Modelling of heavy-tailed, non-Gaussian, skewed records.

Jörg Wegener, Krys Podgórski

Mathematical Statistics Lund, Sweden

jw@maths.lth.se

Motivation

> "Geostatistical data often display non-Gaussian features, such as **heavy tails** or **skewness** (...) and observations that would be 'outliers' under a Gaussian sampling process."
>
> M. Blanca Palacios & Mark F. J. Steel (2006), "Non-Gaussian Bayesian Geostatistical Modeling", *Journal of the American Statistical Association*, Vol. 101, Nr. 474

Precipitation data is a prominent example of observations showing strong non-Gaussian features. Due to the amount of e.g. rainfall being non-negative, the corresponding distributions show strong asymmetries (Figure 1).

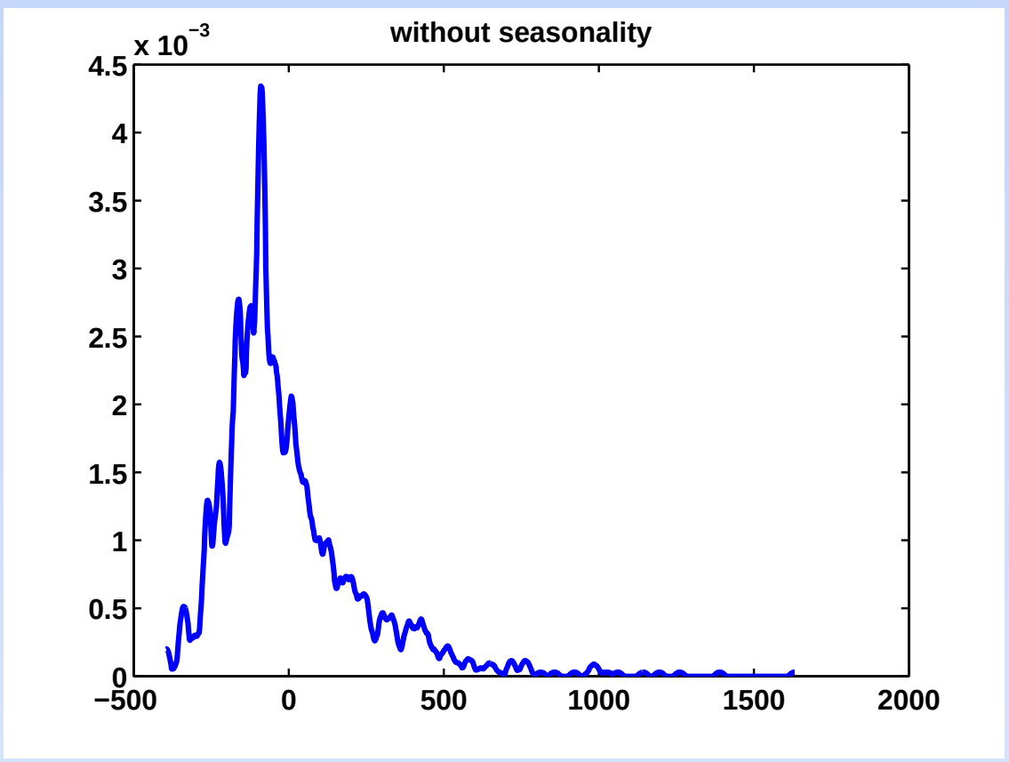

FIGURE 1: Estimated PDF of Sahel rainfall anomalies.

A way of accounting for distributional properties like skewness and (one-sided) heavy tails, is illustrated in the following: By means of a stochastic process (Laplace integral) it is possible to simulate and analyse data which stems from a very flexible class of underlying distributions (Figure 4).

Example I

Rainfall in Africa

Distribution and correlation function of rainfall observations from Sahel region are simulated by means of the Laplace integral model. For a definition of the model, see theory section.

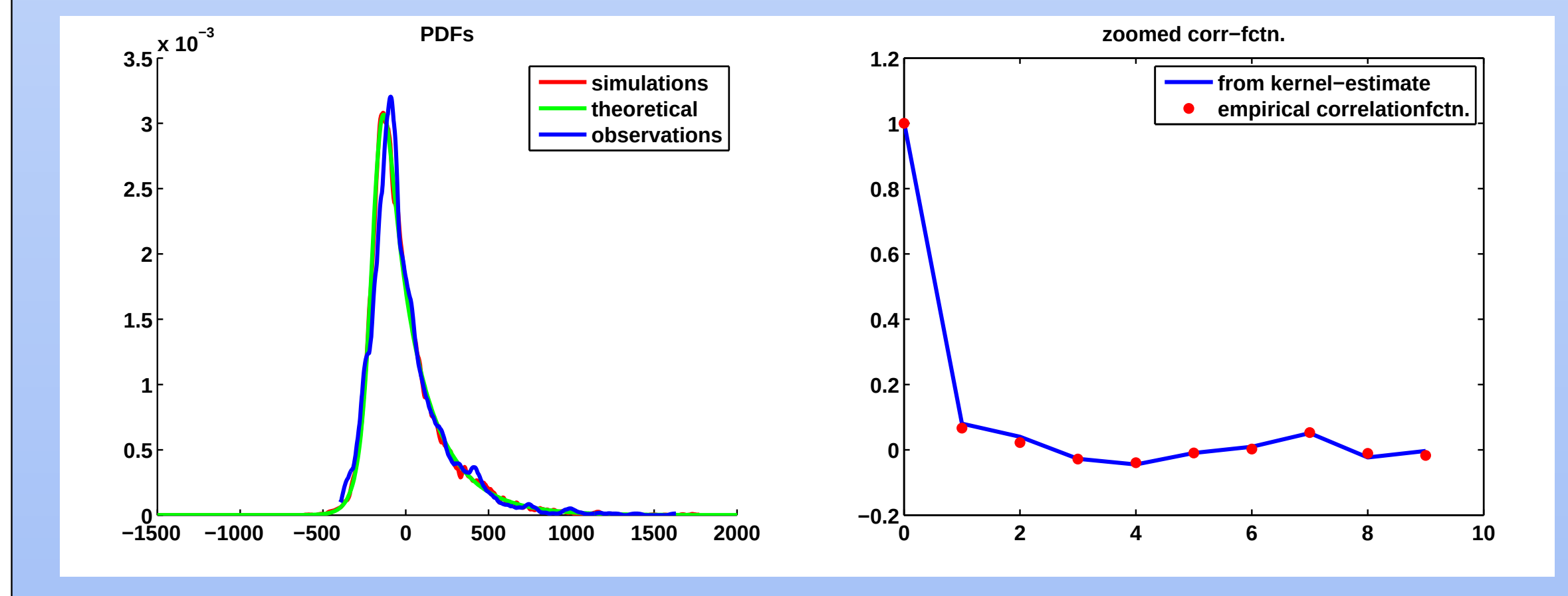

FIGURE 2: Left: Estimated PDFs of observed anomalies (blue) and simulated Laplace integral model (red), compared to fitted generalised Laplace distribution (green). Right: Autocorrelation function of observed anomalies (red dots) and simulated (blue).

Example II

Sealevel pressure in Stockholm

Air pressure observations usually exhibit heavy lower tails, due to higher absolute pressure in storm events compared to high pressure events (Figure 3). Daily sealevel pressure observations from Stockholm are modeled by the Laplace integral technique.

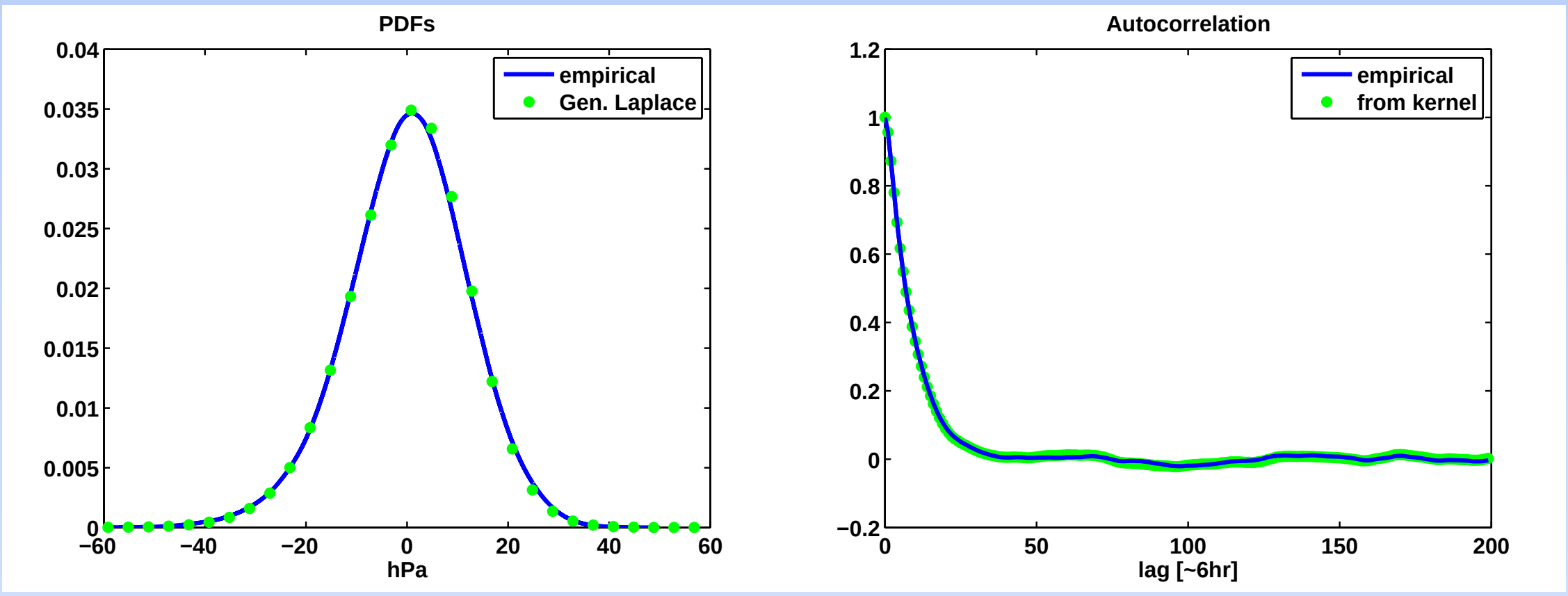

FIGURE 3: Left: PDFs of observed sealevel pressure anomalies (blue) and fitted Laplace integral model. Right: Autocorrelation function of observed anomalies (blue) and by estimating the kernel function (green).

Theory

Although allowing for much heavier than Gaussian tails and asymmetry, the *generalised Laplace distributions* maintain the elegant Gaussian property: Due to finite second moments the spectral theory remains well defined for stochastic Laplace integral processes.

- *Laplace motion* $L(t)$ is defined as a Lévy process with increments having characteristic function (Kotz et al., 2001):

  $$\varphi(u) = \left(1 - i\mu u + \frac{\sigma^2 u^2}{2}\right)^{-1/\nu} ,$$

  with parameters μ (symmetry), σ (scale) and ν (shape).
  The above characteristic function corresponds to the generalised Laplace distributions (Figure 4).

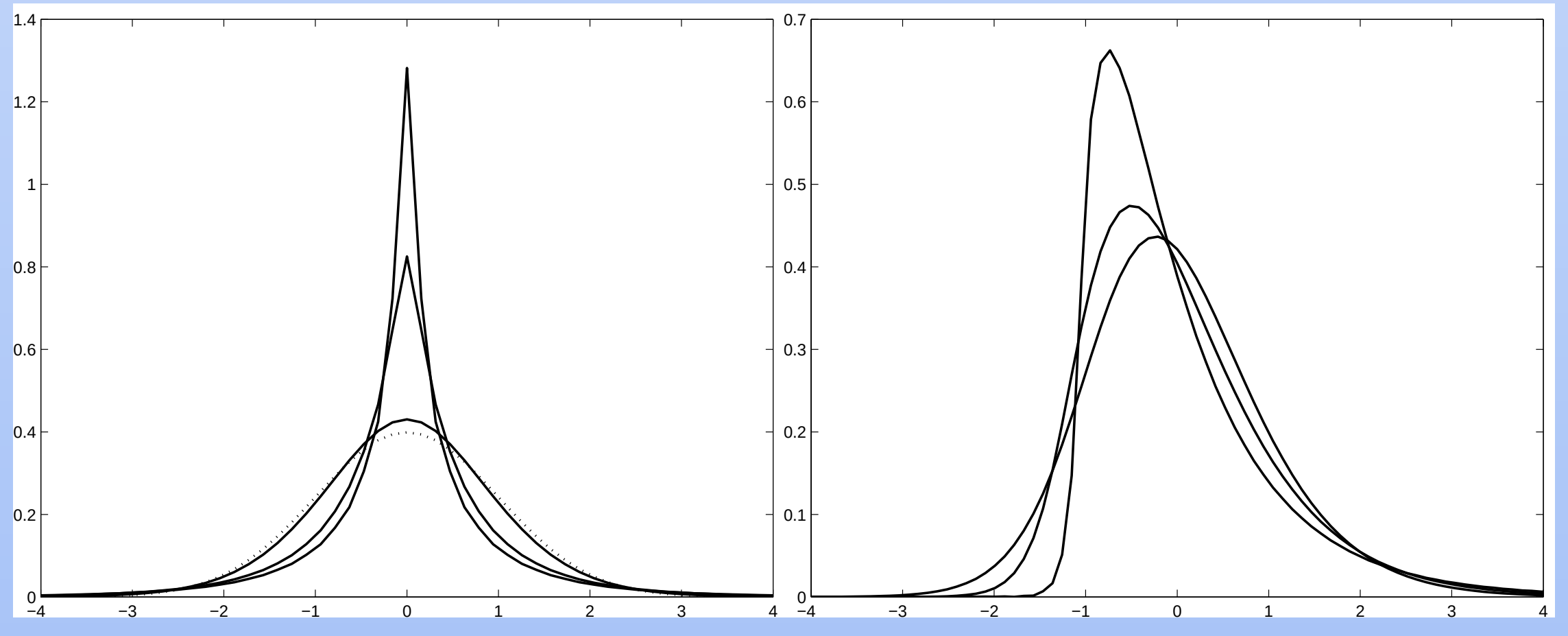

FIGURE 4: Generalised Laplace distributions (solid). Left: Symmetric case, i.e. $\mu = 0$, different values for σ and ν. Standard Normal distribution (dots) for comparison. Right: Skewed case, $\mu = 0.5$, different values for σ and ν.

- The stochastic Laplace measure Λ relates to the Laplace motion for an interval $(a, b] \subseteq \mathbb{R}_+$:

  $$\Lambda(a, b] = L(b) - L(a) .$$

- Laplace integral (Åberg and Podgórski, 2009):

  $$X_t = \int_{\mathcal{X}} f(t - x)\, d\Lambda(x) ,$$

  where:

  $\Lambda(x)$ Laplace measure
  $f(\cdot)$ kernel
  $\mathcal{X}$ here: $\mathbb{R}$
  X_t has characteristic function $\Phi_{X_t}(f(\cdot), \ldots)$

- Estimating kernel function:

  $$\widehat{f}(x) = \sqrt{2\pi}\frac{\mathcal{F}^{-1}\sqrt{\widehat{R}(\omega)}}{\sqrt{\int_{-\infty}^{\infty} \widehat{R}(\omega)\, d\omega}} ,$$

  where:

  $\widehat{R}(\omega)$ estimate of spectrum
  $\mathcal{F}^{-1}$ inverse Fourier transform

- When estimating the kernel function $f(\cdot)$, one gets the **autocorrelation function for free:**

  $$\widehat{r}(\tau) = \widehat{f}(x) * \widehat{f}(x) ,$$

  where:

  $\widehat{r}(\tau)$ autocorrelation function
  τ timelag

Further Application

Bias correction of storm events in a regional climate model

Extremely low sealevel pressure events in the *Rossby Centre Regional Atmosphere Model* are simulated too weak compared to observations (Figure 5). Moreover, pressure distributions are known to show asymmetric tail behaviour (Figure 3) with heavier tails at low pressure. This seems to make the Laplace model an interesting candidate to successfully bias-correct model output towards observations.

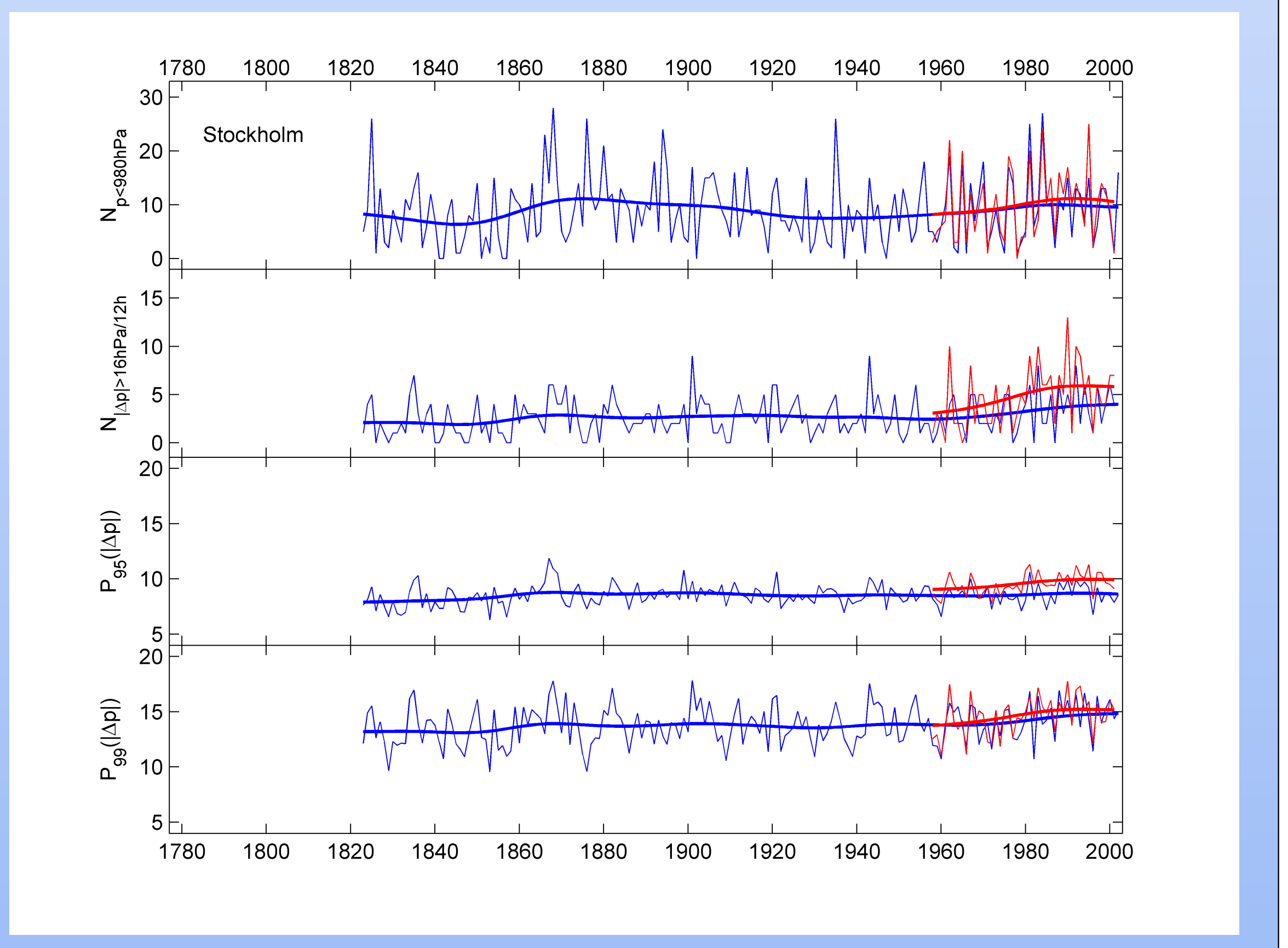

FIGURE 5: Observations of sealevel pressure (blue) and regional climate model simulations (red). Low pressure events ("storms") are underestimated, i.e. biased towards higher pressure. This is mostly apparent from the percentile plots (lower rows).

Courtesy of *Lars Bärring*, SMHI.

References

Åberg, S. and Podgórski, K. (2009). A Class of Non-Gaussian second order Spatio-Temporal Models. *Extremes.* To appear.

Kotz, S., Kozubowski, T., and Podgórski, K. (2001). *The Laplace distribution and generalizations: A revisit with applications to communications, economics, engineering and finance.* Birkhäuser, Boston.