

Differences in activation energy of SOC from different treatments

Lorenzo Menichetti

2024-06-17

Scope of the work

This work was initiated on respiration time series from Spain, France and Bulgaria as a preliminary exploration of how to define and test a set of hypotheses connecting the data streams from field sampling campaigns in Holisoils.

It is ultimately a stratified variance decomposition. With parametric, process based models, Bayesian, but just a **variance decomposition** at the end.

Variance over time

- ▶ The variance in observed respiration caused by temperature can be described by a Arrhenius-like relationship (LLoyd & Taylor)
- ▶ The variance in observed respiration caused by soil moisture can be described by a bitonic relationship (Moyano)
- ▶ Some residual variance can be described by a seasonal term (cos)

Between sites and between treatments variance

We hypothesized that an important component of the observed between sites and between treatments variance can be described by independent parameterizations of the scaling functions, specifically:

- ▶ The effect of initial SOC, so substrate availability, can be described by a site-specific intercept (A parameter)
- ▶ The effect of **different activation energies**, therefore different SOC qualities caused by treatments, can be described by treatment-specific “slopes” (Activation energy, E_a parameter)

Model

$$Rh(t) = \left(A_k \cdot \xi_{temp(t,j)} \cdot \xi_{moist(t,j)} \right) + \xi_{sin(t,j)}$$

Where the index j corresponds to the treatments and the index k corresponds to the plots (minimal statistical unit), while (t) denotes the variation over time of each forcing variable (temperature, moisture, or days of the year)

Temperature scaling term

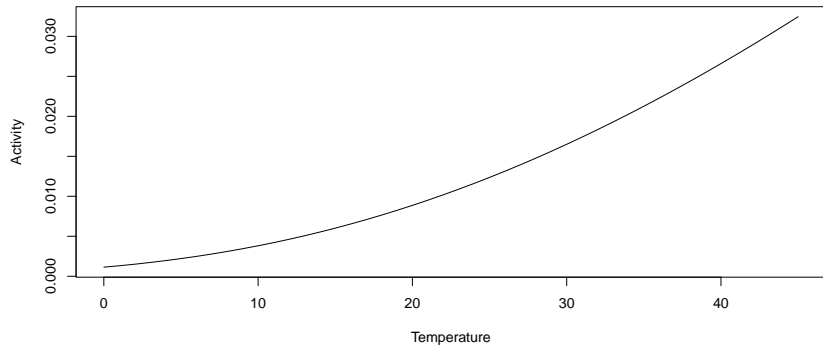
The temperature scaling function assumes an independent activation energy E_a varying by treatments j :

$$\xi_{temp}(t,j) = e^{\frac{-E_{a_j}}{(T(t)+273.15)-T_0}}$$

The two parameters E_a and T_0 are represented by Bayesian priors.

Temperature scaling term

Lloyd-Taylor



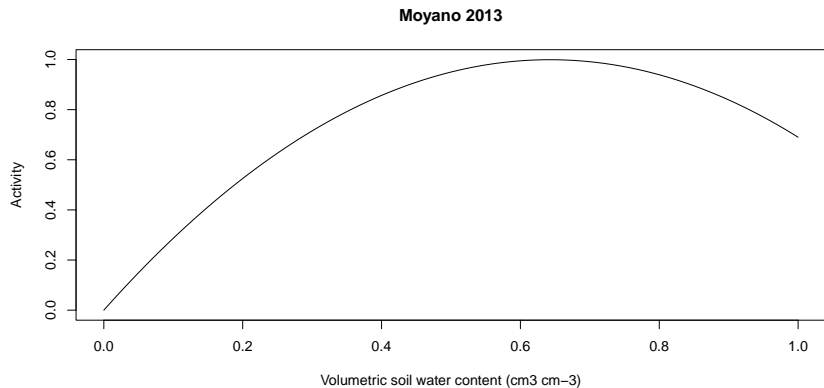
Moisture scaling term

The moisture scaling function had to be taken from a simplified version of Moyano et al. 2013 in the SoilR package. This was due to execution time of the function (with more covariates) in Moyano et al. 2012, one or two orders of magnitudes above. I had to drop SOC, BD and Clay as covariates. Didn't lose almost anything in predictive power.

$$\xi_{moist(t,k)} = a_j \cdot \theta_v(t) - b_j \cdot (\theta_v(t))^2$$

Variance is by treatment.

Moisture scaling term



Seasonality scaling term

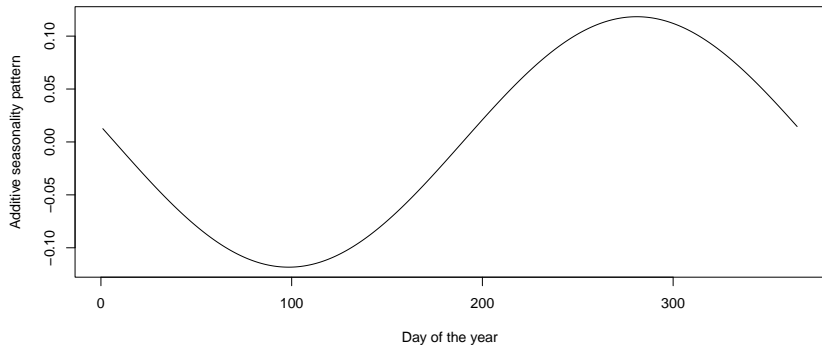
The seasonality scaling functions assume a sine function with period = 365 days, with both amplitude $ampl_j$ (the magnitude) and peak day $peak_j$ (when the function is at its only annual peak) varying as a function of treatments:

$$\xi_{sin(t,k)} = ampl_k \cdot \sin\left(\frac{2 \cdot \pi}{365}\right) \cdot day(t) + \left(\frac{2 \cdot \pi}{365}\right) \cdot (peak_k - 1) - \frac{\pi}{2}$$

Both the terms $ampl_j$ and $peak_j$ are described by Bayesian priors. Variance is by treatment.

Seasonality scaling term

Sine seasonality



Bottom line: stratification

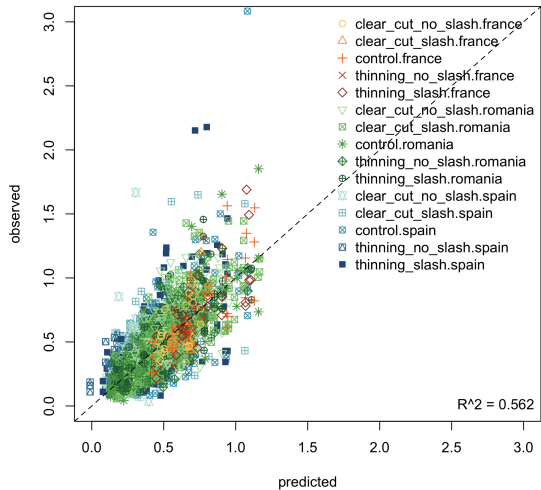
- ▶ The activation energy E_a of the substrate vary with **treatments**
- ▶ The “intercept”, explaining the variance due to SOC, varies with **plots**
- ▶ The two modifiers of the moisture scaling vary with **treatments**
- ▶ The parameters of the seasonal component vary with **treatments**

Implementation (coarsely)

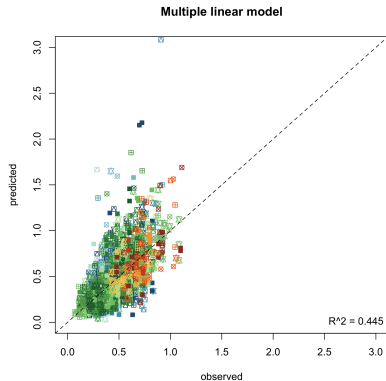
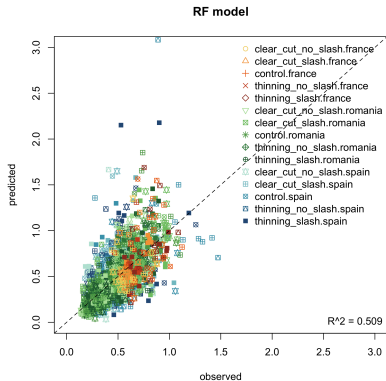
The model is written in Stan (<https://mc-stan.org>), which is an environment for statistical (Bayesian) modeling.

Once the model is written, Stan compiles it before running the sampler, so it is relatively good in terms of performances (it runs from R or Python or directly from Bash, in my case I use the R interface)

Model fit: not too bad

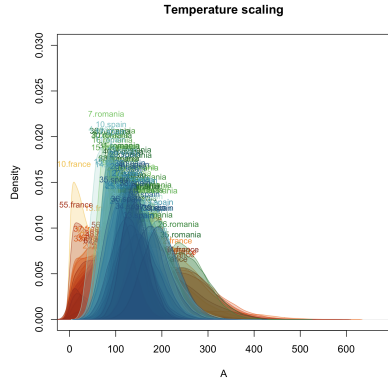
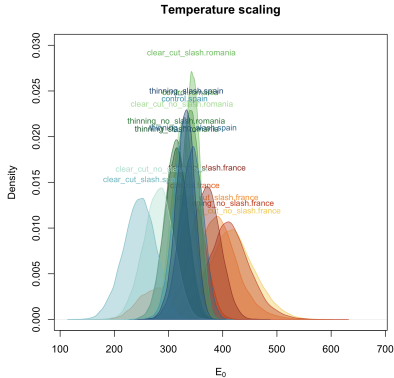


ML Model fit (benchmark)



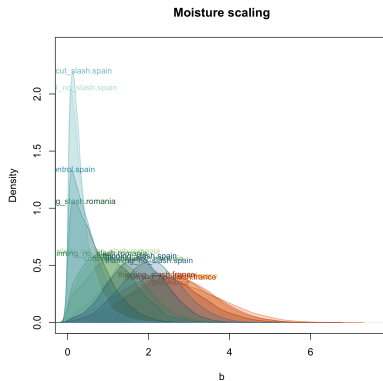
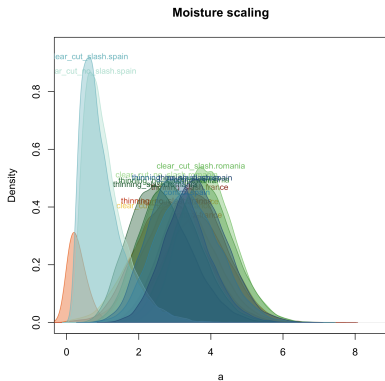
Fitting two machine learning models (random forests and multiple linear regression) earns an even worse fit, 0.51 and 0.45 respectively.

Posteriors and hypothesis testing: temperature



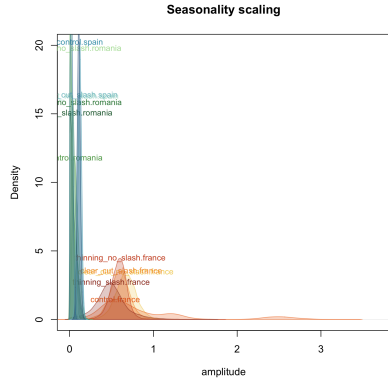
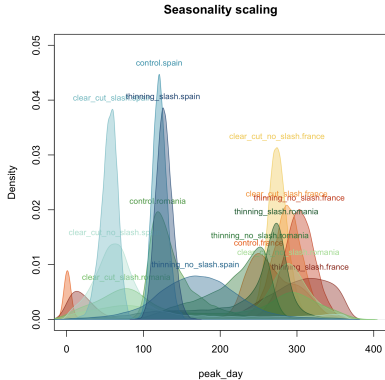
Unsurprisingly, there are some significant **differences between the activation energies (or temperature sensitivity) of the treatments**. This might be the main line of our story.

Posteriors and hypothesis testing: moisture



The parameters of the moisture function also differed by treatment.

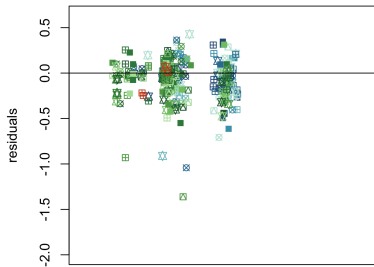
Posteriors and hypothesis testing: seasonality



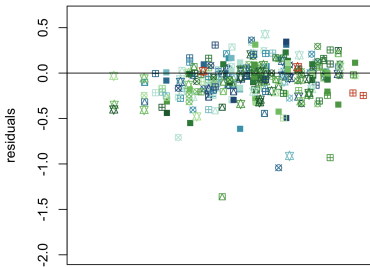
Also the seasonality component differs by treatment.

But adding the seasonality model gained information, from a R^2 of 0.55 to 0.56 so it is not particularly relevant.

Residuals



T1 soil



Soil moisture

Residuals seems distributed in a quite healthy way

Residual variance to be explained?

I tried a RF model to see if there is some more information to use:

residuals \sim *disturbance*, *density*_{trees}, *density*_{quercus},
MeanDBH, *porosity*, *BD*, *pH*, *phosphorous*,
Ntot, *Ctot*, *biom*_{fungi}, *biom*_{bact},
*biom*_{actinobac}, *biom*_{gram_pos}, *biom*_{gram_neg},
*biom*_{tot}, *fungi* : *bact*

There's not much. The fitness of the model is $R^2 = 0.06$.

Next steps

- ▶ Testing for hysteretic phenomena? We might see if including the amplitude of the short term variation in soil moisture can improve the model