From Point Data to Smooth Maps: Spatial Interpolation of Tree Water Deficit

William H. Aeberhard

Swiss Data Science Center (SDSC), ETH Zurich, Switzerland

SDSC

ETH zürich EPFL

TreeNet Annual Meeting
Birmensdorf, Switzerland

November 12, 2024

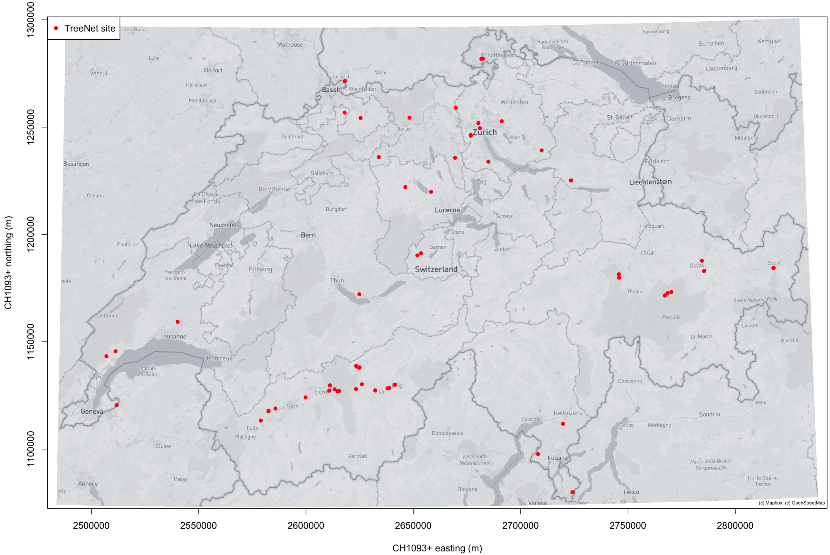
Outline

1. Introduction to Fore[st]cast Project
2. Toy example: spatial prediction of TWD at a given time point
3. Additive model predictions
4. Discussion

Fore[st]cast: Overview

- Main goal: predict tree growth and water deficit (drought stress) in space-time, focus on Switzerland
- TreeNet data: network of dendrometers measuring tree stem diameter, among others, with high precision. 709 series, with measurements every 10 minutes, over 73 sites within Switzerland.
- Tree water deficit (TWD): difference of stem radius (μm) relative to latest highest value.
- Task: spatial "interpolation" of TWD from point data (sites) to a map covering Switzerland.
- Implications for drought prediction and management at national level

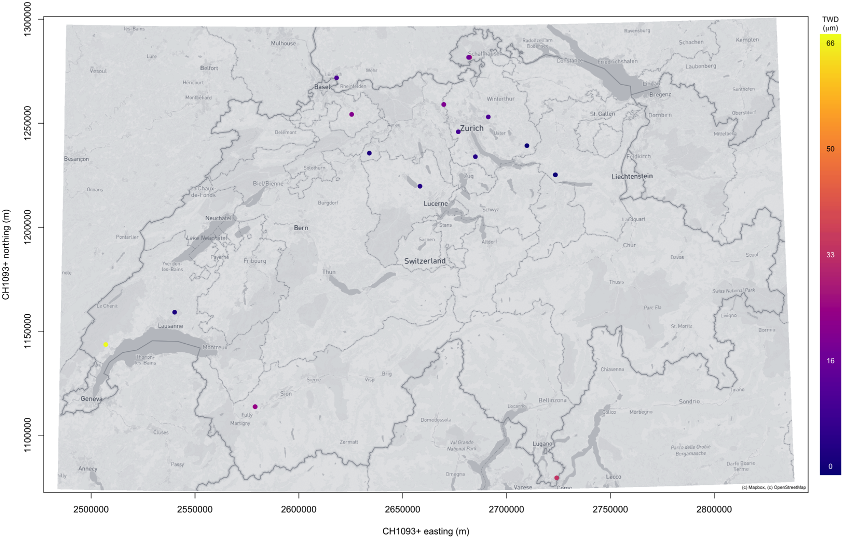
TreeNet sites in Switzerland



Toy example

- Single time point: 2022-07-05 at 12:00
- Single tree species: *Fagus sylvatica*
- TWD observations: 60 dendrometers over 20 sites. Hourly values computed as minimum over the six 10-minute measurements.
- TWD aggregated by site median $\Rightarrow n = 20$
- Covariates: total precipitation and air temperature from COSMO-1E, pointwise mean of 11 ensemble members, hourly values

TreeNet beech TWD values on 2022-07-05T12:00:00



Additive Model

- Additive model (AM): approximate the (unknown) mapping from precipitation amount (P) and air temperature (T), with arbitrary lags, to TWD by a sum of basis expansions:

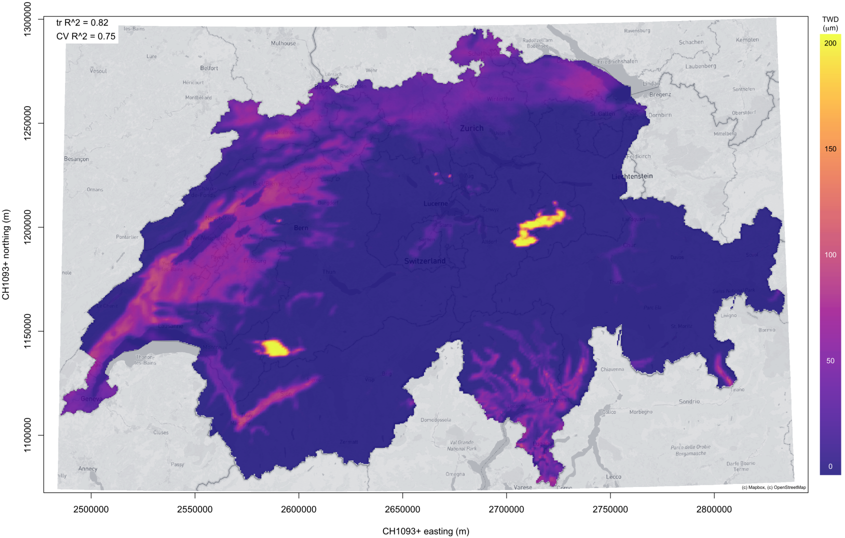
$$\begin{aligned} \mathrm{E}[g(\mathrm{TWD}_{s,t})] &= f(P_{s,1:t}, T_{s,1:t}) \\ &= \sum_{k=1}^{K_P} \sum_{j=1}^{p_k} \beta_{k,j} B_j(P_{s,t-\Delta_k}) + \sum_{k=1}^{K_T} \sum_{j=1}^{p_k} \beta_{k,j} B_j(T_{s,t-\Delta_k}) \end{aligned}$$

where $s = 1, 2, \ldots, n$ refers to a site, t is a given time point, g is a transformation (sigmoid bounded within $[0, 200]$ here), B is a user-specified basis function (thin plate regression splines here), and the lag $\Delta_k = 0, 1, 2, \ldots$ includes summing/averaging over time windows.

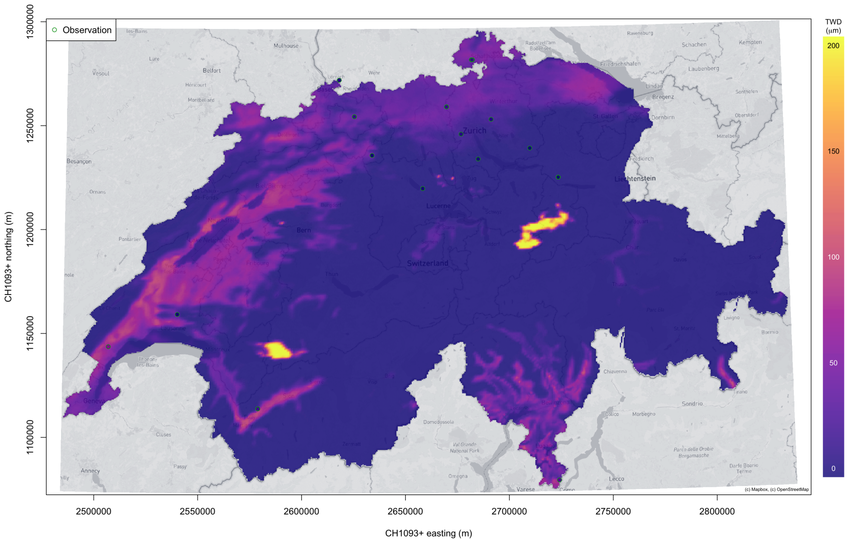
Additive Model (cont'd)

- With p_k set low, the $\beta_{k,j}$'s are estimated by minimizing the mean squared loss (Gaussian negative log-likelihood) penalized by the integrated squared second derivative of the smooth.
- Choosing K_P and K_T and which lags Δ_k to keep is done automatically by a forward stepwise selection procedure, aiming to minimize the leave-one-out cross-validation (CV) mean squared error.
- On the beech toy example, the procedure kept four effects: $P_{s,t-8}$, $P_{s,t-9}$, $T_{s,t-5}$ and $T_{s,t-7}$ (t is at an hourly resolution).

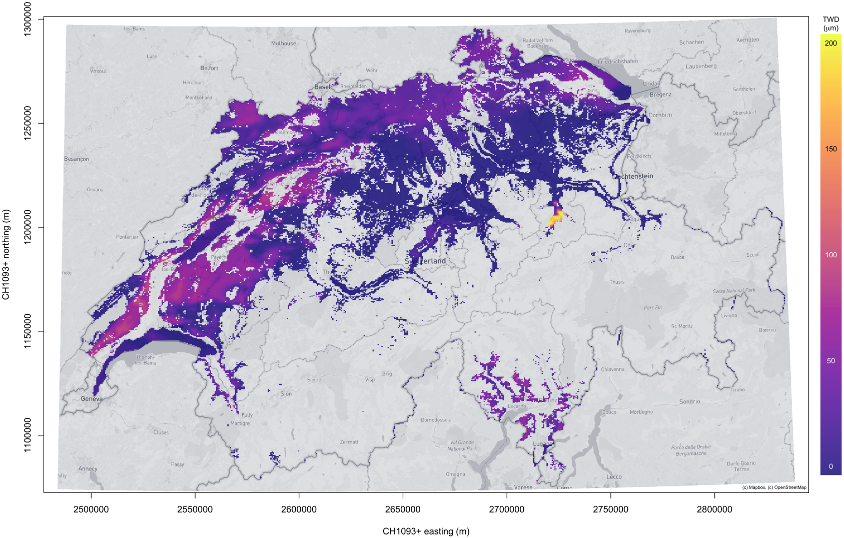
AM predictions



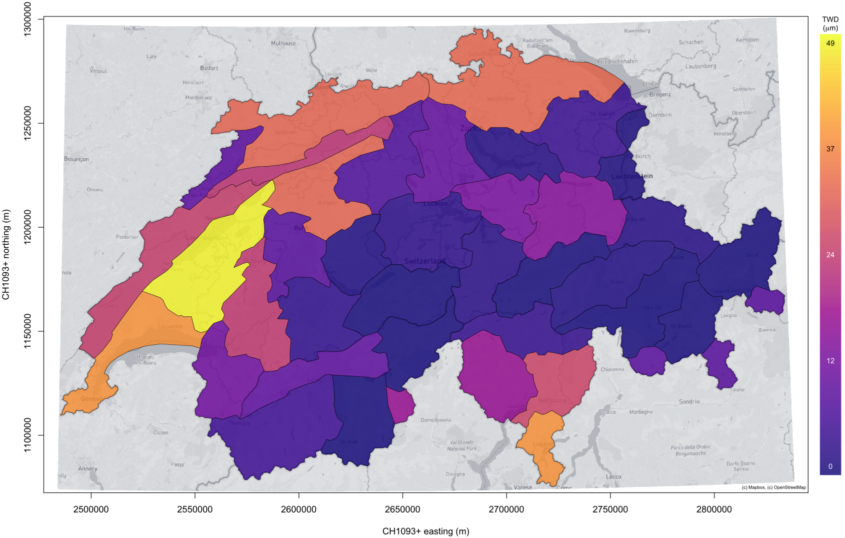
AM predictions with TWD observations overlaid



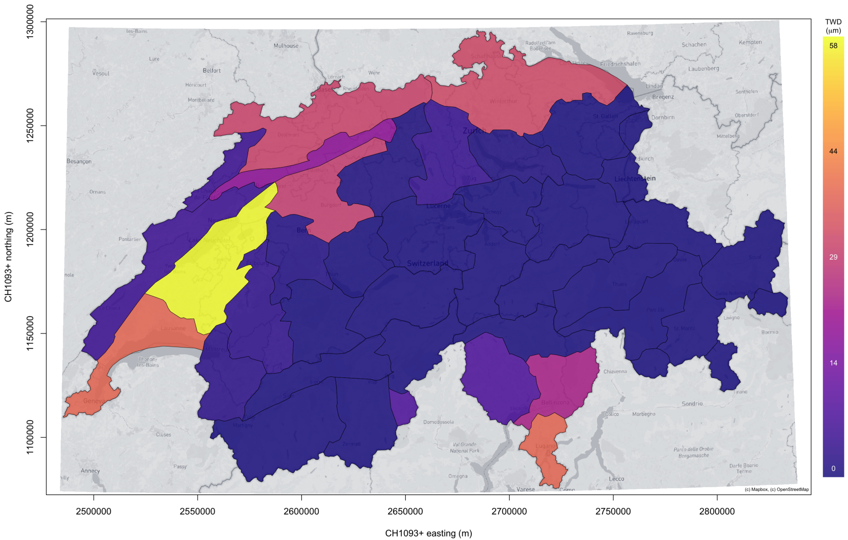
AM predictions with beech presence mask



AM predictions (all), mean over BAFU "large regions"



AM predictions (all), median over BAFU “large regions”



Discussion

- Simple, computationally inexpensive model to extrapolate TWD from point data to smooth maps. Variable selection and prediction is data-driven.
- Can easily include more environmental/meteorological covariates, as long as spatio-temporal resolution is sufficient
- Extrapolation challenging due to small sample size $\Rightarrow$ model in space-time, i.e., consider the mapping between (P, T) and TWD at many locations over time
- Increase coverage of masked maps: model jointly multiple species with normalized TWD.