




The MacLeish Package: R Package for analyzing weather and spatial data

Benjamin S. Baumer¹, Dominique Kelly¹, and Albert Y. Kim¹

DOI:

1 Smith College, Northampton, Massachusetts, USA

Software

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

Submitted:

Published:

License

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

Abstract

The Ada & Archibald MacLeish Field Station is a 250-acre patchwork of forest and pasture land in Whately, MA that presents the opportunity for all members of the Smith College community (students, faculty, staff, and alum) to pursue artistic inquiry, environmental research, outdoor education, and low-impact recreation. The `macleish` R package was created to analyze weather and spatial data based on the MacLeish Field Station.

The package includes the following functionality: multiple layers of map data coded as `sf` package objects, functions to access historical and live weather readings, data from ongoing forest ecology research, and functions to access photos from a PhenoCam (Phenology Camera) Network video camera mounted on the 25-meter tower records a daily photograph of the forest canopy.

By making this data accessible in an R Package with a simple interface, we hope the `macleish` R package will serve as an educational tool that fosters interdisciplinary collaboration within the Smith College community. At the moment, the package is mostly used in the Statistical & Data Science department, but computational confidence is expected to increase in the Biology department and Environmental Science & Policy program. Accordingly we hope the students in these programs not only use the package but also contribute to it. Furthermore, we view this package as a blueprint for other institutions that possess similar research sites to bring the related research communities together.

Introduction

Computational analysis is becoming a crucial component of research in ecology. However, programming presents new challenges for ecological investigators who are unfamiliar with statistical applications. Even researchers that are versed in efficient programming may be restricted by knowledge of code. Therefore, interdisciplinary research between ecology and statistics remains limited in scope, depth, and quality (Visser et al. (2015)).

Here, we create an R Package with a simple interface that access to weather, spatial, and photography data to inspire interdisciplinary collaboration. The `macleish` Package, uses data from the Ada and Archibald MacLeish Field Station.

The Ada and Archibald MacLeish Field Station is a 243-acre patchwork of forest and pasture land on top of a flat ridge in Whately, Massachusetts (See Figure 1). The station is located about 11 miles from Smith College. The community at Smith College, including students, faculty, staff, and alum, pursue artistic inquiry, environmental research, outdoor education, and low-impact recreation at the MacLeish Field Station. Students are integral to the operation and maintenance of the field as stewards for the land, designers of programs, and participants in key decision-making (“MacLeish Field Station,” n.d.)).

Although a diverse ecosystem exists in the MacLeish Field Station, invasive insects and

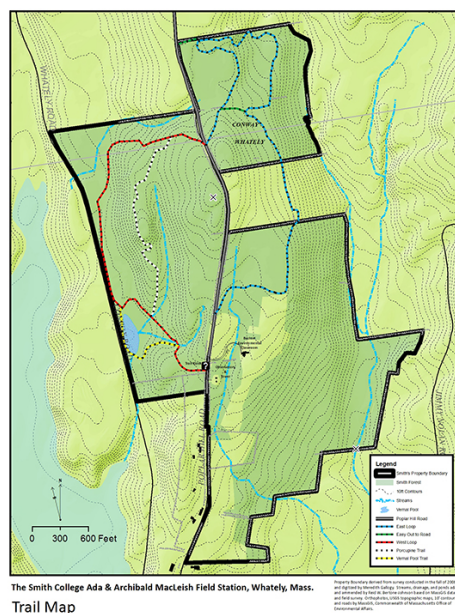


Figure 1: Map of MacLeish Field Station, Whately, Massachusetts, USA.

fungus have drastically altered the forest composition. American chestnut trees were once the most common tree in New England forests. However, due to the lethal chestnut blight, a fungus from Asia, most chestnut trees died because of infection. To harbor the survival of the American chestnut tree, the tree was crossed with the Chinese chestnut tree, which is resistant to the blight, to create a blight-resistant hybrid. White pine trees were another victim of a fungus native to Asia, referred to as the white pine blister rust. Similar to the chestnut blight, the white pine blister is fatal. Although decades of effort to control the white pine blister rust conserved most of the white pine tree population, a combination of logging and the blister rust greatly reduced the composition of forests at MacLeish. Furthermore, the Eastern Hemlock tree is a martyr of the Hemlock Woolly Adelgid, an invasive insect native to Asia. High volumes of woolly adelgid populations can kill a tree within a few years by weighing down the branches and causing the needles to fall off. Finally, the beech scale insect is an invasive insect from Europe. The insect is capable of carrying the fungus *Nectria coccinea*, which is transmitted into the tree when the insects feed on beech sap. *Nectria coccinea* causes the bark of beech trees to blister and crack, making the tree susceptible to further infection by insects and other kinds of fungus. Beech trees infected with the fungus die within six years. Unfortunately, both the Hemlock Woolly adelgid and the beech scale insect are highly resistant to the cold and thrive during the warm. Therefore, trees are threatened by these insects year-round. Similarly, as the climate is projected to warm, the range of both insects will continue to expand (Moise, Wiese, & Sorensen (2021)).

Likewise, other invasive species exist at the MacLeish Field Station. The autumn olive, Japanese barberry, multiflora rose, oriental bittersweet, and wild grape are all invasive species that contend with native Massachusetts plant species. The ability of these invasive species to produce an abundant quantity of seeds, grow quickly, and thrive without natural predators allow them to pose a serious threat to out-compete native species. Conservation research and efforts are ongoing to address the expansion of these invasive species (Moise et al. (2021)).

Within the MacLeish Field Station are a number of stations and administrative buildings

that collect data and foster environmental education. For instance, the Whately and Orchard Weather Stations store historical and collect current meteorological weather data. The Smith College Observatory is frequently used by the astronomy department at Smith College for astronomical research. The Bechtel Environmental Classroom was created by the Environmental Science & Policy program using top-notch sustainable strategies. The classroom provides a meeting space for environmental laboratory and educational use ("MacLeish Field Station," n.d.)). Evidently, the MacLeish Field Station is highly interdisciplinary, appealing to multiple departments at once, and aims to inspire research in the arts, humanities, and natural sciences among the Smith College community. Through the `macleish` Package, we hope to expand the opportunity for members of Smith College to pursue research, contribute to the R Package, and influence other institutions with analogous research sites to create more accessible programming interfaces.

Examples of research conducted at MacLeish by Smith Faculty include Zukswert et al. (2014), Ignace, Fassler, & Bellemare (2018), Guswa (2011), Jackson & Bellemare (2018), Guswa (2011), Bassar, Letcher, Nislow, & Whiteley (2015), Wetzel (2017), Visser et al. (2015), and Moise et al. (2021).

Statement of Need

The MacLeish Field Station is abundant with scientific inquiry, where the Smith College community explore the diverse ecosystem of plants and animals. The `macleish` package offers access to

1. Data from open-source databases
2. interactive maps of the MacLeish Field Station and surrounding area
3. Meteorological reading data from historical and current databases
4. Data on tree diameter of different species infected by the Hemlock Woolly adelgid
5. A database of images of tree canopy over time

Users may generate usable summary statistics and plots using the data from the `macleish` package.

macleish Package Main Functionalities

To use the `macleish` package, first run `install.packages(macleish)` to install the package in R, then run the following code to load the package:

```
library(macleish)
```

Maps of MacLeish

Shapefile-based maps of MacLeish Field Station are available in the `macleish_layers` data object. The following layers will be displayed on the map.

```
names(macleish_layers)
## [1] "landmarks"      "forests"        "streams"
## [4] "challenge_courses" "buildings"      "wetlands"
## [7] "boundary"       "research"       "soil"
## [10] "trails"         NA               NA
## [13] "camp_sites"     "elevation"
```

Running the following code creates an interactive `leaflet` map for users to zoom in and out, identify the Orchard and Whately Weather Stations, locate the area of MacLeish Field Station, and pinpoint other geographical features such as bodies of water, roads,

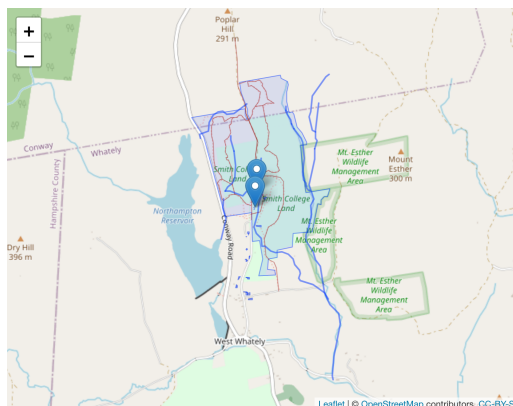


Figure 2: Interactive 'leaflet' map of the MacLeish Field Station.

mountains, trails, and wildlife management areas. Figure 2 displays a static image of the resulting interactive map.

```
map <- leaflet() %>%
  addTiles() %>%
  addPolygons(
    data = macleish_layers[["boundary"]],
    weight = 1, fillOpacity = 0.1
  ) %>%
  addPolygons(
    data = macleish_layers[["buildings"]],
    weight = 1, popup = ~ name
  ) %>%
  addPolylines(
    data = macleish_layers[["trails"]],
    weight = 1, color = "brown",
    popup = ~ name
  ) %>%
  addPolylines(
    data = macleish_layers[["streams"]],
    weight = 2
  ) %>%
  addMarkers(
    data = filter(macleish_layers[["landmarks"]], grepl("Met", Label)),
    popup = ~Label
  )

include_graphics(path = "leaflet.png")
```

Field data

Two tree plot datasets is available through `tree_diameter1` and `tree_diameter2`. The time frame varies between `tree_diameter1` (2010-2015) and `tree_diameter2` (2009-2016), and `tree_diameter2` lacks data on position of individual trees. The datasets present change in tree diameter over a set of years for two plots dominated by the parasites Hemlock Woolly Adelgid. See Tables 1 and 2.

```
tree_diameter1
```

```
tree_diameter1 %>%
  head() %>%
  select(-module) %>%
  kable(
    format = "latex",
    caption = "Tree diameter dataset 1.",
    booktabs = TRUE,
    escape = FALSE,
    linesep = ""
  ) %>%
  kable_styling(
    latex_options = c("hold_position")
  )
```

Table 1: Tree diameter dataset 1.

tag	species	position	semester	year	diameter	notes
2001	Sweet Birch	subcanopy	Fall	2015	NA	Alive in 2009, dead in 2010
2001	Sweet Birch	subcanopy	Fall	2014	NA	Alive in 2009, dead in 2010
2001	Sweet Birch	subcanopy	Fall	2013	NA	Alive in 2009, dead in 2010
2001	Sweet Birch	subcanopy	Fall	2012	NA	Alive in 2009, dead in 2010
2001	Sweet Birch	subcanopy	Fall	2011	NA	Alive in 2009, dead in 2010
2001	Sweet Birch	subcanopy	Fall	2010	NA	Alive in 2009, dead in 2010

```
tree_diameter2

tree_diameter2 %>%
  head() %>%
  select(-module) %>%
  kable(
    format = "latex",
    caption = "Tree diameter dataset 2.",
    booktabs = TRUE,
    escape = FALSE,
    linesep = ""
  ) %>%
  kable_styling(
    latex_options = c("hold_position")
  )
```

Table 2: Tree diameter dataset 2.

tag	species	semester	year	diameter	notes
2171	Eastern Hemlock	Fall	2016	28.2	NA
2171	Eastern Hemlock	Fall	2012	27.4	NA
2171	Eastern Hemlock	Fall	2011	27.2	NA
2171	Eastern Hemlock	Fall	2010	27.1	NA
2171	Eastern Hemlock	Fall	2009	27.0	NA
2172	Sweet Birch	Fall	2016	24.5	NA

For example, using the `tree_diameter1` dataset, the boxplot below demonstrates the distribution of diameter for different species of trees at different positions from 2010-2015.

By viewing an individual species, the distribution of diameter varies based on the position

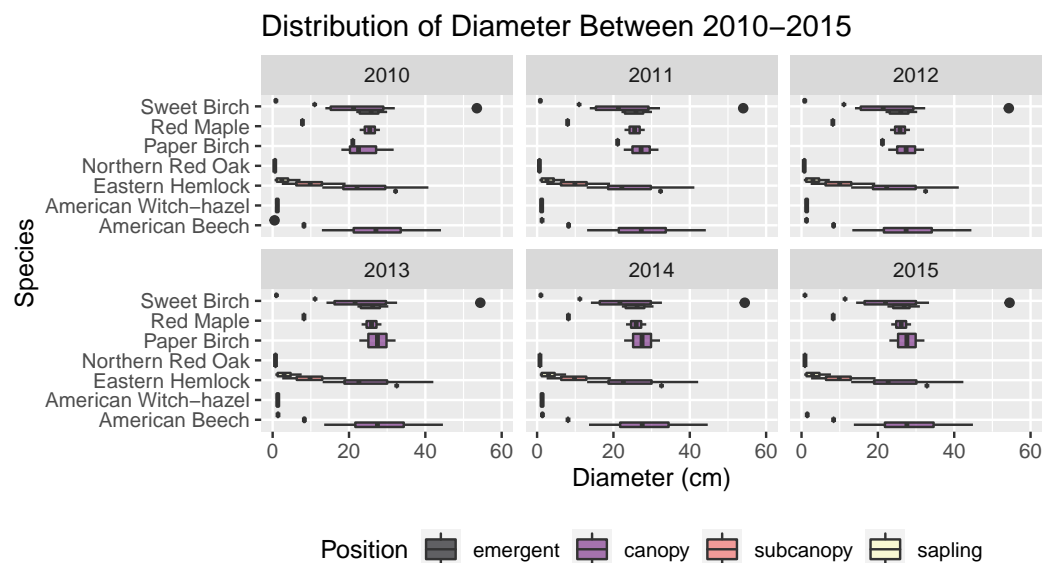


Figure 3: Boxplot of tree diameter depending on species and position between 2010-2015.

of the species. For instance, the median diameter of emergent Eastern Hemlock trees is the greatest, despite having the smallest overall range. However, the overall range of the diameter of the canopy is the greatest.

Below is a box plot of the distribution of diameter based on different tree species from 2009-2012 using the `tree_diameter2` dataset. Using individual data points on top of the box plots reveals the underlying distribution of each group.

Weather data

Weather data from 2015 is available through `whately_2015` and `orchard_2015` data objects. Data was recorded every 10 minutes every day from 2015-01-01 to 2015-31-12.

`whately_2015`

```
## # A tibble: 52,560 x 8
##   when                temperature wind_speed wind_dir rel_humidity
##   <dtm>                <dbl>         <dbl>   <dbl>      <dbl>
## 1 2015-01-01 00:00:00      -9.32           1.40    225.       54.6
## 2 2015-01-01 00:10:00      -9.46           1.51    248.       55.4
## 3 2015-01-01 00:20:00      -9.44           1.62    258.       56.2
## 4 2015-01-01 00:30:00      -9.3           1.14    244.       56.4
## 5 2015-01-01 00:40:00      -9.32           1.22    238.       56.9
## 6 2015-01-01 00:50:00      -9.34           1.09    242.       57.2
## 7 2015-01-01 01:00:00      -9.3           1.17    242.       57.7
## 8 2015-01-01 01:10:00      -9.1           1.31    244.       58.2
## 9 2015-01-01 01:20:00      -9.07           1.31    226.       59.0
## 10 2015-01-01 01:30:00      -8.99           1.81    220        59.3
## # ... with 52,550 more rows, and 3 more variables: pressure <int>,
## #   solar_radiation <dbl>, rainfall <dbl>
```

`orchard_2015`

```
## # A tibble: 52,547 x 9
##   when                temperature wind_speed wind_dir rel_humidity
##   <dtm>                <dbl>         <dbl>   <dbl>      <dbl>
## 1 2015-01-01 00:00:00      -9.62           0.49    216.       62.0
```

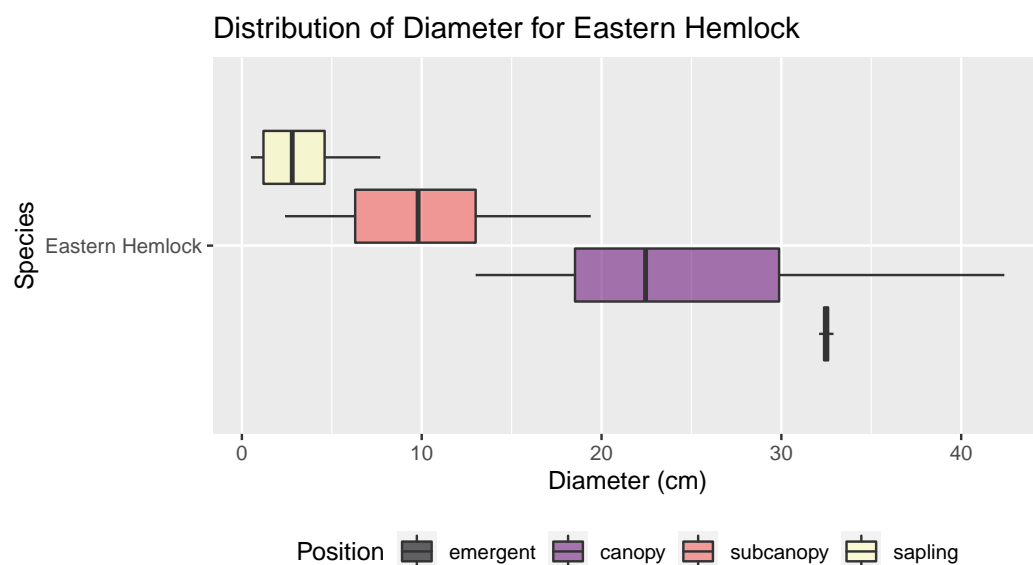


Figure 4: Boxplot of tree diameter of Eastern Hemlock tree depending on position between 2010-2015.

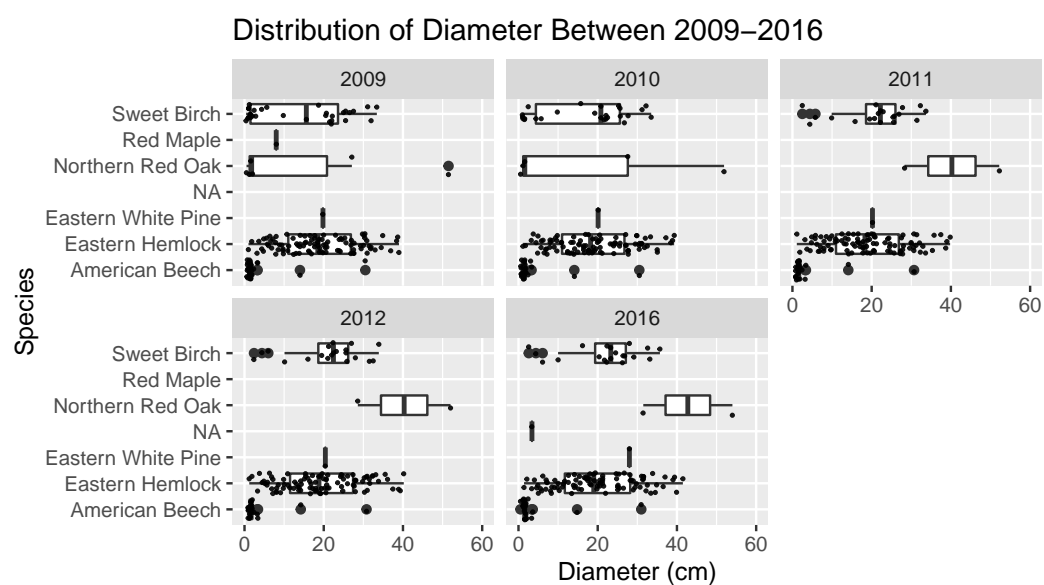


Figure 5: Boxplot of tree diameter depending on species between 2009-2016.



Figure 6: Image of the MacLeish Field Station taken above retrieved from the PhenoCam Network.

```
## 2 2015-01-01 00:10:00      -9.45      0.517      210.      59.1
## 3 2015-01-01 00:20:00      -9.52      0.576      216.      58.7
## 4 2015-01-01 00:30:00      -9.68      0.522      199.      58.6
## 5 2015-01-01 00:40:00      -9.56      0.296      186.      57.3
## 6 2015-01-01 00:50:00      -9.59      0.382      184.      57.6
## 7 2015-01-01 01:00:00      -9.51      0.497      196.      57.0
## 8 2015-01-01 01:10:00      -9.7      0.781      210.      60.5
## 9 2015-01-01 01:20:00      -9.7      0.787      217.      61.1
## 10 2015-01-01 01:30:00     -9.34      0.693      209.      59.8
## # ... with 52,537 more rows, and 4 more variables: pressure <dbl>,
## #   par_density <dbl>, par_total <dbl>, rainfall <dbl>
```

Furthermore, the ETL framework retrieves live and historical (2012-01-03 for Whately and 2014-06-27 for Orchard) weather data from both MacLeish Field Station weather monitors.

Phenocam functions

The PhenoCam Network is an ecosystem phenology camera network that provides automated, near-surface remote sensing of canopy phenology across the northeastern United States and adjacent Canada, including the MacLeish Field Station. High-resolution digital cameras at the MacLeish Field Station take pictures of the canopy and upload them to the PhenoCam server every half hour. The PhenoCam Network permits analysis of archived images, providing an objective means by which canopy phenology can be monitored and quantified at relatively low cost and with minimal expenses.

Figure 6 displays an example of an image retrieved from the PhenoCam Network.

The function `get_macleish_image()` retrieves an image from the PhenoCam Network when the date and time are specified.

Application and Perspectives

The `macleish` package is an R Package with a simple interface that access to weather, spatial, and photography data to inspire interdisciplinary collaboration among statisticians and ecologists. This is a dynamic project that was built with the interest of biologists. It currently supports interdisciplinary collaboration to foster statistical application among the ecology community. More functions and data are expected to be compiled in the package over time that seems fit to users to support an increasing collection of statistical analyses and visualizations.

Acknowledgements

We acknowledge contributions from Jesse Bellemare and Paul Wetzel.

References

- Bassar, R. D., Letcher, B. H., Nislow, K. H., & Whiteley, A. R. (2015). Changes in seasonal climate outpace compensatory density-dependence in eastern brook trout. *Global Change Biology*. Retrieved from <https://onlinelibrary.wiley.com/doi/pdf/10.1111/gcb.13135>
- Guswa, A. J. (2011). Canopy vs. Roots: Produce on and Destruction of Variability in Soil Moisture and Hydrologic Fluxes. *Vadose Zone Journal*. Retrieved from <https://access.onlinelibrary.wiley.com/doi/10.2136/vzj2011.0159>
- Ignace, D. D., Fassler, A., & Bellemare, J. (2018). Decline of a foundation tree species due to invasive insects will trigger net release of soil organic carbon. *Ecosphere*, 9:8. Retrieved from <https://esajournals.onlinelibrary.wiley.com/doi/pdf/10.1002/ecs2.2391>.
- Jackson, M. R., & Bellemare, J. (2018). The potential for indirect negative effects of exotic insect species on a liverwort, *Bazzania trilobata* (Lepidoziaceae), mediated by the decline of a foundation tree species, *Tsuga canadensis* (Pinaceae). *Journal of the Torrey Botanical Society*. doi:<https://doi.org/10.3159/TORREY-D-17-00040.1>
- MacLeish Field Station. (n.d.). Retrieved from <https://www.smith.edu/about-smith/sustainable-smith/macleish>
- Moise, A., Wiese, S., & Sorensen, C. (2021). Land Use History of MacLeish: A Cultural and Ecological Timeline of the Smith College Field Station. Retrieved from <https://storymaps.arcgis.com/stories/14c7392dbd4e42d4a4ef7775cb56062c>
- Visser, M. D., McMahon, S. M., Merow, C., Dixon, P. M., Record, S., & Jongejans, E. (2015). Speeding Up Ecological and Evolutionary Computations in R; Essentials of High Performance Computing for Biologists. *PLOS*. Retrieved from <https://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1004140>
- Wetzel, P. (2017). 2016 Annual Weather Report: Smith College Ada & Archibald MacLeish Field Station Whately, Massachusetts, USA. Retrieved from https://docs.google.com/document/d/1MZc7F7fP5iBrJrjF6_Ovg7FfCs7WoIk_UIFPB4aAcE/edit?usp=sharing
- Zukswert, J. M., Bellemare, J., Rhodes, A. L., Sweezy, T., Gallogly, M., Acevedo, S., & Taylor, R. S. (2014). Forest Community Structure Differs, but not Ecosystem Processes,

25 Years after Eastern Hemlock Removal in an Accidental Experiment. *JSTOR*. Retrieved from <https://www.jstor.org/stable/26454958>