---

title: 'Project 6: Phenological Diversity Trends By Remote Sensing Related Datacubes'

tags:

- Rao's Q Index

- Time-Weighted Dynamic Time Warping

- Landscape Heterogeneity

- Remote Sensing

- Time Series

- Phenology

authors:

- name: Elliot Shayle

orcid: 0009-0008-2994-0887

affiliation: 1

- name: Matteo Marcantonio

orcid: 0000-0003-3896-2355

affiliation: 2

- name: Rocco Labadessa

orcid: XXXX-XXXX-XXXX-XXXX

affiliation: 3

- name: Chiara Richiardi

orcid: 0000-0002-2370-7768

affiliation: 4

- name: Saverio Vicario

orcid: 0000-0003-1140-0483

affiliation: 2

affiliations:

- name: Environmental Informatics, University of Marburg, Deutschhausstr. 12 - 35032, Marburg, Germany

index: 1

- name: CNR - IIA, Via Orabona, 4, Bari, Italy

index: 2

- name: Joint European Research Center (Freelancer scientist), Ispra, Italy

index: 3

- name: ENEA, Strada per Crescentino, 13, Turin, Italy

index: 4

date: 03 April 2024

bibliography: paper.bib

authors\_short: Shayle et al. (2024) Phenological Diversity Trends By Remote Sensing Related Datacubes

group: B-Cubed

event: B-Cubed Hackathon 2024 - Hacking biodiversity data cubes for policy

editor\_options:

markdown:

wrap: 72

---

# Project 6: Phenological Diversity Trends By Remote Sensing Related Datacubes

Authors: E. Shayle (1), M. Marcantonio\* (2), C. Richiardi (3), R.

Labadessa (4), & S. Vicario (4)

1: Environmental Informatics, University of Marburg, Deutschhausstr.

12 - 35032, Marburg, Germany

2: Joint European Research Center (Freelance scientist), Ispra, Italy

3: ENEA, Strada per Crescentino, 13, Turin, Italy

4: CNR - IIA, Via Orabona, 4, Bari, Italy

\* corresponding author

Keywords: Rao's Q Index, Time-Weighted Dynamic Time Warping, Landscape

Heterogeneity, Remote Sensing, Time Series, & Phenology

# Introduction:

The B3 Hackathon brought together computer scientistis and ecologists from a variety of

institutions to rapidly create novel informatics solutions to the

biodiversity challenges facing the planet. We identified that the

addition of time-weighting to the R package "rasterdiv" would be a

worthwhile contribution to the environmental informatics community.

Rasterdiv was created to calculate diversity indices with data of the

class "raster layer". Biodiversity indexes commonly focus on the spatial

component. Here we outline how our extension to the pre-existing

implementation of Rao's diversity indices [@Rocchini2017] can account

for the temporal dimension of data, alongside the relevant biological

context to our extension.

## The Importance of Biodiversity Indices:

Ecosystems presenting a heterogeneous biota have been shown both experimentally and

theoretically to provide greater utility to all the agents which

comprise that ecosystem. This is through the provision of more and more resourches and

varied niches available for species . A high species diversity subsequentlyincreases the value of ecosystem services provided to the communities

surrounding an ecosystem. Heterogeneous ecosystems are typically also

more resilient to disturbances they experience, likely because they have

functional redundancy . Due to the centrality of biodiversity to healthy

ecosystem functioning, quantitative measures of biodiversity are

required to understand how ecosystems are responding to ongoing

environmental changes, such as climate changes and shifting land use.

Remotely sensing applications generate large quantities of data representing the state of Earth surface with almost continuous spatio-temporal resolution. This “big data” require to be summarised in meaningful ways to describe ecosystems trends and patterns such as changes in ecosystem diversity. To this end, information theory has been applied to remote data for many years to derive indices of environmental diversity. For example, Shannon's H value has been widely used as a proxy estimating species diversity , but

can be inadequate when applied to data generated by

remote sensing platforms (e.g. images from Earth observation

satellites). To create quantified data from ecosystems, most analytical

approaches assess discrete points within the ecosystem, such as those

from a quadrat, or pixels in the case of aerial remote sensing datasets.

One limitation is that Shannon's H value is that it does not consider

the distance between each sampled point (whether they are species,

pixel, or any other quantitative abstractions of an observation). This

approach treats all objects within a dataset as equally distant from one

another and can easily saturate if objects within a dataset, even if just slightly, differ.

Rao's Quadratic Diversity Index (Rao's Q) adds space as a trait to its

abstraction of biodiversity by accounting for the distance between

observations within a study site. As a spatially informed alternative to

Shannon's H, Rao's Q has been demonstrated experimentally to offer

greater efficacy when representing biodiversity in aerial remote sensing

datasets [@Rocchini2021], for which pixels are the discrete observation

units. However, Rao's Q remains limited by its inability to assess trait

change over time. Current implementations of the index (for example in the rasterdiv R package) only assess one

snapshot of the data at a time. We set out to overcome this limitation

by incorporating Time-Weighted Dynamic Time Warping (TWDTW) to include

time as a component of the distance variable within Rao's Q.

## The Purpose of (Time-Weighted) Dynamic Time Warping & its Ecological Utility:

Dynamic Time Warping (DTW) is a mathematical approach used to compare

data series when the timing of observations differs. It has been used in

a variety of disciplines. DTW works by finding the smallest distance

between two time series.

However, by flattening the differences in timing, biologically

significant differences can also be obscured, such as when comparing

plant phenology. For instance, many tree species require a minimum

number of Growing Degree Hours (GDH) to commence their springtime

budburst [@Fu2019]. Other ecosystem processes typically need to coincide

with phenological events, so phenology timing represents an important

differentiating factor for time series representing ecosystems with

plants.

The TWDTW approach rectifies this by including a cost to aligning pixels

with greater temporal separation. Therefore, the TWDTW function is less

likely to match the time series to others which exhibit substantially

different phenologies. This has been successfully demonstrated by

[@Maus2016] to classify changing land use patterns in the Brazilian

Amazon, and was a more effective tool than standard DTW when applied to

heterogeneous biological environments like these.

Equation:

![TWDTW Equation from Maus

2016](Figures-Images-USW/TWDTW%20Equation%20from%20Maus%202016.png)

Reproduced from Maus [@Maus2016]. In addition to the standard cost

matrix of the DTW function, they also propose the equation above to

implement a temporal cost. In the equation α is the steepness of the

logistic function used for penalisation of time distance, and β is the

midpoint of the curve. Lastly, $g(ti,tj)$ represents the time elapsed

between the dates evaluated in the match ($ti$ and $tj$ times of the

"i"th and "j"th observations).

In this manuscript, we used optical aerial remote sensing data from Sentinel-2 derived

from a small, grazed grassland site in Calabria, Italy to demonstrate

and evaluate our R-based implementation of phenology into Rao's Q index.

We also evaluate its efficacy in comparison to Shannon's H and

unmodified Rao's Q indices.

# Case Study and Results:

## Implementation within rasterdiv

We implemented this method within the existing `paRao()` function of the

rasterdiv R package employing the `twtwd` function from the `twdtw` R

package [@Maus2019] that is a wrapper for a C++ implementation of TWDTW.

The Rao’s index with twdtw distance calculated over a time-series of imageries can thus be derived using the following R function:

`paRao(x=time.series, time\_vector=time, window=11, alpha=1, na.tolerance=0, method="multidimension", dist\_m="twdtw", simplify=4, np=8)`

The arguments and our input parameters of which are:

`x` An `(X,Y,Z)` raster stack (or cube) of spectral data, where the X

and Y axes represent discrete pixel values, and each layer of the Z axis

is a a different temporal snapshot of the raster layer. In our study,

this is the Sentinel derived time series of our study site in Calabria.

`time\_vector` A vector of dates corresponding to every point in the

raster time series, which must be the same as the `Z` axis from the `x`

variable. All pixels in the input time series must share the same

temporal spacing as the temporal pattern to which it is being compared

(i.e. if the time series has observations on days `c(1, 3, 7, ...)`,

then the pattern it is being compared to must also have observations on

days `c(1, 3, 7, ...)`.

`steepness` A continuous numeric value corresponding to the α variable

from the time-weighting function in Maus [@Maus2016]. Lower or higher

values of α ...increase or decrease?... penalisation for deviations from

the pattern time.

`midpoint` A numeric value corresponding to the β variable from the

time-weighting function in Maus [@Maus2016]. The input data must be of

the unit specified by the `time\_scale` argument (i.e. it should be

expressed in days).

`cycle\_length` A string value. Valid input arguments are "year"

`time\_scale` "day"

Other arguments remain unchanged.

## Study Site Description:

The study site was a small (5 hectare) patch within the Macchia Sacra

Special Protection Area. It was selected as it was suitable for thorough

imaging by drone, as this formed the basis of our ground-truthed

observations of Biodiversity. With the expertise in classification

imparted by a an expert bonatist, we defined 8 types of plant

communities within the study site (Figure 1).

The area is characterized by the

presence of a road on on the north-east part of the site. From the level

of the road the elevation declines to a lower part that features a sharp

canyon running south to west, the result of a previous small stream

which had dried up by the time of our drone survey. This part of the

study site is characterized by hydrophilic vegetation. Between these two

extremes is a small hill which culminates in a plateau. The plateau is

the resting area of a herd of cows which graze in the area. This area is

much dryer and subject to strong pasture pressure and mechanical

disruption, but is more nutrient which, owing to the presence of cow

manure.

## Evaluation of the Efficacy of our Results:

We used 144 Sentinel2 images from HRVPP of Phenological Plant Index

(PPI) covering all images taken during 2023. Each image was 20 by 27

pixels (Figure 1). The PPI index was chosen as it is minimally influenced by soil

signal and the presence of shadows [@karkauskaite2017evaluation]. Using

these data, we applied 3 analytical approaches to measure biodiversity:

The Shannon's Biodiversity index applied on the mean yearly value with 3

significant digits of the PPI trajectory; the Rao's Q index with

different values of α, applied to the same dataset; and the Rao's Q

index with our implementation of the TWDTW function across the full time

series of 144 images. A gross visual inspection of Figure 1 illustrates

the inviability of the Shannon's H index, because using 3 digits, all

pixels had different mean biodiversity values and it was not possible to

classify the ecosystem into different groups. The standard Rao's Q index

correctly identified the main biodiversity hotspot as the plateau atop

the hill, and a secondary hotspot where the road intersects with the

study site. We observed that Rao's Q index does not change, changing

alpha given that all pixels are different. Finally, our new

implementation of distance resulted in two meaningful differences from

the standard Rao's Q index: the road is no longer a secondary hotspot,

and the main biodiversity hotspot moved at the borders between two of

the communities identified by our expert.

![Figure study area here]

![Figure

1](Figures-Images-USW/Figure%201%20Time%20Series%20of%20PPI%20for%20Study%20Site%20V1.1.png)

![Figure

2](Figures-Images-USW/Figure%202%20Results%20Overview%20Index%20Comparison%20V1.0.png)

# Discussion:

In this hackathon, we developed a streamlined method for implementing the Time-Weighted Dynamic Time Warping (TWDTW) algorithm to calculate Rao's quadratic entropy index within the “rasterdiv” R package. This enhancement introduces a temporal dimension to the traditional spatial analysis of landscape diversity. Recognising the dynamic nature of plant communities and ecosystems over time, our method integrates phenological variations into diversity assessments derived from satellite imagery. Importantly, the application of this technique to multiband remotely sensed data from disturbed grasslands has revealed that accounting for phenological cycles can refine diversity indices by filtering out artefacts. For instance, it can help distinguishing between semi-natural habitats and artificial land covers like roads, which lack temporal phenological shifts. Such artificial features tend to form clusters of minimal DTW distances when considering DTW as a inter-voxel distance, leading to lower Rao's index values.

By incorporating temporal dynamics into the rasterdiv R package, we broaden the scope for analysing remotely sensed time series. This advancement enriches the suite of diversity indices obtainable from remote sensing data, potentially offering a more comprehensive understanding of landscape heterogeneity.

# GitHub and Data Repositories:

This manuscript, previous revisions, open source data, and scripts can

all be found on the open source GitHub repository

"Samuel-Green/B-3-Hackathon-Project-6" via the URL:

<https://github.com/Samuel-Green/B-3-Hackathon-Project-6>

# Acknowledgements:

Dr. Quentin Groom of Meise Botanic Garden for co-ordinating the project.

Laura Abraham of Meise Botanic Garden for organising the event.

The European Union's Horizon Europe Research and Innovation Programme

(ID No 101059592) for funding the B3 programme, and thus, this event.

# References: