Change Point and Trend Analyses of Annual Expectile Curves of Tropical Storms

Petra Burdejová
Wolfgang Karl Härdle
Piotr Kokoszka
Qian Xiong

Humboldt–Universität zu Berlin
Colorado State University
http://lvb.wiwi.hu-berlin.de
http://www.stat.colostate.edu

Are we getting stronger extremes?

NATIONAL GEOGRAPHIC Daily News

Home Animals Ancient Energy Environment Travel/Cultures Space/Tech Water Weird News Photos News Video News Blogs

Will U.S. Hurricane Forecasting Models Catch Up to Europe's?

A year after Hurricane Sandy, Europe's forecasting technology is still tops.

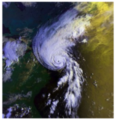
Forbes

INVESTING 4/03/2012 @ 3:50PM | 3,127 views

Damaging Hurricanes Could Impact Energy, Insurance Industries In 2012

Comment Now

by Alan Lammey, Joe Bastardi, Joe D'Aleo, Michael Barak



This image shows Hurricane Bob approaching New England on August 19 at 1226 UTC. This image was produced from data from NOAA-10, provided by NOAA. (Photo credit: Wikipedia)

The 2012 Atlantic Hurricane Season is likely to see an overall decrease in tropical activity as compared to 2011, but with the focus of tropical development closer to the United States.

When evaluating the hurricane season as a whole, it is important to consider more data than just the number of storms that are named. Evaluating in multiple ways. The best way to determine total seasonal activity is by collectively measuring the intensity and duration of named tropical cyclones (both tropical storms and hurricanes), also called the ACE index (Accumulated Cyclone Energy). While the National Hurricane

Bloomberg

By Noah Buhayar - Nov 11, 2013

Typhoon Worse for Philippines Economy Than Sandy for U.S.

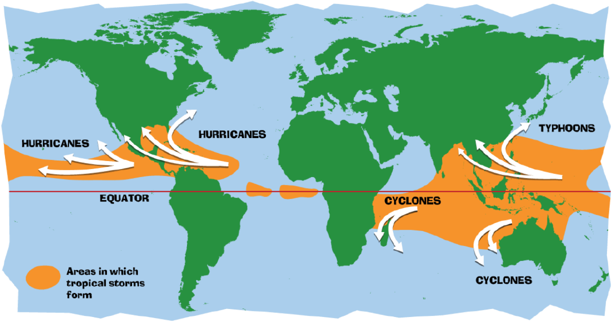
Data

Hurricanes

- North Atlantic
- Years 1851 - 2011

Typhoons

- West Pacific
- Years 1946 - 2010



Data: http://weather.unisys.com Figure: http://www.learnnc.org

Simple trend analysis

- frequency but not intensity
- only values per year
- impact for insurance (Alba & Andrade 2009)

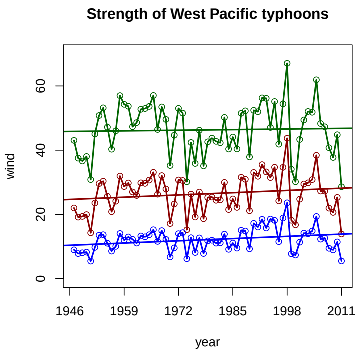


Figure: Annual expectiles for $\tau = 0.25, 0.5, 0.75$ and trend

Complex trend analyses

- strength of wind
- observation every 6 hrs
- use functional data analyses
- trend for every τ-level separately

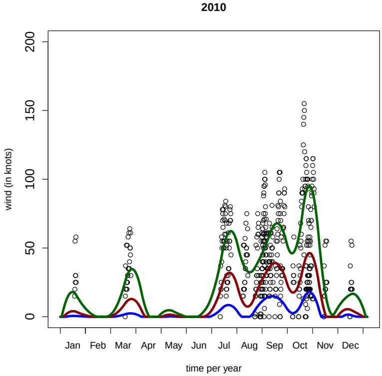


Figure: Typhoon data for 2010 with annual TEC for $\tau = 0.25, 0.5, 0.75$

Challenges

- Reduce dimension
- Retain detail information
- Tail Event Curves (TEC)
- Dynamic (Time-varying)

- Change point testing
- Trend detection

Outline

1. Motivation ✓
2. Quantiles and Expectiles
3. Change point and test trend
4. Conclusion for tropical storms

Quantiles and Expectiles

For r.v. Y obtain tail event measure:

$$q^{\tau} = \arg\min_{\theta} \mathsf{E}\left\{\rho_{\tau}\left(Y-\theta\right)\right\}$$

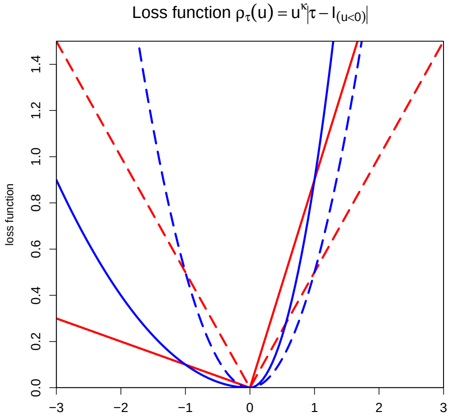
asymmetric loss function

$$\rho_{\tau}\left(u\right) = \left|u\right|^{\alpha}\left|\tau - \mathbf{I}_{\{u<0\}}\right|$$

$\alpha = 1$ for quantiles, $\alpha = 2$ for expectiles

Expectile as quantile

Quantiles and Expectiles



LQRcheck

Figure: Loss function of expectiles and quantiles
for $\tau = 0.5$ (dashed) and $\tau = 0.9$ (solid)

Expectile Curves

Generalized regression τ-expectile

$$e^{\tau}(x) = \arg\min_{\theta} \mathsf{E}\left\{\rho_{\tau}\left(Y - \theta\right) \mid X = x\right\}$$

Expectile $e^{\tau}(x)$ approximated by B-spline basis Definition of B-splines

Penalized splines:

$$\widehat{\alpha} = \arg\min_{\alpha} \sum_{i=1}^{n} \rho_{\tau}\left\{y_i - \alpha^{\top} b(x_i)\right\} + \lambda \alpha^{\top} \Omega \alpha$$

with penalization matrix Ω and shrinkage λ Penalization matrix

Estimation of Expectile Curves

Schnabel and Eilers (2009): iterative LAWS algorithm ▸ LAWS

Schnabel (2011): expectile sheets for joint estimation of curves

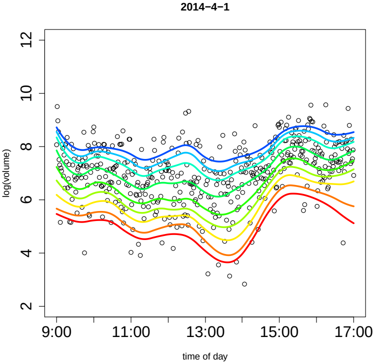


Figure: BAYER trading volume expectiles ($\lambda = 0.5$) on 2014-04-01

Expectiles Curves for Typhoons

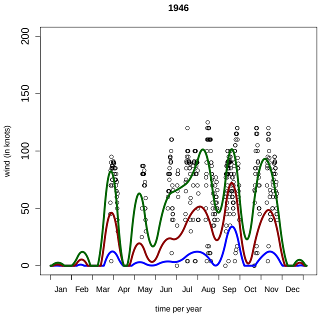


Figure: Every year with expectile curves for $\tau = 0.1, 0.5$ and 0.9.

data_load_hurricanes

Expectiles Curves for Hurricanes

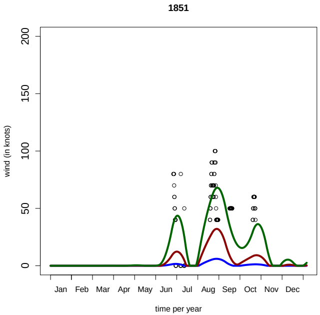


Figure: Every year with expectile curves for $\tau = 0.1, 0.5$ and 0.9.

data_load_hurricanes

Typhoon's curves differences

$e_n^\tau(t)$ — τ-expectile curve for year n

$$\hat{\mu}_k^\tau(t) = \frac{1}{k} \sum_{n=1}^{k} e_n^\tau(t)$$

$$\tilde{\mu}_k^\tau(t) = \frac{1}{N-k} \sum_{n=k+1}^{N} e_n^\tau(t)$$

Normalized differences:

$$P_{\tau,k}(t) = \frac{k(N-k)}{N} \left\{ \hat{\mu}_k^\tau(t) - \tilde{\mu}_k^\tau(t) \right\}$$

$$P_{\tau,k} = \int_0^T P_{\tau,k}^2(t) dt$$

Typhoon's curves differences

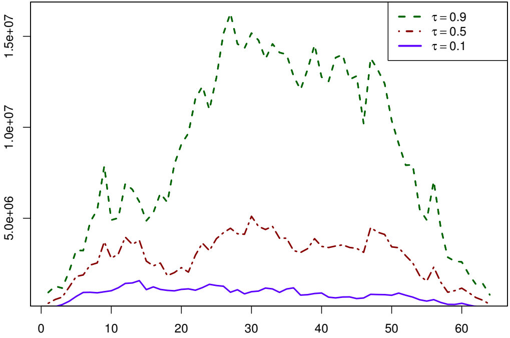


Figure: The squared norms $P_{\tau,k}$ showing the magnitude of change in mean annual pattern for expectile curves of typhoons

P_beta_est

Change point test

Compute func.principal components $v(t)$ and scores

$$\hat{\xi}_{l,n} = \int_0^T \left\{ X_n(t) - \bar{X}_N(t) \right\} \hat{v}_l(t) dt$$

Test for different τ

$$H_0: \text{ no change in the mean function}$$

Test statistics:

$$\hat{S}_d = \frac{1}{N^2} \sum_{l=1}^{d} \frac{1}{\hat{\lambda}_l} \sum_{k=1}^{N} \left(\sum_{1 \le i \le k} \hat{\xi}_{l,i} - \frac{k}{N} \sum_{1 \le i \le k} \hat{\xi}_{l,i} \right)$$

take first d eigenvalues λ_l explaining large proportion of variance

Change point test

d	10	11	12
10%	2.2886	2.4966	2.6862
5%	2.5268	2.7444	2.9490
1%	3.0339	3.2680	3.4911

Table 1: Critical values of the distribution of $\hat{S}_d$

τ	0.1	0.5	0.9
d	10 12	12	
S_d	3.352	3.656	4.508
	***	***	***

Table 2: Values of test statistics S_d for typhoone expectile curve

Test of trend

$$e_n^\tau(t) = \alpha_\tau(t) + n\beta_\tau(t) + \varepsilon_\tau(t)$$

Test for different (fixed) τ

$$H_0 : \beta(t) = 0$$

LS estimator

$$\hat{\beta}(t) = \frac{6}{N(N+1)(N-1)} \sum_{k=1}^{N} (2k - N - 1) e_k^\tau(t).$$

Trend test

Residuals

$$\hat{\varepsilon}_n(t) = X_n(t) - \hat{\alpha}_n(t) - \hat{\beta}_n(t)n,$$

where

$$\hat{\alpha}(t) = \frac{2}{N(N-1)} \sum_{k=1}^{N} (2N + 1 - 3k) X_k(t).$$

Denote $\hat{\lambda}_j$ the eigenvalues of the empirical covariance function

$$\hat{c}(t,s) = \frac{1}{N} \sum_{n=1}^{N} \hat{\varepsilon}_n(t) \hat{\varepsilon}_n(s)$$

Trend test I. (Monte Carlo)

Theorem

Under H_0,

$$\widehat{\Lambda}_N = \frac{N^3}{12} \int_0^1 \left\{ \hat{\beta}(t) \right\}^2 dt \xrightarrow{\mathcal{L}} \Lambda_\infty \stackrel{def}{=} \sum_{j=1}^{\infty} \lambda_j Z_j^2,$$

where $Z_j, j \geq 1$ are normal i.i.d. r.v.

Trend test II. (χ^2-test)

Theorem

Suppose $\mathsf{E}\,\|\varepsilon\|^4 < \infty$ *and* $\lambda_1 > \lambda_2 > \ldots > \lambda_q > \lambda_{q+1} > 0$

Under H_0*,*

$$\widehat{T}_N = \frac{N^3}{12} \sum_{j=1}^{q} \hat{\lambda}_j^{-1} \langle \hat{\beta}, \hat{v}_j \rangle^2 \xrightarrow{\mathcal{L}} \chi_q^2.$$

Finite sample performance

Data generating process for error function

$$B_n(t) = \sqrt{2}\sum_{j=1}^{\infty} Z_{nj}\frac{\sin(j\pi t)}{j\pi} \approx \sqrt{2}\sum_{j=1}^{J} Z_{nj}\frac{\sin(j\pi t)}{j\pi},$$

or

$$\varepsilon(t;\tau) = \sum_{j=1}^{q} \sigma_j(\tau) Z_j \hat{v}_j(t;\tau), \qquad Z_j \sim \text{ iid } (0,1),$$

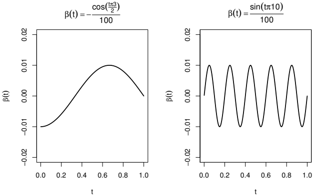
with $\hat{v}_j(t;\tau)$ principal components of expectile curves and $\sigma_j(\tau)$ s.d. of scores

Finite sample performance

$$X_n(t) = \beta(t)n + \varepsilon_n(t).$$

$\tau = 0.1, 0.5$ and 0.9

$$\beta_0(t) = 0; \qquad \beta_1(t) = -\frac{\cos\left(\frac{t\pi 3}{2}\right)}{100}; \qquad \beta_2(t) = \frac{\sin\left(t\pi 20\right)}{100}$$



Finite sample performance

BB	β_0	β_1	β_2
N=30	0.055	0.175	0.136
N=60	0.056	0.967	1.000
N=120	0.064	1.000	1.000

E1	β_0	β_1	β_2
N=30	0.060	0.082	0.078
N=60	0.045	0.438	0.440
N=120	0.042	1.000	1.000

E5	β_0	β_1	β_2
N=30	0.042	0.072	0.060
N=60	0.047	0.435	0.438
N=120	0.044	1.000	1.000

E9	β_0	β_1	β_2
N=30	0.069	0.081	0.091
N=60	0.058	0.435	0.404
N=120	0.042	1.000	1.000

Table 3: Rejection rates of the **Monte Carlo test**. Columns corresponding to β_0 report empirical size, those to β_1 and β_2, empirical power.

Finite sample performance

BB	β_0	β_1	β_2
N=30	0.064	0.344	0.053
N=60	0.058	0.995	0.085
N=120	0.069	1.000	0.238

E1	β_0	β_1	β_2
N=30	0.053	0.071	0.089
N=60	0.058	0.215	0.220
N=120	0.056	0.975	0.971

E5	β_0	β_1	β_2
N=30	0.047	0.065	0.044
N=60	0.064	0.249	0.193
N=120	0.049	0.982	0.898

E9	β_0	β_1	β_2
N=30	0.051	0.075	0.085
N=60	0.065	0.216	0.234
N=120	0.058	0.929	0.967

Table 4: Rejection rates of the **Chi–square test**. Columns corresponding to β_0 report empirical size, those to β_1 and β_2, empirical power.

Tropical storms results

τ	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
typhoons P–value	0.365	0.537	0.545	0.495	0.438	0.381	0.329	0.316	0.269
hurricanes P–value	0.439	0.239	0.133	0.081	0.062	0.047	0.038	0.040	0.055

Table 5: P–values for the Monte Carlo trend test

τ	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
q	10	11	12	12	12	12	12	12	12
typhoons P–value	0.534	0.705	0.722	0.688	0.587	0.466	0.382	0.371	0.453
q	5	5	5	6	6	6	7	7	7
hurricanes P–value	0.069	0.024	0.015	0.006	0.003	0.003	0.004	0.006	0.035

Table 6: P–values for the χ^2 trend test

Conclusions

- Two tests of significance of the slope function provided
- Annual pattern of wind speeds for both hurricanes and typhoons cannot be treated as constant, no matter what expectile level is considered.
- Tests detect a trend in the upper wind speeds of Atlantic hurricanes

Change Point and Trend Analyses of Annual Expectile Curves of Tropical Storms

Petra Burdejová
Wolfgang Karl Härdle
Piotr Kokoszka
Qian Xiong

Humboldt–Universität zu Berlin
Colorado State University
http://lvb.wiwi.hu-berlin.de
http://www.stat.colostate.edu

References

Alba E., Andrade R.
Evaluating the Impact of the Increase in Hurricane Frequency Using an Internal Model. A Simulation Analysis
INEGI, Mexico
Conference paper

Burdejova P., Härdle W., Kokoszka P., Xiong Q.
Change Point and Trend Analyses of Annual Expectile Curves of Tropical Storms
Humboldt University, Discussion Papers 2015
DOI: 10.1093/jjfinec/nbu004

Schnabel, S.K., Eilers, P. H. C.
Optimal expectile smoothing
Computational Statistics and Data Analysis **53**(12): 4168-4177, 2009, DOI: 10.1016/j.csda.2009.05.002

Expectile as quantile

$e_\tau(Y)$ is the τ-quantile of the cdf T, where

$$T(y) = \frac{G(y) - xF(y)}{2\{G(y) - yF(y)\} + \{y - \mu_Y\}}, \tag{1}$$

$$G(y) = \int_{-\infty}^{y} u \, dF(u) \tag{2}$$

Back Quantiles and Expectiles

B-splines

Knot vector $t = (t_1, \dots, t_M)$ as nondecreasing sequence in $[0, 1]$

Control points $P_0, \dots, P_N$

Define i-th B-spline basis function $N_{i,j}$ of order j as

$$N_{i,0}(t) = \begin{cases} 1 & \text{if } t_i < t < t_{i+1} \\ 0 & \text{otherwise} \end{cases}$$

$$N_{i,j}(t) = \frac{t - t_i}{t_{i+j} - t_i} N_{i,j-1}(t) + \frac{t_{i+j+1} - t}{t_{i+j+1} - t_{i+1}} N_{i+1,j-1}(t)$$

$j = 1, \dots, N - M - 1$

Back to Estimation of Expectile Curves

Quantile Curves Penalty

$$\widehat{\alpha} = \arg\min_{\alpha} \sum_{i=1}^{n} \rho_\tau \left\{ y_i - \alpha^\top b(x_i) \right\} + \lambda \alpha^\top \Omega \alpha$$

where $b(x) = (b_1, \ldots, b_K)^\top$ is vector of B-spline basis functions

Denote $\tilde{b}(x) = (\tilde{b}_1(x), \ldots, \tilde{b}_K(x))^\top$ the vector of second derivatives of basis functions

and set

$$\Omega = \int \tilde{b}(x)\tilde{b}(x)^\top \, dx$$

▸ Back to Estimation of Expectile Curves

LAWS estimation

Schnabel and Eilers (2009):

$$\min \sum_{i=1}^{n} w_i(\tau)(y_i - \mu_i)^2$$

where

$$w_i(\tau) = \begin{cases} \tau & \text{if } y_i > \mu_i \\ 1-\tau & \text{if } y_i \le \mu_i, \end{cases}$$

μ_i expected value according to some model.

Iterations:

- fixed weights, closed form solution of weighted regression
- recalculate weights

until convergence criterion met.

Back to Expectiles Curves

LAWS estimation

Example:

Classical linear regression model

$$Y = X\beta + \varepsilon$$

where $\mathrm{E}(\varepsilon | X) = 0$ and $\mu = \mathrm{E}(Y | X) = X\beta$.

$$\widehat{\beta} = \arg\min_{\beta} \sum_{i=1}^{n} w_i (y_i - \mu_i)^2$$

Then:

$$\widehat{\beta} = (X^\top W X)^{-1} X W Y$$

with W diagonal matrix of fixed weights w_i.

Back to Expectile Curves

Functional principal components

$X(t)$ stochastic process on compound interval T
with mean function $\mu(t) = \mathrm{E}\{X(t)\}$
and covariance function $k(s,t) = \mathrm{cov}(X(s), X(t))$

There exist orthogonal sequence of eigenfunctions ϕ_j and eigenvalues λ_i such that $k(s,t) = \sum_{j=1}^{\infty} \lambda_j \phi_j(s)\phi_j(t)$
We can rewrite process as

$$X(t) = \mu(t) + \sum_{j=1}^{\infty} \sqrt{\lambda_j}\kappa_j \phi_j(t)$$

where $\kappa_j = \frac{1}{\sqrt{\lambda_j}} \int X(t)\phi_j(s)ds$, $\mathrm{E}(\kappa_j) = 0$ and $\mathrm{E}(\kappa_j \kappa_K) = \delta_{jk}$.

Back to Change point test