

# ECOGRAPHY

## Software note

### **distantia: an open-source toolset to quantify dissimilarity between multivariate ecological time-series**

Blas M. Benito and H. John B. Birks

Blas M. Benito (<https://orcid.org/0000-0001-5105-7232>) ✉ ([blasbenito@gmail.com](mailto:blasbenito@gmail.com)), Dept of Biological Sciences, Univ. of Bergen, Bergen, Norway. – H. John B. Birks (<https://orcid.org/0000-0002-5891-9859>), Dept of Biological Sciences and Bjerknes Centre for Climate Research, Univ. of Bergen, Bergen, Norway, and Environmental Change Research Centre, Univ. College London, London, UK.

#### Ecography

43: 1–8, 2020

doi: 10.1111/ecog.04895

Subject Editor: Michael Borregaard

Editor-in-Chief: Hanna Tuomisto

Accepted 12 December 2019



There is a large array of methods to extract knowledge and perform ecological forecasting from ecological time-series. However, in spite of its importance for data-mining, pattern-matching and ecological synthesis, methods to assess their similarity are scarce. We introduce *distantia* (v1.0.1), an R package providing general toolset to quantify dissimilarity between ecological time-series, independently of their regularity and number of samples. The functions in *distantia* provide the means to compute dissimilarity scores by time and by shape and assess their significance, evaluate the partial contribution of each variable to dissimilarity, and align or combine sequences by similarity. We evaluate the sensitivity of the dissimilarity metrics implemented in *distantia*, describe its structure and functionality, and showcase its applications with two examples. Particularly, we evaluate how geographic factors drive the dissimilarity between nine pollen sequences dated to the Last Interglacial, and compare the temporal dynamics of climate and enhanced vegetation index of three stands across the range of the European beech. We expect this package may enhance the capabilities of researchers from different fields to explore dissimilarity patterns between multivariate ecological time-series, and aid in generating and testing new hypotheses on why the temporal dynamics of complex-systems changes over space and time.

Keywords: elastic methods, lock-step, time-series, variable importance

#### Background

Multivariate ecological time-series (METS hereafter) are ordered sequences of observations of a set of variables describing the state of an ecological system at given times (Turchin and Taylor 1992). Good examples of METS are the point-data produced by automatic meteorological stations and buoys, and the spatio-temporal data provided by remote sensors, climate simulations or palaeoecological studies, among many others.

There is no shortage of methods to extract knowledge from METS and perform ecological forecasting. Frequentist, Bayesian and machine learning methods can help establish links between abiotic drivers and biotic responses (Thackeray et al. 2016), evaluate ecological memory processes (Ogle et al. 2015), detect critical transitions in



[www.ecography.org](http://www.ecography.org)

© 2020 The Authors. Ecography published by John Wiley & Sons Ltd on behalf of Nordic Society Oikos  
This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

ecosystem dynamics (Dakos et al. 2012), or establish causal links between environmental drivers and biotic responses (Sugihara et al. 2012). However, the comparison of multivariate time-series, in spite of its importance for data-mining, pattern-matching and ecological synthesis, has received far less attention (but see Wang et al. 2013, Górecki and Łuczak 2015), and open-source tools to assess the dissimilarity between METS are relatively scarce.

According to Wang et al. (2013), available methods to compare METS can be divided in those that assess dissimilarity by time (lock-step methods), and those that assess similarity by shape (elastic methods). Lock-step methods compare METS of the same dimensions and only require the computation of a distance (e.g. Euclidean) between paired samples. However, lock-step methods are generally sensitive to differences between METS due to time shifts in temporal patterns (Wang et al. 2013). Elastic measures tackle this issue by assuming that the values of the gradient over which samples have been taken (time, depth) are either unavailable, too uncertain to be taken into account, or irrelevant due to differences in context between the METS (e.g. sites at different latitudes or elevations). The goal of elastic measures is to align the shapes of the compared METS by pairing their most similar samples while maintaining sample order. This particular property makes elastic dissimilarity measures very attractive to analyse METS taken at different temporal resolutions, sites where time shifts are to be expected due to large latitudinal or elevational gradients (Zhang et al. 2004), or data for which defining the time or age of the samples can be expensive or inaccurate, as is often the case of palaeoecological datasets.

A pioneer example of elastic dissimilarity measures applied to ecological data is the sequence slotting algorithm (Birks and Gordon 1985), which aims to compare and combine stratigraphical sequences. The original variants of this algorithm were first implemented as the Fortran programs SLOTSEQ and SEQSLLOT (Gordon and Birks 1974, Birks and Gordon 1985, Clark 1985). It was later extended by Maher (1993) to improve its graphical output, and finally implemented as the Windows program CPLslot (Hounslow and Clark 2016). Sequence slotting has remained restricted to palaeoecology and palaeoclimatology, where it has been used to compare pollen or other fossil stratigraphical sequences (Lotter et al. 1992, Anderson et al. 1994) and palaeoclimatic and palaeomagnetic data (Thompson and Clark 1989, Maher and Thompson 1995). The sequence slotting method shares most of its internal logic with dynamic time warping (DTW, Berndt and Clifford 1994). DTW evaluates the extent to which one time-series has to be reshaped in order to match another reference time-series. The method is implemented in the R packages ‘dtw’ (Giorgino 2009) and ‘Tsdist’ (Mori et al. 2016), and has been used to align remote sensing data with different temporal resolutions (Baumann et al. 2017), cluster areas with similar temporal dynamics (Suominen 2018), and to correlate stratigraphical data from different sediment cores (Trauth et al. 2018).

In this paper we present the R package *distantia* (Benito and Birks 2019), which contributes to the ecosystem of open-source packages aimed at comparing METS by expanding on

the original ideas of the sequence slotting method through new features that may be of interest for both ecologists and palaeoecologists. Among others, the package provides functions to: 1) prepare and transform the data; 2) compute dissimilarity scores based on lock-step and elastic methods; 3) apply a restricted permutation test to assess the significance of the dissimilarity values; 4) evaluate the individual contribution of each variable to the overall dissimilarity of the compared METS.

## Methods and features

### Dissimilarity metrics

The R package *distantia* implements the metric  $\Psi$  (Gordon and Birks 1974; Eq. 1). It computes the dissimilarity between two METS A and B (with lengths  $m$  and  $n$ ) as the sum of distances between their respective samples ( $AB_{\text{between}}$ ), normalized by the sum of distances between the consecutive samples within each sequence ( $AB_{\text{within}}$ ). This normalization allows the comparison of  $\Psi$  values across pairs of sequences with different lengths.

$$\Psi = \frac{AB_{\text{between}} - AB_{\text{within}}}{AB_{\text{within}}} \quad (1)$$

The component  $AB_{\text{between}}$  can be computed in different ways depending on the nature of the data and the objective of the analysis. If the comparison is made by time with a lock-step method, then A and B must have the same length, and the distances between their respective samples with the same index  $i$  are computed through a distance function  $D$  (Eq. 2).

$$AB_{\text{between}} = 2 \times \sum_{i=1}^{m=n} D(A_i, B_i) \quad (2)$$

If A and B have a different number of samples, or the comparison is to be made by shape, an elastic method relying on a dynamic programming algorithm is used. The original formulation (Gordon and Birks 1974) follows an orthogonal search pattern within the distance matrix (Eq. 3, Fig. 1), and aims to intercalate the samples of both sequences (hence its original name, sequence slotting).

$$AB_{\text{between}} = 2 \times \left( D(A_1, B_1) + \sum_{i=2}^m \sum_{j=2}^n \min \left( \begin{array}{l} D(A_{i-1}, B_j) \\ D(A_i, B_{j-1}) \end{array} \right) \right) \quad (3)$$

Additionally, the *distantia* package introduces a diagonal search mode (Eq. 4), which can both intercalate and/or pair together samples of each sequence based on their similarity.

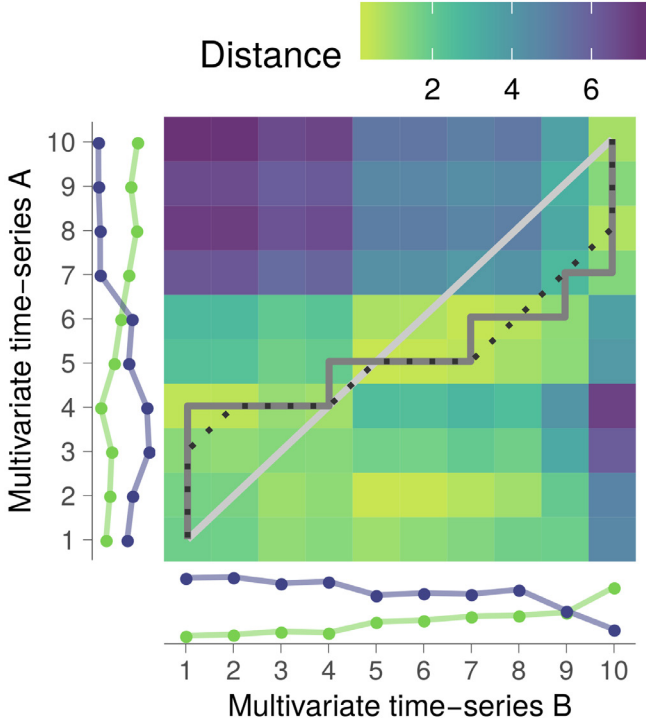


Figure 1. Distance matrix between two METS A (lower panel) and B (left panel) of 10 samples each when using a lock-step method (light grey diagonal line), an elastic measure (dark grey line) and an elastic measure considering diagonals (black dotted line). Straight segments with a length of more than two samples that run parallel to the matrix axes are called ‘blocks’.

$$AB_{\text{between}} = 2 \times \left( D(A_1, B_1) + \sum_{i=2}^m \sum_{j=2}^n \min \begin{pmatrix} D(A_{i-1}, B_j) \\ D(A_i, B_{j-1}) \\ D(A_i, B_j) \end{pmatrix} \right) \quad (4)$$

Equations 3, 4 can generate long straight segments (Fig. 1) called ‘blocks’ in sections where no best-match between samples of each METS can be found. Blocks inflate the value of  $AB_{\text{between}}$  and  $\Psi$  artificially because the distance to the same sample of one of the sequences is counted several times on each block. This issue is particularly problematic when two sequences represent the same system dynamics, but one of them has a lower number of samples due to missing data or differences in sampling resolution. The *distantia* package incorporates an algorithm, activated by the option `ignore.blocks=TRUE`, that identifies straight segments within the least-cost path generated by Eq. 3, 4, which are then ignored during the computation of  $\Psi$ . This option yields more conservative dissimilarity scores when the compared sequences have very different numbers of rows.

Finally, the component  $AB_{\text{within}}$  of Eq. 1, which represents the cumulative sum of dissimilarities between consecutive samples of each sequence, is computed according Eq. 5.

$$AB_{\text{within}} = \sum_{i=2}^m D(A_{i-1}, A_i) + \sum_{i=2}^n D(B_{i-1}, B_i) \quad (5)$$

## Structure and features

The *distantia* package has a modular design, and offers a set of core functions that perform the basic operations required to assess dissimilarity between pairs of METS, and a set of larger functions that combine the operations of the basic functions to perform more complex tasks (Table 1).

The features of the package are applicable to a wide range of ecological problems aiming to answer the question ‘do these sites/times show the same patterns?’. Comparisons by time can be applied to METS produced by meteorological stations, remote sensing platforms and other automatic devices, while comparisons by shape are useful when a temporal shift

Table 1. Main functions of *distantia*.

Name	Function	Input	Output
<code>workflowPsi()</code>	Computes $\Psi$ among two or more METS. Implements lock-step and an elastic methods.	A data frame with two or more METS.	A data frame with sequence names and $\Psi$ values.
<code>workflowPsiHP()</code>	High-performance version of the function above	Same as above (SAA)	SAA
<code>workflowNullPsi()</code>	Computes $\Psi$ on restricted permutations of the input data.	SAA	data frames with null $\Psi$ values and the probability of obtaining a given $\Psi$ under the null hypothesis.
<code>workflowNullPsiHP()</code>	High-performance version of the function above.	SAA	SAA
<code>workflowImportance()</code>	Jackknife approach to compute the contribution of each variable to $\Psi$ .	SAA	Absolute and relative change (%) in $\Psi$ when each variable is removed.
<code>workflowSlotting()</code>	Combines two sequences into one.	A data frame with two METS.	A data frame with the combined sequences.
<code>workflowPartialMatch()</code>	Finds the section of a longer sequence that better matches a shorter sequence.	SAA	A data frame with matching sections of the longer sequence and their $\Psi$ with the short sequence.
<code>workflowTransfer()</code>	Transfers an attribute (generally time or age) from one sequence to another.	SAA	The sequence with the transferred attribute.

between sequences is expected, as it would be the case with palaeoecological or phenological data.

As a novel feature, *distantia* implements the function `workflowImportance()` to quantify the partial contribution of each variable of the compared dataset on their dissimilarity. This feature aims to answer the question ‘why are these sequences different?’ by applying a jackknife approach, which consists of computing partial  $\Psi$  values by removing one variable at a time. This feature is illustrated with an example in the following section.

The package also provides the function `workflowPsiNull()` which applies a restricted permutation test to estimate the probability of finding a given  $\Psi$  value by chance. The permutation occurs at a local temporal scale, by switching randomly selected datapoints, independently by column, with one of their immediate column neighbours. This method assumes that METS observed from the same system must show similar overall trends and local differences resulting from the observational error and the inherent randomness of ecological systems.

## Examples

In this section we describe two examples of the potential applications of *distantia* in different subfields of ecology: 1) comparison of nine central European pollen sequences dated to the Last Interglacial; 2) comparison of the environmental and vegetation dynamics of three plots of *Fagus sylvatica* located across its presence range in Europe. In both cases, dissimilarity was computed with an elastic method considering diagonals and block removal.

### Comparison of pollen sequences dated to the Last Interglacial

The Last Interglacial (hereafter, LIG) was a warm period dated to 129–116 ka BP, and is well represented in the European pollen record (Tzedakis 2007). Here we use *distantia* to: assess dissimilarity between nine unevenly sampled pollen sequences from central Europe (Supplementary material Appendix 1 Table A1); quantify the influence of geographic features on the dissimilarity among sequences; and identify pollen types contributing to the differences between sequences.

$\Psi$  values show that the sites are organized into three groups according to their similarity (Fig. 2a). We fitted the GLM model  $\Psi \sim \text{distance} + \text{difference in elevation} + \text{difference in latitude}$ , and found that difference in elevation is the most relevant predictor of dissimilarity between the pollen sequences (Supplementary material Appendix 1 Table A3; Fig. 2b).

Two sites (Achenhang and Jammertal, Fig. 2c) show a disproportionate dissimilarity ( $\Psi = 2.507$ ,  $p = 0.986$ ) considering their distance (184 km) and elevation difference (~233 m). To better understand this discrepancy we assess the contribution of each pollen type to their dissimilarity with `workflowImportance()`. We find that  $\Psi$  is reduced by 22.7% when *Picea* is removed, followed by *Corylus* (9.91%), *Carpinus* (7%) and

*Abies* (5.76%). These differences are explained by the expansion of *Picea* into the Alps (Achenhang) at the beginning of the LIG (Ravazzi 2002), while Jammertal shows a succession characteristic of a lowland, with an early expansion of *Corylus* and a climate optimum featuring a high abundance of *Carpinus* (Klotz et al. 2003).

### Functional dynamics of three populations of *Fagus sylvatica*

The European beech *Fagus sylvatica* is a drought-sensitive tree dominant in central and western Europe. Here we evaluate the dissimilarity in climate and photosynthetic dynamics between three evenly sampled METS of the same length representing mono-specific stands in Spain (ES), Germany (DE)

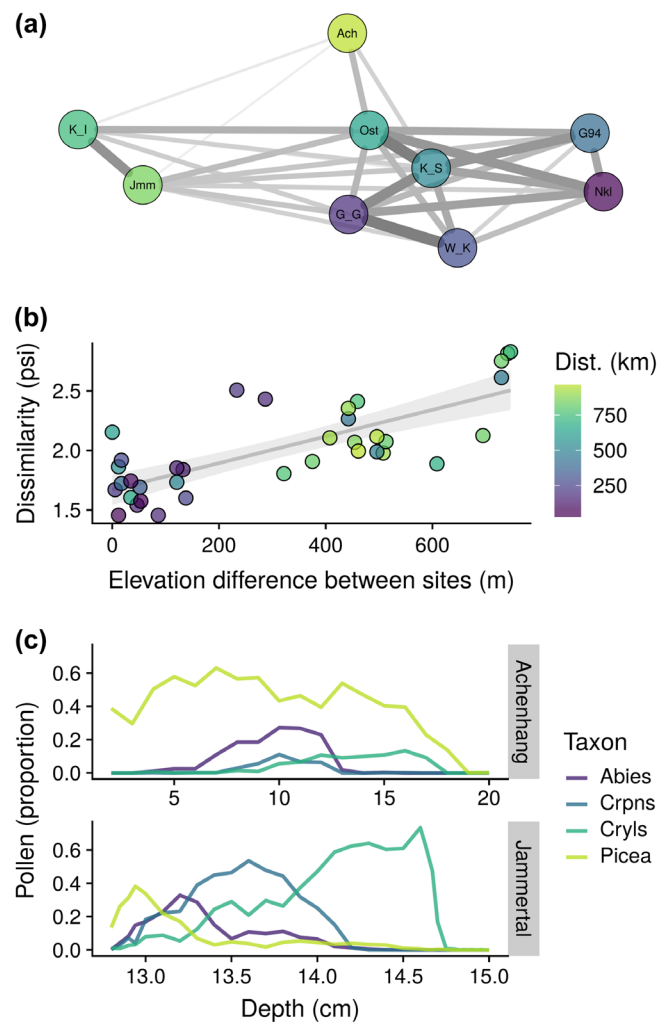


Figure 2. Analysis of similarity/dissimilarity between LIG pollen sequences. (a) Similarity ( $\Psi^{-1}$ ) between the sites Achenhang (Ach), Glowczyn G2 (G\_G), Grobern94 (G94), Jammertal (Jmm), Kletnia Stara (K\_S), Krumbach I (K\_I), Naklo (Nkl), Ostrow (Ost) and Warszawa Kasprzak (W\_K); (b) dissimilarity versus difference in elevation; (c) pollen curves of taxa contributing to the dissimilarity between Achenhang and Jammertal.

Table 2. Dissimilarity between three *Fagus sylvatica* stands in Spain, Germany and Sweden;  $\Psi$ -null is the null expectation of  $\Psi$ , computed on 999 restricted permutations of the input datasets;  $p$  is the proportion of  $\Psi$ -null values that were lower than  $\Psi$ . Columns Temperature, Rainfall and EVI represent  $\Psi$  % reduction when each variable is removed.

Sites		$\Psi$	$\Psi$ -null	$p$	Temperature	Rainfall	EVI
Spain	Germany	1.2845	1.3502	0.075	11.46	11.66	12.71
Spain	Sweden	1.4009	1.4112	0.607	10.18	7.20	10.94
Germany	Sweden	0.8351	1.0060	0.001	-0.70	3.57	10.66

and Sweden (SE; Supplementary material Appendix 1 Section 1.2) selected from the EU Forest database (Mauri et al. 2017). We used the ‘MODISTools’ R package (Tuck et al. 2014) to retrieve EVI time-series (enhanced vegetation index from 2001 to 2018) for these sites and coupled them with climate time-series from the CRU TS v. 4.03 (Climate Research Unit) dataset (Harris et al. 2014).

Using *distancia* we find that stands in DE and SE were more similar between them than with ES. All variables contributed

to dissimilarity between sites, with the exception of climate variables when comparing the DE and SE sites (Table 2).

When grouping the data by year to better understand temporal trends in dissimilarity patterns and variable contribution to dissimilarity, we note that all locations are increasing their dissimilarity throughout time, albeit frequently interrupted by punctuated events, likely linked to cold and warm spells (Fig. 3a). The breakdown of dissimilarity by variable shows that EVI is the variable with the strongest positive

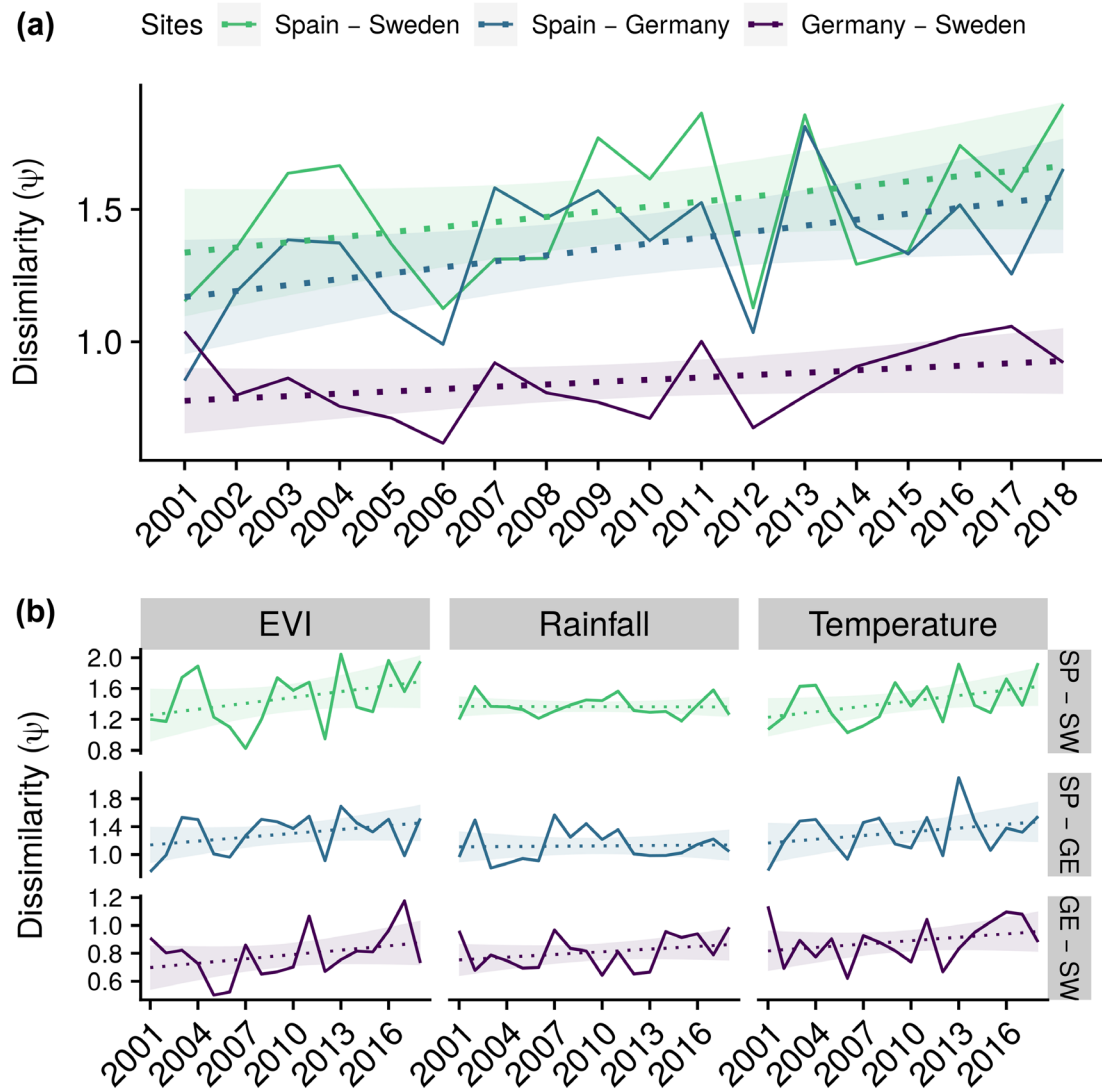


Figure 3. Dissimilarity over time between three stands of *Fagus sylvatica* in Spain (SP), Germany (GE) and Sweden (SE, panel a), and partial contribution of each variable to dissimilarity (panel b).

influence on the increasing dissimilarity, followed by temperature (Fig. 3b).

This trend results from EVI values increasing across sites during the winter and spring months, but only in SE these are increasing during the summer as well. Meanwhile, ES and DE populations show decreased EVI values during these months (Supplementary material Appendix 1 Fig. A10–A12). Increased summer temperature linked to lower rainfall is the likely explanation, but the data available (12 cases per year) are insufficient to establish the statistical significance of these findings.

### Sensitivity analysis of dissimilarity measures

We analyse the sensitivity of the dissimilarity measures available in *distantia* on the example datasets *climate* and *pollenGP* provided with the package (Supplementary material Appendix 1 Section 2 for further details) under two different scenarios: 1) increasing differences in data values of initially identical datasets; 2) increasing differences in the number of rows between otherwise identical datasets. In scenario 1, an increasing number of data-points is randomly selected and modified by adding or subtracting a random percentage of their own value, while in scenario 2, an increasing number of rows is randomly selected and removed from one of the

compared datasets. Each random data modification or row removal is performed 30 times per scenario, and the average and standard deviation of  $\Psi$  is computed across iterations.

Under scenario 1, lock-step and elastic-diagonal methods yield equivalent results across transformations and distance metrics, and both are more sensitive to changes in the data than the elastic-orthogonal methods, especially when the differences between the METS being compared are relatively small (Fig. 4).  $\Psi$  values increase exponentially under scenario 2 when blocks are not removed during the computation of  $AB_{\text{between}}$ , independently of the data transformation or distance metric used. On the other hand, methods removing blocks yield exponential responses with different rates, demonstrating their suitability to assess differences between datasets sampled at different temporal resolutions.

Several recommendations emerge from our sensitivity analysis: 1) each combination of distance metric, data transformation and dissimilarity algorithm yields a particular scale of  $\Psi$  values for a given pool of datasets, and therefore different methods must not be used in the same analysis; 2) Euclidean distances coupled with elastic-diagonal and lock-step methods are more sensitive to small differences between datasets; 3) data transformations, such as the Hellinger transformation commonly applied to pollen data (Birks and Gordon 1985), increase the sensitivity and reduce the bias of dissimilarity

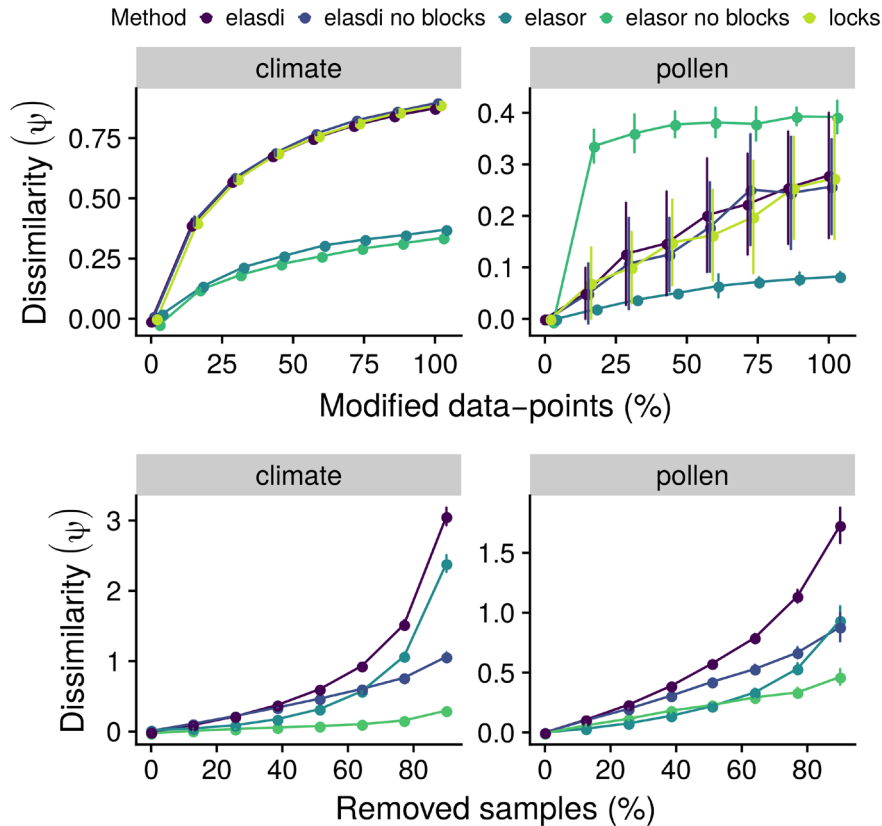


Figure 4. Sensitivity of the dissimilarity measures implemented in *distantia* (abbreviations: elasdi, elastic diagonal; elasdi no blocks, elastic diagonal with block removal; elador, elastic orthogonal; elador no blocks, elastic orthogonal with block removal; locks, lock-step) to increasing number of modified datapoints (scenario 1, upper panel) and decreasing number of rows in one of the datasets (scenario 2, lower panel).

scores; 4) the elastic-diagonal no-blocks method shows the most balanced properties across types of datasets and scenarios of differences between datasets.

## Discussion

Multivariate time-series are the most relevant data format that ecologists have available to analyse and understand the dynamics of complex systems (Boero et al. 2015). However, open-source tools to quantify dissimilarity between METS taken at different times or sites are scarce. This package fills this gap by introducing the R package *distantia*, which provides a general toolset to streamline the assessment of dissimilarity between METS.

We have shown that *distantia* covers an array of dissimilarity metrics useful to analyse different types of METS, including those that are unevenly sampled and have different number of cases. Comparisons by shape (i.e. elastic), are especially relevant when the collected data are unevenly spaced along the sampling dimension, as it is commonly the case with palaeoecological datasets (Willis et al. 2010). However, elastic measures are relevant as well when the importance of the sampling dimension (time in most cases, but others such as elevation or latitude are possible) is context-dependant, even when the data is regularly sampled. Phenological or climatological observations at different latitudes or elevations are a clear example of this, since elastic dissimilarity measures can easily accommodate time-delays in temperature change without penalizing the resulting dissimilarity scores when the objective is to compare dynamics without a focus on synchronicity. On the other hand, comparisons by time (i.e. lock-step) are the natural option when the synchronicity between the studied phenomena is relevant, and the data are captured by automatic devices, or generated by simulations such as global circulation models.

Independently of the nature of the data, *distantia* is designed to simplify the analysis of METS as much as possible. In consequence, functions such as `workflowPsiHP()` can be applied to many sequences at once, facilitating analysis on large published databases such as Neotoma (Goring et al. 2015), remote sensed data such as MODIS products (LP DAAC 2019), or climate data such as CRU TS (Harris et al. 2014), among many others.

We hope this package may enhance the capabilities of researchers from different fields to easily explore dissimilarity patterns between METS, and to generate and test new hypotheses on why the dynamics of complex-systems changes over space and time.

To cite *distantia* or acknowledge its use, cite this Software note as follows, substituting the version of the application that you used for 'version 0':

Benito, B. M. and Birks, H. J. B. 2019. *distantia*: an open-source toolset to quantify dissimilarity between multivariate ecological time-series. – *Ecography* 42: 000–000 (ver. 0).

## Data availability statement

Supplementary material with the code and data used in this paper are available at Zenodo under the doi: 10.5281/zenodo.3520961. R package *distantia* is available via CRAN (<<https://CRAN.R-project.org/package=distantia>>) and GitHub (<<https://github.com/BlasBenito/distantia>>), doi: 10.5281/zenodo.3520959).

*Acknowledgements* – This paper is a contribution to the IGNEX Project. We thank Gavin Simpson and an anonymous reviewer for their contribution in strengthening this manuscript.

*Funding* – BMB and HJBB are supported by FRIMEDBIO (Research Council of Norway) through IGNEX (project 249894). HJBB is also supported by the European Research Council through the grant agreement 74143 – HOPE.

*Author contributions* – BMB and HJBB conceived the ideas; BMB developed the package and led the writing of the manuscript.

## References

- Anderson, N. J. et al. 1994. Spatial heterogeneity of diatom stratigraphy in varved and non-varved sediments of a small, boreal forest lake. – *Aquat. Sci.* 56: 40–58.
- Baumann, M. et al. 2017. Phenology from Landsat when data is scarce: using MODIS and dynamic time-warping to combine multi-year Landsat imagery to derive annual phenology curves. – *Int. J. Appl. Earth Observ. Geoinform.* 54: 72–83.
- Benito, B. M. and Birks, H. J. B. 2019. Assessing dissimilarity between multivariate time series with the R package *distantia*. – R package ver. 102.
- Berndt, D. J. and Clifford, J. 1994. Using dynamic time warping to find patterns in time series. – In: *Proc. 3rd Int. Conf. Knowl. Discov. Data Min.*, pp. 359–370.
- Birks, H. J. B. and Gordon, A. D. 1985. Numerical methods in Quaternary pollen analysis. – Academic Press.
- Boero, F. et al. 2015. Time is an affliction: why ecology cannot be as predictive as physics and why it needs time series. – *J. Sea Res.* 101: 12–18.
- Clark, R. M. 1985. A FORTRAN program for constrained sequence-slotting based on minimum combined path length. – *Comput. Geosci.* 11: 605–617.
- Dakos, V. et al. 2012. Methods for detecting early warnings of critical transitions in time series illustrated using simulated ecological data. – *PLoS One* 7: e41010.
- Giorgino, T. 2009. Computing and visualizing dynamic time warping alignments in R: the *dtw* package. – *J. Stat. Softw.* 31: 1–24.
- Gordon, A. D. and Birks, H. J. B. 1974. Numerical methods in Quaternary palaeoecology. II. Comparison of pollen diagrams. – *New Phytol.* 73: 221–249.
- Górecki, T. and Łuczak, M. 2015. Multivariate time series classification with parametric derivative dynamic time warping. – *Expert Syst. Appl.* 42: 2305–2312.
- Goring, S. et al. 2015. *neotoma*: a programmatic interface to the neotoma paleoecological database. – *Open Quat.* 1: Art. 2.
- Harris, I. et al. 2014. Updated high-resolution grids of monthly climatic observations – the CRU TS3.10 dataset. – *Int. J. Climatol.* 34: 623–642.

- Hounslow, M. W. and Clark, R. M. 2016. CPLSlot, a program for objective correlation between successions using sequence slotting. – doi: 10.13140/RG.2.2.17513.29288
- Klotz, S. et al. 2003. Continental European Eemian and early Würmian climate evolution: comparing signals using different quantitative reconstruction approaches based on pollen. – *Global Planet. Change* 36: 277–294.
- Lotter, A. F. et al. 1992. Rates of change and chronological problems during the late-glacial period. – *Clim. Dyn.* 6: 233–239.
- LP DAAC 2019. MODIS: moderate resolution imaging spectroradiometer. – <<https://lpdaac.usgs.gov/>>.
- Maher, B. A. and Thompson, R. 1995. Paleorainfall reconstructions from pedogenic magnetic susceptibility variations in the Chinese loess and paleosols. – *Quat. Res.* 44: 383–391.
- Maher, L. J. 1993. SLOTDEEP.EXE: manual correlation using a dissimilarity matrix. – *Methods Newslett.* 9: 21–26.
- Mauri, A. et al. 2017. EU-forest, a high-resolution tree occurrence dataset for Europe. – *Sci. Data* 4: 160123.
- Mori, U. et al. 2016. Distance measures for time series in R: the TSdist package. – *R J.* 8: 451.
- Ogle, K. et al. 2015. Quantifying ecological memory in plant and ecosystem processes. – *Ecol. Lett.* 18: 221–235.
- Ravazzi, C. 2002. Late Quaternary history of spruce in southern Europe. – *Rev. Palaeobot. Palynol.* 120: 131–177.
- Sugihara, G. et al. 2012. Detecting causality in complex ecosystems. – *Science* 338: 496–500.
- Suominen, T. 2018. Applying MERIS time series and dynamic time warping for delineating areas with similar temporal behaviour in the northern Baltic Sea. – *Ecol. Indic.* 95: 794–804.
- Thackeray, S. J. et al. 2016. Phenological sensitivity to climate across taxa and trophic levels. – *Nature* 535: 241–245.
- Thompson, R. and Clark, R. M. 1989. Sequence slotting for stratigraphic correlation between cores: theory and practice. – *J. Paleolimnol.* 2: 173–184.
- Trauth, M. H. et al. 2018. Abrupt or gradual? Change point analysis of the late Pleistocene–Holocene climate record from Chew Bahir, southern Ethiopia. – *Quat. Res.* 90: 321–330.
- Tuck, S. L. et al. 2014. MODISTools – downloading and processing MODIS remotely sensed data in R. – *Ecol. Evol.* 4: 4658–4668.
- Turchin, P. and Taylor, A. D. 1992. Complex dynamics in ecological time series. – *Ecology* 73: 289–305.
- Tzedakis, P. 2007. Pollen records, last interglacial of Europe. – In: Elias, S. A. (ed.), *Encyclopedia of Quaternary science*. Elsevier, pp. 2597–2605.
- Wang, X. et al. 2013. Experimental comparison of representation methods and distance measures for time series data. – *Data Min. Knowl. Discov.* 26: 275–309.
- Willis, K. J. et al. 2010. Biodiversity baselines, thresholds and resilience: testing predictions and assumptions using palaeoecological data. – *Trends Ecol. Evol.* 25: 583–591.
- Zhang, X. et al. 2004. Climate controls on vegetation phenological patterns in northern mid- and high latitudes inferred from MODIS data. – *Global Change Biol.* 10: 1133–1145.

Supplementary material (available online as Appendix ecog-04895 at <[www.ecography.org/appendix/ecog-04895](http://www.ecography.org/appendix/ecog-04895)>). Appendix 1.