Towards History-aware Sensitivity Analysis For Time Series

Mouad Yachouti[1] (✉), Guillaume Perrin[2], and Josselin Garnier[1]

[1] CMAP, CNRS, École polytechnique, Institut Polytechnique de Paris, 91120 Palaiseau, France
{mouad.yachouti,josselin.garnier}@polytechnique.edu
[2] COSYS, Université Gustave Eiffel, 14-20 Boulevard Newton, 77447 Marne-la-Vallée, France
guillaume.perrin@univ-eiffel.fr

Abstract. Explaining the outcome of dynamic systems is non-trivial due to the temporal nature and correlation of the input variables. In this work, we propose a framework of history-aware sensitivity analysis for stationary time-series to quantify different memory effects and clarify their roles. For this purpose, we decompose the output time series into non-correlated components, namely the instantaneous component and the memory components. The latter are sorted in decreasing order of variance to reflect the importance of the variables. We highlight the compensation phenomena between the resulting components and illustrate them in the case of independent variables in a linear setting. To enable history-aware explanations, variance-based sensitivity indices are derived from the obtained decomposition. We demonstrate the effectiveness of our methodology in providing insights to explain output time-series in both synthetic and real-world cases.

Keywords: Sensitivity analysis · Time Series · Explainability.

1 Introduction

Large amounts of historical data are collected to monitor complex dynamic systems, especially in industry and other mission-critical applications, such as in nuclear industry, finance, manufacturing, etc. Numerical models have been extensively used to mimic the behaviour of these systems, notably thanks to the success of black-box models such as Gaussian processes [20,21] and deep learning models [14]. Explaining a system's outcome or their surrogates based on the input variables is a crucial need; detecting the most influential variables improves the understanding of the system and its dynamic and, moreover, allows effective maintenance policies and facilitate the compliance with current legislation such as General Data Protection Regulation (GDPR) or regulations to ensure nuclear safety. To this end, rich frameworks, including uncertainty quantification (UQ), explainable AI (XAI), and Sensitivity analysis (SA) among others, have been developed in the last two decades to tackle this problematic and build confidence.

In practice, performing such techniques to assess variables importance for dynamic models is a non-trivial task and requires a history-aware methodology [1] due to the functional nature of the variables and the response, the temporal correlation of the inputs (for example, the presence of a daily periodicity in the data), or lag or delay 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) which is frequent in economic, medical, and environmental data, etc. In this paper, we assume that we are provided with a sample of input multivariate time series and the corresponding sample of outputs. We aim at explaining the latter based on the input variables by quantifying different memory effects of their underlying dynamics, notably, the instantaneous and the linear memory effects.

To extract meaningful knowledge, SA studies how inputs uncertainties are attributed to the uncertainty observed on the output of a model [22]. In particular, the framework of variance-based SA [3] and the so-called Sobol indices [23, 24] allow to decompose the output's variance into different parts and attribute each one to individual (or combined) input variances. For multivariate or time dependent models, performing SA on separate scalar outputs, using univariate Sobol indices for instance, can provide some information on the evolution of the sensitivity over time. This approach, however, introduces redundancy and may miss some dynamic features. To overcome this, a framework is introduced in [13] that consists in extracting orthogonal features and then performing SA on the most informative components individually; for each component, a global sensitivity index can be derived. In a more general setting, in [6], new generalized Sobol indices are constructed for multivariate or functional outputs with an updated pick and freeze estimators. Building on this, [1] focused on the study of global SA for time dependent models of the form $Y_t = f(X, t)$; to improve computational efficiency and avoid redundancy, a combination of Polynomial chaos [8] and Karhunen-Loeve expansions (see for eg. Karhunen (1947) and Loeve (1948)) is used to approximate the output and deduce generalized Sobol indices as in [6] and their equivalence to the work of [13] is proven. Finally, in the context of dependent variables, [25] alleviates the assumption of independence of variables by using copulas [18] and proposes a method for both time dependent models $Y_t = f(X, t)$ and models defined by means of stochastic process $Y_t = f(X_t)$.

Our work is also related to HOFD (hierarchically orthogonal functional decomposition) introduced in [4] to generalize the Hoeffding-Sobol decomposition in case of dependent variables and mGS-PCE (modified Gram-Schmidt based polynomial chaos expansion) proposed in [17]. Both methods are iterative and leverage the ability of the Gram-Schmidt process to handle dependency and correlation between variables. In our case, this process is used to handle, as a first step, the temporal correlation between the instantaneous variables and the memory components, and, as a second step, to handle correlation among the memory components.

Given the limitations of the aforementioned methods on outputs of the form $Y_t = f(X_s, s \leq t)$, we propose a methodology to perform a quantitative variance-based SA in attempt to make a step towards a generalization of the above mentioned framework to time-series. Our approach consists in decomposing the output time-series into a sum of an instantaneous component, a lag component, and a residual component; the purpose of the first two components is to quantify and clarify the roles of the memory effects of the inputs on the output. The components are constructed in a specific order to ensure an orthogonality property and avoid redundancy between variables in case of dependency. To do so, we apply a Gram-Schmidt process to make the variables orthogonal to the already captured effects. Although this construction holds under weak assumptions (for example, it does not require independence), in this work, we restrict ourselves to second-order stationary processes for two main reasons: first, the ability to use sample covariance and variance estimators, and, second, the constant variance under this assumption reflects in a global fashion the behaviour of the resulting components. We evaluate the ability of our method to approximate the true quantities on a toy example and, then, show its performance on a synthetic controlled case and real-world cases where the stationarity and the independence assumptions do not necessarily apply. The source code of the experiments will be made available on GitHub later.

Our contributions are summarized as follows:

- we propose a two-stage approach to quantify memory effects;
- we distinguish the instantaneous effects from the lag effects and attribute to each one a component to reflect the captured dynamic;
- based on the modelled components, we derive simple variance-based indices to measure history-aware variables importance and provide useful post-hoc global explanations.

2 Proposed approach

We consider the scenario where the response variable Y is a function indexed by the time t with values in $\mathcal{Y} \subset \mathbb{R}$. This function is assumed to be the output of a function f depending on the history of $D \geq 1$ observed variables gathered in $X = (X^1, \ldots, X^D)$ such that:

$$\forall t \in \mathbb{N},\, Y_t = f\left((X_s)_{s \leq t}, \varepsilon\right), \tag{1}$$

where ε is an error component (to model observation errors or the fact that Y may also depend on additional but not observed variables for instance). In this work, we make the assumption of second-order stationarity. In particular, under this assumption, the mean and the variance of our processes remain constant, and the covariance function does not depend on the time origin. Without loss of generality, the means are assumed to be zero. In practice, the underlying model f is unknown and we are only provided with a sample of $N > 1$ input-output pairs, noted $\mathcal{D} = \{(x_0, y_0), \ldots, (x_{N-1}, y_{N-1})\}$ so that for each $0 \leq t \leq N-1$, $(x_t, y_t) \in \mathbb{R}^D \times \mathcal{Y}$ is the available observation of (X_t, Y_t). This set is also assumed regularly-sampled.

Given the output Y_t, we seek a decomposition of the form:

$$Y_t = Y_t^{\text{inst}} \oplus Y_t^{\text{lag}} \oplus Y_t^{\text{res}}, \tag{2}$$

where

- Y_t^{inst} is the instantaneous component, ie. a function of the instantaneous input $X_t = (X_t^1, \dots, X_t^D)$,
- Y_t^{lag} is the linear memory component, ie. a linear function of lagged variables $X_{t-1:t-K} = (X_{t-1:t-K}^1, \dots, X_{t-1:t-K}^D)$ where for each $1 \leq d \leq D$, the notation refers to the vector gathering the observations of X^d between $t-K$ and $t-1$ in decreasing order of time,
- Y_t^{res} is a residual that takes into account non-linear memory effects, unobserved variables, noise, etc,
- $\oplus$ denotes an orthogonal sum when considering covariance as inner product.

The memory term Y_t^{lag} is also decomposed as a sum of D orthogonal components:

$$Y_t^{\text{lag}} = Y_t^1 \oplus \cdots \oplus Y_t^D, \tag{3}$$

such that each component Y_t^d quantifies the effects of the d-th variable that cannot be captured by the other components.

To achieve this second decomposition, we propose to proceed iteratively to ensure the orthogonality property, by constructing the instantaneous component first, then, transforming suitably the lagged variables to eliminate the effects already quantified. Inevitably, this strategy imposes an order between variables. However, this order will lead to a hierarchy between variables that reflects their importance; we propose then to sort them in a decreasing order of variances.

2.1 Memory effects quantification

Instantaneous component. The aim of the instantaneous component Y_t^{inst} is to capture the information carried by the instantaneous variables that are in X_t. A large variety of models, from linear models to complex neural networks tailored for time-series, can be used to approach the output Y_t depending on the complexity of the data. Here, we impose a polynomial form to Y_t^{inst}:

$$Y_t^{\text{inst}} = \sum_{p}^{P_r} \beta_p \psi^p(X_t), \tag{4}$$

where $\beta := (\beta_1, \dots, \beta_{P_r})$ is the model's parameter, $\psi(X_t) := (\psi^1(X_t), \dots, \psi^{P_r}(X_t))$ is the vector gathering polynomial functions of the instantaneous variables, where the total order is less than or equal to a specified degree r, and $P_r = \binom{D+r}{D} - 1$ is the number of polynomial terms. We choose orthogonal polynomials: if we denote $m(X_t) := (m^1(X_t), \dots, m^{P_r}(X_t))$ the vector of monomials, and L the triangular matrix resulting from the Cholesky decomposition of the covariance matrix of $m(X_t)$, we have that $\psi(X_t)$ is the solution of:

$$L\psi(X_t) = m(X_t). \tag{5}$$

In case of independent variables, using the former normalization of $\psi(X_t)$, obtaining β is straightforward by projection:

$$\beta = \mathrm{Cov}(\psi(X_t), Y_t). \tag{6}$$

Lag component. Lag effects cannot be extracted based on raw variables since this will induce redundancy and will not result in the sought orthogonality property. To this end, we transform the variables $(X^1, \ldots, X^D)$ into $(\tilde{X}^1, \ldots, \tilde{X}^D)$ iteratively: each input variable is made orthogonal to the instantaneous component and to the already fitted components using the Gram-Schmidt procedure. The lag components are chosen as distributed lag models [7]:

$$\forall d \in \{1, \ldots, D\},\ Y_t^{\mathrm{lag},d} = \sum_{s=1}^{K_d} H_s^d \tilde{X}_{t-s}^d, \tag{7}$$

where $H^d = (H_1^d, \ldots, H_{K_d}^d)$ is a vector parameter called lag weights or filter, K_d is the window size (for simplicity, we consider in the following that all variables have the same size, ie. for all d, $K_d = K$ for $K \geq 1$), and the transformed variables are constructed following the sequence:

$$\begin{aligned}\tilde{X}_{t-s}^{d,t} :=\ & X_{t-s}^d \\ & - \frac{1}{\mathrm{Var}(Y_t^{\mathrm{inst}})} \mathrm{Cov}(X_{t-s}^d, Y_t^{\mathrm{inst}}) Y_t^{\mathrm{inst}} \\ & - \sum_{i=1}^{d-1} \frac{1}{\mathrm{Var}(Y_t^{\mathrm{lag},i})} \mathrm{Cov}(X_{t-s}^d, Y_t^{\mathrm{lag},i}) Y_t^{\mathrm{lag},i}.\end{aligned} \tag{8}$$

We assume that input variables are sorted in decreasing order of variances of the lag models, ie. $\mathrm{Var}(Y_t^{\mathrm{lag},1}) \geq \cdots \geq \mathrm{Var}(Y_t^{\mathrm{lag},D})$. Analogically to the instantaneous parameter estimation, if the input variables are assumed independent, the filters can be obtained by projection:

$$\mathrm{Var}(\tilde{X}_{t-1:t-K}^{d,t}) H^d = \mathrm{Cov}(\tilde{X}_{t-1:t-K}^{d,t}, Y_t). \tag{9}$$

We propose in section 2.4 to estimate them by least squares to handle the case of dependent variables.

2.2 Variance-based time series explainability

The orthogonality obtained by construction ensures that

$$\mathrm{Var}(Y_t) = \mathrm{Var}(Y_t^{\mathrm{inst}}) + \mathrm{Var}(Y_t^{\mathrm{lag}}) + \mathrm{Var}(Y_t^{\mathrm{res}}). \tag{10}$$

In case of independent inputs, the Hoeffding-Sobol [10,23] decomposition can be applied to decompose the instantaneous variance. Efforts have been made to propose generalizations of this decomposition to the case of dependent variables [5, 15] or others indices [9, 16, 19] based on tools from game theory such as the Shapley values. As the focus of this work is on separating the instantaneous and lag contributions, we restrict ourselves to the computation of the total instantaneous index defined as:

$$S^{\text{inst}} := \frac{\text{Var}(Y_t^{\text{inst}})}{\text{Var}(Y_t)}, \tag{11}$$

the total lag index defined as:

$$S^{\text{lag}} := \frac{\text{Var}(Y_t^{\text{lag}})}{\text{Var}(Y_t)}, \tag{12}$$

and the individual lag index of variable d as:

$$S_d^{\text{lag}} := \frac{\text{Var}(Y_t^{\text{lag},d})}{\text{Var}(Y_t^{\text{lag}})}. \tag{13}$$

The indices defined above benefit from different properties. They are constant in case of stationarity, positive thanks to the positivity of the variance, and the individual lag indices sum to one. Finally, since the lag components are obtained iteratively, the lag indices should be interpreted according to the order of variables. For example, the index S_d^{lag} represents the part of variance explained by (or the effects of) the variable d without the effects of the same variable already taken into account in S^{inst} and S_i^{lag} for $i \leq d-1$.

2.3 Instantaneous versus lag compensation phenomena

For a better understanding of the proposed methodology and a better interpretation of the proposed decomposition, we conduct a theoretical analysis on the linear case, namely, the case where the true output Y_t can be written as $Y_t = Y_t^{*,\text{inst}} + Y_t^{*,\text{lag}}$ such that:

$$Y_t^{*,\text{inst}} = \sum_{d=1}^{D} \beta_d^* X_t^d \text{ and } Y_t^{*,\text{lag}} = \sum_{d=1}^{D} H^{*,d\top} X_{t-1:t-K_d}^d \tag{14}$$

and where the variables are Gaussian processes. Considering this setting allows to conduct a similar reasoning for dependent variables. However, for the sake of simplicity, we consider independent auto-regressive (AR) models of order 1 for the statistical properties[3] of the input vector:

$$\forall d \in [|1, D|],\ X_t^d = \rho_d X_{t-1}^d + \varepsilon_t^d \text{ where } \left(\varepsilon_t^d\right)_t \text{ are drawn i.i.d from } \mathcal{N}(0, 1-\rho_d^2).$$

Under these assumptions, the instantaneous component Y_{inst} defined by equation (4) with a degree $r = 1$ is precisely the conditional expectation of Y_t knowing

[3] in particular, in our setting, a zero mean, a variance of one and the covariance between t and $t+s$ being simply ρ^s

$\{X_t^1, \ldots, X_t^D\}$ which can be obtained using the properties of Gaussian conditioning as:

$$Y_t^{\text{inst}} = \mathbb{E}\left[Y_t | X_t^1, \ldots, X_t^D\right] = \beta^\top X_t, \tag{15}$$

where β is the solution to (6) and its closed-form is:

$$\beta = \left(\beta_d^* + \sum_{s=1}^{K_d} H_s^{*,d}\, \rho_d^s\right)_{d=1}^{D} \tag{16}$$

This result highlights the fact that long memory effects can be taken into account via the temporal correlation of the variables. Thus, the modelled lag effects are the remaining effects that are not modelled by the instantaneous variable. For instance, if K (the window memory of filters) is significantly high compared to the correlation length[4] τ of the input processes, the instantaneous component can hardly capture the information of the filters. In the opposite case where the correlation length is high, the instantaneous component will be preponderant.

Similarly, we have that the first lag component, defined by (7) and (8), is the conditional expectation of Y_t knowing $\tilde{X}_{t-1:t-K}^{1,t}$ which is obtained, again, by Gaussian conditioning as:

$$Y_t^{\text{lag},1} = \mathbb{E}\left[Y_t | \tilde{X}_{t-1}^1, \ldots, \tilde{X}_{t-K}^1\right] = H^{1\top} \tilde{X}_{t-1:t-K}^{1,t}, \tag{17}$$

such that:

$$\tilde{X}_{t-s}^{1,t} := X_{t-s}^1 - \rho_1^s \frac{\beta_1^*}{\sum_k \beta_k^{*2}} Y_t^{\text{inst}} \text{ and } H^1 = \tilde{\Sigma}_1^{-1} C_1, \tag{18}$$

where:

$$C_1 = \text{Cov}(\tilde{X}_{t-1:t-K}^{1,t}, Y_t) = \left(\sum_{l=1}^{K} H_l^{*1}\, (\rho_1^{|l-s|} - \rho_1^{l+s})\right)_{s=1}^{K}, \tag{19}$$

and

$$\tilde{\Sigma}_1 = \text{Var}(\tilde{X}_{t-1:t-K}^{1,t}) = \left[\rho_1^{|i-j|} - \frac{\beta_1^{*2}}{\sum_k \beta_k^{*2}} \rho_1^{i+j}\right]_{1 \leq i,j \leq K}. \tag{20}$$

In the univariate case, observe that C_1 is the product of $\tilde{\Sigma}_1$ and $H^{*,1} = (H_1^{*,1}, \ldots, H_K^{*,1})$, and thus the lagged model reduces to:

$$Y_t^{\text{lag},1} = H^{*,1\top} \tilde{X}_{t-1:t-K}^{1,t}. \tag{21}$$

In the multivariate case, by taking $|\rho| \longrightarrow 0$ for short memory processes or $|\rho| \longrightarrow 1$ for long memory processes, expanding the terms depending of ρ to the order 2 allows to observe other compensation phenomena. We refer to section 3.2 for an illustration and numerical results. Finally, the hierarchy between lag components is obtained by brute force by estimation and comparing all the variances at each iteration of the procedure.

[4] The correlation length τ is defined as the time step at which the autocorrelation function drops below $1/e$ and reflects the memory of the process. For an AR(1) process of parameter ρ, we have $\tau = -1/\log(\rho)$.

2.4 Practical estimation of the model parameters

In practice, only estimators of the covariance and the variance are available through the sample D and the independence assumption does not always hold. Hence, assuming that the time-series of our study are second-order stationary, we replace the covariances and variances with sample covariance and sample variance. For versatility, β is obtained by ordinary least squares:

$$\beta \approx \hat{\beta} \in \arg\min_{b \in \mathbb{R}^{P_r}} \sum_{t=K}^{N-1} (y_t - \psi(x_t)^\top b)^2. \tag{22}$$

Then, the filters are estimated iteratively by a penalized least squares to handle notably the case of dependent inputs:

$$H^d \approx \hat{H}^d \in \arg\min_{h \in \mathbb{R}^K} \sum_{t=K}^{T-1} (y_t - \tilde{x}^{d\,\top}_{t-1:t-K} h)^2 + \lambda \|\nabla h\|_2^2 \tag{23}$$

where we remind that $\tilde{x}^d_{t-1:t-K} = (\tilde{x}^d_{t-1}, \ldots, \tilde{x}^d_{t-K})$ and ∇h is the discrete gradient of h defined by:

$$\nabla h = (h_2 - h_1, \ldots, h_K - h_{K-1}). \tag{24}$$

The objective of this penalization is to avoid ill-posed problems and promote the smoothness of the filters obtained. Finally, we list the hyper-parameters our method relies on:

- for each variable d, the lag window K^d (responsible for the memory of the filter),
- the maximum degree r of the instantaneous component (reflects the complexity of the instantaneous model),
- the penalization parameter λ to select smooth filters in case of ill-posed problem.

For their selection, we suggest to perform an expanding window cross-validation[5] where the train set is expanding at each step and the size of the validation set is fixed.

3 Experiments

We illustrate the properties of our method on synthetic data and evaluate its performance on 3 data sets of different sizes and numbers of variables.

[5] `https://scikit-learn.org/stable/modules/generated/sklearn.model_selection.TimeSeriesSplit.html`

3.1 Setting

For our method to be applied efficiently and be as faithful as possible to the data, the hyper-parameters need to be carefully selected. For our experiments, the hyper-parameters are chosen in the following grid based on the cross-validation procedure mentioned above such that the train and validation sets are separated by a gap of 100 time steps:

- the lag window $K \in [10, 25, 50, 75, 100]$ (we impose the same window size for all variables for simplicity);
- the maximum degree $r \in [1, 2, 3, 4, 5]$;
- the penalization parameter $\lambda \in [0, 1, 10, 10^2, 10^3, 10^4]$, with no penalization if $\lambda = 0$.

3.2 Synthetic dataset

We consider two toy data sets which we refer to, respectively, as *linear case* and *general case*. The linear case satisfies the assumptions made in 2.3 and allows to recover numerically the closed-forms of the filters to study empirically the convergence of the estimators. Compared to the linear case, the general case consists of a non-linear short memory component in addition to a linear memory component, which allows to observe the behaviour of the method facing non-linearity. In all cases, the three variables are AR(1) processes. We consider in the linear case two scenarios: short correlation inputs such that the autoregressive parameters ρ_1, ρ_2, ρ_3 are chosen randomly in [0.69, 0.695] and long correlation inputs such that ρ_1, ρ_2, ρ_3 are chosen in [0.95, 0.955] to obtain correlation lengths raging approximatively around 3 time units in the short correlation case and around 20 time units in the long correlation case.

Depending on the case, the definition of Y_t is:

$$Y_t := \sum_{d=1}^{D} \beta_d^* X_t^d + \sum_{d=1}^{D} \sum_{s=1}^{K} H_s^{\text{lin},d} X_{t-s}^d \text{ in the linear case,} \tag{25}$$

$$Y_t := \sum_{p=1}^{P_r} \sum_{s=0}^{5} H_s^{\text{n-l},p}\, m^p(X_{t-s}) + \sum_{d=1}^{D} \sum_{s=1}^{K} H_s^{\text{lin},d} X_{t-s}^d \text{ in the general case} \tag{26}$$

where β^* is chosen in $[-1/10, 1/10]$. The chosen form of the linear filters is:

$$H_t^{\text{lin},d} := \frac{1}{\omega_d \sqrt{1-\zeta_d^2}} C_{\gamma_{d,1}} \exp\left(-650\, \zeta_d\, \omega_d t\right) \sin\left(\frac{650}{\omega_d \sqrt{1-\zeta_d^2}}\, t\right) \mathbf{1}_{t \in [|1,K|]}, \tag{27}$$

and the non-linear filters are:

$$H_t^{\text{n-l},d} := \frac{1}{\omega_d \sqrt{1-\zeta_d^2}} C_{\gamma_{d,2}} \exp\left(-t\right) \mathbf{1}_{t \in [|0,5|]}, \tag{28}$$

where ω_d, ζ_d, $\gamma_{d,1}$ and $\gamma_{d,2}$ (respectively) are chosen randomly in $[5 \cdot 10^{-5}, 5 \cdot 10^{-4}]$, $[0.8, 0.95]$, $[0, 1]$ and $[4 \cdot 10^{-2}, 6 \cdot 10^{-2}]$, and $C_{\gamma_{d,1}}$ and $C_{\gamma_{d,2}}$ are normalizing constants such that the integrals of the corresponding filters sum (respectively) to $\gamma_{d,1}$ and $\gamma_{d,2}$. The figure 1 shows the time evolution of the linear filters used in the linear case while the figure 2 shows the filters that are applied to linear variables X_t^1, X_t^2 and X_t^3 in the general case[6]. As the inputs X^1, X^2 and X^3 have the same amplitudes (their variance is equal to 1), looking at theses filters informs us about the sensitivity hierarchy, in the sense that, the higher the filter's integral for a variable, once normalized by the corresponding correlation length, the more influence we expect that variable to have on the output. This intuition is confirmed by the individual lag indices shown on figures 3 and 4 on which we see that the variable with the most influence on the lag component is X^2 followed by X^3 and then X^1 which is the same order implied by the integrals of linear filters shown on figure 1. Interestingly enough, even if the instantaneous component is more influential in the long correlation case than in the short correlation case, figures 3 and 4 show that the individual lag indices report the same amount of lag importance in both settings. Finally, the pie chart of figure 5 and the filters in figure 2 show that the order of individual lag indices is the same as the order of integrals of filters that are applied to the linear variables.

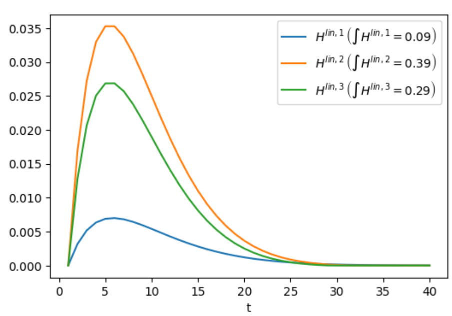

Fig. 1. Linear filters that are applied to linear variables in the linear case.

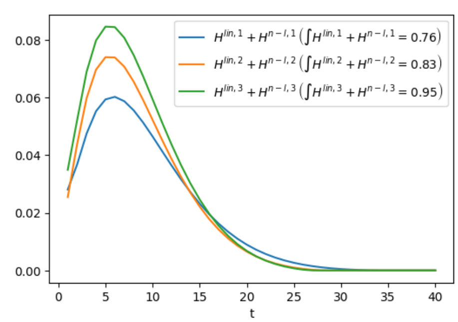

Fig. 2. Filters that are applied to linear variables in the general case.

For each linear scenario, we fit our method on a subset of size N and vary N from 200 to 100000. We save the estimated filters and record the RMSE (root mean squared error) of the estimators over 25 iterations. Since we are interested in assessing the convergence, the filters saved are the filters obtained at the first iteration after fitting the instantaneous model, ie. the filters of all variables are estimated as if the latter were the first ordered variable. The figures 6.(A) and 7.(A) show the estimated filters averaged and their 95% confidence intervals for $N = 100000$. We observe that the estimated filters in the short correlation

[6] In fact, since the monomials in equation (26) include the instantaneous variables X_t^1, X_t^2 and X_t^3, a part of the memory of linear variables is carried by the non-linear filters.

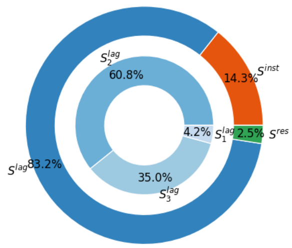

Fig. 3. Pie charts of indices in the short correlation setting

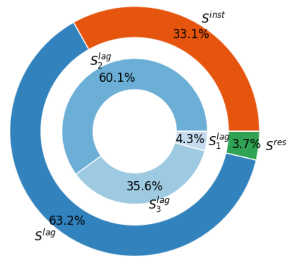

Fig. 4. Pie charts of indices in the long correlation setting

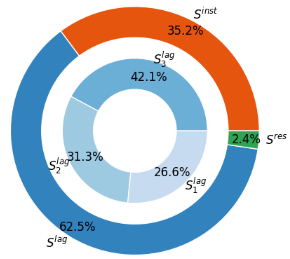

Fig. 5. Pie charts of indices in the general setting

scenario of figure 6.(A) are smoother and have reduced confidence intervals compared to the filters obtained in the long correlation setting. Moreover, the figure 6.(B) shows that in the first setting the convergence is quicker than in the setting of figure 7.(B). Finally, in virtue of the compensation analysis done in section 2.3, figure 6.(C) shows that the lag component is preponderant and control the global dynamic of the output while in the long correlation setting we observe on figure 7.(C) the tendency of the instantaneous component to be in advance and the role of the lag component in capturing the delay effects and correcting the former component.

In the general case, we perform as in the linear setting 25 repetitions and vary N between 200 and 10000. The autoregressive parameters ρ_1, ρ_2 and ρ_3 are taken in $[0.79, 0.795]$. Since non-linearities are involved, a closed-form for the filters is not available; we propose then to record the R2-score and the variances obtained to assess the convergence of the method in this case. The figure 8.(A) shows the estimated filters along with the linear filters. Here, the best lag window selected is 25. We observe that the estimators recover the shape of the linear filters without being equal because of the compensation phenomena. We observe also on figure 8.(B) that the instantaneous component is in advance compared to the lag component. Finally, figures 8.(C) and 8.(D) illustrate the convergence of variances and the R2 scores. We expectedly observe in particular that we do not recover the total variance, nor an R2-score equal to 1 which is explained by the fact that: 1)method's parameters and filters are estimated through a least squares instead of a projection approach, and, 2) the best penalization parameter selected is not zero.

3.3 Real applications datasets

In this section, we apply our methodology to real-world data sets to detect the lag effects and quantify to how much extent they allow a better explanation of the output.

We consider three data sets:

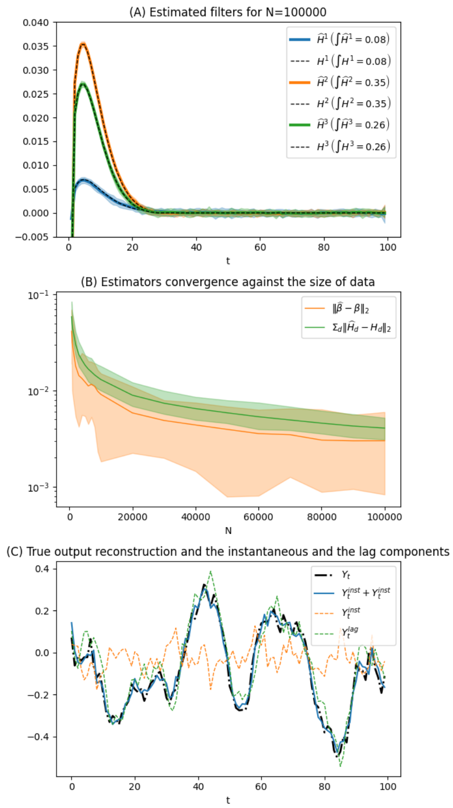

Fig. 6. Short correlation experiments visualization.

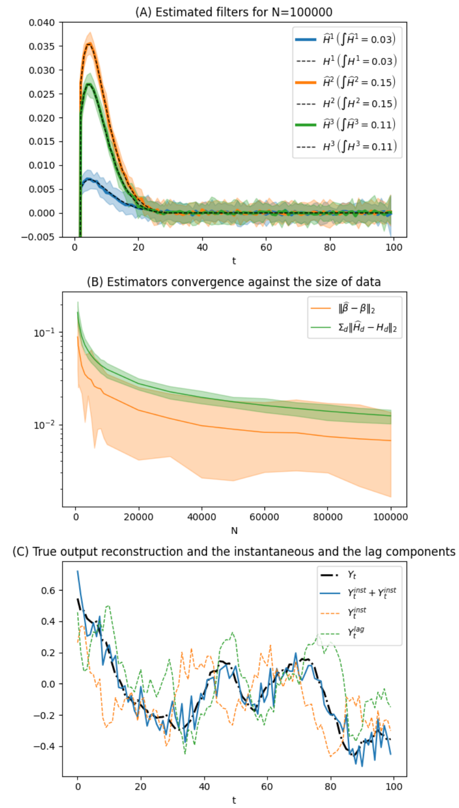

Fig. 7. Long correlation experiments visualization.

- **AirQuality**[7] is a data set provided by UCI Machine Learning repository [12] that has been first introduced in [26]. It consists of hourly measured pollutants data in addition to meteorological data with the same time sampling for different stations across multiple sites in Beijing. In our work, we only focus on one station. Data was split into two sets of size 3000: a train set between March and September 2013 and a test set between March and September 2014. We select `O3` (Ozone concentration) as the target variable to explain based on six variables: `NO2` (nitrogen dioxide concentration), `CO` (carbon monoxide), `TEMP` (temperature), `PRES` (pressure), `DEWP` (dew point temperature) and `WSPM` (wind speed).
- **GEFCom2012**[8] is a data set introduced in [11] of hourly measured wind power production with weather forecasting variables for different wind farms. We select data from one farm (chosen arbitrarily) as follows: the first 4414

[7] https://archive.ics.uci.edu/dataset/501/beijing+multi+site+air+quality+data
[8] http://blog.drhongtao.com/2016/07/gefcom2012-load-forecasting-data.html

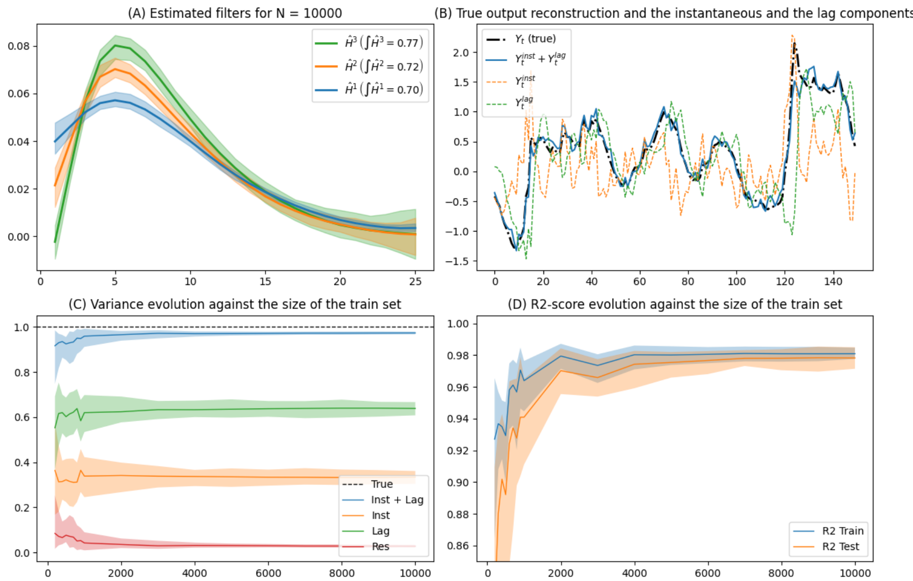

Fig. 8. general synthetic experiments visualization.

time steps are selected as the training set and the last 4414 as the test set. The target is the wind power production; we study the influence of four weather variables: `ws` (wind speed), `wd` (wind direction), `u` (the zonal wind component) and `v` (the meridional wind component).

– **WindPower** is a data set with a higher resolution than the other data sets (every 10 min). The train and test sets are of size 30000. We focus in this work on quantifying the potential delay observed on the power production of a wind farm (target) taking into account the wind speed, its direction and the temperature. In other words, we ask ourselves to how much extent the wind starting blowing at a time t is responsible for the rotation of the wind turbine (and thus the production) at the time $t + h$ where h can be seen as the duration the turbine takes to start. Data is available upon request and we refer to [2] for extensive details.

In our experiments, since only a small subset of values is missing for the AirQuality and WindPower data sets, a linear interpolation is performed to impute them. Train and test sets are selected so that they have the same distribution and the stationary property is relatively fulfilled (the study of non-stationary time series being left for future work). Data are standardized and the best hyper-parameters are selected as explained in section 3.1. We use the R2-score to assess the validity of our model and report the recorded scores in table 1 on which we read that the model produces good results in reconstructing the output and to predict new data. We compute for each dataset the total indices and the individual lag indices defined in section 2.2 and visualize them in the pie charts of figures 9, 10 and 11.

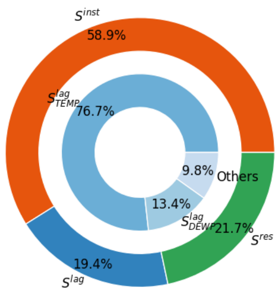

Fig. 9. Pie charts of indices for AirQuality

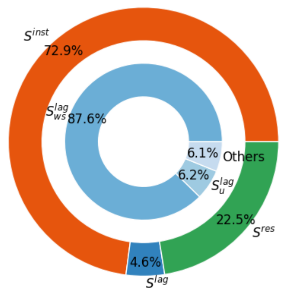

Fig. 10. Pie charts of indices for GEFCom

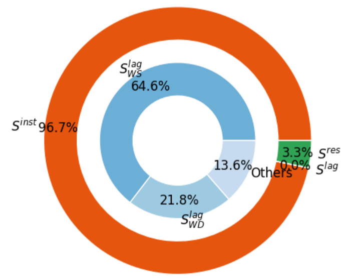

Fig. 11. Pie charts of indices for WindPower

Table 1. Predictive performance metrics.

	R^2_{test}	R^2_{train}	R^2_{inst}	R^2_{lag}
AirQuality	0.72	0.79	0.59	0.2
GEFCom	0.73	0.77	0.73	0.05
WindPower	0.95	0.97	0.97	0.00

We observe that the instantaneous effects are responsible for a substantial part in explaining the output time-series in the three application cases; the lag effects have an influence on the output in the AirQuality and GEFCom cases and no influence for WindPower data. These results are consistent with the R2-scores of table 1. We notice in particular that the lag component helps improving the results obtained by the instantaneous component of around 34% gain from R^2_{inst} to R^2_{train} for AirQuality data for example and in a lesser extent on GEFCom data. For WindPower data however, taking into account the lag effects does not improve (nor deteriorate) the performance of the instantaneous model since the latter captures all the required effects to model the target variable. Finally, we take the example of AirQuality to illustrate the decomposition of lag effects and attribute to each part an influence on the ozone concentration. In this case, the lag effects are due up to 76.7% to the temperature (TEMP) and up to 13.4% to the dew point temperature (DEWP). These two variables are known to be related for instance to solar radiation and humidity which seem to have an effect on the ozone production[9].

[9] see for eg. `https://www.epa.gov/air-trends/trends-ozone-adjusted-weather-conditions`

4 Conclusion and future work

This work proposes a methodology for history-aware sensitivity analysis based on a decomposition of the output time series into three non-correlated components: an instantaneous component, a lag component, and a residual component. This approach provides insights and clarifies the roles of memory effects of the input variables. Our experiments demonstrate the efficiency of our method and its ability to provide plausible explanations in both synthetic and real cases. Future work will focus on enhancing the interpretability of the memory component in presence of dependent variables and extending the framework to handle non-stationary and irregularly-sampled time series (for instance, handling the dependency to the seasons, manage measurement noise and missing data, etc).

Acknowledgments. This work was supported by the Chair Stress Test, RISK Management and Financial Steering, led by Ecole Polytechnique, its Foundation and sponsored by BNP Paribas. Wind power data is provided by Energy4Climate Interdisciplinary Center (E4C) of Institut Polytechnique de Paris, for which the authors would like to express their thanks.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Alexanderian, A., Gremaud, P.A., Smith, R.C.: Variance-based sensitivity analysis for time-dependent processes. Reliability Engineering System Safety **196**, 106722 (2020)
2. Bouche, D., Flamary, R., d'Alché Buc, F., Plougonven, R., Clausel, M., Badosa, J., Drobinski, P.: Wind power predictions from nowcasts to 4-hour forecasts: A learning approach with variable selection. Renewable Energy **211**, 938–947 (2023)
3. Chan, K., Saltelli, A., Tarantola, S.: Sensitivity analysis of model output: Variance-based methods make the difference. In: Winter Simulation Conference Proceedings,. pp. 261–268 (1997)
4. Chastaing, G., Prieur, C., Gamboa, F.: Generalized sobol sensitivity indices for dependent variables: numerical methods (2014), https://arxiv.org/abs/1303.4372
5. Chastaing, G., Gamboa, F., Prieur, C.: Generalized hoeffding-sobol decomposition for dependent variables -application to sensitivity analysis. Electronic Journal of Statistics **6** (12 2011)
6. Gamboa, F., Janon, A., Klein, T., Lagnoux, A.: Sensitivity analysis for multidimensional and functional outputs. Electronic Journal of Statistics **8**(1), 575 – 603 (2014)
7. Gasparrini, A., Armstrong, B., Kenward, M.G.: Distributed lag non-linear models. Statistics in Medicine **29**(21), 2224–2234 (2010)
8. Ghanem, R., Spanos, P.: Stochastic Finite Elements: A Spectral Approach. Civil, Mechanical and Other Engineering Series, Dover Publications (2003)
9. Herin, M., Il Idrissi, M., Chabridon, V., Iooss, B.: Proportional marginal effects for global sensitivity analysis. SIAM/ASA Journal on Uncertainty Quantification **12**(2), 667–692 (2024)

10. Hoeffding, W.: A Class of Statistics with Asymptotically Normal Distribution. The Annals of Mathematical Statistics **19**(3), 293 – 325 (1948)
11. Hong, T., Pinson, P., Fan, S.: Global energy forecasting competition 2012. International Journal of Forecasting **30**(2), 357–363 (2014)
12. arkelle Kelly, Rachel Longjohn, K.N.: The uci machine learning repository (2017), http://archive.ics.uci.edu
13. Lamboni, M., Monod, H., Makowski, D.: Multivariate sensitivity analysis to measure global contribution of input factors in dynamic models. Reliability Engineering System Safety **96**(4), 450–459 (2011)
14. Lara-Benítez, P., Carranza-García, M., Santos, J.C.R.: An experimental review on deep learning architectures for time series forecasting. International journal of neural systems p. 2130001 (2020)
15. Li, G., Rabitz, H., Yelvington, P., Oluwole, O., Bacon, F., Kolb, C., Schoendorf, J.: Global sensitivity analysis for systems with independent and/or correlated inputs. The journal of physical chemistry. A **114**, 6022–32 (05 2010)
16. Li, G., Rabitz, H., Yelvington, P., Oluwole, O., Bacon, F., Kolb, C., Schoendorf, J.: Global sensitivity analysis for systems with independent and/or correlated inputs. The journal of physical chemistry. A **114**, 6022–32 (05 2010)
17. Liu, Z., Choe, Y.: Data-driven sensitivity indices for models with dependent inputs using the polynomial chaos expansion. Structural Safety **88** (03 2018)
18. Nelsen, R.: An Introduction to Copulas. Springer Series in Statistics, Springer New York (2007)
19. Owen, A.B., Prieur, C.: On shapley value for measuring importance of dependent inputs. SIAM/ASA Journal on Uncertainty Quantification **5**(1), 986–1002 (2017)
20. Perrin, G.: Adaptive calibration of a computer code with time-series output. Reliability Engineering System Safety **196**, 106728 (2020)
21. Roberts, S., Osborne, M., Ebden, M., Reece, S., Gibson, N., Aigrain, S.: Gaussian processes for time-series modelling. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences **371**(1984), 20110550 (jan 2013)
22. Saltelli, A.: Sensitivity analysis for importance assessment. Risk Analysis **22**(3), 579–590 (2002)
23. Sobol, I.M.: Sensitivity estimates for nonlinear mathematical models (1993), https://api.semanticscholar.org/CorpusID:115460399
24. Sobol, I.: Global sensitivity indices for nonlinear mathematical models and their monte carlo estimates. Mathematics and Computers in Simulation **55**(1), 271–280 (2001), the Second IMACS Seminar on Monte Carlo Methods
25. Song, S., Bai, Z., Wei, H., Xiao, Y.: Copula-based methods for global sensitivity analysis with correlated random variables and stochastic processes under incomplete probability information. Aerospace Science and Technology **129**, 107811 (2022)
26. Zhang, S., Guo, B., Dong, A., He, J., Xu, Z., Chen, S.X.: Cautionary tales on airquality improvement in beijing. Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences **473** (2017)