A re-analysis of North Atlantic hurricane activity with dynamic latent processes

Mathieu Boudreault[1], Jean-Philippe Boucher[1], Louis-Philippe Caron[2,3]
[1] Université du Québec à Montréal (UQÀM), Montréal, Canada
[2] Catalan Institute of Climate Sciences (IC3), Barcelona, Spain
[2] Swedish Meteorological and Hydrological Institute (SMHI), Nörrkoping, Sweden
Contact: boudreault.mathieu@uqam.ca

Abstract

We analyze the number of Atlantic tropical cyclones using a type of dynamic latent process that generalizes the Poisson regression. It is assumed that the (log) conditional mean of a Poisson random variable is influenced by a set of covariates as well as a latent autoregressive process with Gaussian noise. The autoregressive part of the model automatically accounts for unobserved and unexplained long-term trends, while the noise part of the model allows for additional randomness in the conditional mean. The model is estimated using different tropical cyclone databases (e.g. with/without short-duration hurricanes, with/without likely missed storms), and various climate indices. After removing short-duration storms and adding likely missed storms, the evidence for a long-term trend is very weak. Furthermore, when one or more covariate is used, such as the Atlantic Multidecadal Oscillation (AMO), the long-term trend completely vanishes and becomes statistically non-significant, in accordance with previous studies. However, in all cases, the evidence for additional noise in the conditional mean of the Poisson regression is strong. This indicates that the uncertainty in the predictions of the number of tropical cyclones based on a Poisson regression might be underestimated.

1. Introduction

The occurrence of North Atlantic tropical cyclones varies on many timescales, from intraseasonal to multidecadal. Multiple studies have associated these variations to different climate indices, such as ENSO, the West African Monsoon and the AMO, to name a few. Exactly which of those climate indices are the most significant and which should be included in the modeling of tropical cyclone counts remain unclear. Investigation into this issue is complicated by the lack of reliable cyclone data in the pre-satellite era and by possible non-stationary relationships between the indices and the cyclone count.

Furthermore, some authors have argued that North Atlantic cyclone numbers are sensitive to changes in aerosols (natural or anthropogenic) and greenhouse gases concentration. Whether or not we are already experiencing a detectable change in tropical cyclone frequency due to the latter is still (hotly) debated. In fact, even whether or not the number of cyclones has increased over the last 100 years or so is still unclear and depends (crucially) on how many storms were missed by the observational network in place before the satellite era.

Here, we use a statistical approach to investigate the relationship between various climate indices and North Atlantic tropical cyclones. We expand on previous work done by Villarini et al. (2010), Elsner and Jagger (2008) and Sabbatelli and Mann (2007), amongst others, by expanding the range of climate indices and by using a generalization of the Poisson regression, which accounts for the presence of noise and autoregression.

2. Climate Indices

The different climate indices used as predictors to model the North Atlantic cyclone count are

1. Atlantic Multi-decadal Oscillation (AMO)
2. Atlantic Meridional Mode (AMM)
3. Mean ASO Main Development Region (MDR) sea surface temperature
4. Mean ASO global tropical sea surface temperature
5. Southern Oscillation Index (SOI)
6. Mean JJAS rainfall over the West Sahel (a measure of the intensity of the Sahel monsoon)
7. Mean ASO North Atlantic Oscillation index (NAO)
8. Sunspot numbers (a measure of solar activity)

3. Tropical Cyclone Timeseries

The relationship between climate indices and North Atlantic tropical cyclones is investigated using a variety of timeseries (1878-2012):

1. Total number of tropical cyclones (based on HURDAT)
2. Number of cyclones adjusted to account for possible missing storms (Vecchi et al., 2008)
3. Total number of cyclones that survived >48h (based on HURDAT)
4. Number of cyclones >48h, adjusted for possible missing storms (Landsea et al. (2009)
5. Total number of hurricanes (based on HURDAT)
6. Number of hurricanes, adjusted for possible missing storms (Vecchi et al., 2011)
7. Total number of US landfalling hurricanes

4. Dynamical latent processes

A natural distribution to count events such as cyclones is the Poisson distribution, defined such that the probability of observing n events at time t is

$$\Pr(N_t = n) = \frac{\lambda^n e^{-\lambda}}{n!}$$ where λ is the mean number of events (in this case, constant over time).

However, we can expect that for some periods, the mean number of events can change (such that during a +/- phase of the AMO or the different phases of ENSO). By modifying the original Poisson distribution such that the mean λ becomes a function of time-varying covariates, we obtain a more appropriate (realistic) relationship, a Poisson regression:

$$\Pr(N_t = n) = \frac{\lambda_t^n e^{-\lambda_t}}{n!}$$ where $\lambda_t = \exp(\beta_0 + \beta_1 X_{1,t} + \beta_2 X_{2,t} + \ldots + \beta_p X_{p,t})$ and X_i is a time-varying covariate.

It is likely that λ_t has an additional source of noise or randomness that needs to be accounted for. Moreover, given the nature of the data, there might be some autoregression (autocorrelation) in the data that should also be accounted for. We thus generalized the Poisson regression such that it is more robust to the presence of this noise and autoregression. This is done by modifying λ_t such that

$$\lambda_t = \theta_t \times \exp(\beta_0 + \beta_1 X_{1,t} + \beta_2 X_{2,t} + \ldots + \beta_p X_{p,t})$$

θ_t is a deflator/inflator defined as $\theta_t = \exp(x_t)$

x_t is simply a latent autoregressive model of order 1 (AR(1)) with a zero mean: $x_t = \phi x_{t-1} + \nu \varepsilon_t$

ϕ indicates the level of autoregression (of dependence upon past obs. of the latent process).

ν indicates the level of noise or of additional randomness in the conditional mean of the regression.

ε_t is a standard normal random variable.

This model is also known as a Poisson model with a lognormal autoregressive latent process.

When $\phi = \nu = 0$, we recover the standard Poisson regression (no noise, no autoregression).

At best, we capture an additional noise and/or autoregression that can help improve the predictions. At worst, we recover the results of a standard Poisson regression.

5. Comparing Poisson with dynamic latent process

We start by comparing the dynamic latent process with the Poisson model (without regressors), i.e.

$\lambda_t = \lambda$ vs $\lambda_t = \theta_t \times \exp(\beta_0)$ for each of the 7 timeseries introduced in **3**.

For 6 out of 7 datasets, we find significant autoregression and/or noise (through a likelihood ratio-test). For number of U.S. landfalling hurricanes, pure Poisson model is adequate; other timeseries are better represented with dynamic latent processes.

Noise is statistically significant for most datasets (based on a t-test).

6. Trend in timeseries

Is there a significant trend in the different timeseries presented in **3**?

Use hockey stick function: 0 until time y; positive slope afterwards.

Rationale: hockey stick function might capture an increase in the number of observed storms due to change in observing system (e.g. satellite coverage), climate change, etc.

Hockey stick model: $\lambda_t = \theta_t \times \exp(\beta_0 + \beta_1 \times \max(t - y, 0))$ where y=1960, 1970, 1980, 1990

Significance of hockey stick function (β_1) returns information on significance of trend, from time y onward.

Upward trend is statistically significant for each cutoff year for 6 out of 8 timeseries.
Exceptions: Number of US landfalling hurricanes; number of hurricanes adjusted for missing storms.

Autoregression and noise still look important, thus dynamic latent process may be more appropriate than Poisson regression.

7. Regression onto a single climate index

Can the different cyclone timeseries be somewhat related to the different climate indices?

49 regressions conducted with both approaches: dynamic latent process vs Poisson regression.

Model: $\lambda_t = \theta_t \times \exp(\beta_0 + \beta_1 \times X_t)$, where X_t is the time evolution of one climate index.

AMM, AMO, MDR SST, West Sahel precipitation: Statistically significant positive for all timeseries, except for US landfalling hurricanes.

NAO: Statistically significant (negative) for all timeseries

SOI: Weakly positive, but statistically significant for all timeseries.

Sunspot numbers: Negative, but not statistically significant for all timeseries.

Very similar findings with a Poisson regression, but autoregression and/or noise still significant for about half of the regressions.

8. Trend and climate indices

Is there an additional upward trend, even if we include a climate index?

49 regressions conducted with both approaches: dynamic latent process vs Poisson regression.

Model: $\lambda_t = \theta_t \times \exp(\beta_0 + \beta_1 \times X_t + \beta_2 \times \max(t - 1970, 0))$

We could have used any pivotal year, but 1970 was chosen based on results obtained in **6**.

Check for significance of β_2.

AMM and AMO: trend is significant for total number of cyclones and total number of cyclones after accounting for missing storms.

- Suggest that AMO/AMM + short lived storms explain all increase in number of storms since 1970

No trend in US landfalling hurricanes.

9. Summary, conclusion and future work

We performed an extensive analysis of North Atlantic tropical cyclone timeseries using a generalization of the Poisson regression (dynamical latent process) and a series of climate indices.

Model is robust to the presence of noise and autoregression.

Most significant climate indices: AMM, AMO, MDR SST, NAO, West Sahel precipitation (not all independent)

Weaker, but significant contribution from SOI.

Once we account for AMO/AMM, no additional upward trend (post-1970) for data that exclude short-duration tropical cyclones.

Total US landfalling hurricanes: simple Poisson model sufficient. No upward trend, and only NAO required to explain that timeseries.

Future work: combination of many predictors; impact on prediction intervals: Poisson regression vs dynamic latent process; small sample standard errors (simulation) vs asymptotic (this presentation).