Report on the meeting of the Space/Space-Time group of the workshop: "Extreme events in climate and weather"

Thursday, August 26, 2010

General Discussion

There was a consensus on the need for non–statonary (in time) and non–homogeneous (in space) models of extremes. For concreteness, the group focused on precipitation measured over time at a collection of stations distributed over space.

It is important to develop good models that capture:

(i) changes in spatial dependence as a function of the season

(ii) long–term temporal trends due to global warming.

This is motivated by the need to better predict cumulative areal extremes of precipitation over time. Successful models would have an important impact on insurance (flood, drought, fire), civil engineering (updating the local building codes for the construction of bridges, dams, irrigation systems, planing sewers, etc.), agriculture (predicting drought), and public policy (decisions on funding infrastructure projects such as reservoirs), etc.

Another motivating point is to be able to provide realistic models for statistical down–scaling from global climate models to finer local scales. One deficiency of the current GCM's is the coarse resolution and rather inadequate representation of rainfall events. We feel that a direct statistical study of local station data would be very helpful in building models that can replicate extreme precipitation.

Conditional Extremes

Let $X(s,t)$ denote the precipitation amount at station s during time period t (in certain temporal units, eg over a day, week, month, etc.) A common issue with many current statistical methods dealing with extremes is related to considering the station–wise maximum

$$M_T(s) := \max_{1\le t\le T} X(s,t).$$

In focusing on $M_T(s)$, one essentially discards important data. In particular, the maximum at station s_0 may or may not occur at the same time point as the maximum at another station s. Thus, one has a very limited information about the spatial dependence. This issue has been mentioned a number of times in various contexts throughout the week.

One approach to incorporate more data is as follows. Consider a station s_0.

1. Suppose that the yearly maximum of daily precipitation at s_0 is achieved on day $t(s_0)$.

2. Record the "snap–shot" of daily precipitation at all other stations $X(s, t(s_0))$, at the day $t(s_0)$ when station s_0 achieved its maximum.

3. Collect samples the so-constructed "snap–shots" over a period of say 100 consecutive years.

By focusing on the above sample of surfaces, one essentially conditions on the (yearly) extreme event at a station s_0. The distribution of these surfaces can be interpreted as a distribution of possible storms diving rise to an extreme event at s_0. Such data on the conditional extremes may be used to:

a. Produce a map of the conditional covariances, variances and correlations, e.g.:

$$\mathrm{Cov}(X(s_1, t(s_0)), X(s_2, t(s_0)|M_T(s_0) = X(s_0, t(s_0))).$$

A contour–plot or a heat–map of such quantities may be used as an empirical characterization of the spatial distribution of a storm. Indeed, one expects that stations located nearby to be highly correlated.

b. The samples of such surfaces may be also estimated by usual smoothing and dimension reduction techniques (splines, wavelets, functional PCA) to arrive at more flexible spatial models of "storms".

Spatio–Temporal Extremes: Beyond max–stable models

The methodology of considering extreme scenarios by conditioning on the times when the extremes are achieved motivates certain extensions of the popular Smith storm model. Namely, consider a Poisson point process $(X_i, T_i, \Theta_i, \epsilon_i),\ i \in \mathbb{N}$, in the space $(\mathbb{R}^2 \times \mathbb{R}_+ \times \Theta \times \mathbb{R}_+)$ with intensity $dxdt\nu(d\theta)dy$. The point $(X_i, T_i, \Theta_i, \epsilon_i)$ corresponds to a "storm" originating at location X_i at time T_i, of type Θ_i and "inverse intensity" ϵ_i. The "influence" of this storm at time t and location s would then be modeled by:

$$u_i(s, t) := f(s - X_i, t - T_i, \Theta_i)/\epsilon_i, \tag{1}$$

where f is a suitable non–negative function.

Max–stable models: Under standard integrability conditions on f, the maxima (of all influences at time t and location s)

$$\xi(s, t) := \bigvee_{i \in \mathbb{N}} u_i(s, t),$$

is a max–stable process (space–time random field). It is well–known that one has the *extremal integral* representation:

$$\xi(s,t) = \int^{\mathrm{e}}_{\mathbb{R}^2\times\mathbb{R}\times\Theta} f(s-u,t-\tau,\theta)M_1(du,d\tau,d\theta),$$

where $M_1(\cdot)$ is a $1-$Fréchet sup–measure with control measure $dud\tau\nu(d\theta)$ [1, 2].

The finite–dimensional distributions of $\xi(s,t)$ (in space and time!) are then given as follows:

$$\mathbb{P}\{\xi(s_i,t_i)\le x_i,\ 1\le i\le k\} = \exp\Big\{-\int_{\mathbb{R}^2\times\mathbb{R}\times\Theta}\Big(\max_{1\le i\le k} f(s_i-u,t_i-\tau,\theta)/x_i\Big)dud\tau\nu(d\theta)\Big\}.$$

This is an instance of a *mixed moving maxima* model, which is stationary in space and time. In particular, the resulting random field is ergodic (and mixing) and existing general results [3] can be used to show the strong consistency of empirical functionals of $\xi(s,t)$.

Beyond max–stability As indicated above, a major drawback of max–stable spatial (not spatio–temporal) models is that considering station–wise maxima erases temporal information about extremes present in the data. One possible approach is to consider the entire point process $u_i,\ i\in\mathbb{N}$, where $u_i = u_i(s,t)$ are points in a space of functions given by (1).

Past data can be used, for example, by using functional PCA methods, to identify an adequate low–dimensional *parametric* model for the u_i's. The intensity measure of the underlying Poisson point process may be modeled by using orographic information along with existing deformation or non–parametric smoothing techniques. In fact, one can consider *non–product* intensities, incorporating a seasonal interaction between space and time. For example,

$$\mu(dt,du) = \exp\{-\rho(u,A)(1-\sin(t))\}dudt$$

defines a measure on $\mathbb{R}^2\times\mathbb{R}$, which emphasizes points u close to a region $A\subset\mathbb{R}^2$ in the plane, with the exception of the "Summer" season (corresponding to $t=\pi/2$). More realistic finite–dimensional parameter models can be constructed as linear combinations of functional PCA components. This can give rise to a variety of inference problems.

References

[1] L. de Haan. A spectral representation for max–stable processes. *Annals of Probability*, 12(4):1194–1204, 1984.

[2] S. Stoev and M. S. Taqqu. Extremal stochastic integrals: a parallel between max–stable processes and $\alpha-$stable processes. *Extremes*, 8:237–266, 2005.

[3] Stilian A. Stoev. On the ergodicity and mixing of max-stable processes. *Stochastic Process. Appl.*, 118(9):1679–1705, 2008.