Switching autoregressive models
Application to wind time series

Pierre Ailliot [1]

Joint work with Valérie Monbet[2] and Françoise Pène[1]

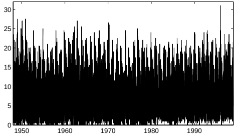
[1]LMBA, UBO
[2]IRMAR, Rennes 1

Stochastic weather generators?

- Stochastic tools for generating sequences of meteorological variables

Historical data

50 years of wind speed at Ouessant



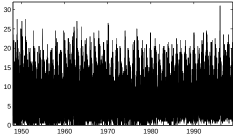
Stochastic model

MS-AR model calibrated on historical data

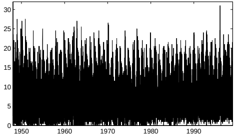
→

Synthetic data

Large number of synthetic wind sequences with statistical properties similar to original data



⋮



Stochastic weather generators?

- Stochastic tools for generating sequences of meteorological variables
- Used in impact studies when climate is involved and not enough data to estimate the quantities of interest

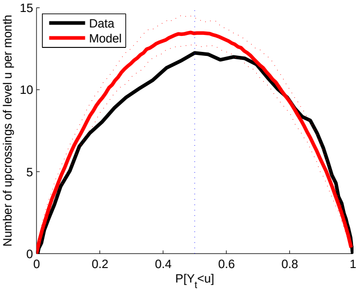
Synthetic weather conditions → System → Distribution of the response

- Most usual applications: hydrology and agriculture
  - Meteorological variables: rainfall, temperature, solar radiation, humidity, wind speed,...
- Why developing wind generators?
  - Renewable energy
  - Coastal erosion
  - ...
- In the first part of this talk, focus on wind speed at a single location
  - Ouessant, 50 years, $\Delta t = 6h$
  - Preprocessing step to remove seasonal components (+ trends, daily components)?
    - In this talk, data blocked by month, results for January (no significant daily component)

- Classical approach for modeling time series
- Simulation of the stationary time series $\{Y_t\}$ in three steps:
  - First step: find a deterministic monotonic function g such that
    $$X_t = g(Y_t)$$
    has (approximatively) Gaussian margins
    - "Normal score transformation": $g = \Phi^{-1} \circ F_Y$
      $F_Y(y) = P[Y_t < y]$, Φ cdf of $\mathcal{N}(0,1)$
    - "Box-cox transformation": $g(y) = \frac{y^\lambda - 1}{\lambda}$
  - Second step: $\{X_t\}$ is a stationary process with (approximately) Gaussian margins. Further assume that $\{X_t\}$ is a Gaussian process and simulate this process.
    - Exact simulation
    - ARMA models
  - Third step: apply the inverse transform g^{-1} to the simulated Gaussian sequence

Transformed Gaussian processes

- \+ Easy to implement
- \+ Can be used for multivariate data (several sites and/or variables)
- \- May be hard to interpret
- \- Can not reproduce some non-linearities?
  - A stationary Gaussian process has a symmetric behavior below α and above $1-\alpha$ quantile
  - The same holds true for transformed Gaussian processes (since g is monotonic)



  - Observed storms have shorter durations than calm periods
- Alternatives?
  - More general transformation $X_t = g(Y_{h(t)})$ with g and h (stochastic?) transformations such that $\{X_t\}$ is a Gaussian process
  - Non-linear time series models

Weather type models

- Weather type models are widely used for meteorological variables
  - Weather type = typical weather situation
    - e.g. dry/wet days, convective/frontal rainfall, cyclonic/anticyclonic conditions,...
    - Generally between 2 and 10 weather types are used; enough to summarize the climate
  - Weather type models characterized by
    - A model for the weather type sequence
    - A model for the meteorological variables conditionally to the weather type
- An example : the chain dependent model (Richardson, 1981)
  - Weather type S_t={Dry,Wet}
    - First order Markov chain with 2 states
  - Rainfall amount R_t conditionally independent in time given the weather type sequence
  - Temperature, solar radiation, wind speed Y_t
    - Residual correlation between successive observations modeled as an AR process with time varying coefficients

$$Y_t = \beta_0^{(S_t)} + \beta_1^{(S_t)} Y_{t-1} + \sigma^{(S_t)} \epsilon_t$$

$(\beta_i^{(s)})$, $(\sigma^{(s)})$ unknown parameters and $\{\epsilon_t\}$ a Gaussian white noise

$$\begin{array}{lccccccccc}
\textit{Rainfall} & \cdots & & R_{t-1} & & R_t & & R_{t+1} & & \cdots \\
 & & & \uparrow & & \uparrow & & \uparrow & & \\
\textit{Weather type} & \cdots & \rightarrow & S_{t-1} & \rightarrow & S_t & \rightarrow & S_{t+1} & \rightarrow & \cdots \\
 & & & \downarrow & & \downarrow & & \downarrow & & \\
\textit{Temperature, wind, ...} & \cdots & \rightarrow & Y_{t-1} & \rightarrow & Y_t & \rightarrow & Y_{t+1} & \rightarrow & \cdots
\end{array}$$

  - Wind dynamics modeled as a mixture of two AR(1) processes
- More sophisticated models for wind time series discussed in the sequel

A MS-AR model for the wind intensity

$$\begin{array}{lcccccccccc} \textit{Weather type} & \cdots & \rightarrow & S_{t-1} & \rightarrow & S_t & \rightarrow & S_{t+1} & \rightarrow & \cdots \\ & & & \downarrow & & \downarrow & & \downarrow & & \\ \textit{Wind speed} & \cdots & \rightarrow & Y_{t-1} & \rightarrow & Y_t & \rightarrow & Y_{t+1} & \rightarrow & \cdots \end{array}$$

- Various strategies available for defining the weather type
  - In Richardson's model $S_t = \{Dry, Wet\}$
    - Probably not the optimal clustering for describing wind conditions!
  - Run clustering algorithm on meteorological variables (e.g. pressure fields,...)
  - Introduce the weather type as a latent (hidden) variable
    - \+ Estimation procedure will find the "optimal" weather type
    - \- Estimation procedure more complicated, simpler models are needed
- Hidden weather type modeled as a first order Markov chain
- Linear Gaussian AR(p) model for the wind speed conditionally to the weather type

$$Y_t = \beta_0^{(S_t)} + \beta_1^{(S_t)} Y_{t-1} + \ldots \beta_p^{(S_t)} Y_{t-p} + \sigma^{(S_t)} \epsilon_t$$

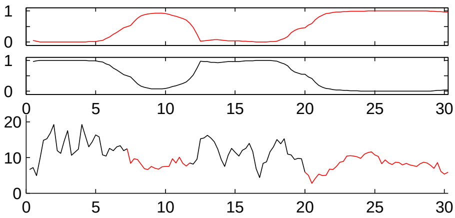
  - $(\beta_i^{(s)})$, $(\sigma^{(s)})$ unknown parameters and $\{\epsilon_t\}$ iid $\mathcal{N}(0,1)$ sequence
- MS-AR: "Markov Switching AutoRegressive" model

- Maximum likelihood estimates (MLE) can be computed with the EM algorithm
- The selected model is interpretable (results for January at Ouessant)

$$Y_t = \begin{cases} 1.46 + 0.79 Y_{t-1} + 1.37\epsilon_t & (S_{t-1} = 1) \\ 2.24 + 0.77 Y_{t-1} + 2.4\epsilon_t & (S_{t-1} = 2) \end{cases} .$$

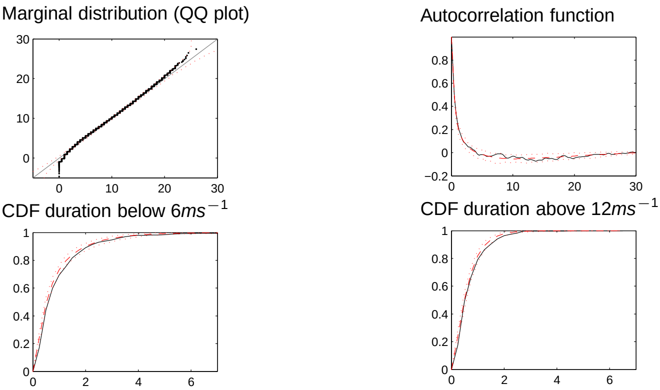
  - **Regime 1** : lower intensity and temporal variability, anticyclonic conditions
  - **Regime 2** : higher intensity and temporal variability, cyclonic conditions
  - Transition matrix : $\begin{bmatrix} 0.98 & 0.02 \\ 0.03 & 0.97 \end{bmatrix}$, stationary distribution $\begin{bmatrix} 0.40 \\ 0.60 \end{bmatrix}$

- Smoothing probabilities $P[S_t | y_1, ..., y_T]$ for Jan. 2000

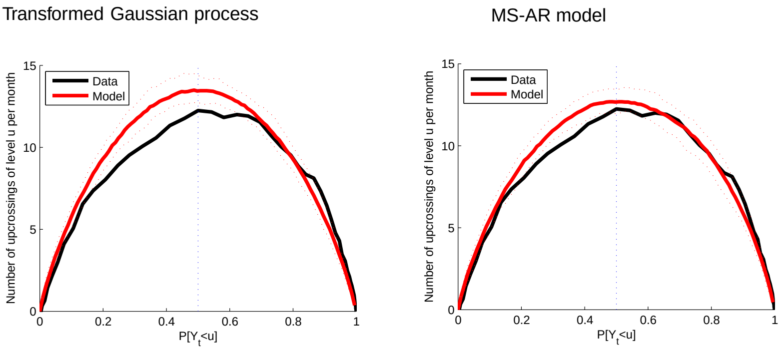


A MS-AR model for the wind intensity

- The model is able to reproduce some important statistical properties of the data



- MS-AR models only partly reproduce the observed asymmetry in up-crossing rate
  - Still too much variability at low level!



- In MS-AR models the regime switchings are independent of past wind conditions
  - The probability of staying in the anticyclonic regime at time t is higher if the wind speed is very low at time $t-1$?
  - Adding this in the model could create more asymmetry?

$$
\begin{array}{lccccccccc}
\textbf{Hidden Regime} & \cdots & \rightarrow & S_{t-1} & \rightarrow & S_t & \rightarrow & S_{t+1} & \rightarrow & \cdots \\
 & & & \downarrow & \nearrow & \downarrow & \nearrow & \downarrow & & \\
\textbf{Observed time series} & \cdots & \rightarrow & Y_{t-1} & \rightarrow & Y_t & \rightarrow & Y_{t+1} & \rightarrow & \cdots
\end{array}
$$

- Logistic link function for the switching probabilities. For $s \in \{1, 2\}$,

$$P(S_t = s | S_{t-1} = s, Y_{t-1} = y_{t-1}) = \pi_-^{(s)} + \frac{1 - \pi_-^{(s)} - \pi_+^{(s)}}{1 + \exp\left(\lambda_0^{(s)} + \lambda_1^{(s)} y_{t-1}\right)}$$

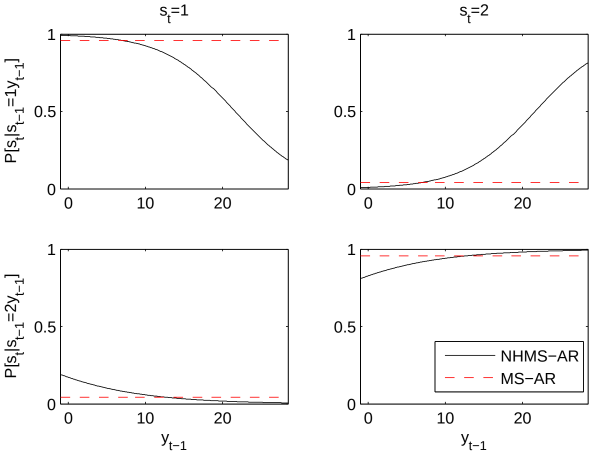
- Linear Gaussian AR(p) model for the wind speed conditionally to the weather type

$$Y_t = \beta_0^{(S_t)} + \beta_1^{(S_t)} Y_{t-1} + \ldots \beta_p^{(S_t)} Y_{t-p} + \sigma^{(S_t)} \epsilon_t$$

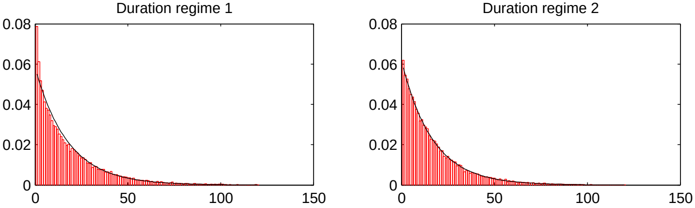
  - $(\beta_i^{(s)})$, $(\sigma^{(s)})$ unknown parameters and (ϵ_t) iid sequence of $\mathcal{N}(0, 1)$ r.v.

A non-homogeneous MS-AR models for the wind intensity

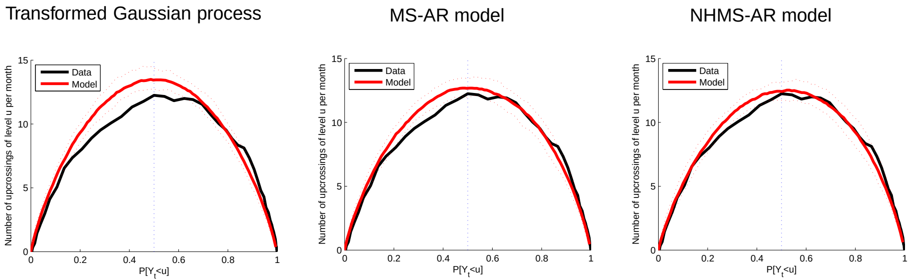
- Model fitted with the EM algorithm
- Similar interpretation for the regimes
  - **Regime 1** : lower intensity and temporal variability, anticyclonic conditions
  - **Regime 2** : higher intensity and temporal variability, cyclonic conditions
- Higher probability of staying in regime 1 if low wind speed



- Durations in the regimes are no more geometric (especially in the first regime)



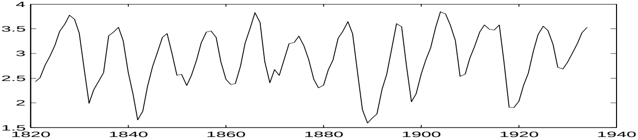
- Create (a little) more asymmetry in the dynamics



A non-homogeneous MS-AR models for the wind intensity

- Similar results for temperature time series in Brest with this model
  - Including non-homogeneous switchings permits a better description of long cold and warm periods
- The model has been generalized to handle
  - Multivariate time series
    - Applied to bivariate wind data ((u, v) components)
  - Non linear autoregressive models
    - e.g. wind direction with von-Mises distribution
  - Other link functions

- Annual number of lynx trapped in the Mackenzie River district (after log transfo.)



- Periodic fluctuations (competition between several species, predator-prey interaction,...)
- Asymmetric cycles (increasing phase are slower than decreasing phase)
- Benchmark for non-linear time series
- SETAR(2) model proposed in Tong (90)

$$Y_t = \begin{cases} 0.51 + 1.23 Y_{t-1} - 0.37 Y_{t-2} + 0.18 \epsilon_t & (Y_{t-2} \leq 3.15) \\ 2.32 + 1.53 Y_{t-1} - 1.27 Y_{t-2} + 0.23 \epsilon_t & (Y_{t-2} > 3.15) \end{cases} .$$

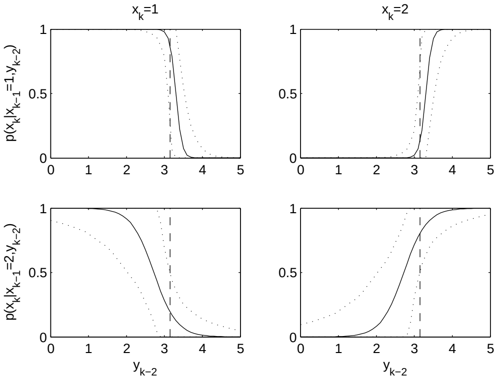
  - First regime describes population increase

- The NHMS-AR model includes the SETAR(2) model as a limit case

$$P(S_t = s | S_{t-1} = s, Y_{t-1} = y_{t-1}) = \pi_-^{(s)} + \frac{1 - \pi_-^{(s)} - \pi_+^{(s)}}{1 + \exp\left(\lambda_0^{(s)} + \lambda_1^{(s)} y_{t-1}\right)}$$

Another example: lynx data

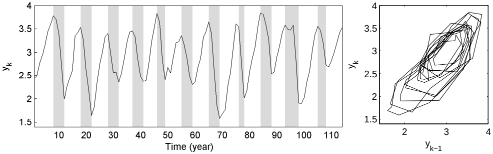
- Fitted NHMS-AR and SETAR(2) models have similar estimates for the AR coefficients
- Transition probabilities



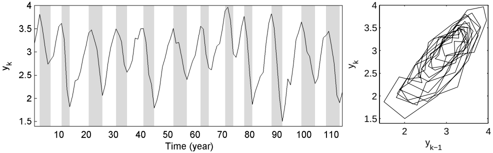
- SETAR(2) model: switchings occur when Y_{t-2} crosses level $u = 3.15$ (dashed vertical line)
- NHMS-AR model: switchings are no more a deterministic function of Y_{t-2}
  - Fuzzy extension of the NHMS-AR model

- The fitted model is able to generate cycles

Data



Simulated sequence



- Consistency of MLEs proven for NHMS-AR models under general conditions
  - Ergodicity of $\{S_t, Y_t\}$
  - Existence of moments for the stationary solution
  - Identifiability
- For the specific model considered previously
  - $Y_t = \beta_0^{(S_t)} + \beta_1^{(S_t)} Y_{t-1} + \sigma^{(S_t)} \epsilon_t$
    - Identifiability constraint: $\left(\beta_0^{(1)}, \beta_1^{(1)}, \sigma^{(1)}\right) \neq \left(\beta_0^{(2)}, \beta_1^{(2)}, \sigma^{(2)}\right)$
  - $S_t \in \{1, 2\}$ and $P(S_t = s | S_{t-1} = s, Y_{t-1} = y_{t-1}) = \pi_-^{(s)} + \frac{1-\pi_-^{(s)}-\pi_+^{(s)}}{1+\exp\left(\lambda_0^{(s)}+\lambda_1^{(s)} y_{t-1}\right)}$
    - Identifiability constraint: $\forall s \in \{1, 2\}, \pi_-^{(s)} = \pi_+^{(s)} = \pi_0$ where $0 < \pi_0 < 1/2$ is a fixed constant
  - Ergodicity and second order moments if $\exists R > 0$ such that $\forall s_0 \in \{1, 2\}, y_0 \in \mathbb{R}^s$

$$\|y_0\| > R \quad \Rightarrow \quad \sum_{s_1=1}^{2} p_{1,\theta}(s_1|s_0, y_0) |\beta_1^{(s_1)}|^2 < 1$$

    - Satisfied in particular if $\forall s \in \{1, 2\}, |\beta_1^{(s)}| < 1$ (the two regimes are stable)
- Under the above assumptions,..., on a set of probability one, the limit values θ of the sequence of random variables $(\hat{\theta}_{n,x_0})_n$ are equal to θ^* up to a permutation of indices.

Non-homogeneous MS-AR models: MLEs are consistent

- Theoretical framework flexible enough to include models with exogenous covariates
- For example non-homogeneous HMMs used for statistical downscaling

$$
\begin{array}{llllllllll}
\textbf{Covariates} & \cdots & \rightarrow & Z_{k-1} & \rightarrow & Z_k & \rightarrow & Z_{k+1} & \rightarrow & \cdots \\
 & & & & \searrow & & \searrow & & \searrow & \\
\textbf{Hidden Regime} & \cdots & \rightarrow & X_{k-1} & \rightarrow & X_k & \rightarrow & X_{k+1} & \rightarrow & \cdots \\
 & & & \downarrow & & \downarrow & & \downarrow & & \\
\textbf{Output time series} & \cdots & & R_{k-1} & & R_k & & R_{k+1} & & \cdots
\end{array}
$$

- Covariates = large scale information
  - Assumed to be an ergodic Markov process
- Output time series = local weather conditions
  - (rainfall,temperature,...) at a meteorological station

Conclusions

- Weather type models provide a flexible and interpretable family of models for meteorological time series
- Models have been developed/validated for generating
  - (rainfall, temperature, solar radiation, humidity, wind speed) simultaneously at a single location
  - Wind speed, wind direction, temperature independently at a single location
  - Rainfall at several locations simultaneously
- More works needed for
  - Developing multi-site and multi-parameters generators
  - Improving some aspects of existing models
    - Interanual variability underestimated ("overdispersion" phenomenon)
    - Probability of long "events" (dry spell, heat wave,...) generally underestimated
- R package under development
- References

  - Ailliot P., Monbet V., (2012), Markov-switching autoregressive models for wind time series. Environmental Modelling & Software, 30, pp 92-101.
  - Ailliot P., Pène F. (2013), Consistency of the maximum likelihood estimate for Non-homogeneous Markov-switching models. Submitted.