Wavelet-Variance-Based Estimation of Latent Time Series Models

Stéphane Guerrier (UCSB)

joint work with

Jan Skaloud (EPFL)
Yannick Stebler (u-blox)
Maria-Pia Victoria-Feser (University of Geneva)

25th of April, 2014

Introduction

General Framework:

- We propose a new estimation method for the parameters of a time series models.
- It is based on the **Generalized Method of Wavelet Moments (GMWM)** of *Guerrier et al, JASA, 2013*.
- The estimator is based on the matching between empirical and model based Wavelet Variances (WV).
- This method is particularly well suited for (Gaussian) latent time series models (or composite stochastic processes) which are employed in many engineering applications.

Setting:

Let $F_{\boldsymbol{\theta}}$ be the model associated to the univariate time series with Gaussian innovations $(Y_t; t \in \mathbb{Z})$ that is stationary or non-stationary but with stationary backward differences of order d.

Latent time series models

Problem statement:

- Many processes of interest cannot be **"observed" directly (or perfectly)**, examples include:
  - concentration of a chemical substance,
  - population of a species,
  - activity of brain cells.
- It is required to employ a sort of "device" to record it (say (y_t)).
- The observed time series (y_t) is composed of (at least) two latent (unobserved) processes say (p_t) ("true" process; $F_{\boldsymbol{\theta}_1}$) and (e_t) (measurement error; $F_{\boldsymbol{\theta}_2}$).
- We have that $y_t = p_t + e_t$.

Latent time series models

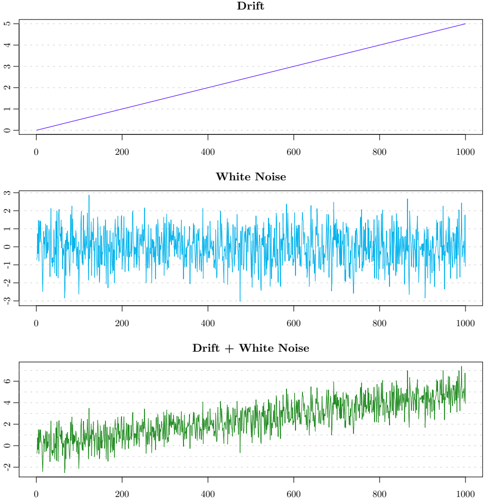
Problem statement:

- Researcher's interest generally lies in either estimating $\boldsymbol{\theta}_1$ or in the recovery of (p_t).
- The measurement error (or (p_t)) can also be itself composed of several latent processes.
- This is often the case in engineering due to the "physics of device" employed, examples include:
  - **inertial sensors**,
  - radio-astronomical instrumentation,
  - gas monitoring spectrometers.

Outline

1. What is a latent time series?
2. Existing methods
3. The Wavelet Variance
4. The Generalized Method of Wavelet Moments
5. Simulations
6. Examples of applications in engineering: inertial sensors calibration
    - Virtual reality
    - 3D maps
    - Search and rescue
    - Robots (e.g. autonomous underwater vehicles)
7. Extensions:
    - Robust Generalized Method of Wavelet Moments
    - Model selection

An easy latent time series model



Remarks:

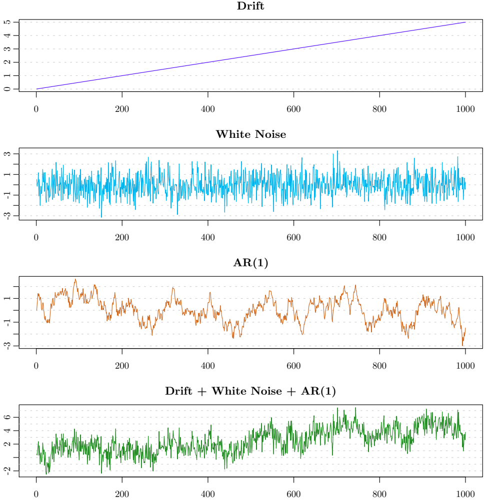
- Simple linear regression model:

$$y_t = \omega t + \varepsilon_t$$

$$\varepsilon_t \overset{iid}{\sim} \mathcal{N}\left(0, \sigma^2\right)$$

- MLE is perfectly fine.
- **What if we add an AR(1) process?**

Adding an autoregressive process



Remarks:

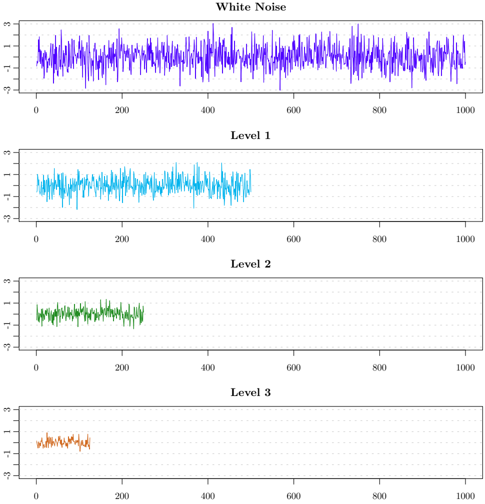
- Not a linear regression model but a **state space model**.
- Computing the likelihood is not an easy task (Kalman filter).
- **MLE (in fact EM-KF) fails**.

Estimation of Latent Time Series Models

Existing methods:

- Transforming into a "non-latent" model (e.g. ARMA)
  - Does not work in general.
  - Tends to diverge when one latent time series is "close" to unit root.
  - Difficult to "inverse".
- MLE of an associated state-space (possibly using EM algorithm)
  - Computationally intensive.
  - Systematically diverges with "complex" models.
  - A lot of work is needed for a new model (see *Stebler et al, Meas Sci Tech, 2012*).
- "Graphical" method used in engineering
  - Limited to a few possible models.
  - Not consistent in general (see *Guerrier et al, 2013b*).
  - "Inefficient" (see *Guerrier et al, 2013b*).

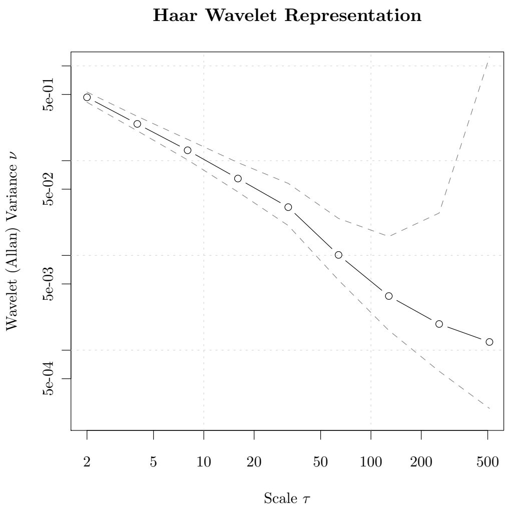
Looking differently at a time series using the Allan variance



Remarks:

- Proposed by *Allan, Proc. IEEE, 1966*
- Originally intended to study the stability of atomic clocks.
- Allan Variance is closely related to the (haar) Wavelet Variance.

The Allan (wavelet) variance



Initial idea:

Match the WV:

- Exploit the relationship that exists between the model $F_{\boldsymbol{\theta}}$ and the WV $\boldsymbol{\nu}(\boldsymbol{\theta})$ (i.e. mapping $\boldsymbol{\theta} \mapsto \boldsymbol{\nu}(\boldsymbol{\theta})$).
- "Inverse" this mapping by minimising some discrepancy between empirical WV ($\hat{\boldsymbol{\nu}}$) and the theoretical WV $\boldsymbol{\nu}(\boldsymbol{\theta})$.
- This should provide an approximation of the point $\boldsymbol{\theta}(\hat{\boldsymbol{\nu}})$.

Wavelet Variance

Empirical WV:

- The WV (ν_{τ_j}) is the **variance of wavelet coefficients** for the scale τ_j.
- Wavelet coefficients ($\overline{W}_{j,t}$) are weighted averages computed on the series Y_t.
- The weights are called wavelet filters h_j: e.g. the Haar wavelet filter ($d = 1$).
- The wavelet filters give non nil weights to observations at a given lag, called scale (τ_j) (windows size of length L_j). Hence there are as many WV as there are scales.
- The wavelet filters can be computed on consecutive windows, or on overlapping windows (to get $\widetilde{W}_{j,t}$ using $\tilde{h}_j$). Overlapping windows lead to more efficient estimators.

Wavelet Variance

Definition:

The Wavelet Variance (WV) is the variance of the wavelet coefficients, i.e.

$$\nu_{\tau_j} = \text{var}\left(\widetilde{W}_{j,t}\right), \quad \text{where} \quad \widetilde{W}_{j,t} = \sum_{l=0}^{L_j-1} \tilde{h}_{j,l} y_{t-l}, \; t \in \mathbb{Z}$$

and where $\tilde{h}_{j,l}$ wavelet filters (based on MODWT).

Remark:

The Allan variance is a special case of the WV, in fact $\sigma_{\bar{y}}(\tau) = 2\nu_\tau$ where ν_τ is based on Haar wavelet.

Estimation of the WV

MODWT estimator:

A consistent estimator for ν_{τ_j} is given by the MODWT estimator defined in *Percival, Biometrika, 1995*

$$\hat{\nu}(\tau_j) = \frac{1}{M(T_j)} \sum_{t \in T_j} \widetilde{W}_{j,t}^2$$

Theorem: asymptotic normality

Serroukh et al, JASA, 2000 show that under suitable conditions

$$\sqrt{M(T_j)} \left(\hat{\nu}(\tau_j) - \nu_{\tau_j}\right) \xmapsto[T \to \infty]{\mathcal{D}} \mathcal{N}\left(0, S_{W_j}(0)\right)$$

A more general theorem

Theorem: multivariate extension

We extended this result to the multivariate case and demonstrated that under some regularity conditions

$$\sqrt{T}\left(\hat{\boldsymbol{\nu}} - \mathbb{E}[\hat{\boldsymbol{\nu}}]\right) \xrightarrow[T\to\infty]{\mathcal{D}} \mathcal{N}\left(0, \mathbf{V}_{\hat{\boldsymbol{\nu}}}\right)$$

where $\mathbf{V}_{\hat{\boldsymbol{\nu}}} = [\sigma^2_{kl}]_{k,l=1,\ldots,J}$.

Remark:

- This theorem generalises *Serroukh et al, JASA, 2000* result and enables to compute the (asymptotic) covariance between the WV (or the AV) at two different scales.
- We proposed an estimator for σ^2_{kl}.

Theoretical WV

WV implied by F_{θ}:

Given a model $F_{\boldsymbol{\theta}}$ one can compute the theoretical WV as:

$$\nu_{\tau_j} = f\left(\boldsymbol{\theta}\right) = \int_{-1/2}^{1/2} |\widetilde{H}_j(f)|^2 S_{F_{\boldsymbol{\theta}}}(f) df$$

Example:

Consider an AR(1):

$$Y_t = \phi_1 Y_{t-1} + \varepsilon_t, \;\; \varepsilon_t \overset{iid}{\sim} \mathcal{N}\left(0, \sigma_\varepsilon^2\right), \;\; t = 1, \ldots, T$$

The theoretical (haar) WV of such process is given by

$$\nu_{\tau_j} = \frac{\left(\frac{\tau_j}{2} - 3\phi_1 - \frac{\tau_j \phi_1^2}{2} + 4\phi_1^{\frac{\tau_j}{2}+1} - \phi_1^{\tau_j+1}\right)\sigma_\varepsilon^2}{\frac{\tau_j^2}{8}\left(1-\phi_1\right)^2\left(1-\phi_1^2\right)}$$

WV of latent time series models

A very useful property:

Suppose we have

$$Y_t = X_t^{(1)} + \ldots + X_t^{(k)}$$

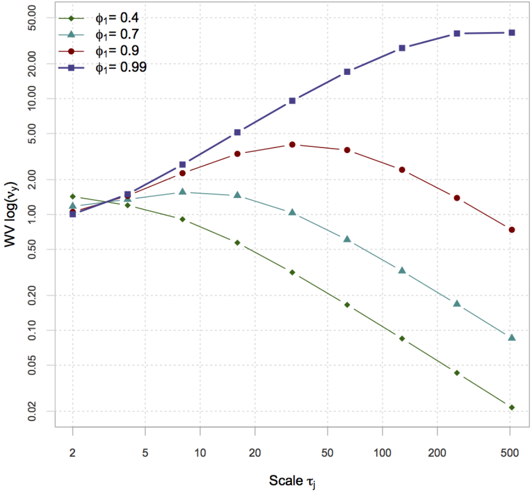
then the PSD of Y_t is

$$S_{Y_t} = S_{X_t^{(1)}} + \ldots + S_{X_t^{(k)}}$$

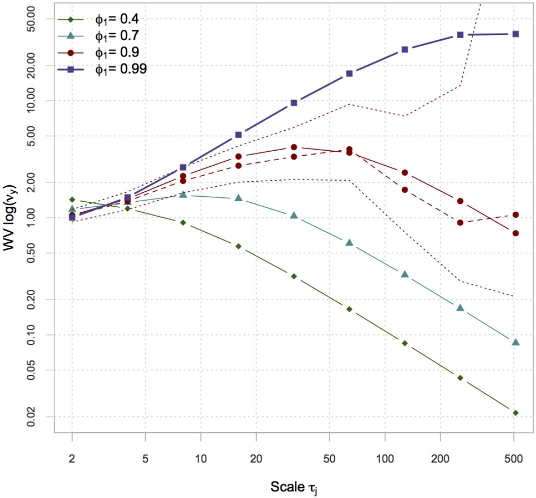
so the WV of Y_t is given by

$$\nu_{Y_t,\tau_j} = \int_{-1/2}^{1/2} |\widetilde{H}_j(f)|^2 \left(\sum_{i=1}^{k} S_{X_t^{(i)}} \right) df = \sum_{i=1}^{k} \nu_{X_t^{(i)},\tau_j}$$

Principle of the GMWM



Principle of the GMWM



The GMWM estimator

Definition:

The GMWM estimator is the solution of the following optimisation problem

$$\hat{\boldsymbol{\theta}} = \underset{\boldsymbol{\theta} \in \boldsymbol{\Theta}}{\text{argmin}} \ (\hat{\boldsymbol{\nu}} - \boldsymbol{\nu}(\boldsymbol{\theta}))^T \boldsymbol{\Omega} (\hat{\boldsymbol{\nu}} - \boldsymbol{\nu}(\boldsymbol{\theta}))$$

in which $\boldsymbol{\Omega}$, a positive definite weighting matrix, is chosen in a suitable manner such that the above quadratic form is convex.

Theorem: Consistency

$\hat{\boldsymbol{\theta}}$ is a consistent estimator of $\boldsymbol{\theta}$ (under some regularity conditions) for a large class of (latent) models.

Asymptotic distribution of $\hat{\theta}$

Theorem: asymptotic normality

We showed that (under some regularity conditions)

$$\sqrt{T}\left(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0\right) \xrightarrow[T\to\infty]{\mathcal{D}} \mathcal{N}\left(0, \mathbf{V}_{\hat{\boldsymbol{\theta}}}\right)$$

where $\mathbf{V}_{\hat{\theta}} = \mathbf{B}\mathbf{V}_{\hat{\boldsymbol{\nu}}}\mathbf{B}^T$, $\mathbf{B} = \left(\mathbf{D}^T\boldsymbol{\Omega}\mathbf{D}\right)^{-1}\mathbf{D}^T\boldsymbol{\Omega}$, $\mathbf{D} = \partial\boldsymbol{\nu}(\boldsymbol{\theta})/\partial\boldsymbol{\theta}^T$ and $\mathbf{V}_{\hat{\boldsymbol{\nu}}}$ is the asymptotic covariance matrix of $\hat{\boldsymbol{\nu}}$.

Choosing Ω:

The most efficient estimator is (asymptotically) obtained by choosing $\boldsymbol{\Omega} = \mathbf{V}_{\hat{\boldsymbol{\nu}}}^{-1}$, leading then to $\mathbf{V}_{\hat{\theta}} = (\mathbf{D}^T\mathbf{V}_{\hat{\boldsymbol{\nu}}}^{-1}\mathbf{D})^{-1}$.

Modified Forms of the GMWM

Improving the GMWM:

Simulation studies demonstrated that GMWM estimators can present some bias in finite samples. We propose 2 modified versions to improve the GMWM performance:

- augmented GMWM (add other moments),
- bias correction (based on indirect inference) $\rightarrow$ not today.

Augmented GMWM:

- We consider the augmented moment vector $\boldsymbol{\nu}^{+}(\boldsymbol{\theta}) = [\boldsymbol{m}(\boldsymbol{\theta})\ \boldsymbol{\nu}(\boldsymbol{\theta})]$ where $\boldsymbol{m}(\boldsymbol{\theta})$ is a vector of (functions of) moments of Y_t.
- *Guerrier et al, 2013b*, show that the efficiency of the GMWM can be greatly improved if one adds other moments of the process.

Simulations

Setting:

The model "White Noise + Gauss Markov + Drift" is in fact commonly used in many engineering applications (e.g. inertial navigation). We simulate $B = 100$ signals of length $T = 6000$ from this model with $\boldsymbol{\theta}_0 = (\sigma_{WN}, \sigma_{GM}, \beta, \omega) = (2, 4, 0.05, 0.005)$

Comparison with ML (EM-KF)

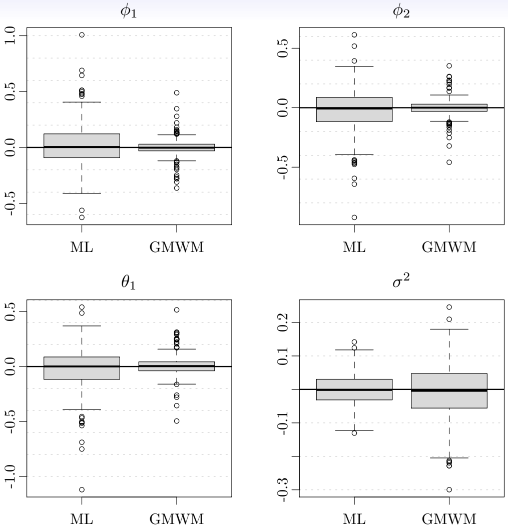
	GMWM		EM-KF	
	RMSE	R-RMSE	RMSE	R-RMSE
σ^2_{GM}	0.96	0.06	74.57	4.66
β	$4.63 \cdot 10^{-3}$	0.09	0.04	0.85
σ^2_{WN}	0.11	0.03	0.16	0.04
ω	$2.79 \cdot 10^{-4}$	0.06	0.12	23.58

ARMA models

Simulation Setting: ARMA(2,1)

- $Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \theta_1 \varepsilon_{t-1} + \varepsilon_t$.
- $\varepsilon \overset{iid}{\sim} \mathcal{N}\left(0, \sigma^2\right)$.
- $\phi_1 = 0.8$, $\phi_2 = 0.1$, $\theta_1 = 0.3$ and $\sigma^2 = 1$.
- Sample size $T = 10^3$.
- Bootstrap replications $B = 500$.

ARMA models



ϕ_1
1.0
0.5
0.0
-0.5
ML
GMWM
ϕ_2
0.5
0.0
-0.5
ML
GMWM
θ_1
0.5
0.0
-0.5
-1.0
ML
GMWM
σ^2
0.2
0.1
-0.1
-0.3
ML
GMWM

Case Study: Inertial Sensors

Inertial Measurement Unit (IMU):

An IMU is composed of accelerometers and gyroscopes. Very widely employed as part of "integrated" navigation systems, a few examples:

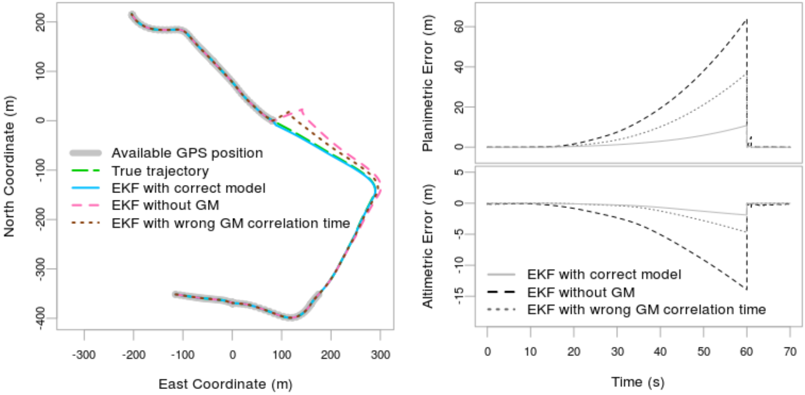
- Smartphones
- Robots
- (Small) Unmanned Aerial Vehicles (e.g. 3D modelling, search and rescue, etc)
- Autonomous underwater vehicles (e.g. water quality monitoring)
- Sport (e.g. ski, swimming, etc)
- Virtual reality (e.g. video games, movies, etc)

GMWM to calibrate IMUs

IMU are subject to measurement errors...

1. IMUs are corrupted by **stochastic errors** (typically modelled as latent time series).
2. These errors are modelled using stochastic processes in a "navigation filter".
3. The "quality" of the estimated error model has a **great impact on the "navigation" performance** (e.g. orientation, position, ...).
4. There exist (at least) three methods to estimate IMU errors, **the GMWM is one of them**.
5. **GMWM was designed for such problems and performs very well.**

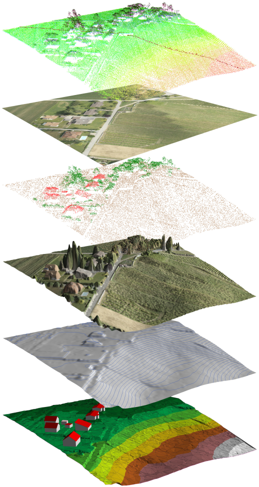
Effect on position of error model



Helicopter-based mobile mapping

Ecole Polytechnique Fédérale de Lausanne (EPFL)

scan2map™
HANDHELD AIRBORNE MAPPING SYSTEM



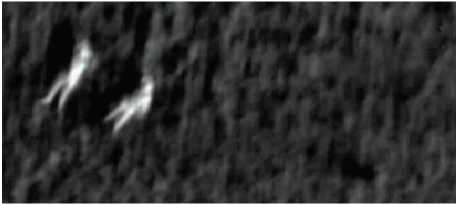
Helicopter-based mobile mapping

Ecole Polytechnique Fédérale de Lausanne (EPFL)

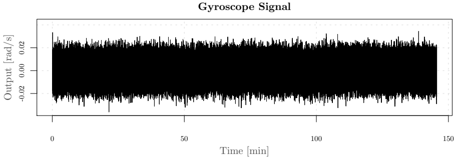
3D Map of the Matterhorn (Switzerland)

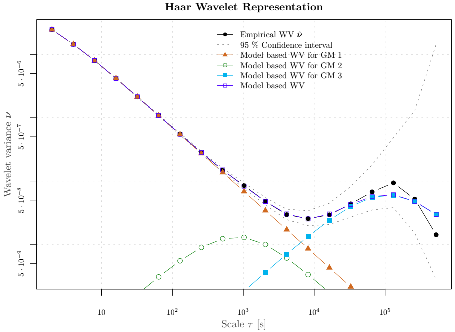
UAV for search and rescue operations

Institut de Geomàtica (Barcelona), EPFL, ...

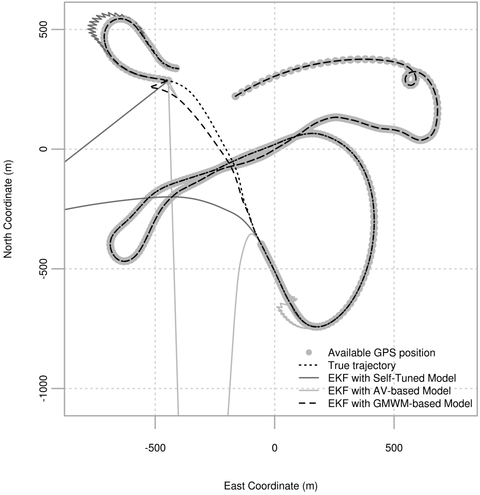


IMU calibration using the GMWM





Comparing results with benchmark methods



Extensions

Robust GMWM:

- We proposed a robust estimator of the WV.
- This extends the results of *Mondal and Percival, Ann I Stat Math, 2010*.
- It can be shown that the influence function of the GMWM is proportional to the influence function of the WV (since it is an "indirect" estimator, see e.g. *Genton & Ronchetti, JASA, 2003*).
- Therefore we proposed a robust GMWM estimator (bounded influence function) which is consistent and asymptotically normally distributed (see *Guerrier et al, 2013c)*.
- This estimator provides a flexible and simple framework to estimate in a robust fashion many time series models such as ARMA models.

Extensions

Model selection for GMWM estimators:

- We proposed the Wavelet Variance Information Criterion (WVIC) which is defined as:

$$\text{WVIC} = \mathbb{E}\left[\mathbb{E}_0\left[\left(\hat{\boldsymbol{\nu}}^0 - \boldsymbol{\nu}(\hat{\boldsymbol{\theta}})\right)^T \boldsymbol{\Omega}\left(\hat{\boldsymbol{\nu}}^0 - \boldsymbol{\nu}(\hat{\boldsymbol{\theta}})\right)\right]\right].$$

- Using the "optimism" theorem of *Guerrier & Victoria-Feser, 2013* we derived an unbiased and consistent estimator of the WVIC.
- This model selection criterion was specifically developed for the GMWM estimation technique.
- An automatic algorithm for selecting the error model of inertial sensor based on the WVIC was proposed in *Guerrier et al, ION, 2013*.

Current research

Latent model with time varying parameters:

- Project in collaboration with several researchers from the Swiss Center for Affective Sciences.
- The objective is to extend the GMWM approach to models with time varying parameters with the purpose of describing electroencephalogram time series.
- Consider for example the model:

$$
\begin{aligned}
Y_t &= X_t + Z_t \;\; \text{where} \\
X_t &= \phi X_{t-1} + \epsilon_t, \;\; \epsilon_t \sim \mathcal{N}\left(0, \sigma^2\right) \\
Z_t &= \psi_t Z_{t-1} + \varepsilon_t \;\; \varepsilon_t \sim \mathcal{N}\left(0, \gamma_t^2\right)
\end{aligned}
$$

Conclusions

The GMWM estimator...

- is able to model signals of **complex spectral structure** (where classical methods fail!).
- can be computed (in some situations) using **computationally inexpensive algorithm** (unlike EM-KF approach) even with very large sample sizes.
- is **consistent** (unlike many methods used in the engineering community) and asymptotically **normally distributed**.
- can be used as a robust estimator for time series models.
- can be applied to model the error behaviour of inertial sensors and **considerably increases navigation accuracy**.

Thank you very much for your attention

Main References on the GMWM

Methodology-related:

- Guerrier, S., Skaloud, J., Stebler, Y. and Victoria-Feser, M.P. Wavelet Variance based Estimation for Composite Stochastic Processes. Journal of the American Statistical Association, 2013
- Guerrier, S., Stebler, Y., Skaloud, J. and Victoria-Feser,M.-P. Limits of the Allan Variance and Optimal Tuning of Wavelet Variance based Estimators. Submitted working paper, 2013
- Guerrier, S., Molinari, R. and Victoria-Feser,M.-P. Estimation of Time Series Models via Robust Wavelet Variance. Submitted working paper, 2013

Engineering-related:

- Stebler, Y., Guerrier, S., Skaloud, J. and Victoria-Feser, M.-P., The Generalized Method of Wavelet Moments for Inertial Navigation Filter Design, IEEE Transactions on Aerospace and Electronic Systems (to appear in print), 2014
- Stebler, Y., Guerrier, S., Skaloud, J. and Victoria-Feser, M.-P., Study and Modelling of MEMS-based Inertial Sensors Operating in Dynamic Conditions, Submitted working paper
- Stebler, Y., Guerrier, S., Skaloud, J., and Victoria-Feser, M.P.: A Framework for Inertial Sensor Calibration Using Complex Stochastic Error Models. Position Location and Navigation Symposium (PLANS), IEEE/ION (2012)
- Stebler, Y., Guerrier, S., Skaloud, J., and Victoria-Feser, M.P.: Constrained expectation-maximization algorithm for stochastic inertial error modeling: study of feasibility. Measurement Science and Technology 22(8) (2011)