History-aware approaches for time-series analysis

Mouad Yachouti

joint work with Guillaume Perrin and Josselin Garnier

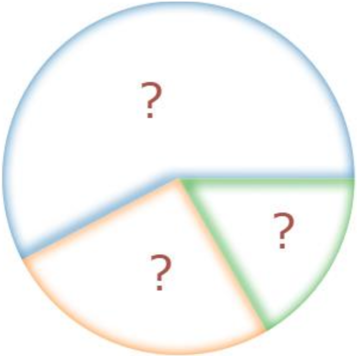
ETICS - Évian-les-Bains - October 10th, 2025

Sensitivity analysis

- **Inputs**: observations of D processes gathered in $X_t = (X_t^1, \dots, X_t^D)$;

- **Output:** $\forall t \in \mathbb{N},\ Y_t = f\left((X_s)_{s \leq t}, \varepsilon\right)$ where f is a function and ε is an error component.

$$\mathrm{Var}(Y_t) = ? + ? + ?$$



- **Inputs**: observations of D processes gathered in $X_t = (X_t^1, \ldots, X_t^D)$;
- **Output:** $\forall t \in \mathbb{N},\ Y_t = f\left((X_s)_{s \leq t}, \varepsilon\right)$ where f is a function and ε is an error component.
- **Assumptions:**
  - f is unknown,
  - zero means,
  - finite second-order moments,
  - only access to $\mathcal{D} = \{(x_0, y_0), \ldots, (x_{N-1}, y_{N-1})\}$ of $N > 1$ input-output pairs,
  - stationarity.

Aim: construct orthogonal[1] decomposition of the form:

$$Y_t = Y_t^{\text{inst}} \oplus Y_t^{\text{lag}} \oplus Y_t^{\text{res}} \tag{1}$$

$$\text{and} \quad Y_t^{\text{lag}} = Y_t^{\text{lag},1} \oplus \cdots \oplus Y_t^{\text{lag},D} \tag{2}$$

where:

- Y_t^{inst} (the instantaneous component)[2] is a function of $X_t = (X_t^1, \ldots, X_t^D)$,
- Y_t^{lag} (the lag component):
  $\forall d \in [|1, D|]$, $Y_t^{\text{lag},d}$ is a function $X_{t-1:t-K}^d = (X_{t-1}^d, \ldots, X_{t-K}^d)$,
- Y_t^{res} is a residual
  (for eg. non-linear memory effects, unobserved variables, noise, etc).

[1] considering the covariance as inner product.

[2] HDMR decomposition (Hoeffding-Sobol for eg.) if independent variables; game-theoretic indices (for eg, PME Herin et al. 2024) if dependent.

Orthogonality by construction implies:

$$\mathrm{Var}(Y_t) = \mathrm{Var}(Y_t^{\mathrm{inst}}) + \mathrm{Var}(Y_t^{\mathrm{lag}}) + \mathrm{Var}(Y_t^{\mathrm{res}}). \tag{3}$$

We define:

- the instantaneous index:

$$S^{\mathrm{inst}} := \frac{\mathrm{Var}(Y_t^{\mathrm{inst}})}{\mathrm{Var}(Y_t)}, \tag{4}$$

- the total lag and the individual lag indices:

$$S^{\mathrm{lag}} := \frac{\mathrm{Var}(Y_t^{\mathrm{lag}})}{\mathrm{Var}(Y_t)}, \quad S_d^{\mathrm{lag}} := \frac{\mathrm{Var}(Y_t^{\mathrm{lag},d})}{\mathrm{Var}(Y_t^{\mathrm{lag}})}. \tag{5}$$

A polynomial form for the instantaneous component:

$$Y_t^{\text{inst}} = \sum_{p}^{P_r} \beta_p \psi^p(X_t), \tag{6}$$

where:

- $\beta := (\beta_1, \ldots, \beta_{P_r})$;
- $\psi(X_t) := (\psi^1(X_t), \ldots, \psi^{P_r}(X_t))$ is a vector of polynomial functions of X_t;
- $P_r = \binom{D+r}{D} - 1$ is the number of polynomial terms,
- r is the maximal degree.

A polynomial form for the instantaneous component:

$$Y_t^{\text{inst}} = \sum_{p}^{P_r} \beta_p \psi^p(X_t), \tag{6}$$

where:

- $\beta := (\beta_1, \dots, \beta_{P_r})$;
- $\psi(X_t) := (\psi^1(X_t), \dots, \psi^{P_r}(X_t))$ is a vector of polynomial functions of X_t;
- $P_r = \binom{D+r}{D} - 1$ is the number of polynomial terms,
- r is the maximal degree.

In practice:

$$\beta \approx \hat{\beta} \in \arg\min_{b \in \mathbb{R}^{P_r}} \sum_{t=K}^{N-1} (y_t - \psi(x_t)^\top b)^2. \tag{7}$$

Gram-Schmidt process to transform the variables $(X^1, \ldots, X^D)$ into $(\tilde{X}^1, \ldots, \tilde{X}^D)$:

$$\begin{aligned} \forall d \in [|1, D|],\ \tilde{X}^{d,t}_{t-s} := X^d_{t-s} & \\ & - \frac{1}{\mathrm{Var}(Y^{\mathrm{inst}}_t)} \mathrm{Cov}(X^d_{t-s}, Y^{\mathrm{inst}}_t) Y^{\mathrm{inst}}_t \\ & - \sum_{i=1}^{d-1} \frac{1}{\mathrm{Var}(Y^{\mathrm{lag},i}_t)} \mathrm{Cov}(X^d_{t-s}, Y^{\mathrm{lag},i}_t) Y^{\mathrm{lag},i}_t, \end{aligned} \tag{8}$$

where the lag components $Y^{\mathrm{lag},1}_t, \ldots, Y^{\mathrm{lag},D}_t$ are chosen as distributed lag models (Gasparrini, Armstrong, and Kenward 2010):

$$\forall d \in \{1, \ldots, D\},\ Y^{\mathrm{lag},d}_t = \sum_{s=1}^{K} H^d_s \tilde{X}^d_{t-s}, \tag{9}$$

Gram-Schmidt process to transform the variables $(X^1, \dots, X^D)$ into $(\tilde{X}^1, \dots, \tilde{X}^D)$:

$$\forall d \in [\|1, D\|],\ \tilde{X}^{d,t}_{t-s} := X^d_{t-s} - \frac{1}{\mathrm{Var}(Y^{\mathrm{inst}}_t)} \mathrm{Cov}(X^d_{t-s}, Y^{\mathrm{inst}}_t) Y^{\mathrm{inst}}_t - \sum_{i=1}^{d-1} \frac{1}{\mathrm{Var}(Y^{\mathrm{lag},i}_t)} \mathrm{Cov}(X^d_{t-s}, Y^{\mathrm{lag},i}_t) Y^{\mathrm{lag},i}_t, \tag{8}$$

where the lag components $Y^{\mathrm{lag},1}_t, \dots, Y^{\mathrm{lag},D}_t$ are chosen as distributed lag models (Gasparrini, Armstrong, and Kenward 2010):

$$\forall d \in \{1, \dots, D\},\ Y^{\mathrm{lag},d}_t = \sum_{s=1}^{K} H^d_s \tilde{X}^d_{t-s}, \tag{9}$$

Without loss of generality, $\mathrm{Var}(Y^{\mathrm{lag},1}_t) \geq \dots \geq \mathrm{Var}(Y^{\mathrm{lag},D}_t)$.

In practice:

$$H^d \approx \hat{H}^d \in \operatorname{arg\,min}_{h \in \mathbb{R}^K} \sum_{t=K}^{T-1} (y_t - \tilde{x}_{t-1:t-K}^{d\,\top} h)^2 + \lambda^2 \|\nabla h\|_2^2 \tag{10}$$

where ∇h is the discrete gradient of h defined by:

$$\nabla h = \left(h_2 - h_1, \ldots, h_K - h_{K-1}\right). \tag{11}$$

Instantaneous, yet memorizes: compensation phenomena between the instantaneous part and the lag one due to temporal correlation of the processes.

Illustration: Gaussian univariate case

- $X_t = \rho X_{t-1} + \varepsilon_t \quad ; \quad (\varepsilon_t)_t$ drawn i.i.d from $\mathcal{N}(0, 1-\rho^2)$,
- $Y_t = Y_t^{*,\text{inst}} + Y_t^{*,\text{lag}}$.

$$Y_t^{*,\text{inst}} = \beta^* X_t \longrightarrow Y_t^{\text{inst}} = \mathbb{E}\left[Y_t | X_t\right] = \left(\beta^* + \sum_{s=1}^{K} H_s^* \rho^s\right) X_t \tag{12}$$

$$Y_t^{*,\text{lag}} = {H^*}^\top X_{t-1:t-K} \longrightarrow Y_t^{\text{lag}} = \mathbb{E}\left[Y_t | \tilde{X}_{t-1:t-K}\right] = {H^*}^\top \tilde{X}^t_{t-1:t-K} \tag{13}$$

$$\text{such that} \quad \tilde{X}^t_{t-s} = X_{t-s} - \frac{1}{\beta}\rho^s Y_t^{\text{inst}} = X_{t-s} - \rho^s X_t.$$

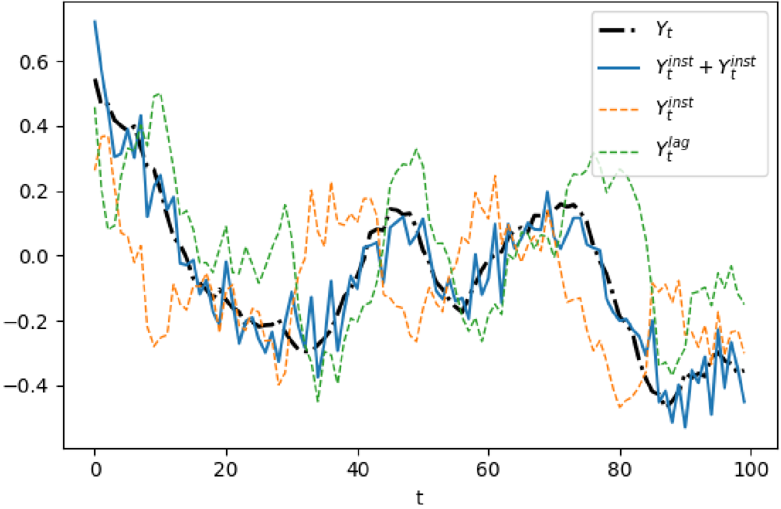


Figure 2: True output (black), reconstructed signal (blue), and inst (orange) and lag components (green).

Name	Output	Inputs
AirQuality [3]	O3 concentration	`NO2`, `CO`, `TEMP` (temperature), `PRES` (pressure), `DEWP` (dew point temp.), `WSPM` (wind speed)
GEFCom [4]	Wind Power production	`ws` (wind speed), `u` (zonal wind), `v` (meridional wind)
WindPower [5]	Wind Power production	`WS` (wind speed), `WD` (wind direction) , `T` (temperature)

Table 1: Datasets overview.

	D	N	K	r	λ	time (sec)	R^2_{test}	R^2_{train}	R^2_{inst}	R^2_{lag}
AirQuality	6	3000	50	2	10	0.441	0.718	0.787	0.589	0.199
GEFCom	3	4415	75	3	100	0.332	0.735	0.772	0.729	0.051
WindPower	3	30000	10	3	0	1.636	0.951	0.967	0.967	0.0002

Table 2: Data size, selected hyper-parameters, training times and performance metrics.

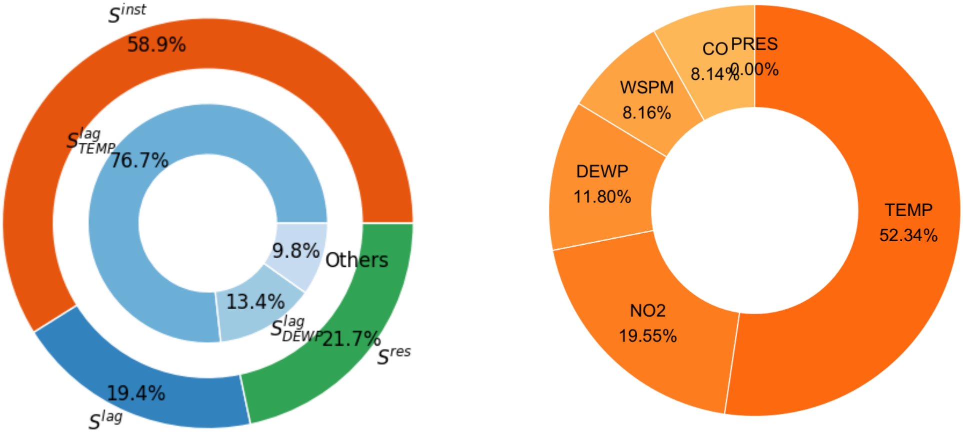


Figure 3: Donut charts of indices (instantaneous, lag, residual) for AirQuality: ours (left) and PME (right)

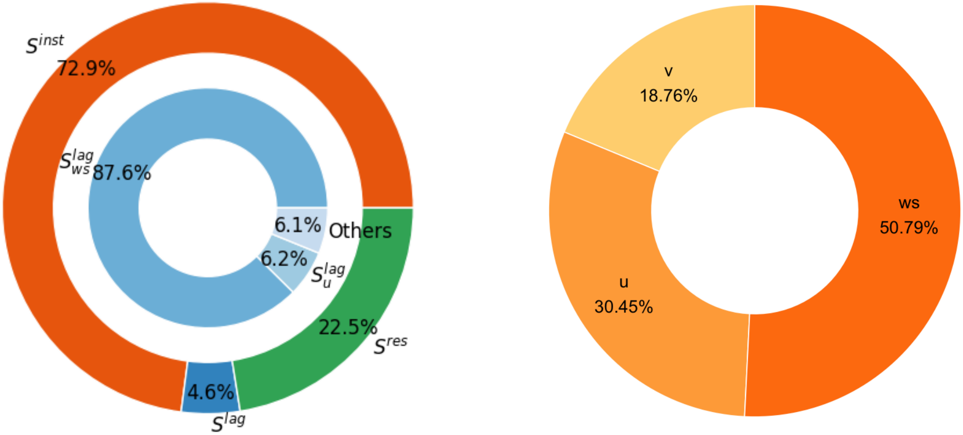


Figure 4: Donut charts of indices (instantaneous, lag, residual) for GEFCom: ours (left) and PME (right)

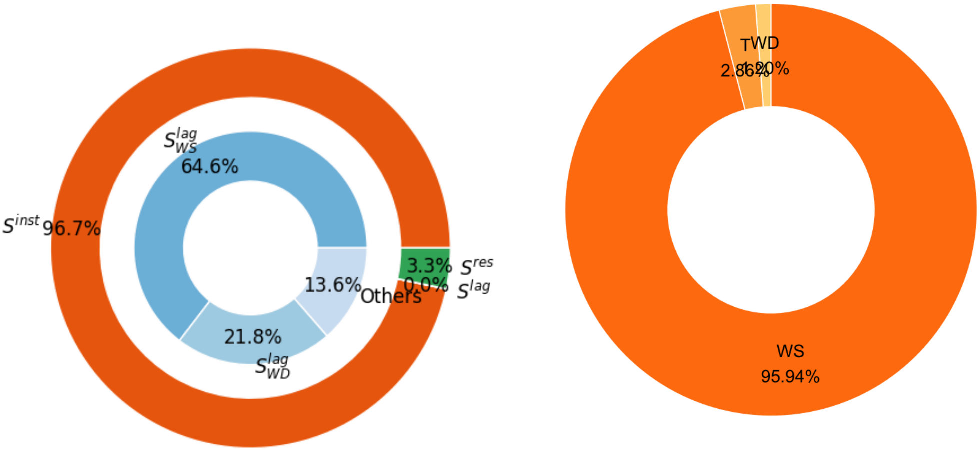


Figure 5: donut-charts of indices for WindPower

Inversion

Example: Gaussian Linear Univariate case

- input: $X_t = \rho X_{t-1} + \varepsilon_t \quad ; \quad (\varepsilon_t)_t$ drawn i.i.d from $\mathcal{N}(0, 1-\rho^2)$,
- historical model: $Y_t^{\text{hist}} = H^\top X_{t-K+1:t} + \eta_t$,
- instantaneous model: $Y_t^{\text{inst}} = h\, X_t + \eta_t$,
- $(\eta_t)_t$ drawn i.i.d from $\mathcal{N}(0, r^2)$.

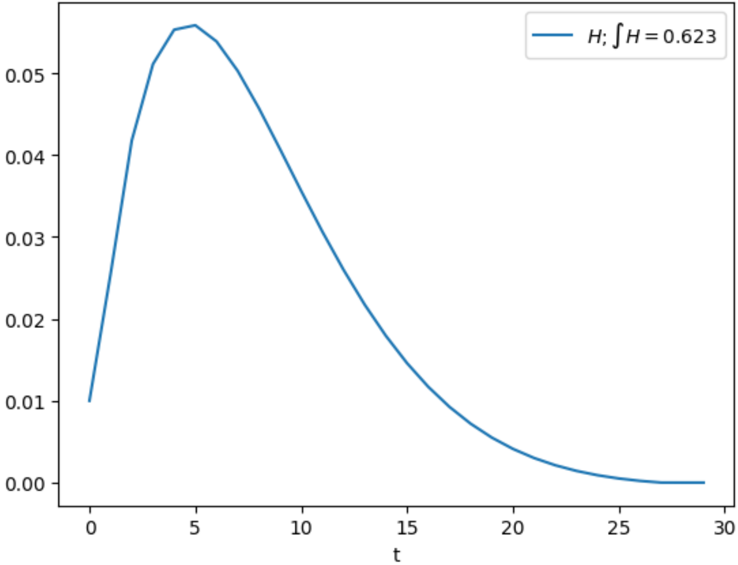


Figure 6: Filter visualisation.

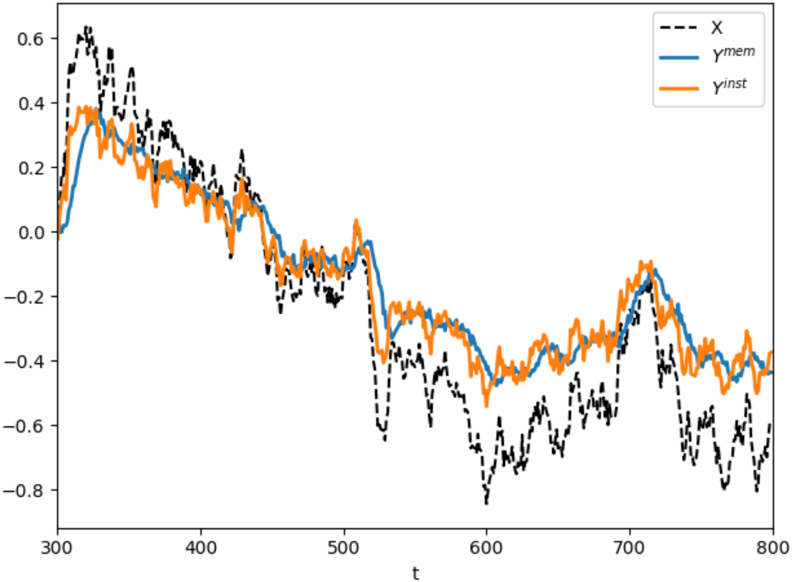


Figure 7: Data visualisation.

Aim: For $t \in [|400, 600|]$ reconstruct X_t given a sample of outputs $y_{0:799}$ for example, ie. provide an estimation of $X_t | Y_{0:799} = y_{0:799}$.

Bayesian formalism:

$$f_{X_t|Y=y} \propto f_{Y|X_t=(\cdot)}(y) \times f_{X_t} \ . \tag{14}$$

Under the aforementioned assumptions:

$$\mathcal{L}(X_t | Y^{\text{inst}}_{0:799} = y_{0:799}) = \mathcal{N}(\mu_t^{\text{inst}}, \sigma_t^{2,\text{inst}}), \tag{15}$$

$$\text{and } \mathcal{L}(X_t | Y^{\text{hist}}_{0:799} = y_{0:799}) = \mathcal{N}(\mu_t^{\text{hist}}, \sigma_t^{2,\text{hist}}). \tag{16}$$

Closed form expressions are available for μ_t^{inst}, $\sigma_t^{2,\text{inst}}$, μ_t^{hist} and $\sigma_t^{2,\text{hist}}$.

Assumptions:

- Data is composed of instantaneous responses $y_{0:799} = y_{0:799}^{\text{inst}}$,
- An instantaneous model Y_t^{inst} is calibrated on a separate dataset.

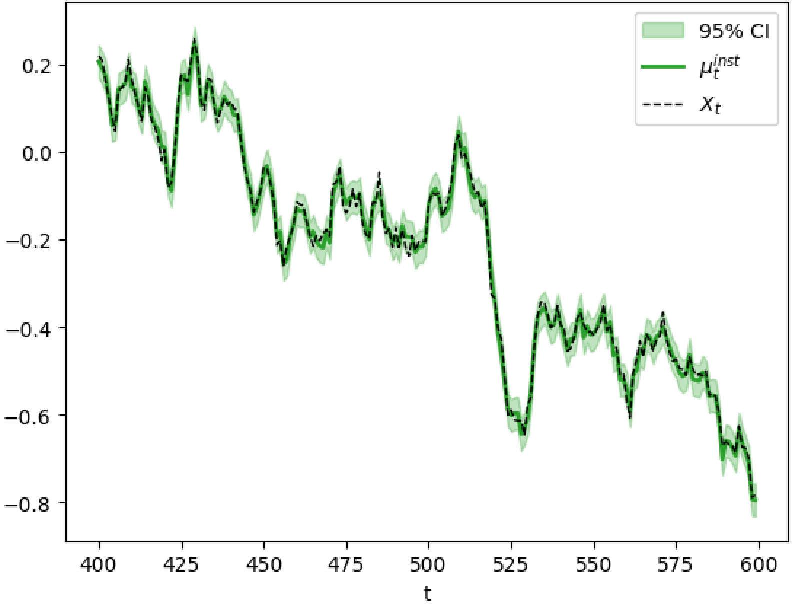


Figure 8: Instantaneous-based inversion of instantaneous outputs – $\mathrm{E}\left[(X_t - \mu_t^{\mathrm{inst}})^2\right] = 0.0002$

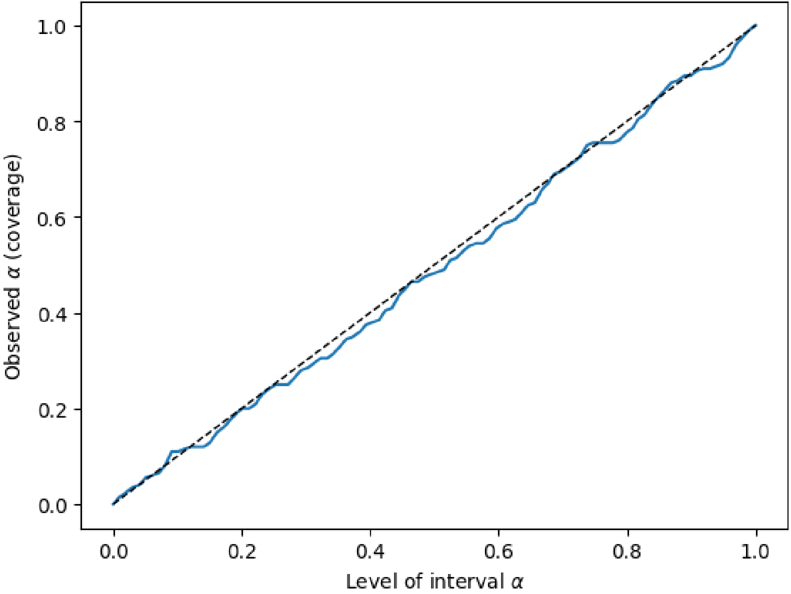


Figure 9: Accuracy of credibility intervals – αIAE = 0.013

4. Setting

5. Instantaneous output

6. Historical output

Assumptions:

- Data is composed of historical responses $y_{0:799} = y^{\text{hist}}_{0:799}$;
- Two models are calibrated on a separate dataset: an instantaneous model Y^{inst} and a historical model Y^{hist}.

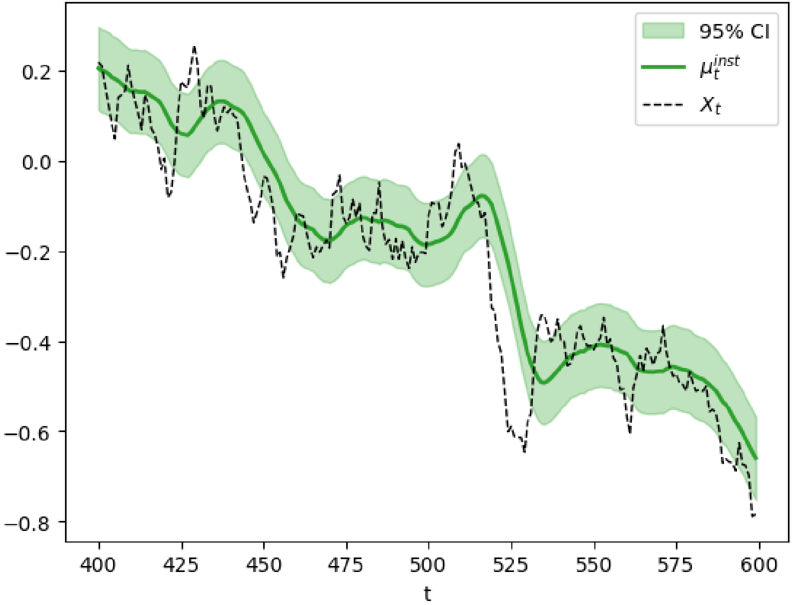


Figure 10: Instantaneous-based inversion of historical outputs – $\mathrm{E}\left[(X_t - \mu_t^{\text{inst}})^2\right] = 0.01$

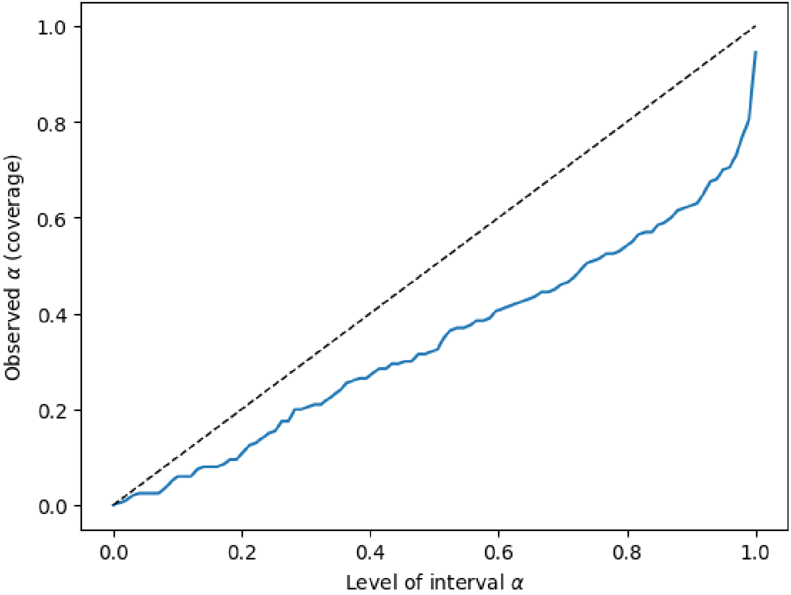


Figure 11: Accuracy of credibility intervals – αIAE = 0.159

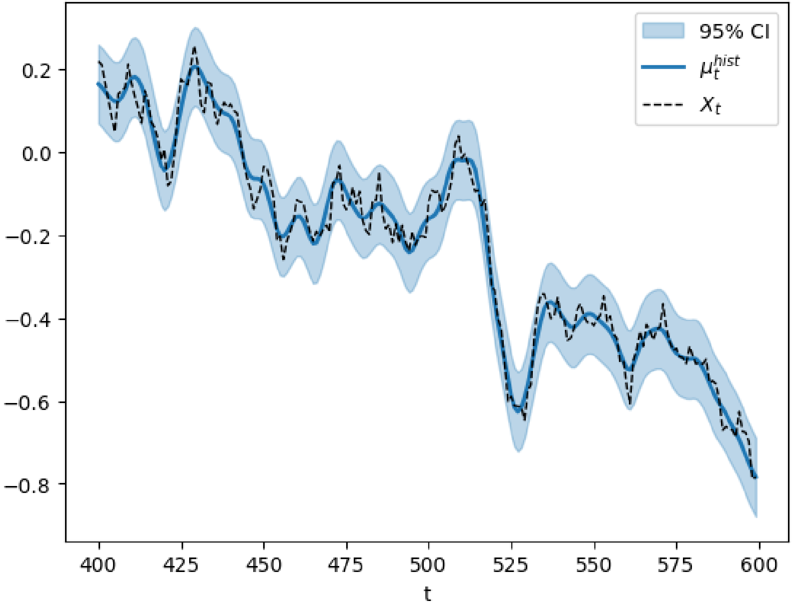


Figure 12: History-based inversion of historical outputs – $\mathrm{E}\left[X_t - \mu_t^{\text{hist}})^2\right] = 0.001$

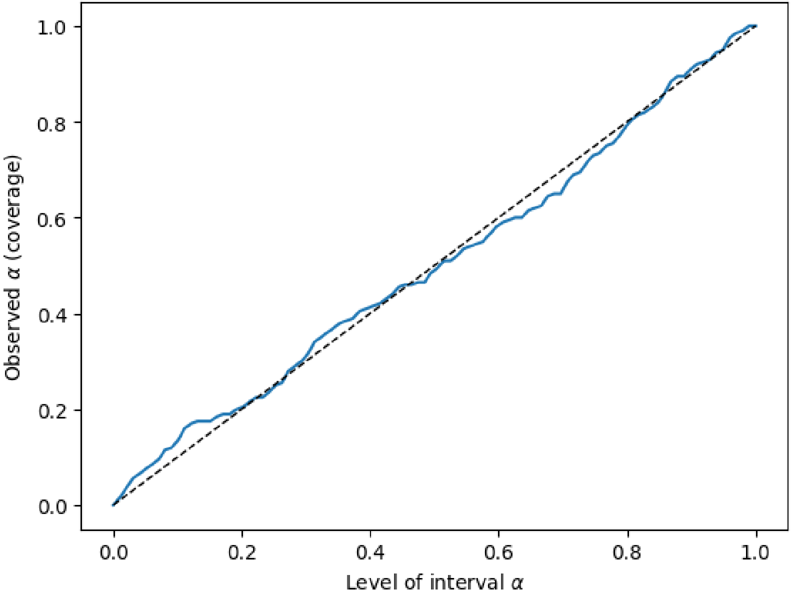


Figure 13: Accuracy of credibility intervals – αIAE = 0.017

Conclusion

- A framework of sensitivity analysis is presented for stationary time-series;
- The output is decomposed into three non-correlated components; history-aware variance-based sensitivity indices are defined.
- The methodology is applied to provide insights to explain output time-series in real cases.
- We highlighted the limitation of instantaneous-based inversion in recovering historical data.

Pre-print available on: `https://hal.science/hal-05031234`
(QR code below)

Future work will focus on:

- Handle irregularly-sampled and/or missing data;
- Applying the methodology to other data, models, and other quantities of interest (for eg. failure probabilities);
- Handle non-stationarity.
- Non-linear inversion.

Thank you for listening!

Questions?

Acknowledgment: This research is supported by the Chair Stress Test, RISK Management and Financial Steering, led by Ecole Polytechnique, its Foundation and sponsored by BNP Paribas. WindPower data set is provided by Energy4Climate Interdisciplinary Center (E4C) of Institut Polytechnique de Paris, for which we would like to express thanks.

References

- Bouche, Dimitri et al. (2023). "Wind power predictions from nowcasts to 4-hour forecasts: A learning approach with variable selection". In: *Renewable Energy* 211, pp. 938–947. ISSN: 0960-1481.
- Gasparrini, A., B. Armstrong, and M. G. Kenward (2010). "Distributed lag non-linear models". In: *Statistics in Medicine* 29.21, pp. 2224–2234.
- Herin, Margot et al. (2024). "Proportional marginal effects for global sensitivity analysis". In: *SIAM/ASA Journal on Uncertainty Quantification* 12.2, pp. 667–692.
- Hong, Tao, Pierre Pinson, and Shu Fan (2014). "Global Energy Forecasting Competition 2012". In: *International Journal of Forecasting* 30.2, pp. 357–363. ISSN: 0169-2070.
- Zhang, Shuyi et al. (2017). "Cautionary tales on air-quality improvement in Beijing". In: *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences* 473.