Bidimensional analysis of Wind Energy forecasts including a new Temporal Distortion Index

Laura Frias Paredes (lfiias@cener.com) / Martin Gaston Romeo (National Renewable Energy Centre (CENER) / Fermín Mallor (Public University of Navarre, Spain) / Teresa León (University of Valencia, Spain)

Abstract

The main goal of this work is to provide a deeper insight in the analysis of timing errors which leads to the proposal of a new methodology for its control and measure. A new methodology based on Dynamic Time Warping, is proposed to be considered in the estimation of accuracy as attribute of forecast quality. A new dissimilarity measure, the Temporal Distortion Index, among time series is introduced to complement the traditional verification measures found in the literature. Furthermore we provide a bi-criteria perspective to the problem of comparing different forecasts. The methodology is illustrated with several examples including a real case.

Introduction

Most of the verification methods compare pairs of time series, observations and predictions, and verification measures look at "vertical" distances between the two series. However, when the possibility of one series shows a misalignment in time with respect to the other is considered, a "horizontal" distance or time-distance becomes necessary. We could think of a model that correctlypredicts the events although not "on time" but with a certain time lag. It also may happen that the model is capable of identifying correctly the occurrence of an event but not its duration. Figure 1 intends to illustrate these situations showing the prediction of wind energy generated by LocalPred model currently used by the National Renewable Energy Centre (CENER) [1, 2], orange line, together with the corresponding measured power (black line).

The first and second squares show time lags in the prediction. The model collects the real behavior but for a time ahead of the real series. On the other hand, the two last squares contain events that have been well identified but with less duration that in the real data. Therefore, the model predicts most of the events but not always on time and the estimation of their duration is in another accuracy. In these cases it would be necessary a method allowing elastic shifting of the time axis, to accommodate sequences that are similar but out of phase. Dynamic Time Warping (DTW) was introduced in [3] and [4] as a dynamic programming (DP) based time-normalization algorithm for spoken word recognition. The purpose of DTW was to eliminate temporal differences between two speech patterns by warping the time axis of one of them so that the maximum coincidence with the other was attained. Our methodology is based on the DTW principles which obtain the optimal alignment of two time series by applying dynamic optimization to a shortest path problem. In this problem nodes represent possible temporal pairings between the two series, while possible transitions and distances are defined by a recursive function. This function manages which temporal leaps are allowed and its associated cost to reach a new coupling. In this work different applications suitable adaptations of the DTW basic algorithm have been proposed giving rise to a variety of dynamic-time warping-based techniques for the alignment of forecast and observed renewable energy data series. Therefore, in this work we also present a bicriteria analysis for the estimation of the accuracy as an attribute of forecast quality. A new dissimilarity measure among time series is presented, the Temporal Distortion Index (TDI), which complements the traditional verification measures found in the literature. This measure minimizes the effects of shifting and distortion in time by allowing an adaptive transformation of the time series. The simultaneous consideration of both error measures, TDI and a traditional verification measure, leads to a bi-criteria perspective when comparing different forecasts.

Control and measurement of the timing error component

We work with two time series, the test (or query) series $T = (T_1, T_2, ..., T_N)$ and the reference series $R = (R_1, R_2, ..., R_M)$ with $S = (S_1, S_2, ..., S_N)$, a new series called aligned series and

The test series is transformed, using the optimal path, in a new series called aligned series and denoted by $S' = (S_1, S_2, ..., S_N)$ with a smaller vertical difference to the reference series than the test series. The resulting aligned series is the one, among all those series associated to an optimal path, providing a minor misalignment of the test series. Equivalentiy, the aligned series is the one whose associated optimal path is the closest to the identity path. Hence, a global measure of the temporal distortion carried out in the test series, in order to obtain the aligned series, is provided by the area between the resulting optimal path and the identity path, which is denoted by TDI (Temporal Distortion Index). So, this parameter will serve to describe the temporal misalignment of the error:

$$ TDI = \sum_{i=1}^{N} |y_i - x_i| $$

where $x_i$ is the $i$-th value of the reference series and $y_i$ is the $i$-th value of the test series.

Methodology to obtain the bi-dimensional measure for the forecast

In this section we test the methodology by using real data series: a real wind energy series is compared with a wind energy forecast. Furthermore, we provide a deeper insight in the bi-dimensional measure of the error by considering different recursive formulas.

The time series consist of, on one hand, three days of wind energy production in a wind farm located in the north of Spain and, on the other hand, the respective energy prediction made 72h in advance. In this day a Pareto Frontier [5] is obtained. This Pareto Frontier better characterizes the error in the forecast than only one value or pair of values. Nevertheless, in order to simplify the assessment of forecast errors and for ease the comparison between different methods one or few representative points should be selected. More precisely, in the plot of the Pareto Frontier the first error assessment will reduce much static error but at some point the marginal gain will drop, giving an angle in the graph. A representative bidimensional error assessment should be chosen at this zone, and hence the "elbow criterion".

Conclusions and future lines

- This work proposes a comprehensive methodology to analyze and evaluate these temporary mismatches. A new index of temporal error, named Temporal Distortion Index, TDI, has been defined. It is used as first component of a two-dimensional error vector, BE, whose second component is a static error measurement, as the MAE. The information provided by this vector allows for a joint analysis of errors in time and scale. This process is based on Dynamic Time Warping techniques and uses recursion formulas which control the types of changes that can be made in the time axis. In addition, the availability of different formulas of recurrence leads the use of Pareto Frontiers to assess the behavior of the prediction models.
- For future developments the methodology proposed here can be applied to compare different prediction models in both wind energy and solar radiation. The comparison can be done in terms of temporal distortion and in terms of scale distortion. When extended data are available, it could be possible to analyze what model outperforms others and in what circumstances.
- From a technical point of view, it is necessary a deeper analysis of the recurrence functions to obtain parametrized families of functions whose parameters allow the control of the temporal distortion magnitude. By varying such parameters it would be possible to create Pareto Frontiers more dense which would be used as characteristic error curves of the methods.

References