Problems in using Beals' index to detect species trends in incomplete floristic monitoring data (Reply to Bruelheide et al. (2020))

Erik Christensen | Björn Christensen | Sören Christensen

1Probsteiherhagen, Germany
2Faculty of Business Management, Institute of Statistics and Operations Research, University of Applied Sciences Kiel, Kiel, Germany
3Mathematical Department, Christian-Albrechts-University Kiel, Kiel, Germany

Correspondence
Sören Christensen, Mathematical Department, Christian-Albrechts-University Kiel, Ludewig-Meyn-Str. 4 Kiel, Germany.
Email: christensen@math.uni-kiel.de

Abstract
Bruelheide et al. (Diversity and Distributions, 26, 2020, 782) explored repeated habitat mapping data to identify floristic changes over time on the basis of two surveys. Because of the incompleteness of the data, they utilized the Beals’ index based on the aggregated data from both surveys as a statistical tool for the analysis. The aim of this note is to illustrate problems of this approach, which in particular is shown to produce a systematic underestimation of species decrease (and—potentially less relevant in practice—increase). A specific set of model cases will be introduced to show the effects of unjustified usage of the Beals’ index in this specific form.

KEYWORDS
Beals’ index, bias, biodiversity change, floristic monitoring, incomplete data, modelling approach, species extinction, species trends

1 | THE RAW DATA AND THE AIM OF BRUELHEIDE ET AL. (2020)

Bruelheide et al. (2020) studied repeated habitat mapping data from biotopes in Schleswig-Holstein, Germany (surveys SH1 and SH4), to identify floristic changes over a period of about 30 years (p. 2). The authors admit that the statistical examination of this material is a very ambitious task, and they explain in detail the methodological problems. The main obstacles are (a) the incompleteness of the species lists because of shifts in attention between the two mappings, (b) an unknown observer bias (p. 4) and (c) the lacking of an identity function between the polygons of the two data sets (p. 3). The authors tried to deal with these problems by modifying the raw data and by selecting the intersecting polygons of the first and the second surveys (p. 3).

We do not comment the problems (b) and (c), but focus on (a).

2 | THE INCOMPLETENESS OF DATA AND THE BEALS’ INDEX

The incompleteness of the data was the greatest problem. To identify floristic changes over time, the ratio of relative frequencies of a species i in the second and the first surveys is a suitable quantity to consider. However, a shift in attention between the first and the second inventory of the biotopes generated incompleteness of data, that is gaps in the occurrence matrices and undersized relative frequencies of the affected species.

Thus, Bruelheide et al. searching for a method to fill the gaps in the occurrence matrices used the Beals’ index (Beals, 1984), which utilizes species composition of habitats to calculate “occurrence probabilities” (p. 4) for the species, and by this means provides information about the relative frequencies. Not only the presence of species i, but also the co-occurrences with other species determine...
the Beals’ index of species $i$ in polygon $p$. Even the lack of species $i$ in polygon $p$ can create a Beals’ value greater than 0 due to observations in other polygons. Because the Beals’ index fills gaps, it seems to be the solution for the problem of incomplete data. However, using this method for the two matrices separately does not solve the problem (see model case 1 and Table 1 “normal Beals”).

### 3 | THE PROBLEM OF USING THE INTEGRAL BEALS’ INDEX

In this situation, Bruelheide et al. (2020) proposed the idea of pooling the matrices of both samples as basis for calculating the Beals’ index (we will call it “integral Beals’ index” in the following). Bruelheide et al. identified the precondition to use integral Beals’ index, namely “that the species co-occurrence matrix is static in time” (p. 4), and assumed that this is fulfilled. However, if the co-occurrence matrix is static in time, the relative frequencies of the species must be static in time, too, and therefore there are no trends of species. This creates contradictions to the approach of the authors:

- Because Bruelheide et al. found trends, the assumption (or the calculation) must be wrong.
- The assumption does not reflect reality: Transformation in presences and compositions of species in monitored areas with consequences for the co-occurrence matrix is a normal process, at least in a period of decades. Immigration and extinction of species, changes due to environmental conditions (i.e. nitrogen input, climatic changes), competition, succession and diseases, all these have effects on the vegetation and implicate changes of species presences and composition over time. Each species has its own responses to the different environmental factors.
- Bruelheide et al. want to explore trends of species, but they use a procedure which is based on the assumption that there are no trends.

Therefore, there is no mathematical justification of using the Beals’ index in this way. All calculations of monitoring data, which are based on the integral Beals’ index, are scientifically questionable. The integral Beals’ approach gives reason to expect biases in the trends of species.

### 4 | WHAT ARE THE CONSEQUENCES OF USING THE INTEGRAL BEALS’ INDEX?

We now describe potential undesirable consequences of an improper application of the integral Beals’ index. To this end, we have to explore how the normal vs. the integral Beals’ index works. Using real data, these effects can scarcely be detected because they become invisible in the plethora of the many polygons and their various combinations of species. In order to circumvent this problem, we use simple models to understand the driving effects and compare the values of normal and integral Beals’ index, if the occurrence matrices have gaps. We will concentrate on these model cases: (a) overlooking a species, (b) appearing of a new invasive species and (c) extinction of a species.

Our model cases shall be as small as possible to focus on the aim alone and keep all other values constant. Thus, we choose only two polygons, monitoring in two surveys, and distinguish different cases for the occurrence and non-occurrence of species $i$.

Note that the Beals’ index delivers an estimator for the occurrence probability. The arithmetic mean of normal Beals’ values therefore has a certain similarity to the relative frequency of species. In our model, the arithmetic mean of the normal Beals’ values and the relative frequencies of species are equal.

#### 4.1 | Model case 1: Overlooking a species (see Table 1)

We first consider two polygons with the same type of habitat including species $i$ in a first survey S1 and the second survey S2, too, but in S1 we assume that species $i$ was overlooked in both polygons. In addition, there are species $j1-j20$ which occur just in polygon 1 or in polygon 2, resp., or in both polygons, the same in both surveys, all without being overlooked (Table 1). The relative frequencies and normal Beals’ values (i.e. the Beals’ index separately for S1 and S2) of species $i$ of surveys S1 and S2 are 0% vs. 100%. Because of the assumption that there was no shifting in presences of species $i$ from S1 to S2, these values cannot describe the reality. In contrast, the integral Beals’ values are 50% vs. 52.5%, which seems suitable because this indicates that there is nearly no trend, but 50% “occurrence probability” in S2 is not a reasonable measure of a species that appears in all polygons of S2. As expected, using the integral Beals’ index lowers the differences due to overlooking of species.

#### 4.2 | Model case 2: Appearing of a new invasive species (see Table 1)

A different situation is associated with the same figure in Table 1: A new invasive species $i$ appears in the study area, which was not
recorded in the monitoring S1, but was recorded in the monitoring S2. Relative frequencies and normal Beals’ values of species i of surveys S1 and S2 are 0% vs. 100%, which is suitable. But the integral Beals’ index has nearly equal values. Thus, the integral Beals’ index gives a false indication that this species was already present in the earlier period, and the description of the quick and complete settlement of the invasive species by integral Beals’ value 52.5% seems unsuitable.

### 4.3 Model case 3: Extinction of a species (see Table 2)

To model the situation of the extinction of a species, we reverse Table 1: In S1, species i was present in both polygons, but in the subsequent survey S2, it has died out, and all the other species are assumed to persist. The relative frequency of species i declines from 100% to 0%, so does the normal Beals’ index, which seems suitable. But the integral Beals’ index drops from 52.5% to 50% only, with no indication of the extinction of species i in S2.

Model case 3 shows that the integral Beals’ index might lead to a strong underestimation of species decrease. It also shows that even the total loss of a species may lead to a nearly constant integral Beals’ index.

Bruelheide et al. (2020) point to an unexpected result of their survey, viz “the absence of a trend for highly threatened species” (p. 9). Here we find a hypothesis to explain it.

The model cases describe extreme situations with noticeable values. In the real matrices, these extreme values will of course not emerge, but the model cases supply general perception for the real situations. They make it possible to explain how the incorrect use of integral Beals’ index may affect (a) trends of species and (b) gaps by different attentions.

There are several cases of combinations of species trends and shifts in attention:

- If a species shows no trend, and no shift in attention between the mappings occurred, the integral Beals’ index provides correct

### Table 2 Matrix of monitoring the polygons 1 and 2 in surveys S1 and S2 with relative frequencies and normal and integral Beals’ values of species i (see text model case 3). X presence

| Species i    | S1 | S2 |
|--------------|----|----|
|              | polygon 1 | polygon 2 | polygon 1 | polygon 2 |
| Species j11 | X  | X  |    |    |
| Species j11 | X  | X  |    |    |
| Species j21  | X  | X  |    |    |
| Species j21  | X  | X  |    |    |
| Relative frequency | 100% | 0%  |    |    |
| Normal Beals’ value | 100% | 100% | 0.0% | 0.0% |
| Integral Beals’ value | 52.5% | 52.5% | 50.0% | 50.0% |

The special problem is that the integral Beals’ index makes no difference between gaps in occurrence matrices because of overlooking a species, extinction of a species or not yet arriving of a species. Often there are no hints which conditions referring to trend or shift in attention are given. Therefore, the integral Beals’ index seems to be unsuitable for most research questions considered in Bruelheide et al. (2020).

Note that the integral Beals’ index can deliver biased results even for complete monitoring data sets. Incompleteness of data is a further source of bias.

Thus, the use of the integral Beals’ index creates insecure, often biased and in some cases strange results referring to trends and to the so-called “occurrence probability.”

If there is no shift in attention, the results will be more suitable, the lesser the data differ from the assumption to be “static in time.”

The higher the deviation is from the assumption “static in time,” the more unsuitable the results will be. The general effect is a flattening of trends, thus an underestimation of trends.

Another point is the “occurrence probability.” The reader will expect this to be a theoretical measure which has a connection to concrete presences and their relative frequency. However, the incorrect use of the Beals’ index produces “occurrence probabilities” which can distinctly differ from the real relative frequencies of species.

Bruelheide et al. aimed for an ambitious goal, namely to use incomplete monitoring data to detect species trends. The obvious incompleteness of the data was the reason for the authors to use the integral Beals’ index. They intended to fill the gaps by using the Beals’ values as substitute for relative frequencies. But this approach implies flattening of trends, systematic underestimation of species decrease (or increase) and a potential misassessment of the presences of those species which invade or become extinct. Therefore, the “unexpected result” concerning “the absence of a trend for highly threatened species…” (p. 9) can be explained by this statistical procedure.

The model approach used for this reply reveals the different patterns of the normal and the integral Beals’ index, when used for
monitoring data, and sometimes can help to decide which results of Bruelheide et al. (2020) are realistic and which are not.

ACKNOWLEDGEMENTS
We give many thanks to Helge Bruelheide for answering our questions and to Michael Breuer (Kiel), Ulf Friedrichsdorf (Eutin), Ulrich Mierwald (Kiel) and Hans-Ulrich Piontkowski (Eckernförde) for reviewing the manuscript and discussions.

PEER REVIEW
The peer review history for this article is available at https://publons.com/publon/10.1111/ddi.13276.

DATA AVAILABILITY
No data have been used.

ORCID
Sören Christensen https://orcid.org/0000-0002-8598-7507

REFERENCES
Beals, E. W. (1984). Bray-Curtis ordination: An effective strategy for analysis of multivariate ecological data. Advances in Ecological Research, 14, 55.
Bruelheide, H., Jansen, F., Jandt, U., Bernhardt-Römermann, M., Bonn, A., Bowler, D., Dengler, J., Eichenberg, D., Grescho, V., Harter, D., Jugelt, M., Kellner, S., Ludwig, M., Wesche, K., & Lütt, S. (2020). Using incomplete floristic monitoring data from habitat mapping programmes to detect species trends. Diversity and Distributions, 26(7), 782–794. https://doi.org/10.1111/ddi.13058

BIOSKETCHES
Erik Christensen is biologist and mathematician, chairman of the NGO “AG Geobotanik in Schleswig-Holstein & Hamburg,” with research interest in descriptive and conceptional models of species–area relationship and models to estimate the real species richness, in recording the flora and its changing, and all aspects of nature conservation. Björn Christensen is professor for statistics and mathematics with a focus on applied quantitative methods. Sören Christensen is professor for probability and statistics mainly interested in stochastic processes with applications, for example in natural resource management.

How to cite this article: Christensen E, Christensen B, Christensen S. Problems in using Beals' index to detect species trends in incomplete floristic monitoring data (Reply to Bruelheide et al. (2020)). Divers Distrib. 2021;27:1324–1327. https://doi.org/10.1111/ddi.13276