Draft Manuscript of Dynamic Time Warping Within a Phenological Context:

# Introduction:

Introduce RasterDiv briefly

## Importance of biodiversity indices:

* Heterogeneous ecosystems have been shown experimentally and theoretically to be much more resilient to perturbations and typically provide more ecosystem services.
  + There are many ways to measure biodiversity
  + Which is especially important as we need to understand how ecosystems are changing in response to global changes
* Shannon’s H value has been widely used as a proxy for biological diversity, but is not always the most applicable for understanding datasets generated by remote sensing
  + It does not consider the distance between type of instances of observed objects (e.g., species or digital numbers)
  + Thus, treating all different objects as equally different (distant)
* Which is why the Rao’s Quadratic Diversity Index is used
  + Rao’s Q incorporates the distance between separate pixels
  + Experimentally, this has demonstrated the efficacy of Rao’s Q over Shannon’s H in assessing satellite and other aerial remote sensing datasets (Rocchini et al., 2021)
  + Adds traits to the syntetistation of biolgical diverstiy
* Rao’s Q index is still not ideal as a tool to assess changes of triats in time
  + It works on only one snapshot at a time
  + By incorporating TWDTW into a Rao’s Q function, we can incorporate phenology patterns to understand how biodiversity has differed over time
  + Rao’s index only allows for one value (time wise)
    - Addition of the TWDTW function modifies the distance metric to also incorporate changes of the trait through time

## Background to Dynamic Time Warping (DTW) and Time Weighted Dynamic Time Warping (TWDTW):

* Can be used for comparison of geographical features throughout time
* DTW can be used to make comparisons of ecosystems with remote sensing data [examples given in Maus 2016] (Maus et al., 2016)
* In order to make scientifically valid comparisons in ecological remote sensing, the time series of the ecosystem data needs to be appropriately considered
  + Typical DTW disregards the time series – it minimize the difference to find the minimal distance between trajectories, but from a phenological perspective, this is important data which can’t be disregarded
    - Otherwise, time series which are phenologically and biologically different will be classed as equivalent time series
* Factors like Growing Degree Hours (GDH) influence when trees will have budburst
  + The timing of this is important as other ecosystem processes typically coincide with this (Fu et al., 2019)
* Classic DTW doesn’t take into account land use change [Maus]
* Maus introduces TWDTW, which adds a cost from aligning pixels which are more temporally separated
  + Therefore, the function is less likely to match the time series to others which exhibit substantially different phenologies

## Ecological benefits of the TWDTW approach:

* In the Maus paper, they introduce TWDTW within the context of the Brazilian Amazon
* They point out that soybean crops have a specific phenological cycle, which could be matched to other land use types *were it not for temporal bounds*
* Their method adds seasonal bounds to the time series data so that a time series can’t be warped beyond when it’s phenophase should start or end
* In their figure 1, they illustrate this within the context of the Brazilian Amazon
  + Using a section of the Amazon that was deforested, used as pasture, then used for soybeans
  + Without phenological bounding, the forest section of the time series could’ve been matched to the soybean section
* Maus experimentally found that the traditional DTW approach performed worst at identifying land use
* Here we present our R code to implement phenology into Rao’s Q index applied to optical remotely sensed data. We also evaluate its efficacy than Shannon and classical Rao’s index using a small portion of a grassland in Calabria, Italy.

# Methods:

## Applicability and constraints of the approach

* All time series must be complete – there can be no missing or unmatched NA values
* Saverio has used a Gaussian Process function to fill in the gaps in the time series [link to study or refer to methodological justification later on]
  + Alternatively, our technical implementation may offer to the user the option to remove that time series from the data cube

A few words on the moving windows approach on categories

A few words on TWDTW

A focus on how multiple time series can be summarised in one distance number

Explain with words the figure linked by Saverio of how 2 curves vary with one another

A few words on steepness and midpoint parameters of the warping (these can be user defined)(Saverio shared a relevant paper, find which one and read it)

Describe weighting in time warping using the function given in the help page of the twdtw R package

# Results:

When we get the data processed

# Discussion:

Later on

References:

Fu, Y. H., Piao, S., Zhou, X., Geng, X., Hao, F., Vitasse, Y., & Janssens, I. A. (2019). Short photoperiod reduces the temperature sensitivity of leaf-out in saplings of Fagus sylvatica but not in horse chestnut. *Global Change Biology*, *25*(5), 1696–1703. https://doi.org/10.1111/gcb.14599

Maus, V., Câmara, G., Cartaxo, R., Sanchez, A., Ramos, F. M., & Queiroz, G. R. De. (2016). *Time-Weighted Dynamic Time Warping Method for Land-Use and Land-Cover Mapping*. 1–11. https://doi.org/10.1109/JSTARS.2016.2517118

Rocchini, D., Thouverai, E., Marcantonio, M., Iannacito, M., Da Re, D., Torresani, M., Bacaro, G., Bazzichetto, M., Bernardi, A., Foody, G. M., Furrer, R., Kleijn, D., Larsen, S., Lenoir, J., Malavasi, M., Marchetto, E., Messori, F., Montaghi, A., Moudrý, V., … Wegmann, M. (2021). rasterdiv—An Information Theory tailored R package for measuring ecosystem heterogeneity from space: To the origin and back. *Methods in Ecology and Evolution*, *12*(6), 1093–1102. https://doi.org/https://doi.org/10.1111/2041-210X.13583