Enabling Time Series Sensitivity Analysis with Iterative Variance Orthogonal Decomposition

Mouad Yachouti
CMAP, CNRS, École polytechnique, Institut Polytechnique de Paris, 91120 Palaiseau, France

Guillaume Perrin
COSYS, Université Gustave Eiffel, 14-20 Boulevard Newton, 77447 Marne-la-Vallée, France

Josselin Garnier
CMAP, CNRS, École polytechnique, Institut Polytechnique de Paris, 91120 Palaiseau, France

In many critical areas, large amounts of historical data are collected to monitor complex dynamic systems, such as in nuclear industry, finance, manufacturing, etc. To ensure that these systems function properly, a response variable (output) is observed along with environment variables (inputs) that may have an influence on its outcome. Explaining the latter with respect to the input variables has then become a crucial need. Answering this question is non-trivial due to the functional nature of the variables and the response [4], the temporal correlation of the inputs (for example, the presence of a daily periodicity in the data), or memory effects (ie. the fact that the impact of an event at a time t is observed at a time $t + \tau$ for a non-negligible response time τ).

To this end, Sensitivity Analysis (SA) provides powerful tools to engineers and practitioners. In particular, the framework of variance-based SA allows to link the output's variance to the individual (or combined) inputs variances and interpret them as contributions to the total variance. Generalizations to functional and temporal outputs have been the subjects of many works (see, for example, [4, 1]).

In this work, we propose a decomposition procedure for time series to enable a quantitative variance-based SA that clarifies the role of memory effects. The methodology is a two-stage approach. First, a linear model taking into account only the instantaneous input variables and their polynomial transforms is fitted to approach the output. Then, iteratively, for each input variable, a distributed-lag model [5] is fitted to take into account its memory effects while ensuring, by construction, the orthogonality to the already fitted models. This allows to decompose the total variance of the output as the sum of the variances of the resulting components.

The proposed framework is illustrated on multiple toy examples, and, then, applied to a real-world application case of wind power production [3, 2].

This research is supported by the Stress Test, Financial Steering and Risk management Chair hosted by the Center of Applied Mathematics (CMAP) between Ecole Polytechnique, BNP Paribas, and Fondation de l'Ecole Polytechnique. Wind power data is provided by the Energy4Climate Interdisciplinary Center (E4C) of Institut Polytechnique de Paris, for which the authors would like to express their thanks.

References:

[1] Alen Alexanderian, Pierre A. Gremaud, and Ralph C. Smith. Variance-based sensitivity analysis for time-dependent processes. *Reliability Engineering System Safety*, 196:106722, 2020.

[2] Dimitri Bouche, Rémi Flamary, Florence d'Alché Buc, Riwal Plougonven, Marianne Clausel, Jordi Badosa, and Philippe Drobinski. Wind power predictions from nowcasts to 4-hour forecasts: A learning approach with variable selection. *Renewable Energy*, 211:938–947, 2023.

[3] Aurore Dupré, Philippe Drobinski, Jordi Badosa, Christian Briard, and Peter Tankov. The economic value of wind energy nowcasting. *Energies*, 13(20), 2020.

[4] Fabrice Gamboa, Alexandre Janon, Thierry Klein, and Agnès Lagnoux. Sensitivity analysis for multidimensional and functional outputs. *Electronic Journal of Statistics*, 8(1):575 – 603, 2014.

[5] A. Gasparrini, B. Armstrong, and M. G. Kenward. Distributed lag non-linear models. *Statistics in Medicine*, 29(21):2224–2234, 2010.

[Mouad Yachouti; CMAP, École polytechnique, 91120 Palaiseau, France]
[`mouad.yachouti@polytechnique.edu` – `https://github.com/ymouad/`]