Satellite data to improve the accuracy of statistical models for wind resource assessment

Fabio Veronesi[1*], Stefano Grassi[1], Martin Raubal[1]

[1]Institute of Cartography and Geoinformation - ETH Zürich; Stefano-Franscini-Platz 5, 8093 Zurich

[*] Corresponding author: fveronesi@ethz.ch

Introduction

Wind resource assessment includes numerous algorithms to estimate wind speed in locations where direct measurements are not available. These fall into two broad categories: numerical wind flow and statistical models. Numerical wind flow models can be divided by level of complexity: they range from Jackson-Hunt models (e.g. WaSP, Jackson and Hunt, 1975), which are generally considered the industry standard, to more complex models such as Reynolds-averaged Navier-Stokes (RANS; Landberg et al., 2003) and Numerical Weather Prediction (NWP; Bailey & Beaucage, 2012) models. These methods work by solving physical equations, which range from mass and momentum conservation (Jackson-Hunt models) to more complex models that include the full range of computational fluid-dynamics (RANS), and other that also take into account energy exchanges between atmosphere and pedosphere (NWP, Bailey & Beaucage, 2012). These methods are proven to be very accurate in modelling the wind resource, but they are also extremely time and computationally demanding. Examples from literature shows that while NWP methods can achieve RMSEs (root mean square errors) around 0.5 m/s, they also need to run for weeks to map relatively small areas (Douak et al., 2013).

An alternative are statistical models, which run much faster that numerical wind flow models, since they use an approach that is fundamentally different from the latter. They estimate wind speed by computing its correlation with various environmental predictors: such as elevation, slope, air temperature and pressure. The algorithm takes the wind measurements and tries to find patterns that correlate wind speed with the predictors in ways that it can then use to estimate the wind resource in locations where no observations are available, this process is referred to as training. Traditionally, statistical methods were not considered sufficiently accurate to compete with the methods described above. Research studies published in literature show that statistical methods rarely achieve RMSEs below 1 m/s (Luo et al.,2008; Forest et al., 2011;Douak et al., 2013). However, statistical methods can improve by employing more accurate environmental predictors. The advancements in remote sensing in the last decades have provided increasingly higher resolution satellite data, which can be used to improve the accuracy of statistical wind resource assessment algorithms. In this research we tested an approach for estimating wind speed using statistical learning (Veronesi et al., 2015) for creating a 1 Km resolution wind speed map of Switzerland at 10 m above the ground level.

Approach

We collected 10 min averaged wind speed data over a time window of 5 years from the network of weather stations (n=161) of Meteoswiss (Federal Office of Meteorology and Climatology, 2015). For each weather station we fitted Weibull distributions to the observations and calculated their corresponding shape and scale parameters. This statistical learning method works by spatially estimating the two parameters of the Weibull distribution in order to provide users with the full wind distribution in each estimated location.

In order to increase the accuracy of the map we collected satellite raster data from various sources. From NASA Surface meteorology and Solar Energy (NASA), we downloaded the following average data over the last 22 years: solar radiation, air temperature, relative humidity, atmospheric pressure, wind speed at 50 m above ground. From NOAA (National Oceanic and Atmospheric Administration) we downloaded monthly averages for soil temperature, evaporation, precipitation, runoff and moisture, plus weekly NDVI (Normalized Difference Vegetation Index). Finally from METEOSAT (CM SAF) we downloaded the following datasets: soil evaporation, vertically integrated water vapour, latent heat flux, vertically integrated liquid water, near surface specific humidity, precipitation, surface albedo, surface downward longwave radiation, surface incoming shortwave radiation, surface net longwave and shortwave radiation, surface outgoing longwave radiation, surface radiation budget and near surface wind speed. Each of these data were collected for time intervals of at least 5 years, or the maximum amount of data available. In total around 6000 predictors have been used.

Main body of abstract

Switzerland is characterized by a complex topography due to presence of the Alps in the central and southern part of the country, and the Jura range in the north-western part, whereas the rest is characterized by hilly terrain. In addition, large areas are covered by forests or towns of different sizes, whereas the areas occupied by arable and flat terrains are a minor percentage of the total. This case study is a challenging test compared to previous analyses (Veronesi et al., 2015) as we want to test the ability of the developed model to assess the wind resource in complex environments.

We employed two statistical algorithms to map the wind resource of Switzerland. The first is a shrinkage technique, LASSO, which aims at reducing the amount of predictors keeping only the most correlated with wind speed. From this first step we extracted 140 predictors, which is a more manageable number for the analysis. At this point we used Random Forest, which is a statistical learning algorithm that uses ensembles of regression trees, to estimate the two Weibull parameters at 1 km of resolution. To test the accuracy of the method we used a 5-folds cross validation. Basically we divided the dataset into five equal parts (or folds), each containing around 28 weather stations randomly selected. Then the process works by using four of these parts to train the algorithm and estimate the Weibull parameters in the stations included in the fifth fold (test set), which were excluded from the training process. This same procedure is repeated until each fold is used for testing. Since this process is based on random selection, we repeated it 100 times in order to have more robust results. The validation indicates that this method is able to achieve a good level of accuracy, with a RMSE of 0.66 m/s and a MAE (Mean Absolute Error) of 0.46.

Statistical methods have the advantage over numerical wind flow models of being able to assess their own accuracy. During the training process, where the algorithm compares the environmental predictors with the wind speed observations, the algorithm is able to determine the amount of variation that occurs in the wind speed in relation to the variation in the predictors. During the estimation phase, where the values of the predictors are similar to the training data the algorithm will produce estimates with low uncertainty. However, in areas where the predictors have very different values from the training set, or present contrasting information, the algorithm will find it more difficult to estimate wind speed accurately. This is something we can measure through the map uncertainty, which basically allows us to determine in which areas the map is less reliable. Numerical wind flow models are not able to determine the local uncertainty in the map; however, it does not mean they are perfect. The starting data for both approaches are sparse weather stations, and thus all the algorithms will have more problems estimating the wind resource in areas farther away from the weather observations.

Conclusions

This study clearly demonstrates that modern statistical algorithms, coupled with remotely sensed environmental data, can produce level of accuracy comparable with numerical wind flow models. These data allow the statistical algorithm to be trained considering not only the local topography, which governs the local speed changes caused by mass and momentum conservation, but also the spatial pattern of the energy fluxes between soil and atmosphere. These fluxes are involved in major changes in wind speed and therefore are crucial for reaching a good estimation accuracy. Only by taking these additional satellite data into account statistical algorithms have a chance of achieving accuracy levels comparable with numerical wind flow models. Moreover, statistical methods are markedly faster than the other algorithms, meaning they can be potentially employed to create global wind meso-scale maps relatively quickly. Finally, statistical method can provide a measure of their own mapping precision, meaning that we have information regarding the spatial pattern of the error, and not just the average validation error provided by the cross-validation. This information is crucial for wind farm planning, since it allows to pin point with absolute precision the locations on the map in which the algorithm is less accurate. Site-specific uncertainty is not provided by numerical wind flow models, even though they are affected by it nonetheless. This information can be used to identify sites for which more data are needed from sites in which the wind map can be trusted.

Learning objectives

In this research we argue that recent advancements in remote sensing have provided ways of substantially increasing the accuracy of statistical models for wind resource assessments. These algorithms are widely available in various software applications, they can run in a fraction of the time needed by any numerical wind flow model and are able to obtain reasonably good results. For this reason we think it is now time to research new ways for coupling numerical wind flow models, such as the popular WaSP, with statistical models in order to increase the accuracy and the spatial resolution of wind resource maps, including micro-scale modelling, while maintaining the computational time and resources into acceptable limits.

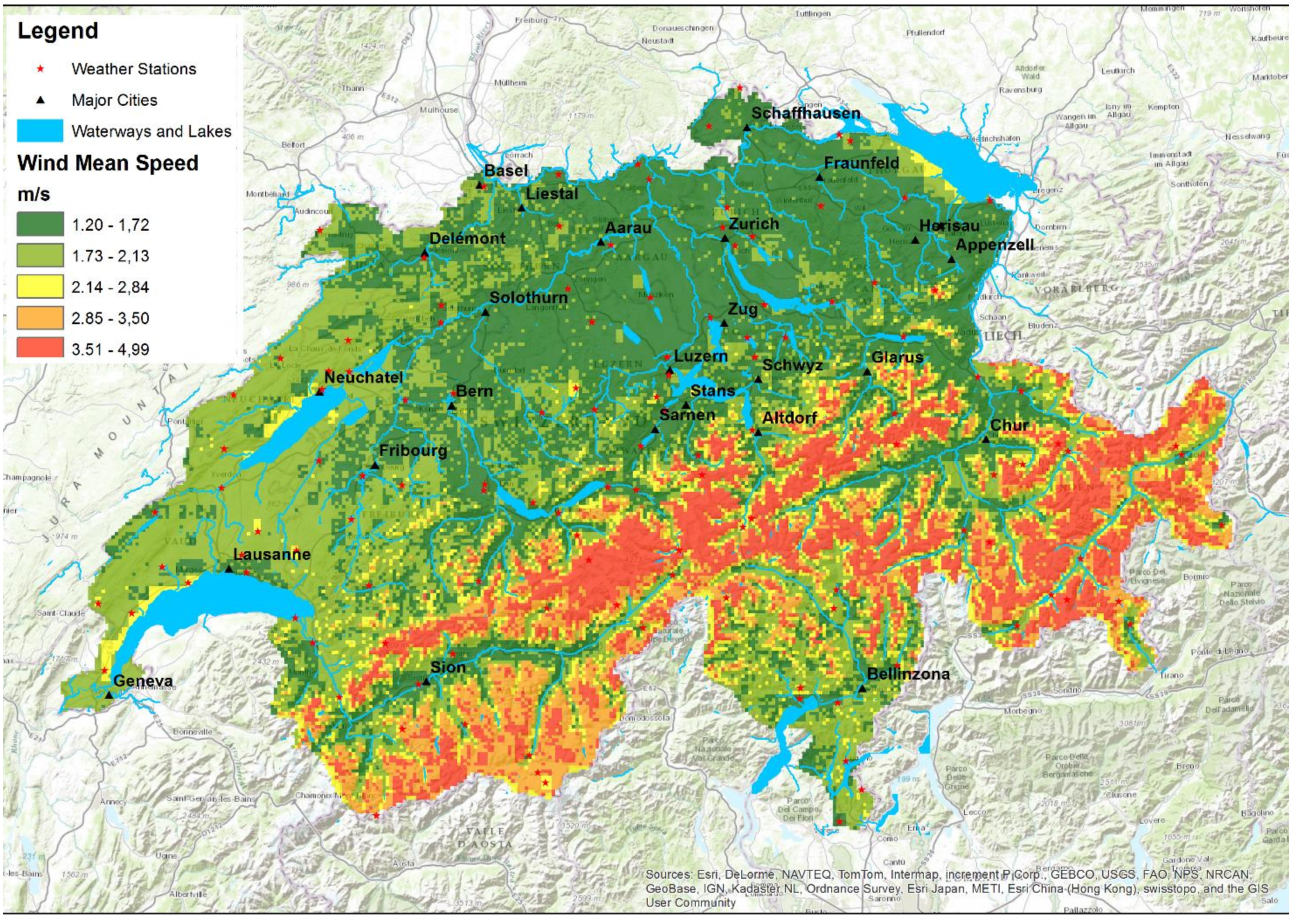

Figure 1: Wind mean speed map of Switzerland. This map shows also the location of the weather stations used to train the algorithm (red stars).

References

Bailey, B. H., & Beaucage, P. (2012). *Wind resource assessment: a practical guide to developing a wind project*. M. Brower (Ed.). John Wiley & Sons.

Beaucage P, Brower MC, Tensen J (2014) Evaluation of four numerical wind flow models for wind resource mapping. *Wind Energy* **17**(2):197-208

CM SAF, The Satellite Application Facility on Climate Monitoring. http://www.cmsaf.eu/EN/Home/home_node.html - Last accessed: 27.04.2015

Douak F., Melgani F., and Benoudjit N. Kernel ridge regression with active learning for wind speed prediction. *Applied Energy* **103** (2013): 328-340.

Foresti L, Tuia D, Kanevski M, Pozdnoukhov A (2011) Learning wind fields with multiple kernels. *Stochastic Environmental Research and Risk Assessment* **25**(1):51-66

Federal Office of Meteorology and Climatology (2015). http://www.meteoswiss.admin.ch/ - Last accessed: 27.04.2015

Jackson PS, Hunt JCR (1975) Turbulent wind flow over a low hill. Quarterly Journal of the Royal Meteorological Society 101(430):929-955

Landberg, L., Myllerup, L., Rathmann, O., Petersen, E. L., Jørgensen, B. H., Badger, J., & Mortensen, N. G. (2003). Wind resource estimation—an overview. Wind Energy, 6(3), 261-271.

Luo W, Taylor MC, Parker SR (2008) A comparison of spatial interpolation methods to estimate continuous wind speed surfaces using irregularly distributed data from England and Wales. *International Journal of Climatology* **28**:947-959

National Oceanic and Atmospheric Administration (NOAA), National Centers for Environmental Information. http://www.ncdc.noaa.gov/data-access - Last accessed: 27.04.2015

NASA, (2014). Surface meteorology and Solar Energy (SSE) Release 6.0. Available at: https://eosweb.larc.nasa.gov/project/sse/SSE6Methodology.pdf – Last accessed: 27.04.2015

Veronesi F, Grassi S, Raubal M & Hurni L (2015). Statistical learning approach for wind speed distribution mapping: the UK as a case study. In book: *Geographic information science as an enabler of smarter cities and communities*, Publisher: AGILE, Editors: F. Bacao, M. Y. Santos, and M. Painho