




Climate variability indices for ecological and crop models in R: the `climatrends` package

Kauê de Sousa ^{1,2}, Jacob van Etten ², Magne Neby ¹, and Svein Ø. Solberg ¹

¹ Department of Agricultural Sciences, Inland Norway University of Applied Sciences, Hamar, Norway ² Digital Inclusion, Bioversity International, Montpellier, France

DOI: [10.21105/joss.04405](https://doi.org/10.21105/joss.04405)

Software

- [Review](#) 
- [Repository](#) 
- [Archive](#) 

Editor: [Kristina Riemer](#) 

Reviewers:

- [@dlebauer](#)
- [@istfer](#)

Submitted: 23 February 2022

Published: 02 May 2023

License

Authors of papers retain copyright and release the work under a Creative Commons Attribution 4.0 International License ([CC BY 4.0](#)).

Summary

Abiotic factors play an important role in most ecological and crop systems that depend on certain levels of temperature, light and precipitation to initiate important physiological events ([Schulze et al., 2019](#)). Understanding how these factors drive the physiological processes is a key approach to provide recommendations for adaptation and biodiversity conservation in applied ecology studies. The package `climatrends` aims to provide the methods in R ([R Core Team, 2020](#)) to compute precipitation and temperature indices that serve as input for climate and crop models ([Kehel et al., 2016](#); [van Etten et al., 2019](#)), trends in climate change ([Aguilar et al., 2005](#); [de Sousa et al., 2018](#)) and applied ecology ([Liu & El-Kassaby, 2018](#); [Prentice et al., 1992](#)).

Statement of need

Reproducibility, the ability to repeat the analysis, and replicability, the ability to repeat an experiment ([Stevens, 2017](#)), are key to performing collaborative scientific research ([Munafò et al., 2017](#); [Powers & Hampton, 2019](#)). It allows scientists to re-perform analysis after a long hiatus and peers to validate analysis and get new insights using original or new data. This is still a gap in most of the studies in agriculture and ecology. `climatrends` addresses this specific issue. The package originates from a set of scripts to compute climate indices in our previous studies ([de Sousa et al., 2018](#); [van Etten et al., 2019](#)). Building up on the interest in expanding the analysis to other regions and to enable reproducible and replicable studies among different research groups within the CGIAR (<https://www.cgiar.org>) and partner institutions we developed `climatrends`. Most of the package functions take into account the heterogeneity of testing sites (locations), dates and seasons, a common characteristic of decentralized agricultural trials ([van Etten et al., 2019](#)). Further development was made to enable time series analysis with fixed periods of time and locations ([de Sousa et al., 2018](#)). The package `climatrends` computes temperature, precipitation, crop growing and crop stress indices that were validated by previous studies on climatology and crop science ([Aguilar et al., 2005](#); [Challinor et al., 2016](#); [Kehel et al., 2016](#)). The indices are described in the package's documentation. `climatrends` was released on CRAN in February 2020 and was the first package to provide methods for agroclimatic indices in R, being able to deal with site heterogeneity and time series analysis. `climatrends` is also unique in its integration with API Client data like NASA POWER ([Sparks, 2018](#)), CHIRPS ([de Sousa et al., 2020](#); [Funk et al., 2015](#)) and AgERA5 ([Brown et al., 2023](#); [Hersbach et al., 2020](#)). Other packages like `pollen` ([Nowosad, 2019](#)) and `croppgrowdays` ([Baker & Mortlock, 2021](#)) are limited to growing degree days, whereas the package `agroclim` ([Serrano-Notivol, 2022](#)) provides temperature and precipitation indices but is limited to zoning areas. Currently, `climatrends` is part of the CRAN Task View in Agriculture (<https://cran.r-project.org/web/views/Agriculture.html>) and

is complementary to `agroclim` and `cropprowdays`.

The package contains eight main functions (Table 1), with a default method for numeric 'vector' and additional methods implemented via the package methods (R Core Team, 2020) for classes 'matrix' (or array), 'data.frame', and 'sf' (of geometry POINT or POLYGON) (Pebesma, 2018). The last two methods are designed to fetch data using API Client, currently from the packages `nasapower` (Sparks, 2018) and `chirps` (de Sousa et al., 2020).

Table 1: Main functions available in `climatrends`.

Function	Definition
<code>cumdrought()</code>	Returns a vector with the cumulative sum of the maximum length of dry spell (MLDS)
<code>cumrain()</code>	Returns a vector with the cumulative sum of the maximum length of wet spell (MLWS)
<code>crop_sensitive()</code>	Compute crop sensitive indices
<code>ETo()</code>	Reference evapotranspiration using the Blaney-Criddle method
<code>GDD()</code>	Compute growing degree-days
<code>late_frost()</code>	Compute the occurrence of late-spring frost
<code>rainfall()</code>	Precipitation indices
<code>temperature()</code>	Temperature indices

Application: a case study with common bean on-farm trials

During five growing seasons (from 2015 to 2017) in Nicaragua, van Etten *et al.* (2019) conducted a crowdsourced on-farm trial experiment following the *tricot* approach (van Etten *et al.*, 2016) testing 11 common bean varieties (*Phaseolus vulgaris* L.) as incomplete blocks of three randomly allocated to 842 blocks. A Plackett–Luce model was used to analyse the data. The model estimates the *worth*, the probability of each variety to outperform all the others in the set (Luce, 1959; Plackett, 1975; Turner *et al.*, 2020). Here we reproduce part of this analysis on the calculation and application of the climate indices. The approach here is slightly different because it considers the growing-degree days from planting date to maturity and add new indices to illustrate the package implementation. The data is available as `data("commonbean", package = "climatrends")`.

We estimate the crop phenological stages based on the growing degree-days using the function `GDD()`. For common bean, we define 900 degree-days, from planting date to maturity (de Medeiros *et al.*, 2016). The input data is an array with the MODIS temperature data (Wan *et al.*, 2015), the vector with planting dates (`cbean$planting_date`), the required amount of degree-days passed to the argument `degree.days` and the character string 'ndays' specifying that the function must return the values as number of days. `GDD()` computes the degree-days for the time series and return the length of the vector where the accumulated growing degree-days reached the pre-defined threshold of 900 degree-days.

The degree-days spanned from 54 to 100 days as shown in Fig. 1a. For simplicity we take the average per season and use this vector to compute the temperature indices.

```
library("climatrends")
library("PlackettLuce")
library("tidyverse")

data("commonbean", package = "climatrends")

cbean <- commonbean[[1]]
modis <- commonbean[[2]]
```

```
# number of days required to accumulate gdd from planting date to maturity
gdd <- GDD(modis,
  day.one = cbean$planting_date,
  degree.days = 900,
  return.as = "ndays")
```

```
# gdd to the cbean data and take the average gdd per season
cbean %<>%
  mutate(gdd = gdd$gdd) %>%
  group_by(season) %>%
  mutate(gdd_season = as.integer(mean(gdd)))
```

To compute the temperature indices we use the function `temperature()`. In van Etten (2019), a forward variable selection was applied to retain the most representative covariates based on the deviance reduction. This analysis retained the maximum night temperature (`maxNT`) as the most representative covariate. To illustrate how the Plackett-Luce trees can grow in complexity as we add more indices, we included the summer days (`SU`, number of days with maximum day temperature $> 30^{\circ}\text{C}$) together with `maxNT`.

```
# temperature indices from planting date to the
# number of days required to accumulate the gdd in each season
temp <- temperature(modis,
  day.one = cbean$planting_date,
  span = cbean$gdd_season)
```

```
cbean <- cbind(cbean, temp)
```

```
# fit a Plackett-Luce tree
plt <- pltree(G ~ maxNT + SU, data = cbean, minsize = 50)
```

Seasonal distribution of `maxNT` captured for each incomplete block in this experiment is shown in Fig. 1b. The data has a bimodal distribution which is reflected in the splitting value (18.7°C) for the Plackett-Luce trees in Fig. 1c. The upper node splits with 49 summer days (`SU`). We can interpret these results as that the on-farm performance of common bean varieties is driven by heat accumulation of diurnal temperature above 30°C (in this case $>70\%$ of the growing days) and warmer nights ($> 18.7^{\circ}\text{C}$).

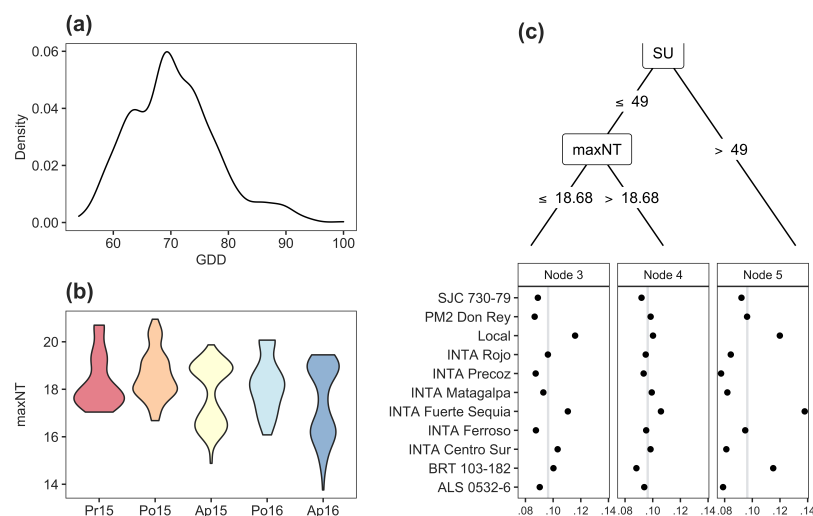


Figure 1: Fig. 1. Application of `climatrends` functions to support the analysis of decentralized on-farm trial data. (A) Days required to reach 900 growing-degree days from planting date calculated using the function `GDD()`. (B) Maximum night temperature (°C) distributed across seasons computed using the function `temperature()`. (C) Plackett-Luce Tree showing the probability that a given variety outperforms the other varieties (axys X) in three different nodes splitted with the summer days (day temperature > 30 °C) and maximum night temperature (°C). Note: the first season (primera, Pr) spans from May to August, the second (postrera, Po) from September to October, and the third (apante, Ap) from November to January.

Further development

The package can support the integration with other datasets as they become available in R via API client packages. In the future, new indices related to the physiology of crops could be implemented. Integration with daily high resolution climate data and seasonal forecast for East Africa from Tomorrow.io's Comprehensive Bespoke Atmospheric Model (CBAM) is under development. To explore the latest functionalities of `climatrends`, please check the package's updates at CRAN (<https://cran.r-project.org/package=climatrends>).

Acknowledgements

This work was supported by The Nordic Joint Committee for Agricultural and Food Research (grant 202100-2817). Additional support was provided by the projects Accelerated Varietal Improvement and Seed Systems in Africa (AVISA, INV-009649) and 1000FARMS (INV-031561) supported by the Bill & Melinda Gates Foundation. The views expressed in this document cannot be taken to reflect the official opinions of these organizations.

References

- Aguilar, E., Peterson, T. C., Obando, P. R., Frutos, R., Retana, J. A., Solera, M., Soley, J., García, I. G., Araujo, R. M., Santos, A. R., Valle, V. E., Brunet, M., Aguilar, L., Álvarez, L., Bautista, M., Castañón, C., Herrera, L., Ruano, E., Sinay, J. J., ... Mayorga, R. (2005). Changes in precipitation and temperature extremes in Central America and northern South America, 1961–2003. *Journal of Geophysical Research*, 110(D23), D23107. <https://doi.org/10.1029/2005JD006119>

- Baker, P., & Mortlock, M. (2021). *cropgrowsdays: Crop Growing Degree Days and Agrometeorological Calculations*. <https://CRAN.R-project.org/package=cropgrowsdays>
- Brown, D., de Sousa, K., & van Etten, J. (2023). ag5Tools: An R package for downloading and extracting agrometeorological data from the AgERA5 database. *SoftwareX*, 21, 101267. <https://doi.org/10.1016/j.softx.2022.101267>
- Challinor, A. J., Koehler, A.-K., Ramirez-Villegas, J., Whitfield, S., & Das, B. (2016). Current warming will reduce yields unless maize breeding and seed systems adapt immediately. *Nature Climate Change*, 6(10), 954–958. <https://doi.org/10.1038/nclimate3061>
- de Medeiros, G. A., Daniel, L. A., & Fengler, F. H. (2016). Growth, Development, and Water Consumption of Irrigated Bean Crop Related to Growing Degree-Days on Different Soil Tillage Systems in Southeast Brazil. *International Journal of Agronomy*, 2016, 8065985. <https://doi.org/10.1155/2016/8065985>
- de Sousa, K., Casanoves, F., Sellare, J., Ospina, A., Suchini, J. G., Aguilar, A., & Mercado, L. (2018). How climate awareness influences farmers' adaptation decisions in Central America? *Journal of Rural Studies*, 64, 11–19. <https://doi.org/10.1016/j.jrurstud.2018.09.018>
- de Sousa, K., Sparks, A., Ashmall, W., van Etten, J., & Solberg, S. (2020). chirps: API Client for the CHIRPS Precipitation Data in R. *Journal of Open Source Software*, 5(51), 2419. <https://doi.org/10.21105/joss.02419>
- Funk, C., Peterson, P., Landsfeld, M., Pedreros, D., Verdin, J., Shukla, S., Husak, G., Rowland, J., Harrison, L., Hoell, A., & Michaelsen, J. (2015). The climate hazards infrared precipitation with stations—a new environmental record for monitoring extremes. *Scientific Data*, 2, 150066. <https://doi.org/10.1038/sdata.2015.66>
- Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., ... Thépaut, J.-N. (2020). The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, 146(730), 1999–2049. <https://doi.org/10.1002/qj.3803>
- Kehel, Z., Crossa, J., & Reynolds, M. (2016). Identifying Climate Patterns during the Crop-Growing Cycle from 30 Years of CIMMYT Elite Spring Wheat International Yield Trials. In A. Bari, A. B. Damania, M. Mackay, & S. Dayanandan (Eds.), *Applied mathematics and omics to assess crop genetic resources for climate change adaptive traits* (pp. 151–174). CRC Press. <https://doi.org/10.1201/9781315369006-13>
- Liu, Y., & El-Kassaby, Y. A. (2018). Evapotranspiration and favorable growing degree-days are key to tree height growth and ecosystem functioning: Meta-analyses of Pacific Northwest historical data. *Scientific Reports*, 8228. <https://doi.org/10.1038/s41598-018-26681-1>
- Luce, R. D. (1959). *Individual Choice Behavior* (p. 153). Courier Corporation.
- Munafò, M. R., Nosek, B. A., Bishop, D. V. M., Button, K. S., Chambers, C. D., Percie du Sert, N., Simonsohn, U., Wagenmakers, E.-J., Ware, J. J., & Ioannidis, J. P. A. (2017). A manifesto for reproducible science. *Nature Human Behaviour*, 1(1), 21. <https://doi.org/10.1038/s41562-016-0021>
- Nowosad, J. (2019). *pollen: Analysis of Aerobiological Data*. <https://CRAN.R-project.org/package=pollen>
- Pebesma, E. (2018). Simple Features for R: Standardized Support for Spatial Vector Data. *The R Journal*, 10(1), 439–446. <https://doi.org/10.32614/RJ-2018-009>
- Plackett, R. L. (1975). The Analysis of Permutations. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 24(2), 193–202. <https://doi.org/10.2307/2346567>

- Powers, S. M., & Hampton, S. E. (2019). Open science, reproducibility, and transparency in ecology. *Ecological Applications*, 29(1), e01822. <https://doi.org/10.1002/eap.1822>
- Prentice, I. C., Cramer, W., Harrison, S. P., Leemans, R., Monserud, R. A., & Solomon, A. M. (1992). Special Paper: A Global Biome Model Based on Plant Physiology and Dominance, Soil Properties and Climate. *Journal of Biogeography*, 19(2), 117. <https://doi.org/10.2307/2845499>
- R Core Team. (2020). *R: A language and environment for statistical computing. version 4.0.0*. CRAN R Project. <https://www.r-project.org/>
- Schulze, E.-D., Beck, E., Buchmann, N., Clemens, S., Müller-Hohenstein, K., & Scherer-Lorenzen, M. (2019). *Plant Ecology* (Second Edi, p. 925). Springer. <https://doi.org/10.1007/978-3-662-56233-8>
- Serrano-Notivoli, R. (2022). *agroclim: Climatic Indices for Agriculture*. <https://CRAN.R-project.org/package=agroclim>
- Sparks, A. H. (2018). nasapower: A NASA POWER Global Meteorology, Surface Solar Energy and Climatology Data Client for R. *Journal of Open Source Software*, 3(30), 1035. <https://doi.org/10.21105/joss.01035>
- Stevens, J. R. (2017). Replicability and Reproducibility in Comparative Psychology. *Frontiers in Psychology*, 8, 862. <https://doi.org/10.3389/fpsyg.2017.00862>
- Turner, H. L., van Etten, J., Firth, D., & Kosmidis, I. (2020). Modelling rankings in R: the Plackett-Luce package. *Computational Statistics*. <https://doi.org/10.1007/s00180-020-00959-3>
- van Etten, J., Beza, E., Calderer, L., van Duijvendijk, K., Fadda, C., Fantahun, B., Kidane, Y. G., van de Gevel, J., Gupta, A., Mengistu, D. K., Kiambi, D., Mathur, P. N., Mercado, L., Mittra, S., Mollel, M. J., Rosas, J. C., Steinke, J., Suchini, J. G., & Zimmerer, K. S. (2016). First experiences with a novel farmer citizen science approach: crowdsourcing participatory variety selection through on-farm triadic comparisons of technologies (tricot). *Experimental Agriculture*, 55(S1), 275–296. <https://doi.org/10.1017/S0014479716000739>
- van Etten, J., de Sousa, K., Aguilar, A., Barrios, M., Coto, A., Dell'Acqua, M., Fadda, C., Gebrehawaryat, Y., van de Gevel, J., Gupta, A., Kiros, A. Y., Madriz, B., Mathur, P., Mengistu, D. K., Mercado, L., Nurhisen Mohammed, J., Paliwal, A., Pè, M. E., Quirós, C. F., ... Steinke, J. (2019). Crop variety management for climate adaptation supported by citizen science. *Proceedings of the National Academy of Sciences*, 116(10), 4194–4199. <https://doi.org/10.1073/pnas.1813720116>
- Wan, Z., Hook, S., & Hulley, G. (2015). MYD11A1 MODIS/Aqua Land Surface Temperature/Emissivity Daily L3 Global 1km SIN Grid V006 [Data set]. *NASA EOSDIS LP DAAC*. <https://doi.org/10.5067/MODIS/MYD11A1.006>