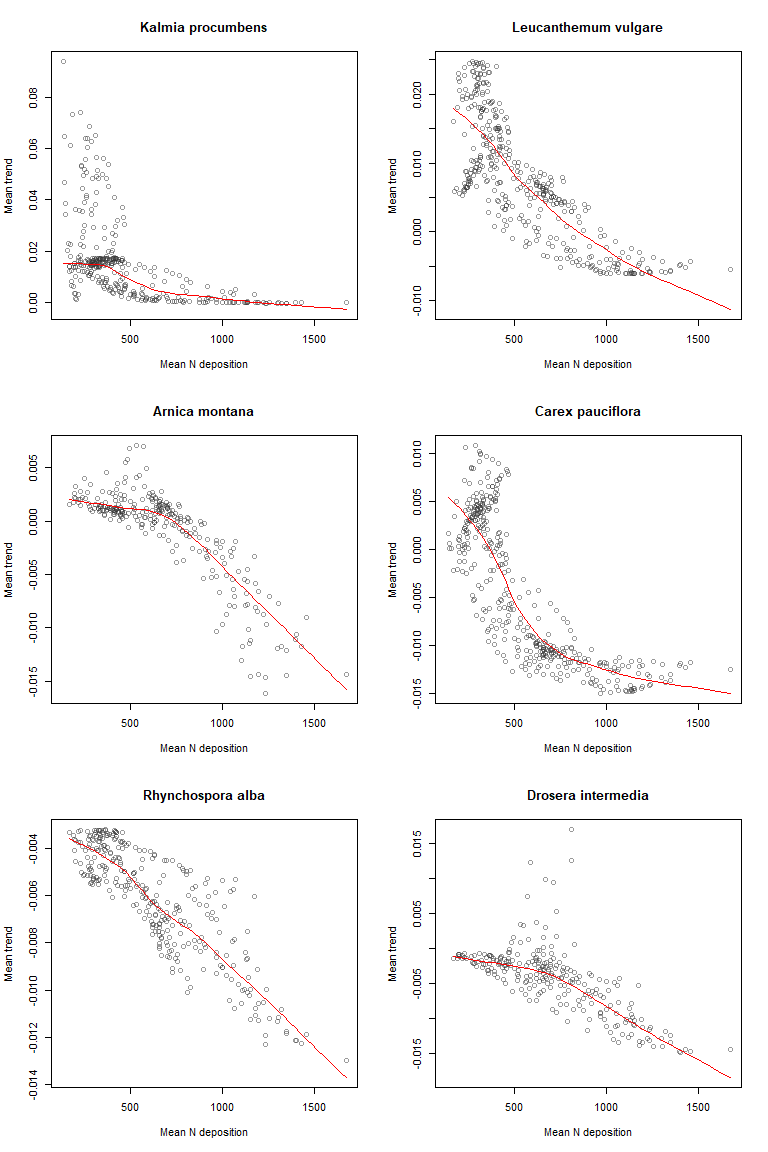
|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Developed | Agricultural | Forest | Barren | Mire | Glacier | Freshwater | Marine |
| Ulmus glabra | 2.01 | 1.89 | 2.25 | 2.63 | 1.82 | 3.63 | 2.38 | 1.89 |
| Papaver radicatum | -0.06 | -0.05 | -0.08 | -0.18 | -0.03 | -0.18 | -0.08 | -0.13 |
| Kalmia procumbens | 1.78 | 1.58 | 2.02 | 2.20 | 2.59 | 2.52 | 1.94 | 2.62 |
| Ranunculus glacialis | 0.82 | 0.54 | 1.34 | 2.58 | 1.01 | 2.48 | 3.59 | 0.62 |
| Primula vulgaris | 0.48 | 0.50 | 0.21 | 0.30 | 0.32 | 0.20 | 0.16 | 1.04 |
| Thelypteris palustris | -0.23 | -0.20 | -0.20 | -0.29 | -0.12 | NaN | -0.17 | -0.30 |
| Moneses uniflora | -0.49 | -0.51 | -0.29 | 0.05 | -0.09 | 0.01 | -0.26 | 0.03 |
| Leucanthemum vulgare | 0.54 | 0.86 | 0.89 | 0.95 | 1.55 | 0.80 | 0.84 | 1.29 |
| Erica cinerea | 0.09 | 0.08 | 0.12 | 0.27 | 0.34 | 0.09 | 0.10 | 0.37 |
| Carex vesicaria | -0.56 | -0.56 | -0.53 | -0.38 | -0.57 | -0.30 | -0.53 | -0.27 |
| Arnica montana | -0.14 | -0.08 | -0.02 | -0.04 | 0.18 | 0.06 | -0.01 | -0.33 |
| Carex pauciflora | -0.69 | -0.46 | -0.36 | -0.08 | 0.11 | 0.02 | -0.29 | 0.30 |
| Dactylorhiza incarnata | 0.01 | -0.02 | -0.01 | 0.04 | -0.10 | 0.20 | -0.03 | 0.06 |
| Rhynchospora alba | -0.70 | -0.65 | -0.59 | -0.48 | -0.47 | -0.38 | -0.57 | -0.62 |
| Rhynchospora fusca | -0.33 | -0.23 | -0.33 | -0.27 | -0.14 | -0.22 | -0.33 | -0.31 |
| Drosera anglica | 0.62 | 0.74 | 0.88 | 1.13 | 1.04 | 1.37 | 1.06 | 1.45 |
| Drosera intermedia | -0.48 | -0.37 | -0.46 | -0.43 | -0.25 | -0.25 | -0.42 | -0.60 |



*Figure 3. Trends for selected indicator species (difference in predicted frequency between 1950 and 2019) in relation to selected climate-related variables.*

The trends of some species may be related to different aspects of climate (Figure 3), which varies more systematically than land cover on broad geographical scales. Several species have higher trends in areas with short growing seasons within their distributions (*Ulmus glabra*, *Leucanthemum vulgare*, *Thelypteris palustris*, *Rhyncospora alba*, *Drosera intermedia* and *Arnica montana*, i.e. areas towards the north or higher elevations. This may indicate a climate-driven change in distribution: northern areas are known to warm faster (“Arctic amplification”), potentially allowing the species to expand northwards as growing conditions improve. The lower, and sometimes negative, trends in warmer areas may be due to increased competition from other native or introduced species (suspected for e.g. *Ulmus glabra*; Henriksen and Hilmo 2015), but also changing land use in the south, such as increased intensification of agricultural areas and regrowth of marginal areas (e.g. for *Arnica montana*). There are also examples of species with tendencies to more positive trends in areas with longer growing seasons (*Papaver radicatum*, *Primula vulgaris*).

For most species there was no clear variation in trend with total annual precipitation alone, but there are a few examples of contrasting trends that seem to be related to the geographical distribution of the species. For example, among the open lowland indicators the predominantly oceanic species *Erica cinerea* and *Primula vulgaris* show increasing trends with total annual precipitation (Figure 3), while other species with broader and more continental distributions, show less positive (*Laucanthemum vulgare*) or even negative trends (*Arnica montana*) with higher precipitation. Among the wetland indicators, the two widespread Carex species seem to respond in opposite ways: *Carex pauciflora* shows lower trends with higher precipitation, while (the slightly more continental) *Carex vesicaria* shows higher (less negative) trends with higher precipitation. These mixed results suggest that other aspects of water than total precipitation may be important, such as ground-related hydrology. For *Thelypteris palustris* we see that in areas with a high topographic wetness index, the trends have been less negative than in areas in drier parts of the landscape (Figure 3). With a finer spatial resolution, these patterns may have been even clearer, perhaps for other wetland species as well. Finally, for a few species with generally positive trends (*Leucanthemum vulgare* and *Ulmus glabra*), high snow cover in May seems to reduce the positive trends (Figure 3), suggesting that snow cover may limit the growing season and thus hamper the expansion of these species in snow-rich areas.



*Figure 4. Mean trends per municipality 1950-2019 for selected species in relation to mean annual nitrogen deposition 1950-2011.*

Several of the indicator species appear to have very clear negative relationships to patterns in nitrogen deposition (Figure 4). This may suggest that nitrogen has a strong effect on many of the species, possibly through competitive interactions. However, it is not perfectly clear that the observed patterns are in fact due to nutrient enrichment. In Norway, nitrogen deposition is strongly related to climatic patterns. Deposition is highest in the southwest, where the climate is warmest and wettest. As a consequence, relationships between species trends and N deposition patterns may resemble relationships between trends and climate variables (depending on the distribution of the species). In fact, all of the patterns plotted in Figure 4 to some degree resemble the patterns for total annual precipitation (although the patterns for precipitation appear less clear due to the much higher spatial resolution). Moreover, because of the relationship between N deposition and beneficial climate, N deposition may also be confounded with high land-use intensity in agricultural and (semi-)urban areas. The low trends for *Leuchantemum vulgare* and the negative trends for *Arnica montana* at high N deposition levels may be due to stronger land-use intensification in highly productive agricultural regions, as well as abandonment of less productive lands in the same regions.

In conclusion, although the effects of pressures have not been formally tested statistically, there are indications in the model results of negative effects of various pressures on several of the species (Table 4). For some species, this study finds few indications of expected effects, such as climatic effects on the alpine species *Ranunculus glacialis* and land use effects on the open lowland species *Primula vulgaris*. However, many of the indicated effects are in line with expectations.

Table 4. Overview of pressures that may affect indicator species negatively, as suggested by expert interpretation (not statistical tests) of the results of this study. Land use: drainage (D), intensification of agriculture/forestry (I). Climate: increased temperature/growing season (T), increased precipitation (P), reduced snow cover (S). Pollution: nitrogen deposition (N).

|  |  |  |  |
| --- | --- | --- | --- |
|  | Land use | Climate | Pollution |
| Ulmus glabra |  |  |  |
| Papaver radicatum |  | T/S |  |
| Kalmia procumbens |  |  | N |
| Ranunculus glacialis |  |  |  |
| Primula vulgaris |  |  |  |
| Thelypteris palustris | D | T |  |
| Moneses uniflora | I |  | N |
| Leucanthemum vulgare | I | T/P | N |
| Erica cinerea |  |  |  |
| Carex vesicaria | D |  |  |
| Arnica montana | I | T/P | N |
| Carex pauciflora | I | T/P | N |
| Dactylorhiza incarnata |  |  |  |
| Rhynchospora alba | I | T/P | N |
| Rhynchospora fusca |  |  | N |
| Drosera anglica |  |  |  |
| Drosera intermedia |  | T | N |

# 5. Perspectives and conclusions

The current analyses have produced more precise estimates of indicator states than previous methods. The analyses have also revealed some new patterns, as discussed above. We should of course beware that with such noisy data, we run the risk of identifying patterns in the noise rather than in the signal. The uneven sampling in space and time, with lower sampling in remote areas and early years, is particularly challenging. To reduce sensitivity to sampling bias, the models were deliberately made smoother than default (gamma=3) for all species. This reduces the risk of detecting false trends, but at the same time may hide real trends for some species.

One opportunity to obtain more information from the data could be to fine-tune the GAMs for each species individually. Adding elevation may improve predictions for several species, but given the coarse spatial resolution of the NI, where predictions are averaged across municipalities with large elevation ranges, it is not clear that this added complexity is worthwhile. Another possibility is further data treatment, such as other resolution in space and time and separate treatment of different data types (e.g. single observations, checklist data, systematic data), although some data types (e.g. checklists) are only available in recent times.

There are several alternative approaches for modelling population states and trends, and as computing power develops, increasingly complex modelling of high-resolution observation and population processes at high resolution (e.g. Wikle & Hooten 2010) will be feasible for many species. A key issue for further progress is improved representation of observation processes. Appropriate modelling of the different kinds of observation practices (corresponding to the different data types above), will most likely also improve the underlying population process models.

## Conclusions

The analyses above suggest that GBIF data largely represent NI vascular plant indicators in the right environment. The models produce sensible patterns, both in terms of the geographical distributions of species, and in terms of temporal trends and spatial variation in trends, and current methods produce more precise estimates than previous methods. The assessment of estimated trends and pressures suggest realistic relationships between species trends and relevant pressures.

These results come with some warnings. There is substantial unexplained variation for all species, which may be related to the representation and separation of process variation and observation error as well as a lack of high-quality representation of covariates. Some presumptive patterns within the range of the data may be caused by observation error, and predictions (far) outside the range of observations should be interpreted with great caution. Finally, based on the current analysis it may be difficult to interpret pressures independently. For instance, land use is related to temperature (e.g. growing season) and N deposition is related to both climate and land use. To address these challenges, we will need higher resolution, local observational studies and experiments. The present data and analyses nevertheless produced results that should be of some value in building and interpreting a high-level biodiversity indicator like the Nature Index.

# References

Henriksen S. and Hilmo O. (eds.) 2015. Norwegian red list for species 2015. Norwegian Biodiversity Information Centre.  
Horvath P., Halvorsen R., Stordal F., Tallaksen L.M., Tang H. and Bryn A. 2019. Distribution modelling of vegetation types based on area frame survey data. Applied Vegetation Science 22(4):547-560. Lund E., Aas W., Høgåsen T. and Larssen T. 2012. Overskridelser av tålegrenser for forsuring og nitrogen for Norge–oppdatering med perioden 2007–2011. NIVA-rapport 6448-2012. Wikle C.K. 2003. Hierarchical models for predicting the spread of ecological processes. Ecology 84:1382-1394. Wikle C.K. and Hooten M. 2010. A general science-based framework for dynamical spatio-temporal models. Test 19:417-451. Wood S.N. 2017. Generalized additive models: an introduction with R. CRC press.

# Appendix 1: R code for indicator estimation

# Script for estimation of indicator values and references states for 17 vascular plants  
# for the Nature Index 2020  
  
######################################  
# Tools  
######################################  
  
# NIcalc pacakge for reading/writing to NI database  
install.packages(pkgs =   
 "https://github.com/NINAnor/NIcalc/releases/download/0.0.1.2/NIcalc\_0.0.1.2.zip",   
 repos = NULL)  
  
# Other libraries needed:  
# rgbif, rgdal, sp, raster, rio, mgcv  
  
  
#################################################  
# Data download and preparation  
#################################################  
  
  
### Indicator species names  
art <- c("alm","fjellvalmue","greplyng","issoleie","kusymre","myrtelg","olavsstake","prestekrage",  
 "purpurlyng","sennegras","solblom","sveltstarr","engmarihand","hvitmyrak","brunmyrak","smalsoldogg","dikesoldogg")  
species <- c("Ulmus glabra","Papaver radicatum","Kalmia procumbens","Ranunculus glacialis","Primula vulgaris",  
 "Thelypteris palustris","Moneses uniflora","Leucanthemum vulgare","Erica cinerea","Carex vesicaria",  
 "Arnica montana","Carex pauciflora","Dactylorhiza incarnata","Rhynchospora alba","Rhynchospora fusca",  
 "Drosera anglica","Drosera intermedia")  
  
  
### Retreive and store old data from NI database  
library(NIcalc)  
NIcalc::getToken(username=myUserName, # Specify your user name  
 password=myPassword, # Specify your password  
 url="https://www8.nina.no/NaturindeksNiCalc")source('species.r')  
myIndicators <- NIcalc::getIndicators()   
myIndicators <- myIndicators[myIndicators$id!=3,] # removing one indicator which is estimated with other data and methods (3: Alge på bjørk)  
myIndicators <- cbind(myIndicators,species=species)  
oldIndicatorData <- list()  
for(i in 1:length(species))  
{  
 oldIndicatorData[[i]] <- NIcalc::getIndicatorValues( indicatorID = myIndicators$id[i])   
}  
names(oldIndicatorData) <- myIndicators$name  
save(oldIndicatorData,file="oldIndicatorData")  
  
  
### Download occurrence data from GBIF  
# This part is based on the extensively commented instructions for  
# asynchronous downloading of GBIF data by Anders Finstad, available here:  
# https://gbif-europe.github.io/nordic\_oikos\_2018\_r/s3\_gbif\_demo/3.x\_async\_download\_gbif.Rmd  
# This allows downloading larger data sets, and citation of the download with a single doi.   
  
# Libraries  
library(rgbif)  
library(rio)  
library(sp)  
library(raster)  
  
# Set up a user profile at GBIF (https://www.gbif.org), then  
# run this once to store your GBIF user credentials to your R session   
options(gbif\_user=rstudioapi::askForPassword("my gbif username"))  
options(gbif\_email=rstudioapi::askForPassword("my registred gbif e-mail"))  
options(gbif\_pwd=rstudioapi::askForPassword("my gbif password"))  
  
# Download all vascular plants  
# Find a taxonkey - get list of gbif keys to filter download  
key <- name\_suggest(q='Plantae', rank='kingdom')$key[1]   
  
# Get download key for all occurrences of plants with coordinates in Norway  
download\_key <-   
 occ\_download(  
 paste('taxonKey',key,sep=" = "),  
 'hasCoordinate = TRUE',  
 "country = NO",  
 type = "and"  
 ) %>%   
 occ\_download\_meta  
  
# Download data  
path <- "C:/Users/oskarpaa/Documents/R/R/NI 2020"  
occ\_download\_get(key=download\_key$key,path=path)  
  
# Citation - copy into documentation  
paste("GBIF Occurrence Download", download\_key[2], "accessed via GBIF.org on", Sys.Date())  
  
# Open data and extract into data frame  
# Get a list of the files within the archive by using "list=TRUE" in the unzip function  
download\_path <- paste(path,"/",download\_key$key,".zip",sep="")  
archive\_files <- unzip(download\_path, files = "NULL", list = T)   
archive\_files  
  
# # Get the occurrence.txt file in as a dataframe (using import from rio)  
# sp <- import(unzip(download\_path,files="occurrence.txt"),header=T,sep="\t")  
# dim(sp)  
  
# Data import from downloaded occurrence file  
sp <- import("occurrence.txt") #,header=T,sep="\t")  
dim(sp)  
head(sp)  
  
# Select fields  
# Only species, geographical coordinates (with uncertainty/precision) and time (year, month, day) are  
# needed for the modelling, but other fields may be useful for error checking etc.  
selectedFields <- c( "institutionID","collectionID","catalogNumber",  
 "basisOfRecord","contributor",  
 "species","scientificName","taxonID","taxonKey",  
 "year","month","day",  
 "countryCode","county","municipality",  
 "decimalLongitude","decimalLatitude","coordinateUncertaintyInMeters","coordinatePrecision")   
sp <- sp[,selectedFields]  
dim(sp)  
  
# Remove obs with missing dates and/or coordinates (shouldn't be necessary when "has coordinate"=TRUE, but quite a few long and lat are missing...)  
sp <- sp[complete.cases(sp[,c("year","month","day","decimalLongitude","decimalLatitude")]),]  
dim(sp)  
  
# Convert lat-long coordinates to coordinate system of Norway raster  
occ\_points <- data.frame(x=sp$decimalLongitude,y=sp$decimalLatitude)  
occ\_points <- SpatialPoints(occ\_points,proj4string=CRS("+proj=longlat +datum=WGS84"))  
occ\_UTM33 <- spTransform(occ\_points,CRS("+proj=utm +zone=33 ellps=GRS80 +units=m"))  
sp$x <- occ\_UTM33$x  
sp$y <- occ\_UTM33$y  
  
# Creating a spatial points data frame, and adding unique date (better than eventDate, which may be misleading if occurrences with month only are recorded on the first day)  
#load("occurrences cleaned") # for high-precision data  
occ\_UTM33 <- SpatialPointsDataFrame(data.frame(x=sp$x,y=sp$y),data=data.frame(sp))  
occ\_UTM33$year\_month\_day <- paste(occ\_UTM33$year,occ\_UTM33$month,occ\_UTM33$day)  
  
  
  
### Rasterize occurrence data  
  
library(raster)  
  
# Set time intervals  
start.year <- seq(1820,2015,by=5) # 5-year intervals from 1820 to 2020  
  
# Rasterize sampling effort in time intervals  
norway <- raster("Data/Raster/Norway.tif") # Background raster of Norway (all values = 1)  
samp\_ras <- stack(norway)  
for(i in 1:length(start.year))  
{  
 print(start.year[i])  
 records.in.interval <- (occ\_UTM33$year >= start.year[i]) & (occ\_UTM33$year < start.year[i]+5)  
 print(table(records.in.interval))  
 if(!any(records.in.interval)) samp\_ras[[i]] <- norway\*0  
 else samp\_ras[[i]] <- norway\*rasterize(occ\_UTM33[records.in.interval,],norway,field="year\_month\_day",fun=function(x,...){length(unique(x))},background=0) # raster with counts of sampling effort in each cell of norway  
}  
names(samp\_ras) <- paste("t",start.year,sep=".")  
save(samp\_ras,file="Data/Raster/samp\_ras\_all") # unfiltered data  
  
# Rasterize species records in time intervals  
source("species.r")  
occ\_ras\_list <- list()  
#occ\_species\_list <- list()  
for(j in 1:length(species))  
{  
 key <- name\_suggest(q=species[j], rank='species')$key  
 occ\_species <- occ\_UTM33[occ\_UTM33$taxonKey%in%key,]  
 cat(species[j],nrow(occ\_species),"\n")  
# occ\_species\_list[[j]] <- occ\_species  
 occ\_ras\_list[[j]] <- stack(norway)  
 for(i in 1:length(start.year))  
 {  
 records.in.interval <- (occ\_species$year >= start.year[i]) & (occ\_species$year < start.year[i]+5)  
 cat(start.year[i],sum(records.in.interval),"\n")  
 if(!any(records.in.interval)) occ\_ras\_list[[j]][[i]] <- norway\*0  
 else occ\_ras\_list[[j]][[i]] <- norway\*rasterize(occ\_species[records.in.interval,],norway,field="year\_month\_day",fun=function(x,...){length(unique(x))},background=0) # raster with counts of occurrences in each cell of norway  
 }  
 names(occ\_ras\_list[[j]]) <- paste("t",start.year,sep=".")  
}  
names(occ\_ras\_list) <- species  
save(occ\_ras\_list,file="Data/Raster/occ\_ras\_list\_all") # unfiltered data  
  
  
# Build training data sets for distribution modelling with  
# all data (no filter on precision), for all times with continuous sampling (1820 onwards)  
load("Data/Raster/occ\_ras\_list\_all") # unfiltered data  
load("Data/Raster/samp\_ras\_all") # unfiltered data  
year <- seq(1820,2015,by=5)  
yr <- paste("t.",year,sep="")  
for(j in 1:length(species))  
{  
 cat(species[j],"\n")  
 for(k in 1:length(year))  
 {  
 cat(year[k],"\n")  
   
 # Extract occurrence and sampling rasters  
 o\_ras <- occ\_ras\_list[[j]][[yr[k]]]  
 s\_ras <- samp\_ras[[yr[k]]]  
   
 # Take the occurrence cells as presences  
 presences <- which(values(o\_ras)>0)  
   
 # Take cells with sampling events of some species  
 # but without occurrence observations of this particular species  
 absences <- which((values(s\_ras)>0) & (values(o\_ras)==0))  
   
 # # Sample absences if too many?  
 # absences\_sample <- sample(absences,size=length(presences)) # sample of same number of absence cells as presence cells  
   
 # Combine presences, absences and environmental data  
 selected <- c(presences,absences)  
 #selected <- c(presences,absences\_sample)  
 xy <- raster::coordinates(o\_ras)[selected, ]  
 data <- data.frame(xy,Y=values(o\_ras)[selected],logS=log(values(s\_ras)[selected]),  
 year=rep(year[k],nrow(xy)))  
 # presence <- as.numeric(data$Y>0)  
 # data <- cbind(presence,data) # MIAmaxent wants presence as the first column  
   
 if(k==1) training\_data <- data  
 else training\_data <- rbind(training\_data,data)  
 }  
 # # Convert discrete environmental predictors to factor variables  
 # training\_data$ar50artype <- factor(training\_data$ar50artype)  
 # training\_data$geonorge123 <- factor(training\_data$geonorge123)  
   
 # Save training data  
 save(training\_data,file=paste("Data/Regression data/",species[j],"\_training\_data\_all",sep=""))  
}  
  
  
  
################################################  
# Modelling  
################################################  
# GAM spacetime model with geographic coordinates (x and y in meters), year (5-year intervals) and interactions (x:year and y:year).  
# Sampling effort as offset. No environmental covariates.  
  
# GAM: all data (no spatial precision filtering) and smoothing parameter gamma=3  
library(mgcv)  
gam.results <- list()  
for(j in 1:length(species))  
{  
 cat(species[j],"\n")  
   
 # Training data for species from GBIF, see NI 2020 plants dataprep.r  
 load(paste("Data/Regression data/",species[j],"\_training\_data\_all",sep=""))  
 d <- training\_data[,c("Y","x","y","year","logS")]  
 m <- gam(Y~ti(x)+ti(y)+ti(year)+ti(x,year)+ti(y,year),  
 data=d,gamma=3,family=poisson,offset=logS,select=TRUE)  
 gam.results[[j]] <- m  
}  
save(gam.results,file="Results/gam.results.all.gamma3")  
  
  
  
################################################  
# Prediction: NI indicator value estimation  
################################################  
  
# GAM predictions: municipality maps for specific NI years  
library(rgdal)  
kommune.poly <- readOGR(dsn="Data/Shapefiles",layer="Norway municipalities",encoding="")  
year <- c(1900,1950,1990,2000,2010,2014,2019)  
xy <- coordinates(norway)  
x <- y <- norway  
values(x) <- xy[,1]  
values(y) <- xy[,2]  
NIGAM\_All.list <- list()  
pdf("GAMplotsMunicAll.pdf")  
# load("Results/gam.results.all.gamma3") # Uncomment to load results from previous steps, if entire script is not run continuously  
for(j in 1:length(species))  
{  
 print(species[j])  
 NIGAM\_All.list[[j]] <- list()  
 for(i in 1:length(year))  
 {  
 print(year[i])  
 pred.ras <- stack(year[i]\*norway,x,y,log(norway))  
 names(pred.ras)[1:5] <- c("year","x","y","logS")  
 pred.dat <- as.data.frame(values(pred.ras))  
 p <- norway  
 pred <- predict(gam.results[[j]],pred.dat,se.fit=TRUE,type="response") # The ordinary raster prediction does not always work: p <- predict(pred.ras,glm.results[[j]],type="response")  
 values(p) <- as.vector(pred$fit)  
 p.poly <- extract(p,kommune.poly,sp=TRUE,fun=mean,na.rm=T,weights=TRUE,normalizeWeights=TRUE)  
 b <- seq(0,0.1,by=0.001)  
 n <- length(b)  
 plot(p.poly,col=rev(terrain.colors(n-1))[cut(p.poly$layer,breaks=b)],main=paste(species[j],year[i]))  
 values(p) <- as.vector(pred$se.fit)  
 p.poly.se <- extract(p,kommune.poly,sp=TRUE,fun=mean,na.rm=T,weights=TRUE,normalizeWeights=TRUE)  
 NIGAM\_All.list[[j]][[i]] <- list(p=p.poly,p.se=p.poly.se)  
 }  
}  
dev.off()  
names(NIGAM\_All.list) <- species  
  
# Build data frame with new NI indicator values (GAM predictions per municipality)  
#load("oldIndicatorData") # Uncomment to load results from previous steps, if entire script is not run continuously  
newIndicatorData <- oldIndicatorData  
for(j in 1:length(species))  
{  
 print(species[j])  
   
 # Make data objects of old and new indicator sets for further manipulation  
 old <- oldIndicatorData[[j]]$indicatorValues  
 new <- newIndicatorData[[j]]$indicatorValues  
  
 selected.year <- old$yearName=="Referanseverdi"  
 oldref <- old[selected.year,]  
 old$ref <- oldref$verdi[match(old$areaName,oldref$areaName)]  
  
 newref.poly <- NIGAM\_All.list[[j]][[1]]$p  
 newref.se.poly <- NIGAM\_All.list[[j]][[1]]$p.se  
 r <- match(newref.poly$NAVN,oldref$areaName)  
 new$verdi <- NA  
 new$verdi[selected.year][r[!is.na(r)]] <- newref.poly$layer[!is.na(r)]  
 new$verdiSE <- NA  
 new$verdiSE[selected.year][r[!is.na(r)]] <- newref.se.poly$layer[!is.na(r)]  
 newref <- new[new$yearName=="Referanseverdi",]  
 new$ref <- newref$verdi[match(new$areaName,newref$areaName)]  
 for(i in 2:(length(year)-1))  
 {  
 selected.year <- old$yearName==as.character(year[i])  
 oldval <- old[selected.year,]  
 newval <- NIGAM\_All.list[[j]][[i]]$p  
 newval.se <- NIGAM\_All.list[[j]][[i]]$p.se  
 o <- match(newval$NAVN,oldval$areaName)  
 new$verdi[selected.year][o[!is.na(o)]] <- newval$layer[!is.na(o)]  
 new$verdiSE[selected.year][o[!is.na(o)]] <- newval.se$layer[!is.na(o)]  
 }  
 def <- old$ref/old$ref # indicator value for definition area (1) or not (NA) - used to remove new predictions outside definition area in next three lines  
 new$ref <- new$ref\*def  
 new$verdi <- new$verdi\*def  
 new$verdiSE <- new$verdiSE\*def  
 newIndicatorData[[j]]$indicatorValues <- new  
}  
  
# Create distribution objects and format data for import to NI base (based on code from NIcalc vignette)  
library(NIcalc)  
updatedIndicatorData <- oldIndicatorData  
for(j in 1:length(species))  
{  
 print(species[j])  
  
 # Create distributions  
 d <- newIndicatorData[[j]]$indicatorValues  
 myData <- data.frame(estimatedStates = d$verdi,  
 standardErrors = d$verdiSE)  
 logNormalParams <- NIcalc::normal2Lognormal(mean = myData$estimatedStates,   
 sd = myData$standardErrors)  
 myData$muLogNormal <- logNormalParams$mean   
 myData$sigmaLogNormal <- logNormalParams$sd  
   
 ddd <- NULL  
 for (i in 1:dim(myData)[[1]])  
 {  
 if(any(is.na(myData[i,c("muLogNormal","sigmaLogNormal")]))) {ddd[i] <- NA; next}  
 ddd[i] <- list(NIcalc::makeDistribution(  
 input = "logNormal",  
 distParams = list(mean = myData$muLogNormal[i],  
 sd = myData$sigmaLogNormal[i])))   
 }  
 myData$distrObjects <- ddd  
 myData$areaIDs <- d$areaId  
 myData$years <- d$yearName  
 myData$Datatype <- d$datatypeId  
   
 rowsWithNAs <- which(is.na(myData$estimatedStates))  
 rowsWithoutNAs <- which(!is.na(myData$estimatedStates))  
 if(any(rowsWithNAs))  
 {  
 for(i in rowsWithNAs)  
 {  
 updatedIndicatorData[[j]] <- NIcalc::setIndicatorValues(updatedIndicatorData[[j]],  
 areaId = myData$areaIDs[i],  
 years = myData$years[i],  
 est = myData$estimatedStates[i],  
 lower = myData$estimatedStates[i]-myData$standardErrors[i],  
 upper = myData$estimatedStates[i]+myData$standardErrors[i])  
 }  
 }  
 for(i in rowsWithoutNAs)  
 {  
 updatedIndicatorData[[j]] <- NIcalc::setIndicatorValues(updatedIndicatorData[[j]],   
 areaId = myData$areaIDs[i],   
 years = myData$years[i],   
 distribution = myData$distrObjects[[i]],  
 datatype = myData$Datatype[i])  
 }   
}  
  
# Write updated indicator data to NI database  
# (unitOfMeasurement needs to be updated manually. New units: "Funnsannsynlighet i 1x1km-ruter")  
# Get token for writing to NI database if script is not run continuously (see above).  
# Code below commented to avoid accidential overwriting of data in the NI database.  
# Uncomment to write to database:  
# for(j in 1:length(species))  
# {  
# print(species[j])  
# NIcalc::writeIndicatorValues(updatedIndicatorData[[j]])  
# }  
  
# Check uploaded data by downloading the same data sets  
for(j in 1:length(species))  
{  
 print(species[j])  
 d1 <- updatedIndicatorData[[j]]$indicatorValues  
 indicatorData <- NIcalc::getIndicatorValues(indicatorID = myIndicators$id[myIndicators$species==species[j]])   
 d2 <- indicatorData$indicatorValues  
 print(head(d1))  
 print(head(d2))  
 check\_all <- data.frame(d1$verdi,d2$verdi)  
 check\_all$check <- check\_all[,1]/check\_all[,2]  
 print("Check all = 1")  
 print(summary(check\_all$check))  
}

# Appendix 2: Summaries of GAMs

names(gam.results) <- species  
lapply(gam.results,summary)

## $`Ulmus glabra`  
##   
## Family: poisson   
## Link function: log   
##   
## Formula:  
## Y ~ ti(x) + ti(y) + ti(year) + ti(x, year) + ti(y, year)  
##   
## Parametric coefficients:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) -6.744 0.260 -25.93 <2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Approximate significance of smooth terms:  
## edf Ref.df Chi.sq p-value   
## ti(x) 2.233 4 346.42 <2e-16 \*\*\*  
## ti(y) 3.994 4 109.57 <2e-16 \*\*\*  
## ti(year) 2.771 4 107.97 <2e-16 \*\*\*  
## ti(x,year) 5.402 16 75.93 <2e-16 \*\*\*  
## ti(y,year) 12.074 16 198.19 <2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## R-sq.(adj) = 0.0426 Deviance explained = 13.1%  
## UBRE = -0.80458 Scale est. = 1 n = 271396  
##   
## $`Papaver radicatum`  
##   
## Family: poisson   
## Link function: log   
##   
## Formula:  
## Y ~ ti(x) + ti(y) + ti(year) + ti(x, year) + ti(y, year)  
##   
## Parametric coefficients:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) -12.3044 0.4337 -28.37 <2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Approximate significance of smooth terms:  
## edf Ref.df Chi.sq p-value   
## ti(x) 3.998334 4 102.860 < 2e-16 \*\*\*  
## ti(y) 3.887569 4 106.278 < 2e-16 \*\*\*  
## ti(year) 2.338899 4 16.369 9.16e-06 \*\*\*  
## ti(x,year) 10.445312 16 55.269 4.28e-09 \*\*\*  
## ti(y,year) 0.002864 16 0.001 0.808   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## R-sq.(adj) = 0.0301 Deviance explained = 29.4%  
## UBRE = -0.99178 Scale est. = 1 n = 271396  
##   
## $`Kalmia procumbens`  
##   
## Family: poisson   
## Link function: log   
##   
## Formula:  
## Y ~ ti(x) + ti(y) + ti(year) + ti(x, year) + ti(y, year)  
##   
## Parametric coefficients:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) -5.63219 0.02759 -204.2 <2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Approximate significance of smooth terms:  
## edf Ref.df Chi.sq p-value   
## ti(x) 2.824 4 275.59 < 2e-16 \*\*\*  
## ti(y) 3.858 4 1346.19 < 2e-16 \*\*\*  
## ti(year) 2.829 4 267.44 < 2e-16 \*\*\*  
## ti(x,year) 2.345 16 43.52 6.08e-14 \*\*\*  
## ti(y,year) 7.245 16 146.14 < 2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## R-sq.(adj) = -0.0147 Deviance explained = 13.1%  
## UBRE = -0.89679 Scale est. = 1 n = 271396  
##   
## $`Ranunculus glacialis`  
##   
## Family: poisson   
## Link function: log   
##   
## Formula:  
## Y ~ ti(x) + ti(y) + ti(year) + ti(x, year) + ti(y, year)  
##   
## Parametric coefficients:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) -7.75276 0.09873 -78.52 <2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Approximate significance of smooth terms:  
## edf Ref.df Chi.sq p-value   
## ti(x) 2.999 4 441.70 < 2e-16 \*\*\*  
## ti(y) 3.985 4 1157.84 < 2e-16 \*\*\*  
## ti(year) 1.467 4 29.54 1.96e-09 \*\*\*  
## ti(x,year) 9.652 16 135.33 < 2e-16 \*\*\*  
## ti(y,year) 13.065 16 276.32 < 2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## R-sq.(adj) = 0.022 Deviance explained = 20.4%  
## UBRE = -0.93746 Scale est. = 1 n = 271396  
##   
## $`Primula vulgaris`  
##   
## Family: poisson   
## Link function: log   
##   
## Formula:  
## Y ~ ti(x) + ti(y) + ti(year) + ti(x, year) + ti(y, year)  
##   
## Parametric coefficients:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) -10.8300 0.5675 -19.08 <2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Approximate significance of smooth terms:  
## edf Ref.df Chi.sq p-value   
## ti(x) 1.9675 4 1762.063 < 2e-16 \*\*\*  
## ti(y) 2.2392 4 1064.481 < 2e-16 \*\*\*  
## ti(year) 2.9493 4 62.571 < 2e-16 \*\*\*  
## ti(x,year) 0.3872 16 1.603 0.0374 \*   
## ti(y,year) 1.8409 16 19.089 1.24e-05 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## R-sq.(adj) = -0.00109 Deviance explained = 32.3%  
## UBRE = -0.94919 Scale est. = 1 n = 271396  
##   
## $`Thelypteris palustris`  
##   
## Family: poisson   
## Link function: log   
##   
## Formula:  
## Y ~ ti(x) + ti(y) + ti(year) + ti(x, year) + ti(y, year)  
##   
## Parametric coefficients:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) -9.5805 0.2651 -36.13 <2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Approximate significance of smooth terms:  
## edf Ref.df Chi.sq p-value   
## ti(x) 2.483584 4 113.821 < 2e-16 \*\*\*  
## ti(y) 1.787053 4 117.260 < 2e-16 \*\*\*  
## ti(year) 1.958562 4 24.690 5.46e-11 \*\*\*  
## ti(x,year) 0.008009 16 0.007 0.32777   
## ti(y,year) 0.652338 16 3.997 0.00178 \*\*   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## R-sq.(adj) = 0.00628 Deviance explained = 14.2%  
## UBRE = -0.98443 Scale est. = 1 n = 271396  
##   
## $`Moneses uniflora`  
##   
## Family: poisson   
## Link function: log   
##   
## Formula:  
## Y ~ ti(x) + ti(y) + ti(year) + ti(x, year) + ti(y, year)  
##   
## Parametric coefficients:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) -5.66789 0.02661 -213 <2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Approximate significance of smooth terms:  
## edf Ref.df Chi.sq p-value   
## ti(x) 1.043 4 644.13 < 2e-16 \*\*\*  
## ti(y) 3.986 4 1522.44 < 2e-16 \*\*\*  
## ti(year) 2.599 4 239.03 < 2e-16 \*\*\*  
## ti(x,year) 6.831 16 38.25 5.77e-09 \*\*\*  
## ti(y,year) 6.615 16 327.79 < 2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## R-sq.(adj) = -0.0178 Deviance explained = 10%  
## UBRE = -0.90386 Scale est. = 1 n = 271396  
##   
## $`Leucanthemum vulgare`  
##   
## Family: poisson   
## Link function: log   
##   
## Formula:  
## Y ~ ti(x) + ti(y) + ti(year) + ti(x, year) + ti(y, year)  
##   
## Parametric coefficients:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) -4.31603 0.01511 -285.7 <2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Approximate significance of smooth terms:  
## edf Ref.df Chi.sq p-value   
## ti(x) 3.806 4 875.10 <2e-16 \*\*\*  
## ti(y) 2.749 4 85.14 <2e-16 \*\*\*  
## ti(year) 3.379 4 805.79 <2e-16 \*\*\*  
## ti(x,year) 2.119 16 101.08 <2e-16 \*\*\*  
## ti(y,year) 11.493 16 270.04 <2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## R-sq.(adj) = 0.0101 Deviance explained = 6.87%  
## UBRE = -0.76225 Scale est. = 1 n = 271396  
##   
## $`Erica cinerea`  
##   
## Family: poisson   
## Link function: log   
##   
## Formula:  
## Y ~ ti(x) + ti(y) + ti(year) + ti(x, year) + ti(y, year)  
##   
## Parametric coefficients:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) -13.613 0.654 -20.82 <2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Approximate significance of smooth terms:  
## edf Ref.df Chi.sq p-value   
## ti(x) 1.22807 4 1145.688 < 2e-16 \*\*\*  
## ti(y) 2.08044 4 531.848 < 2e-16 \*\*\*  
## ti(year) 0.01223 4 0.021 0.10290   
## ti(x,year) 0.01281 16 0.023 0.08750 .   
## ti(y,year) 1.14062 16 5.824 0.00377 \*\*   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## R-sq.(adj) = -0.531 Deviance explained = 45.9%  
## UBRE = -0.9769 Scale est. = 1 n = 271396  
##   
## $`Carex vesicaria`  
##   
## Family: poisson   
## Link function: log   
##   
## Formula:  
## Y ~ ti(x) + ti(y) + ti(year) + ti(x, year) + ti(y, year)  
##   
## Parametric coefficients:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) -5.52407 0.02563 -215.5 <2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Approximate significance of smooth terms:  
## edf Ref.df Chi.sq p-value   
## ti(x) 3.973 4 948.97 < 2e-16 \*\*\*  
## ti(y) 3.681 4 584.45 < 2e-16 \*\*\*  
## ti(year) 2.641 4 318.67 < 2e-16 \*\*\*  
## ti(x,year) 7.048 16 34.86 6.72e-08 \*\*\*  
## ti(y,year) 6.987 16 53.26 2.34e-12 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## R-sq.(adj) = -0.031 Deviance explained = 7.12%  
## UBRE = -0.89356 Scale est. = 1 n = 271396  
##   
## $`Arnica montana`  
##   
## Family: poisson   
## Link function: log   
##   
## Formula:  
## Y ~ ti(x) + ti(y) + ti(year) + ti(x, year) + ti(y, year)  
##   
## Parametric coefficients:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) -12.092 3.406 -3.55 0.000385 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Approximate significance of smooth terms:  
## edf Ref.df Chi.sq p-value   
## ti(x) 2.941 4 331.846 < 2e-16 \*\*\*  
## ti(y) 2.982 4 567.086 < 2e-16 \*\*\*  
## ti(year) 1.525 4 7.022 5.65e-08 \*\*\*  
## ti(x,year) 6.780 16 147.472 < 2e-16 \*\*\*  
## ti(y,year) 2.935 16 101.257 < 2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## R-sq.(adj) = -0.0163 Deviance explained = 9.01%  
## UBRE = -0.91265 Scale est. = 1 n = 271396  
##   
## $`Carex pauciflora`  
##   
## Family: poisson   
## Link function: log   
##   
## Formula:  
## Y ~ ti(x) + ti(y) + ti(year) + ti(x, year) + ti(y, year)  
##   
## Parametric coefficients:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) -4.70377 0.01629 -288.7 <2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Approximate significance of smooth terms:  
## edf Ref.df Chi.sq p-value   
## ti(x) 2.710 4 36.14 1.99e-09 \*\*\*  
## ti(y) 3.996 4 490.62 < 2e-16 \*\*\*  
## ti(year) 3.853 4 979.20 < 2e-16 \*\*\*  
## ti(x,year) 2.990 16 34.61 3.27e-08 \*\*\*  
## ti(y,year) 9.113 16 421.16 < 2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## R-sq.(adj) = -0.0535 Deviance explained = 5.41%  
## UBRE = -0.81749 Scale est. = 1 n = 271396  
##   
## $`Dactylorhiza incarnata`  
##   
## Family: poisson   
## Link function: log   
##   
## Formula:  
## Y ~ ti(x) + ti(y) + ti(year) + ti(x, year) + ti(y, year)  
##   
## Parametric coefficients:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) -6.41999 0.05741 -111.8 <2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Approximate significance of smooth terms:  
## edf Ref.df Chi.sq p-value   
## ti(x) 2.986 4 151.13 < 2e-16 \*\*\*  
## ti(y) 2.719 4 772.73 < 2e-16 \*\*\*  
## ti(year) 2.533 4 96.84 < 2e-16 \*\*\*  
## ti(x,year) 1.542 16 21.82 3.09e-07 \*\*\*  
## ti(y,year) 5.439 16 82.94 < 2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## R-sq.(adj) = 0.00252 Deviance explained = 12.4%  
## UBRE = -0.93901 Scale est. = 1 n = 271396  
##   
## $`Rhynchospora alba`  
##   
## Family: poisson   
## Link function: log   
##   
## Formula:  
## Y ~ ti(x) + ti(y) + ti(year) + ti(x, year) + ti(y, year)  
##   
## Parametric coefficients:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) -7.7430 0.4943 -15.66 <2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Approximate significance of smooth terms:  
## edf Ref.df Chi.sq p-value   
## ti(x) 2.901 4 270.59 < 2e-16 \*\*\*  
## ti(y) 2.847 4 544.84 < 2e-16 \*\*\*  
## ti(year) 2.897 4 218.02 < 2e-16 \*\*\*  
## ti(x,year) 3.833 16 116.87 < 2e-16 \*\*\*  
## ti(y,year) 2.991 16 32.91 4.22e-08 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## R-sq.(adj) = -0.035 Deviance explained = 7.69%  
## UBRE = -0.8933 Scale est. = 1 n = 271396  
##   
## $`Rhynchospora fusca`  
##   
## Family: poisson   
## Link function: log   
##   
## Formula:  
## Y ~ ti(x) + ti(y) + ti(year) + ti(x, year) + ti(y, year)  
##   
## Parametric coefficients:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) -84.04 34.30 -2.45 0.0143 \*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Approximate significance of smooth terms:  
## edf Ref.df Chi.sq p-value   
## ti(x) 1.827 4 23.35 2.2e-06 \*\*\*  
## ti(y) 3.870 4 352.43 < 2e-16 \*\*\*  
## ti(year) 2.266 4 76.69 < 2e-16 \*\*\*  
## ti(x,year) 4.763 16 166.46 < 2e-16 \*\*\*  
## ti(y,year) 6.216 16 76.67 < 2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## R-sq.(adj) = -0.0193 Deviance explained = 12.5%  
## UBRE = -0.96251 Scale est. = 1 n = 271396  
##   
## $`Drosera anglica`  
##   
## Family: poisson   
## Link function: log   
##   
## Formula:  
## Y ~ ti(x) + ti(y) + ti(year) + ti(x, year) + ti(y, year)  
##   
## Parametric coefficients:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) -5.51125 0.02387 -230.9 <2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Approximate significance of smooth terms:  
## edf Ref.df Chi.sq p-value   
## ti(x) 1.007 4 57.05 <2e-16 \*\*\*  
## ti(y) 3.169 4 449.22 <2e-16 \*\*\*  
## ti(year) 3.892 4 493.96 <2e-16 \*\*\*  
## ti(x,year) 8.621 16 81.18 <2e-16 \*\*\*  
## ti(y,year) 10.171 16 128.83 <2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## R-sq.(adj) = 0.00483 Deviance explained = 6.68%  
## UBRE = -0.90162 Scale est. = 1 n = 271396  
##   
## $`Drosera intermedia`  
##   
## Family: poisson   
## Link function: log   
##   
## Formula:  
## Y ~ ti(x) + ti(y) + ti(year) + ti(x, year) + ti(y, year)  
##   
## Parametric coefficients:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) -8.2056 0.1298 -63.2 <2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Approximate significance of smooth terms:  
## edf Ref.df Chi.sq p-value   
## ti(x) 2.059 4 151.89 < 2e-16 \*\*\*  
## ti(y) 1.016 4 700.50 < 2e-16 \*\*\*  
## ti(year) 2.261 4 21.69 2.2e-11 \*\*\*  
## ti(x,year) 6.207 16 299.98 < 2e-16 \*\*\*  
## ti(y,year) 3.255 16 81.04 < 2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## R-sq.(adj) = -0.0287 Deviance explained = 15%  
## UBRE = -0.95124 Scale est. = 1 n = 271396