IMPLEMENTATION OF A VIRTUAL RESEARCH ENVIRONMENT (VRE) TO THE STUDY OF FOREST ENVIRONMENTS

EGU General Assembly 2025
Vienna, Austria & Online
27 April–2 May 2025

Costafreda-Aumedes S.[1], Tagliolato P.[2], Giusti R[1]., Iannuccili M[1]., Matteucci G.[3], Mazzenga F[3]., Messeri A[1], Oggioni A.[2].

[1]CNR – Institute of BioEconomy (IBE), Via Madonna del Piano 10, 50019, Sesto Fiorentino. Italy
[2] CNR - Institute for Electromagnetic Sensing of the Environment, (IREA), Via Alfonso Corti 12, I-20133, Milano, Italy
[3]CNR – Institute of BioEconomy (IBE), Via dei Taurini 19, 00185, Roma. Italy

INTRODUCTION:

The monitoring and study of forests is essential to understand their condition, dynamics and to adopt optimal management to ensure their sustainability. In this framework, Virtual Research Environments (VRE) are online research platforms that allow easy access to the available FAIR data, to find smart solutions and to support decision making. Here we show the VRE - Essential Variables (within the ITINERIS project) which offers several tools to develop open and reproducible science in the perspective of the two global frameworks of Essential Biodiversity Variables (EBVs) and Essential Climate Variables (ECVs), which are known to be critical for plant and soil biogeochemical processes (e.g., tree growth, soil mineralization, water fluxes, litter decomposition). In the VRE, which also includes an R package and a Shiny interface, only some of the variables are considered: among the EBVs, phenology and species distribution; among the ECVs, surface air temperature, precipitation, and relative humidity.

STUDY AREA:

Collelongo - Selva Piana (https://deims.org/9b1d144a-dc37-4b0e-8cda-1dda1d7667da) is a long-term experimental site of 3000-ha mature beech forest (*Fagus sylvatica L.*; >125 years old), one of the founding sites of the Italian ICP Forests network and also part of the eLTER and AnaEE international Research Infrastructures.

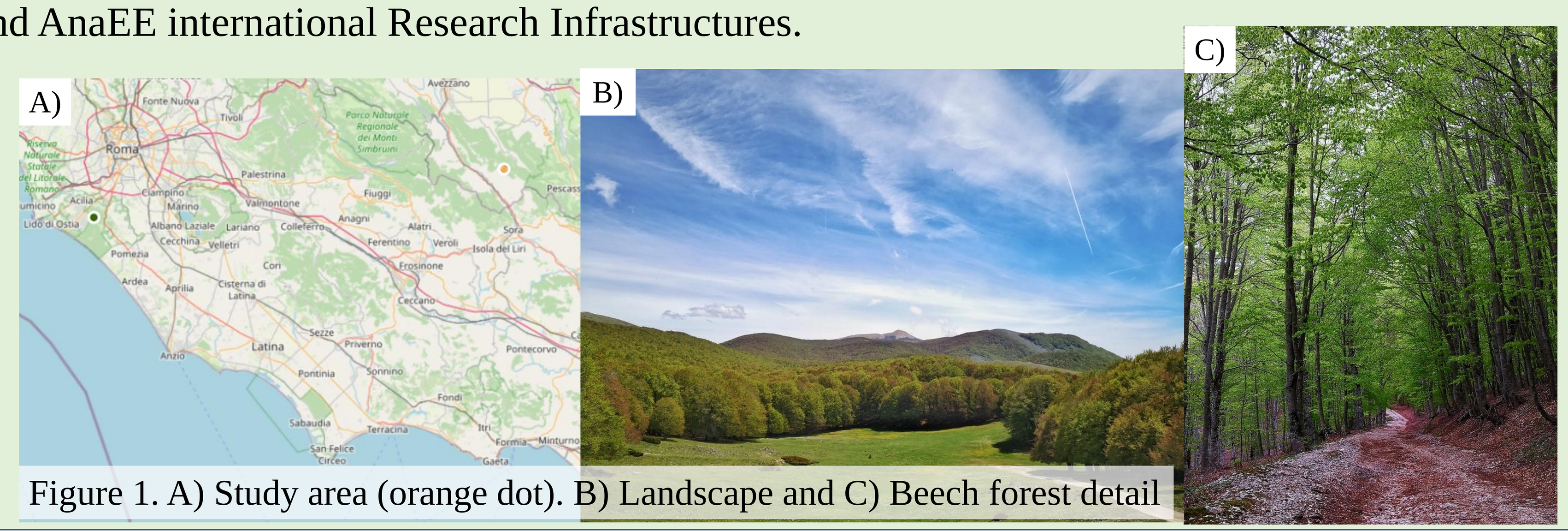

Figure 1. A) Study area (orange dot). B) Landscape and C) Beech forest detail

VRE – Essential Variables:

The ITINERIS.VRE-EV compiles and elaborates existing repositories of terrestrial, marine and freshwater study areas belonging to the eLTER network. The existing datasets are downloaded from data-sharing repositories such as Zenodo or Pangaea, commonly used as a repository by eLTER sites for share dataset, as well as from entities such as Copernicus, PEP725, CEDA, ESA or GBIF, among others. Access to these data is provided through the functions of the ReLTER package (DOI 10.5281/zenodo.7584997, Oggioni et al., 2025).

The functions developed in an R package and the Shiny interface allow users to explore and download datasets, available in different repositories, that contribute to the EBVs or ECVs. In particular, through the Shiny application (Fig. 2A), by selecting the eLTER site of interest and the essential variable, the user can obtain a table containing the relevant datasets and, if desired, download and use them within the VRE by relying on R or JupyterHub (Fig. 2B).

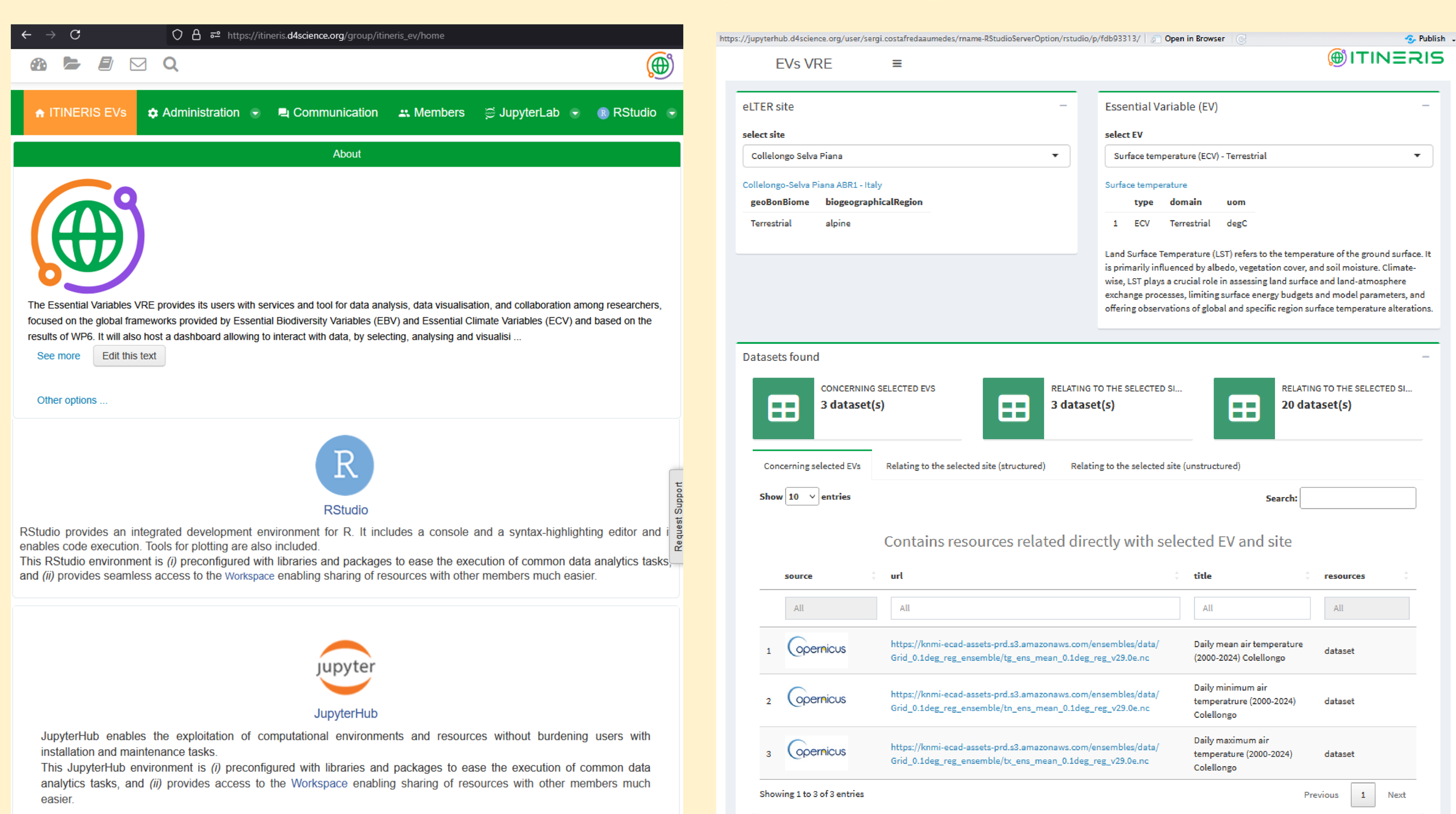

Figure 2. Itineris.D4Science Virtual Research Environment and Shinny App

PROCEDURE:

Terminal 1 | Script_EGU.ipynb

BODY OF THE SCRIPT

Reading the required libraries and sources

```
[3]: sapply(c("data.table", "dplyr", "readxl", "caret", "tidyverse", "MASS", "ReLTER"),
            require, character.only = TRUE)
```

data.table: TRUE **dplyr:** TRUE **readxl:** TRUE **caret:** TRUE **tidyverse:** TRUE **MASS:** TRUE **ReLTER:** TRUE

Constructing the functions to summarize the original data

Seasonaly merge the frequency of Atmospherical Circulation Types

```
[4]: func_season <- function(tab_case)
{
  # Splitting the table by month
  split_tab <- split(tab_case, c(1,1,2,2,2,3,3,3,4,4,4,1))

  # Seasonal frequencies
  summ_tab <- lapply(1:4, function(x) split_tab[[x]][3:length(split_tab[[x]])] %>%
                       colSums() %>%
                       t %>%
                       as.data.frame) %>%
    rbindlist %>%
    data.frame(Year = tab_case[1, 1],
               Season = c("DJF", "MAM", "JJA", "SON"),
               .) %>%
    reshape(idvar = c("Anno"), timevar = "Season", direction = "wide")

  # Adding the names of the columns
  names(summ_tab) <- gsub("X", "", names(summ_tab))

  # Showing the results
  summ_tab
}
```

Daily to monthly frequency

```
[5]: ff <- function(circo, circo_cases, x)
{
  aa = table((circo %>%
              filter(Anno == circo_cases[x, "Anno"] %>% unlist %>% unname,
                     Mese == circo_cases[x, "Mese"] %>% unlist %>% unname))$WT) %>%
    as.data.frame

  bb = t(aa$Freq) %>% as.data.frame; names(bb) <- aa$Var1

  bb
}
```

Calling the ITINERIS shiny app and download the data

```
ITINERIS.EVsVRE::runShinnyApp()
```

Data lecture

```
[6]: circo_pct <- read_excel("Weather types_ up to 2019.xlsx", sheet = 1)
circo_san <- read_excel("Weather types_ up to 2019.xlsx", sheet = 2)

gpp <- fread("_GPP_Reco_NEE_2000_2015.csv")
```

Data preparation

```
[7]: circo_cases_pct <- circo_pct %>% distinct(Anno, Mese)

# PCT (SAN NOT SHOWED)
tab_summ_month_pct <- lapply(1:nrow(circo_cases_pct),
       function(x) ff(circo_pct, circo_cases_pct, x)) %>%
  rbindlist(fill = T) %>%
  cbind(circo_cases_pct, .) %>%
  relocate('1', .after = "Mese") %>%
  mutate_if(is.numeric, ~replace(., is.na(.), 0))

tab_summ_season_pct <- lapply(1:40,
                    function(x) func_season((tab_summ_month_pct %>% group_by(Anno) %>% group_split())[[x]])) %>%
  rbindlist

# MERGING GPP and CIRCO SUMMARIES
gpp_season_pct <- merge(gpp, tab_summ_season_pct, by.x = "year", by.y = "Anno") %>% as.data.frame()

# Renaming the data
names(gpp_season_pct)[2:6] <- c("DT_RECO", "DT_GPP", "NT_GPP", "NT_RECO", "NEE")
names(gpp_season_pct)[7:length(names(gpp_season_pct))] <-
  gsub("\\.", "_", paste0("WT_", names(gpp_season_pct)[7:length(names(gpp_season_pct))]))
```

Data analysis

```
[8]: ## NT_GPP_PCT
sel_gppNT_pct <- c("WT_1_SON", "WT_6_DJF", "WT_4_JJA", "WT_7_JJA", "WT_4_SON", "WT_5_SON")
season_gppNT_pct <- lm(NT_GPP ~ ., data = gpp_season_pct[, c(which(names(gpp_season_pct) %in% c("NT_GPP", sel_gppNT_pct)))])
summary(season_gppNT_pct)

## NEE_PCT
sel_NEE_pct <- c("WT_4_JJA", "WT_7_JJA", "WT_4_SON", "WT_5_SON")
season_NEE_pct <- lm(NEE ~ ., data = gpp_season_pct[, c(which(names(gpp_season_pct) %in% c("NEE", sel_NEE_pct)))])
summary(season_NEE_pct)

## RECO_PCT
sel_RECO_pct <- c("WT_4_JJA", "WT_5_DJF", "WT_7_DJF")
season_RECO_pct <- lm(NT_RECO ~ ., data = gpp_season_pct[, c(which(names(gpp_season_pct) %in% c("NT_RECO", sel_RECO_pct)))])
summary(season_RECO_pct)
```

RESULTS:

WT 2, 4, 7, 9 (Cyclonic); WT 1, 3, 5 (Anticyclonic); WT 6, 8 (Zonal)

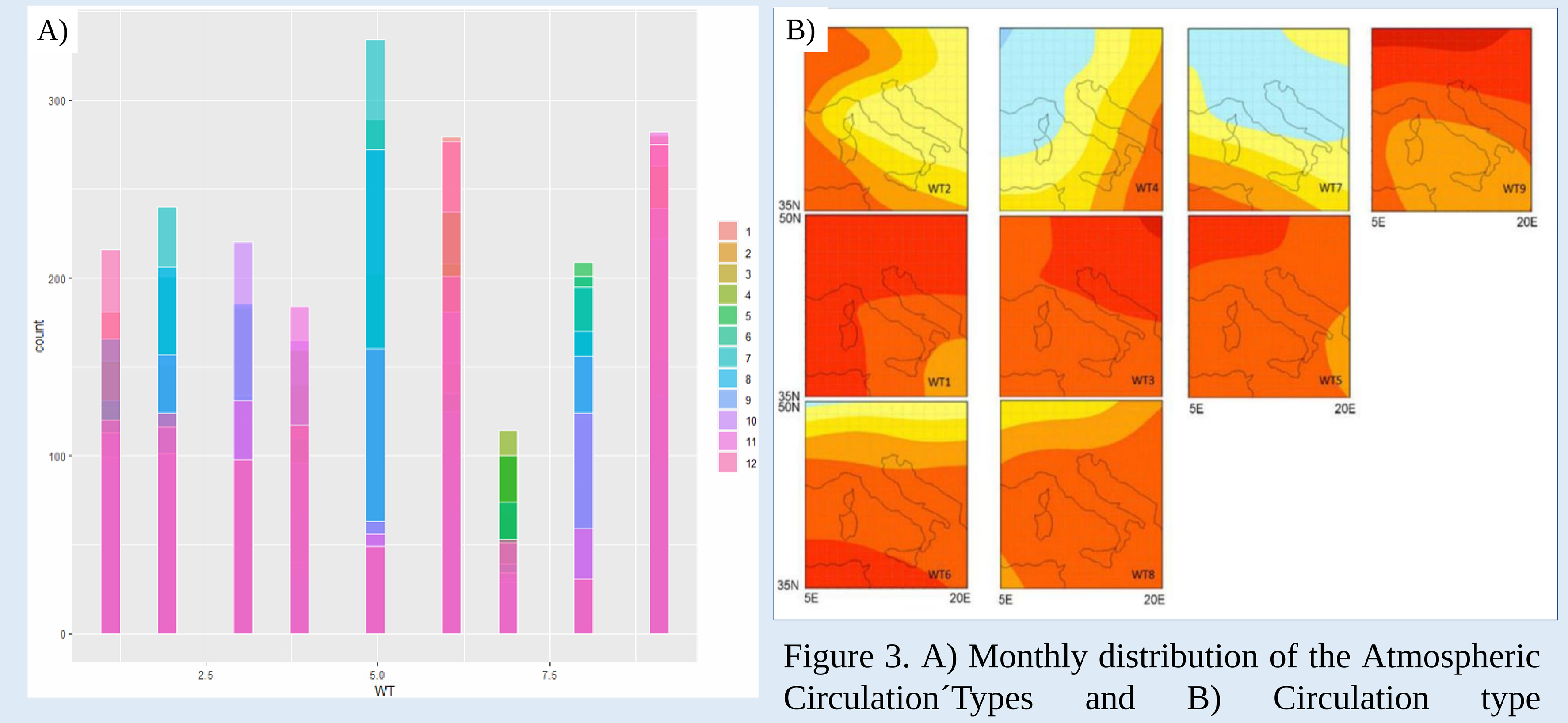

Figure 3. A) Monthly distribution of the Atmospheric Circulation´Types and B) Circulation type classification for Italy based on principal component methods(PCT) for precipitation.

1

Results of the best linear regression model between the annual GPP and the seasonal trend of the atmospheric circulation types. WT4 and WT7 determine positive precipitation anomalies. On the contrary, WT5 determines negative precipitation anomalies in all three months.

```
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   294.08     501.36   0.587  0.57191
WT_6_DJF      -24.85      12.00  -2.072  0.06817 .
WT_4_JJA     -137.20      35.00  -3.920  0.00351 **
WT_7_JJA      114.80      35.42   3.241  0.01014 *
WT_1_SON       43.25      22.25   1.944  0.08377 .
WT_4_SON       79.94      19.14   4.177  0.00239 **
WT_5_SON       99.81      34.58   2.886  0.01799 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 205.5 on 9 degrees of freedom
Multiple R-squared:  0.8133,    Adjusted R-squared:  0.6889
F-statistic: 6.536 on 6 and 9 DF,  p-value: 0.006736
```

2

```
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)    90.49     237.45   0.381 0.710379
WT_4_JJA      119.13      21.45   5.555 0.000171 ***
WT_7_JJA     -105.78      24.78  -4.269 0.001323 **
WT_4_SON      -52.17      11.16  -4.675 0.000677 ***
WT_5_SON      -51.45      18.89  -2.724 0.019775 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 145.3 on 11 degrees of freedom
Multiple R-squared:  0.7999,    Adjusted R-squared:  0.7272
F-statistic: 10.99 on 4 and 11 DF,  p-value: 0.0007745
```

Results of the best linear regression model between the annual NPP and the seasonal trend of the atmospheric circulation types. WT4 and WT7 determine positive precipitation anomalies. On the contrary, WT5 determines negative precipitation anomalies in all three months.

3

Results of the best linear regression model between the annual RECO and the seasonal trend of the atmospheric circulation types. WT4 and WT7 determine positive precipitation anomalies. On the contrary, WT5 determines negative precipitation anomalies in all three months.

```
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) 1250.417     66.566  18.785 2.90e-10 ***
WT_5_DJF     -47.471      7.088  -6.697 2.20e-05 ***
WT_7_DJF     -13.780      6.559  -2.101   0.0574 .
WT_4_JJA     -51.181      8.032  -6.372 3.55e-05 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 52.21 on 12 degrees of freedom
Multiple R-squared:  0.8486,    Adjusted R-squared:  0.8107
F-statistic: 22.42 on 3 and 12 DF,  p-value: 3.298e-05
```

ACKNOWLEDGEMENTS:

EU - Next Generation EU Mission 4 “Education and Research” - Component 2: “From research to business” - Investment 3.1: “Fund for the realisation of an integrated system of research and innovation infrastructures” - Project IR0000032 – ITINERIS - Italian Integrated Environmental Research Infrastructures System - CUP B53C22002150006.

BIBLIOGRAPHY:

Oggioni, A., Silver, M., Tagliolato, P., & Karnieli, A. (2025). ReLTER: An R interface for environmental observation in long term ecological research. Ecological Informatics, 85, 102915. https://doi.org/10.1016/j.ecoinf.2024.102915