Geophysical Research Abstracts
Vol. 18, EGU2016-16698-2, 2016
EGU General Assembly 2016

Revisiting the extended spring indices using gridded weather data and machine learning

Hamed Mehdipoor, Emma Izquierdo-Verdiguier, and Raul Zurita-Milla
Faculty of Geo-Information Science and Earth Observation (ITC), University of Twente, Enschede, the Netherlands

The extended spring indices or SI-x [1] have been successfully used to predict the timing of spring onset at continental scales. The SI-x models were created by combining lilac and honeysuckle volunteered phenological observations, temperature data (from weather stations) and latitudinal information. More precisely, these models use a linear regression to predict the day of year of first leaf and first bloom for these two indicator species.

In this contribution we revisit both the data and the method used to calibrate the SI-x models to check whether the addition of new input data or the use of non-linear regression methods could lead to improments in the model outputs. In particular, we use a recently published dataset [2] of volunteered observations on cloned and common lilac over longer period of time (1980-2014) and we replace the weather station data by 54 features derived from Daymet [3], which provides 1 by 1 km gridded estimates of daily weather parameters (maximum and minimum temperatures, precipitation, water vapor pressure, solar radiation, day length, snow water equivalent) for North America. These features consist of both daily weather values and their long- and short-term accumulations and elevation. we also replace the original linear regression by a non-linear method. Specifically, we use random forests to both identify the most important features and to predict the day of year of the first leaf of cloned and common lilacs.

Preliminary results confirm the importance of the SI-x features (maximum and minimum temperatures and day length). However, our results show that snow water equivalent and water vapor pressure are also necessary to properly model leaf onset. Regarding the predictions, our results indicate that Random Forests yield comparable results to those produced by the SI-x models (in terms of root mean square error –RMSE). For cloned and common lilac, the models predict the day of year of leafing with 16 and 15 days of accuracy respectively. Further research should focus on extensively comparing the features used by both modelling approaches and on analyzing spring onset patterns over continental United States.

References

1. Schwartz, M.D., T.R. Ault, and J.L. Betancourt, Spring onset variations and trends in the continental United States: past and regional assessment using temperature-based indices. International Journal of Climatology, 2013. 33(13): p. 2917-2922.
2. Rosemartin, A.H., et al., Lilac and honeysuckle phenology data 1956–2014. Scientific Data, 2015. 2: p. 150038.
3. Thornton, P.E., et al. Daymet: Daily Surface Weather Data on a 1-km Grid for North America, Version 2. 2014.