Estimation of Locally Stationary Processes and its Application to Climate Modeling

Shreyan Ganguly, Department of Statistics, The Ohio State University

Peter F. Craigmile, Department of Statistics, The Ohio State University

ABSTRACT

In the analysis of climate, it is common to build non-stationary spatio-temporal processes, based on assuming a random walk behavior over time for the error process. Random walk models may be a poor description for the temporal dynamics, leading to inaccurate uncertainty quantification. Likewise, assuming stationarity in time may also not be a reasonable assumption, especially under climate change. In our ongoing research, we present a class of time-varying autoregressive processes that are stationary in space, but locally stationary in time. We demonstrate how to parameterize the time-varying model parameters in terms of a transformation of basis functions. We present some properties of parameter estimates when the process is observed at a finite collection of spatial locations, and apply our methodology to a spatio-temporal analysis of temperature.

BACKGROUND

Stationarity has always played a major role in the modeling of time series. But the weak stationarity assumption is not always favorable while modeling spatio-temporal data, even after detrending and deseasonalization. Locally stationary processes capture non-stationarity of the data in the case where there is a gradual change in the stochastic properties of the process. In this situation the classical characteristics of the process, such as the covariance function at some lag k, the spectral density at some frequency λ, of the parameters for an AR(p) process are curves which change slowly over time. The idea of having a local approximation of a stationary process was first suggested by Priestley (1965). Dahlhaus (1997) introduced the class of locally stationary processes having a time varying spectral representation or, alternatively, an infinite time-varying moving average representation. The theory of locally stationary processes allows for rigorous inference with time series and spatio-temporal data.

MOTIVATION AND OBJECTIVE

Modelling temperatures has become imperative with the advent of Global warming. Temperature data has inherent non-stationary elements present in the form for trend and seasonality. But modeling the residual series even after removing these can be challenging as the autoregressive parameters might exhibit a time varying nature. Assuming stationarity of the process in time and fixed autoregressive parameters, for example, may lead to inaccurate predictions and invalid uncertainty quantification. Moreover, a good model also takes into account the spatial correlations between the different locations in the region of interest. Our proposed model takes into account both these aspects. Moreover, the parameterization of the time varying autoregressive parameters by taking a non-linear transformation of a basis function is done to ensure local stationarity of the process. We run a simulation study to verify that our model indeed belongs to the class of model described in Dahlhaus (2000). Finally, we use our model to estimate the AR parameters from the detrended and deseasonalized GHCNM temperature data of 31 different sites in and around the state of Ohio, USA from 1905 to 2004. The estimates are obtained in a Bayesian setting and compared to the site by site estimates of the univariate TVAR(1) parameter curves of all 31 sites.

MODEL

Let $\boldsymbol{Y}_{t,T} = (Y_{t,T}(s_1), Y_{t,T}(s_2), \cdots, Y_{t,T}(s_d))'$ be a vector of measurements for d locations. Our model for $\boldsymbol{Y}_{t,T}$ is

$$\boldsymbol{Y}_{t,T} - \boldsymbol{\mu}\left(\frac{t}{T}\right) = \boldsymbol{\Phi}\left(\frac{t}{T}\right)\left[\boldsymbol{Y}_{t-1,T} - \boldsymbol{\mu}\left(\frac{t-1}{T}\right)\right] + \boldsymbol{\epsilon}_{t,T}, \quad t = 1, \cdots, T. \quad (1)$$

The mean vector is $\boldsymbol{\mu}\left(\frac{t}{T}\right) = \boldsymbol{Z}(t)_{d\times p} \cdot (\gamma(1), \cdots, \gamma(p))'$ where $\boldsymbol{Z}(t)$ is a matrix containing p dimensional basis vectors for all d locations.

The error process $\boldsymbol{\epsilon}_{t,T} = (\epsilon_{t,T}(s_1), \cdots, \epsilon_{t,T}(s_d))'$ is mean zero with variance $\boldsymbol{\Sigma}^{\epsilon}_{t,t',T}(s,s')$ where

$$\boldsymbol{\Sigma}^{\epsilon}_{t,t',T}(s,s') = (\sigma^{\epsilon}_{t,t',T}(s,s')) = \begin{cases} 0 \text{ if } t \neq t' \\ \sigma^2 \exp\left\{-\frac{|s-s'|}{\tau}\right\} \text{ if } t = t'. \end{cases} \quad s, s' \in D.$$

where D is the domain of interest in $\mathbb{R}^2$.

The spatially varying AR parameters $\boldsymbol{\Phi}(u) = \text{diag}\left[\phi(s_1,u), \cdots, \phi(s_d,u)\right]$, are such that

$$g(\phi(s,u)) = \text{logit}\left(\frac{\phi(s,u)+1}{2}\right) = \eta(s,u) = \boldsymbol{X}(u)'_{1\times p} \cdot \boldsymbol{\beta}(s)_{p\times 1} \quad (2)$$

for $u = t/T \in (0,1]$. (2) ensures that the process $\boldsymbol{Y}_{t,T}$ is locally stationary of the form

$$\boldsymbol{Y}_{t,T} = \boldsymbol{\mu}\left(\frac{t}{T}\right) + \int_{-\pi}^{\pi} \exp(i\lambda t) A^0_{t,T}(s,\lambda) d\xi(s,\lambda)$$

where $A^0_{t,T}(s,\lambda)$ is a spatially varying transfer function. As we have

$$\det(\boldsymbol{\Phi}(u,z)) = \det(I_d - \boldsymbol{\Phi}(u)\cdot z) \neq 0, \quad \forall |z| \leq 1 + c \text{ with } c > 0, z \in \mathbb{C},$$

our model is locally stationary in time.

The time varying spectral density of the process is

$$f_{s,s'}(u,\lambda) = \frac{1}{2\pi}\boldsymbol{\Phi}(u, \exp(i\lambda))^{-1}\Sigma^{\epsilon}_{t,t',T}(s,s')\left[\boldsymbol{\Phi}(u,\exp(-i\lambda))'\right]^{-1}$$

$$= \frac{\sigma^2}{2\pi}\frac{\exp\left\{-\frac{|s-s'|}{\tau}\right\}}{(1-\phi(s,u)\exp(i\lambda))\cdot(1-\phi(s',u)\exp(-i\lambda))}.$$

ASYMPTOTIC RESULTS

Under certain assumptions (Dahlhaus 2000), the asymptotic distribution of the MLE, $\hat{\theta}_T$, of the parameter vector θ, minimizing the conditional log-likelihood

$$\mathcal{L}(\theta) = \prod_{t=1}^{T} f_{\boldsymbol{Y}_{t,T}|\boldsymbol{Y}_{t-1,T}}(\boldsymbol{Y}_{t,T}|\boldsymbol{Y}_{t-1,T}, \theta) \quad \text{where}$$

$$f_{\boldsymbol{Y}_{t,T}|\boldsymbol{Y}_{t-1,T}}(\boldsymbol{Y}_{t,T}|\boldsymbol{Y}_{t-1,T},\theta) \sim \mathcal{N}\left[\boldsymbol{\mu}\left(\frac{t}{T}\right) + \boldsymbol{\Phi}\left(\frac{t}{T}\right)\left[\boldsymbol{Y}_{t-1,T} - \boldsymbol{\mu}\left(\frac{t-1}{T}\right)\right], \boldsymbol{\Sigma}^{\epsilon}_{t,T}(s)\right]$$

is given by

$$\sqrt{T}(\hat{\theta}_T - \theta_0) \Rightarrow \mathcal{N}(0, \Gamma^{-1}V\Gamma^{-1}), \quad (3)$$

as $T \to \infty$, where θ_0 is the true value of $\theta \in \Theta \subset \mathbb{R}^k, \Theta$ compact,

$$\Gamma_{ij} = \frac{1}{4\pi}\int_0^1\int_{-\pi}^{\pi} Tr\{(f - f_{\theta_0})\nabla_{ij} f^{-1}_{\theta_0}\} d\lambda du$$
$$-\frac{1}{4\pi}\int_0^1\int_{-\pi}^{\pi} Tr\{(\nabla_i f_{\theta_0})(\nabla_j f^{-1}_{\theta_0})\} d\lambda du$$
$$+\frac{1}{4\pi}\int_0^1 \nabla^2_{ij}\left\{(\mu(u) - \mu_{\theta_0}(u))' f^{-1}_{\theta_0}(u,0)(\mu(u) - \mu_{\theta_0}(u))\right\} du,$$

and

$$V_{ij} = \frac{1}{4\pi}\int_0^1\int_{-\pi}^{\pi} Tr\{f(\nabla_i f^{-1}_{\theta}) f(\nabla_j f^{-1}_{\theta})\} d\lambda du$$
$$+\frac{1}{2\pi}\int_0^1\int_{-\pi}^{\pi}\left[\nabla_i\{(\mu(u) - \mu_{\theta_0}(u))' f^{-1}_{\theta_0}(u,0)\}\right] f(u,0)$$
$$\times\left[\nabla_j\{f^{-1}_{\theta_0}(u,0)(\mu(u) - \mu_{\theta_0}(u))\}\right] du.$$

When the model is corrctly specified, $f_{\theta} = f_{\theta_0}, \mu = \mu_{\theta_0}, \Gamma = V$ with

$$\Gamma_{ij} = \frac{1}{4\pi}\int_0^1\int_{-\pi}^{\pi} Tr\{f_{\theta_0}(\nabla_i f^{-1}_{\theta_0}) f_{\theta_0}(\nabla_j f^{-1}_{\theta_0})\} d\lambda du$$
$$+\frac{1}{2\pi}\int_0^1 (\nabla_i \mu_{\theta_0}(u))' f^{-1}_{\theta_0}(u,0)(\nabla_j \mu_{\theta_0}(u)) du.$$

SIMULATION

For a practical validation of the fact that the asymptotic distribution of the MLE of the parameter vector θ looks like (3), we carry out a simulation study.

We know, for any estimate $\hat{\theta}_{jT}$ of $\theta_j \in \theta$, $MSE(\hat{\theta}_{jT}) = BIAS(\hat{\theta}_{jT})^2 + Var(\hat{\theta}_{jT})$. Under the correct model specifications, as $T \to \infty$, $BIAS(\hat{\theta}_{jT}) \to 0$ and thus $MSE(\hat{\theta}_{jT}) \to Var(\hat{\theta}_{jT})$. Thus, the MSE's are calculated for each of the parameters in $\theta = (\gamma, \beta(s_1), \cdots, \beta(s_d), \sigma^2, \tau)$ for varying time points and compared to the respective asymptotic variances.

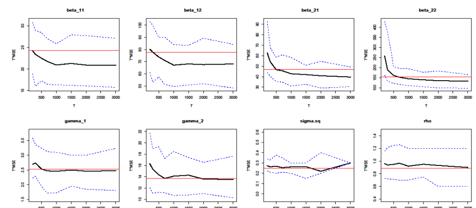
The simulation is done for a two dimensional spatio-temporal model with varying lengths of time T = 200, …, 3000 with the true parameter vector

$\theta_0 = (\gamma_0, \beta_0(s_1), \beta_0(s_2), \sigma_0^2, \rho_0 = e^{\{-1/\tau_0\}})'$ where $\gamma_0 = (1,2)', \beta_0(s_1) = (1, 0.5)', \ \beta_0(s_2) = (2, 0.5)', \ \sigma_0^2 = 0.5$ and $\rho_0 = 0.2$.

The process vector $\boldsymbol{Y}$ is generated 500 times from the true distribution and for each such $\boldsymbol{Y}$, the parameters in θ are estimated. An approximate value of the MSE is calculated as

$$MSE(\hat{\theta}_{jT}) \approx \frac{1}{500}\sum_{i=1}^{500}(\hat{\theta}_{jT_{(i)}} - \theta_{j0})^2, \quad \theta_{j0} \in \theta_0$$

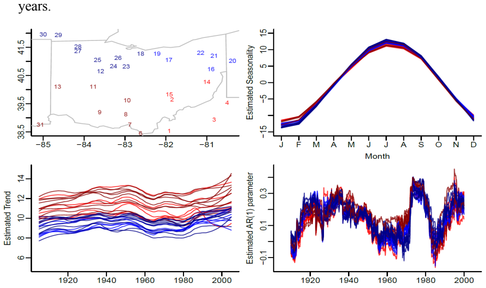
Finally, the above experiment is replicated 100 times and 95% bootstrap intervals for the MSEs are calculated as a measure of uncertainty.



APPLICATION

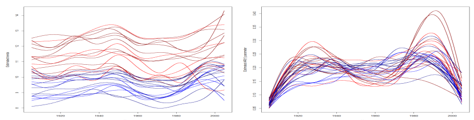
Data description: The Global Historical Climatology Network-Monthly (GHCN-M) dataset, released in 1992, has been an internationally recognized source of data for the study of observed variability and change in land surface temperature. It provides monthly mean temperature data for 7280 stations from 226 countries and territories, ongoing monthly updates of more than 2000 stations to support monitoring of current and evolving climate conditions, and homogeneity adjustments to remove non-climatic influences that can bias the observed temperature record (Peterson and Vose 1997). We pick a subset of 31 locations in and around the state of Ohio, USA and model the behavior of the mean temperature over a period of 100 years, from 1905 to 2004.

The sites are colored according to regions of warmer and cooler temperatures. Preliminary EDA shows the naïve trend and seasonality estimates below along with the windowed estimates for the TVAR(1) coefficients for all 31 sites after detrending and deseasonalizing. The trend estimates are obtained by fitting a local polynomial regression model and the seasonal estimates are from an ANOVA model fit. The AR(1) parameters are lag-window estimates of window length 10 years.



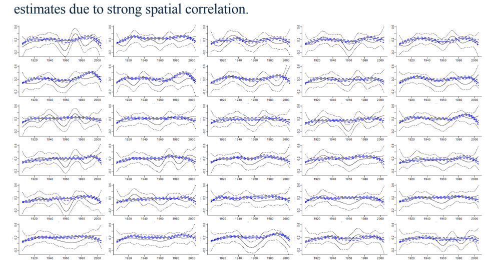
RESULTS

We now fit our model to the deseasonalized data for each site and estimate the trend components and the time varying AR parameters. A spline basis with nodes every 10 years (approximately) is used for both the estimation of the trend and the estimation of the time-varying AR parameters. Firstly, estimates of the time varying AR parameters for each of the 31 sites are obtained by minimizing the univariate conditional log-likelihood corresponding to each site. These estimates are then used as starting values in a Bayesian model to estimate the time varying AR parameters.



It is evident that the parametric model smooths out the heavy fluctuations of the AR parameters seen in the windowed estimates. Overall, the trend and the AR estimates coincide with the naïve estimates. The AR parameters for sites in southern Ohio vary more than the regions in northern Ohio indicating higher temperature fluctuations.

Below , we see a comparison of the uncertainty around the Bayes estimates for the $\phi(s,u)$'s, $u \in (0,1]$, for each $s = s_1, \cdots, s_{30}$ in our model with the uncertainty around the univariate site by site estimates of $\phi(s,u)$'s . As our model learns from both space and time, it has a much narrower confidence bound around the estimates which suggests better predictability, but is much smoother than the naïve estimates due to strong spatial correlation.



FUTURE WORK

- Improve the Bayesian hierarchical model to make spatial predictions and temporal forecasts.
- Extend the model to a general spatio-temporal model, formulating the corresponding theory.
- Compare with other contemporary models to see the competence of the model in terms of uncertainty quantification.

REFERENCES

Priestley, Maurice B. (1965), Evolutionary spectra and non-stationary processes, *Journal of the Royal Statistical Society, Series B (Methodological), 27,* 204-237.

Dahlhaus, Rainer. (1997), Fitting time series models to nonstationary processes, *The Annals of Statistics*, **25**, 1-37.

Dahlhaus, R. (1996), Maximum likelihood estimation and model selection for locally stationary processes, *Journal of Nonparametric Statistics*, **6**, 171-191.

Dahlhaus, Rainer. (2000), A likelihood approximation for locally stationary processes, *The Annals of Statistics*, **28**, 1762-1794.

Peterson, Thomas C., and Russell S. Vose. (1997), An overview of the Global Historical Climatology Network temperature database, *Bulletin of the American Meteorological Society*, **78**, 2837-2849.

Data source: https://www.ncdc.noaa.gov/ghcnm/v3.php