Regional climate model assessment

via spatio-temporal modeling

Peter F. Craigmile

http://www.stat.osu.edu/~pfc/

Inferential Challenges for Large Spatio-Temporal Data Structures

Banff International Research Station workshop, Banff, Canada

3-8 Dec 2017

Adapted from research carried out with

Peter Guttorp (Univ. of Washington and the Norwegian Computing Center)

and Veronica Berrocal (University of Michigan)

- P. F. Craigmile and P. Guttorp (2011). Space-time modeling of trends in temperature series. *Journal of Time Series Analysis*, **32**, 378-395.
- V. J. Berrocal, P. F. Craigmile, and P. Guttorp (2012). Regional climate model assessment using statistical upscaling and downscaling techniques. *Environmetrics*, **23**, 482-492.
- P. F. Craigmile and P. Guttorp (2013). Can a regional climate model reproduce observed extreme temperatures? *Statistica*, **73**, 103-122.

My research is supported by the US National Science Foundation under grants NSF-DMS-1407604 and NSF-SES-1424481. Travel funding by STATMOS.

Investigating and assessing our climate system

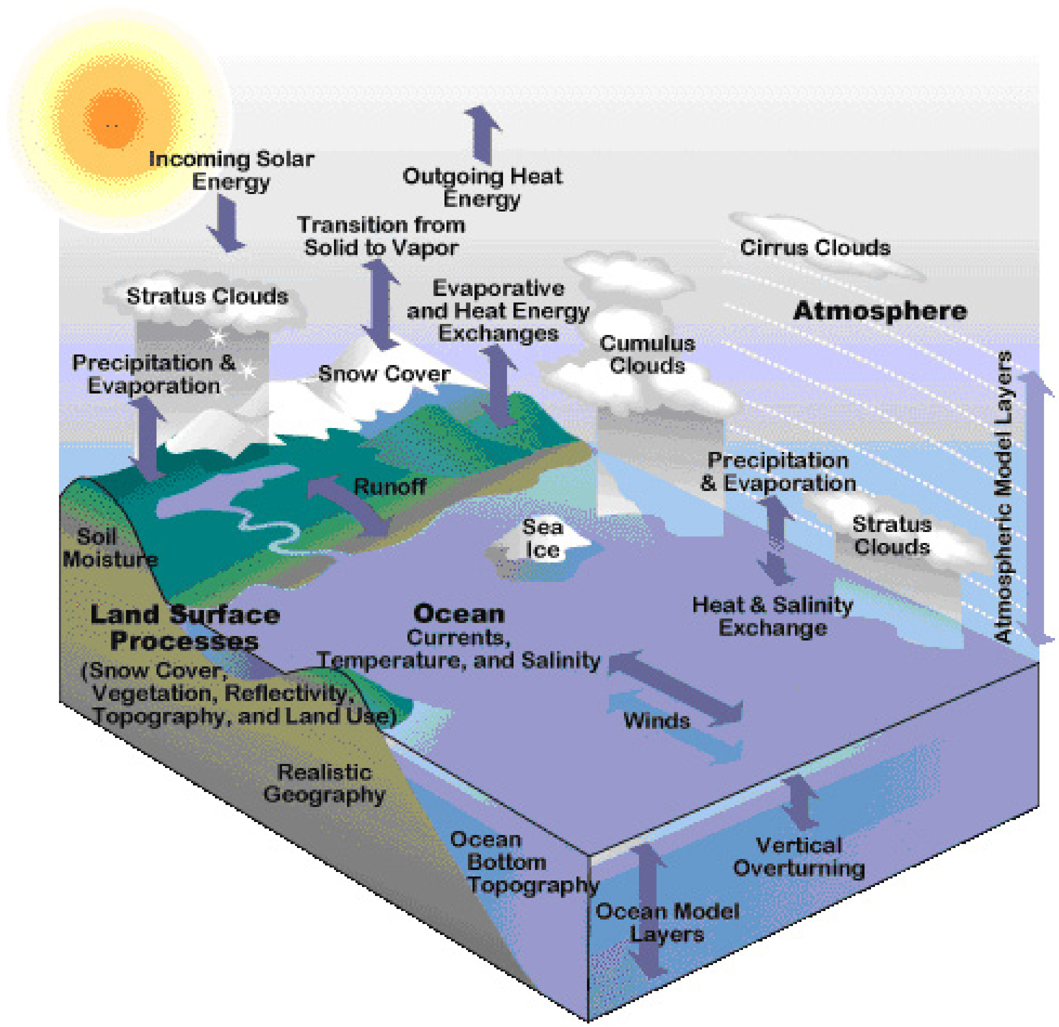
In order to adapt to a changing climate, policymakers need information about what to expect for the climate system.

Examples of local information:

- **observational measurements**
- **regional climate models (RCMs)**
- **proxy measurements** (e.g., paleoclimate – for another talk!)

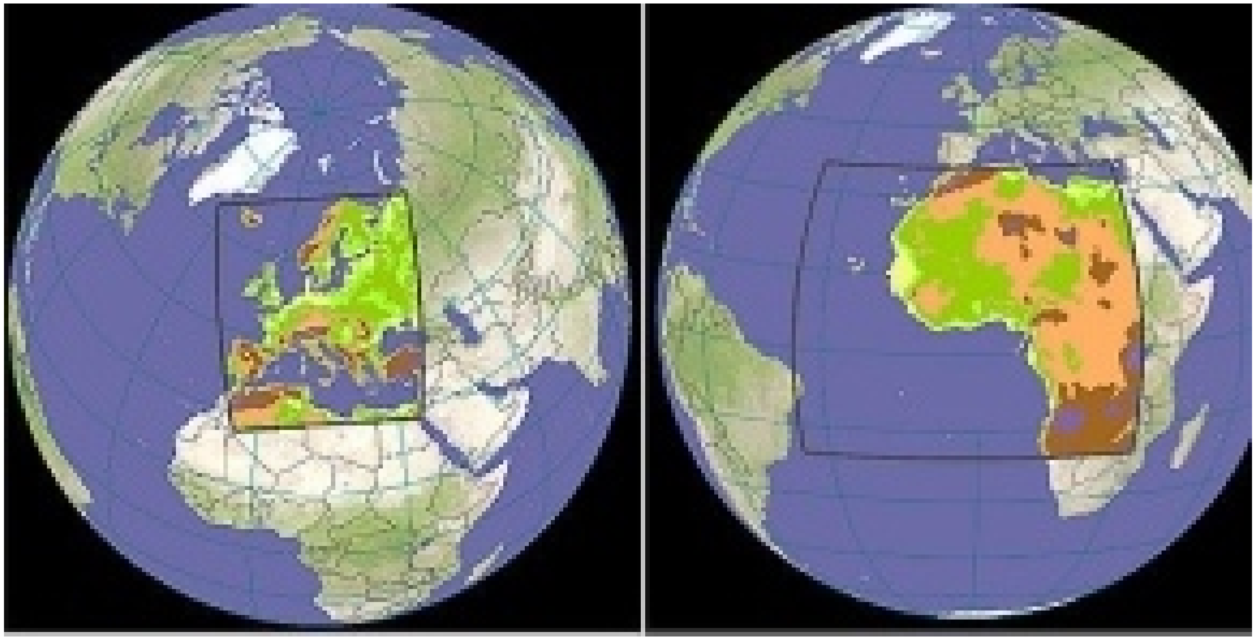
Q: How well do regional models **reproduce** *observed climate*?

Climate models



Source: http://www.windows2universe.org/earth/climate/cli_models2.html

Regional climate models



Source: http://www.smhi.se/en/research/research-news/combined-science-on-climate-models-1.14533

- RCMs are a downscaled global circulation/climate model (GCM).

  Mathematical model that describes, using partial differential equations, the temporal evolution of climate, oceans, atmosphere, ice, and land-use processes over a gridded spatial domain of interest.

Regional climate models

- RCMs typically operate on relatively **small areas**.
  - Within these small areas, there are more spatial locations than from a GCM (more data! more information?)
- RCMs need to use **boundary values** for the global distribution of the atmosphere, oceans, etc. (typically drawn from a GCM).
- Thus, there are **two** sources of variation to consider:
  1. Inadequacies in the RCM;
  2. Inadequacies in the boundary values.
- We run models based on historical data (**a re-analysis**) to try to reduce the second source of variation [Samuelsson et al., 2011].

Regional climate model assessment

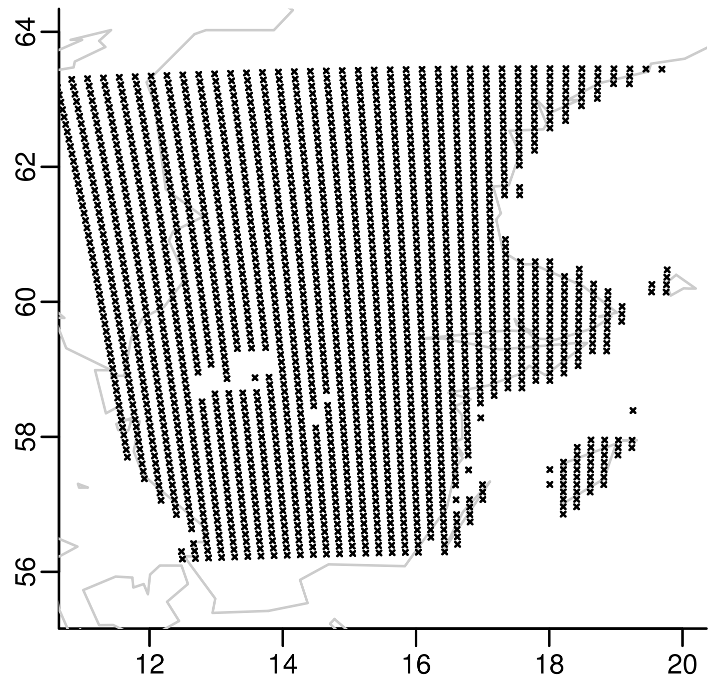
The assessment of **regional climate models (RCMs)** using observations is a non-trivial task.

Climate, being the distribution of weather and other climatic factors over long time periods [Rossow et al., 2005, Guttorp and Xu, 2011], cannot be measured directly.

Rather, a variety of quantities (including weather) are measured, and usually their **long-term averages** are compared to the model output.

More accurately, we argue one should compare the **distribution** of observations and model output, on **comparable spatial** and **temporal** scales.

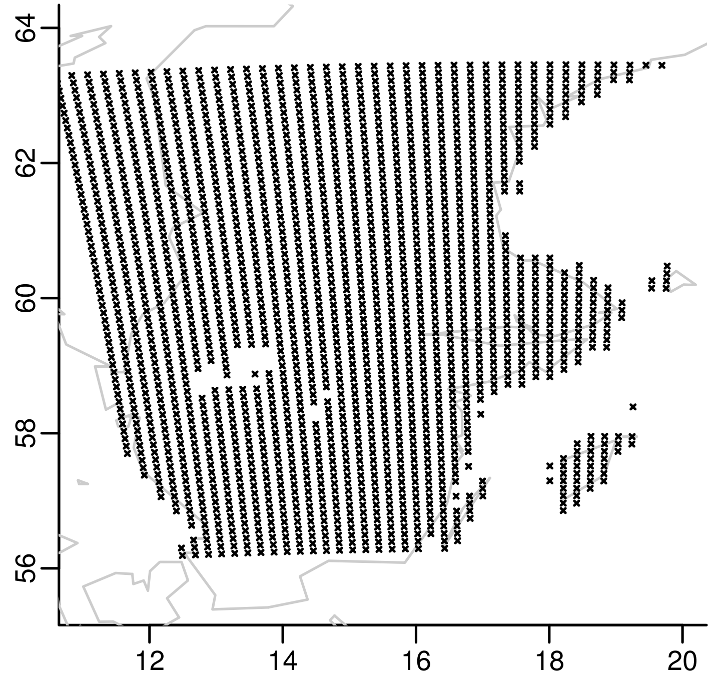
Example: SMHI RCM output



A control run of the Swedish Meteorological and Hydrological Institute (SMHI) regional climate model RCA version 3 [Samuelsson et al., 2011]:

- Run using boundary conditions given by the ERA40 reanalysis [Uppala et al., 2005] in the earlier years, and the ERA-INTERIM reanalysis [Uppala et al., 2008] in the later years.

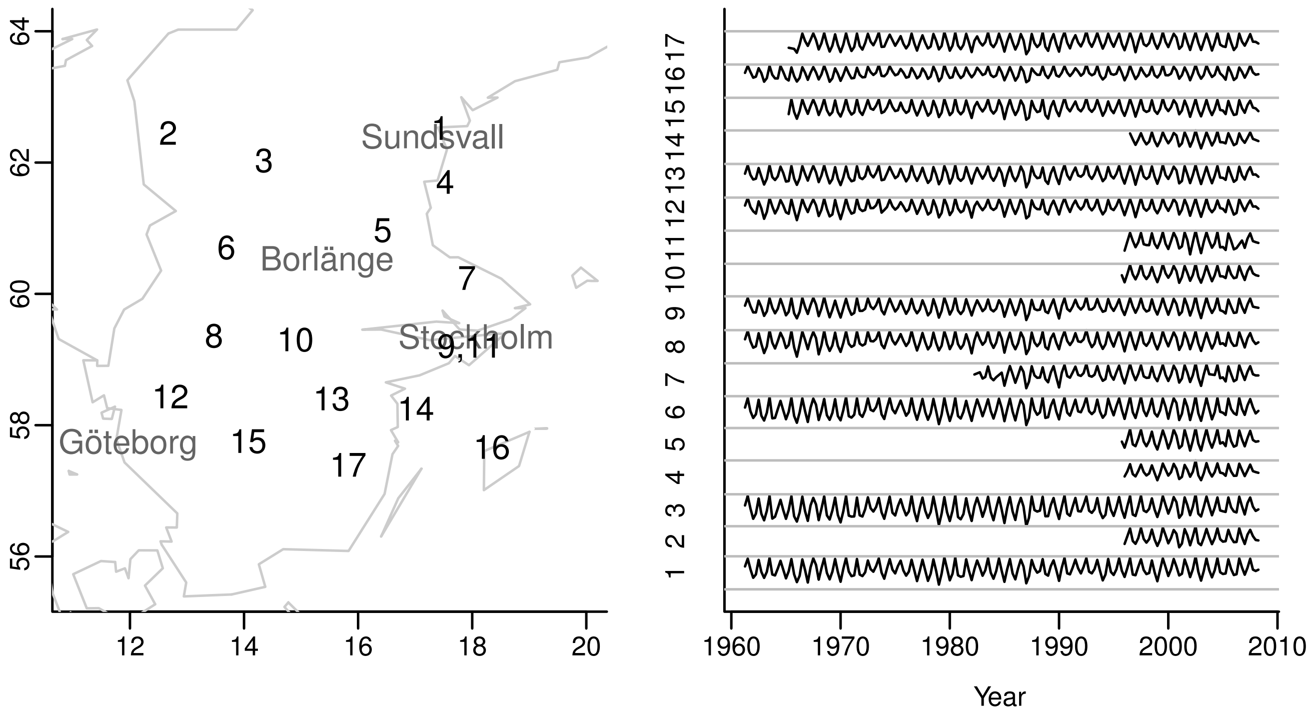
SMHI RCM output, cont.



Fractioned to different land types – we restrict our analysis to the 2 meter temperature given for the **open land** and **snow land covers**.

Available from Dec 1, 1962 to Nov 30, 2007 with a **temporal resolution** of 7.5 minutes. 12.5 km $\times$ 12.5 km **spatial resolution**.

Observational data



- Daily synoptic observations from 17 sites in an area of south central Sweden (Also from Swedish Meteorological and Hydrological Institute, SMHI).

Two analyses

1. A comparison of **seasonal averages** (DJF, MAM, JAJ, SON) of **daily mean temperature** [Berrocal et al., 2012].

2. A comparison of **seasonal minima** (DJF, MAM, JAJ, SON) of **daily mean temperature** [Craigmile and Guttorp, 2013].

In each case we fit (Bayesian hierarchical) **statistical models** to the

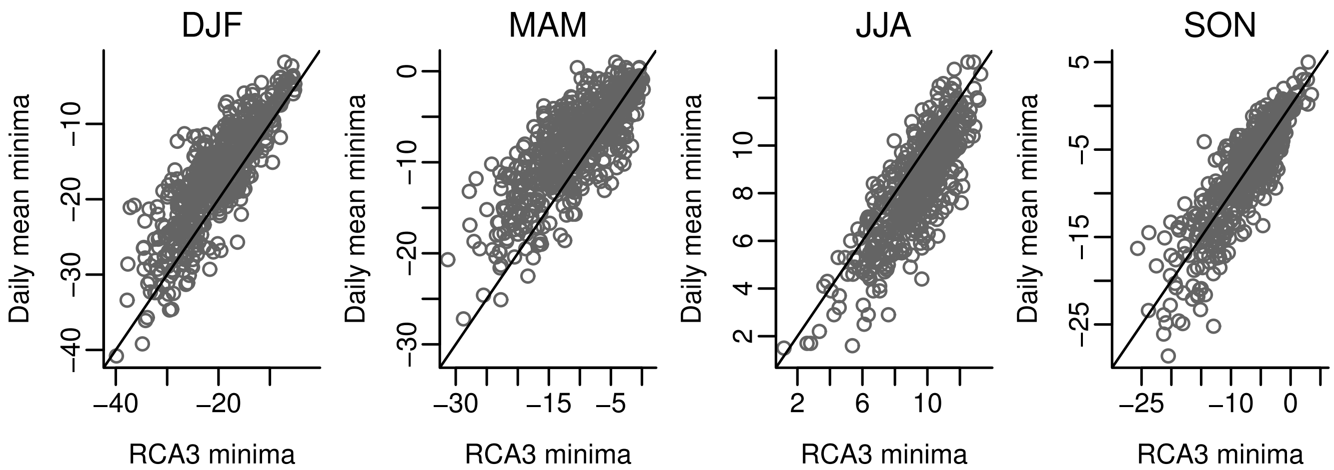
1. Observational data from reserved stations (point referenced), and

2. RCM output, observed on grids.

We **infer** upon the parameters in these models to learn about climate.

We will demonstrate with the second case – analyzing seasonal minima.

Non spatio-temporal comparisons



Seasonal minima for the station data and RCM agree best in the autumn (SON).

In the winter (DJF) and spring (MAM), the observed minima tend to be slightly higher than is observed for the RCM.

For the summer (JJA), this relationship is reversed – we observe cooler minima than is predicted by the model.

Investigating shifts in the distribution

To compare distributions, we use **Doksum's shift function** [Doksum, 1974].

We find a shift function $\Delta(\cdot)$ such that for random variables

$$X \text{ (RCM) and } Y \text{ (observations)}$$

we have $X + \Delta(X) \sim Y$.

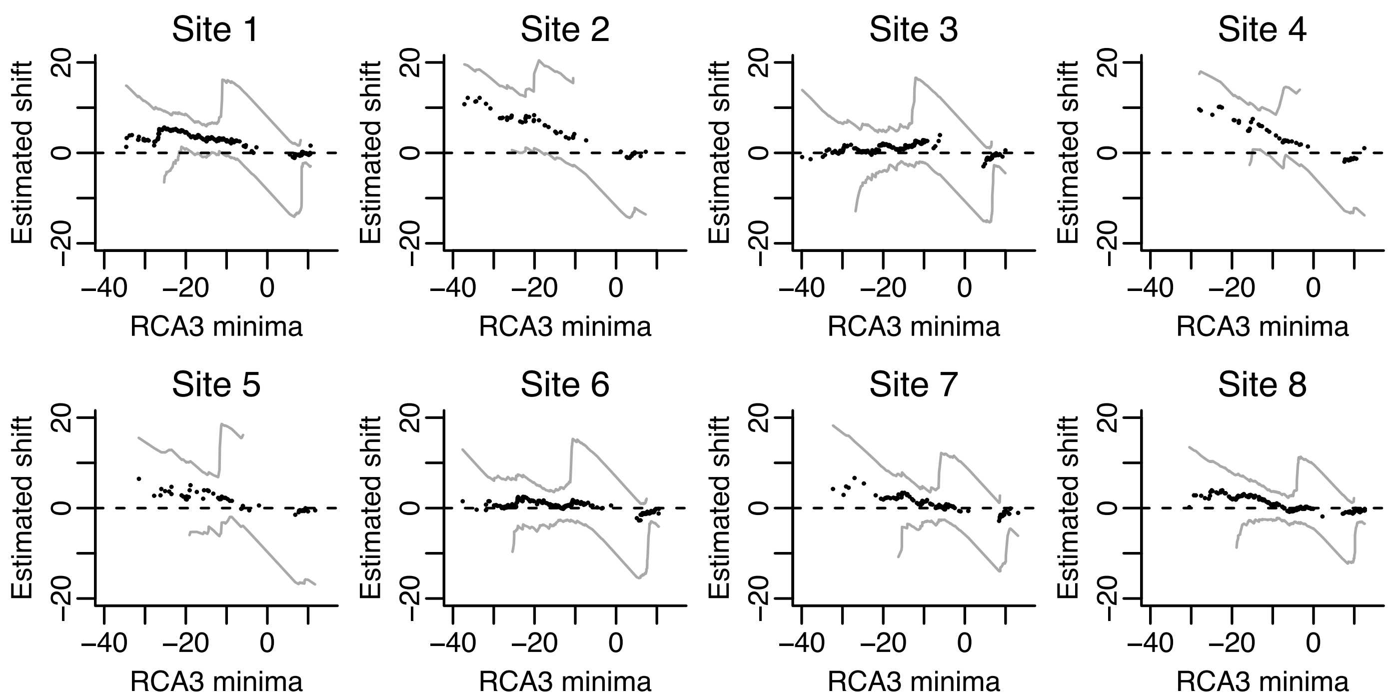
Our **estimate** of $\Delta(x)$ is

$$\widehat{\Delta}(x) = G_m^{-1}(F_n(x)) - x,$$

where F_n (resp. G_m) is the empirical distribution function of X (resp. Y)

Can also construct simultaneous confidence bands for $\Delta(\cdot)$ [Doksum and Sievers, 1976].

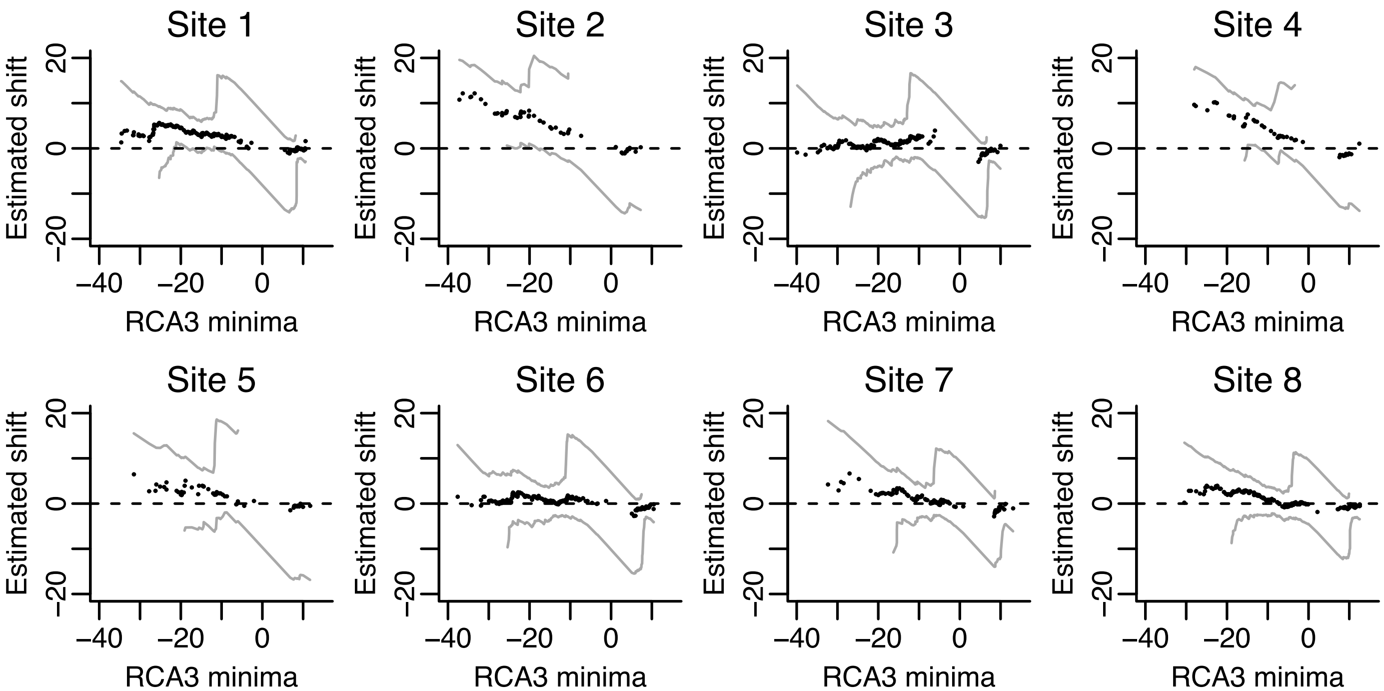
Investigating shifts in the distribution, continued



Anomalous behavior of the RCM around 0°C.

E.g., for grid square containing Station/Site 1 there are no RCA minima between about -2 and 5 °, but there are observed minima in that range of values.

Investigating shifts in the distribution, continued



Minima around 0°C for the RCM are most likely in the autumn (SON).

Less likely for the spring (MAM) and summer, and unlikely in the winter (DJF).

See Nikulin et al. [2011] for an example of the spatial estimation of shifts using the gridded E-OBS data product [Haylock et al., 2008] (with some caveats).

Extreme value theory-based comparisons

Use **(marginal) extreme value theory (EVT)** to analyze and model the seasonal minima for both the

station data

and the

RCM output.

[See Wang et al., 2016, for another example of this.]

Example: modeling the station data

At location $\boldsymbol{s} \in D$, let $Z_t(\boldsymbol{s})$ denote the **block minima** in year $y_t(\boldsymbol{s})$ and season $d_t(\boldsymbol{s})$ (taking on values 1: DJF; 2: MAM; 3: JJA; 4: SON), for time index $t = 1, \ldots, N(\boldsymbol{s})$.

Modeling the **negative of the minima** we suppose

$$[-Z_t(\boldsymbol{s})] \sim \text{GEV}(\widetilde{\mu}_t(\boldsymbol{s}), \sigma_t(\boldsymbol{s}), \xi_t(\boldsymbol{s})),$$

conditionally independent over $\boldsymbol{s}$ and t.

The GEV parameters

$\widetilde{\mu}_t(\boldsymbol{s}) \in \mathbb{R}$: **location parameter** indicating values which the distribution of the negative minima are concentrated around.

$\sigma_t(\boldsymbol{s}) > 0$: **scale parameter** defining the spread of the distribution.

$\xi_t(\boldsymbol{s})$: **shape parameter**. The tails of the GEV distributions are heavier for higher values of the shape parameter (A negative shape parameter leads to bounded tails; otherwise the tails of the distribution are unbounded.)

Interpreting the GEV parameters

We think GEV parameters as **describing climate**, with the changes in the parameters indicating seasonal differences and possible trends.

Given the climate, the model technically assumes that weather at different stations is **conditionally independent**.

- A **oversimplification**, since typically events of extremely cold air arise from arctic air moving south during a high pressure situation.
- So given that one station is extremely cold, it is more likely that another is.

We will critique this later!

Modeling the location parameter

Spatio-temporal model for $\{\mu_t(\boldsymbol{s}) = -\widetilde{\mu}_t(\boldsymbol{s})\}$:

$$\mu_t(\boldsymbol{s}) \;=\; \beta_0(\boldsymbol{s}) + \beta_1(\boldsymbol{s})(y_t(\boldsymbol{s}) - 1960) + \sum_{d=2}^{4} \beta_d(\boldsymbol{s}) I(d_t(\boldsymbol{s}) = d).$$

where $I(\cdot)$ is the indicator function.

We assume each $\{\beta_j(\boldsymbol{s})\}$ are independent **Gaussian processes** each with mean λ_j and isotropic covariance, $\mathrm{cov}(\beta_j(\boldsymbol{s}), \beta_j(\boldsymbol{s}+\boldsymbol{h})) = \tau_j \exp(-||\boldsymbol{h}||/\phi_j)$.

Here $\tau_j > 0$ is the (partial) sill parameter, $\phi_j > 0$ is the range parameter, and $||\cdot||$ denotes the Euclidean norm.

Modeling the scale and shape parameters

Assume that the **scale parameters** vary over space, but are constant in time:

$\sigma_t(\boldsymbol{s}) = \sigma(\boldsymbol{s})$ for all t and $\boldsymbol{s}$.

We suppose $\{\log \sigma(\boldsymbol{s})\}$ is a **Gaussian process** with mean λ_σ and isotropic covariance $\mathrm{cov}(\log \sigma(\boldsymbol{s}), \log \sigma(\boldsymbol{s} + \boldsymbol{h})) = \tau_\sigma \exp(-||\boldsymbol{h}||/\phi_\sigma)$.

Our assumption of a **constant shape parameter**, ξ, is an oversimplification, but is reasonable [e.g. Cooley et al., 2007, Sang and Gelfand, 2010].

Bayesian inference

Our **parameters** of interest are

$$\boldsymbol{\theta} = \Big(\{\beta_j(\boldsymbol{s}) : \boldsymbol{s} \in D, j = 0, \ldots, 4\}, \{\log \sigma(\boldsymbol{s}) : \boldsymbol{s} \in D\}, \xi, \\ \{\lambda_j : j = 0, \ldots, 4\}, \{\tau_j\}, \{\phi_j\}, \lambda_\sigma, \tau_\sigma, \phi_\sigma\Big)^T.$$

With the exception of the hyperparameters for the **shape** parameter ξ and the **spatial range** parameters, we assume **vague priors**.

For the **range** parameters we use the gamma prior choice of Craigmile and Guttorp [2011], who modeled daily mean temperature from the same synoptic stations.

The posterior

The **posterior distribution** of $\boldsymbol{\theta}$ given the data is not available in closed form.

We use a Markov chain Monte Carlo (MCMC) algorithm to sample from the posterior distribution.

The algorithm used is adapted from Mannshardt et al. [2013].

Fit one model to the observational data; another model to the RCM output.

The results are **robust** to **minor changes** in this choice of prior distribution.

Approximations required to fit the RCM model

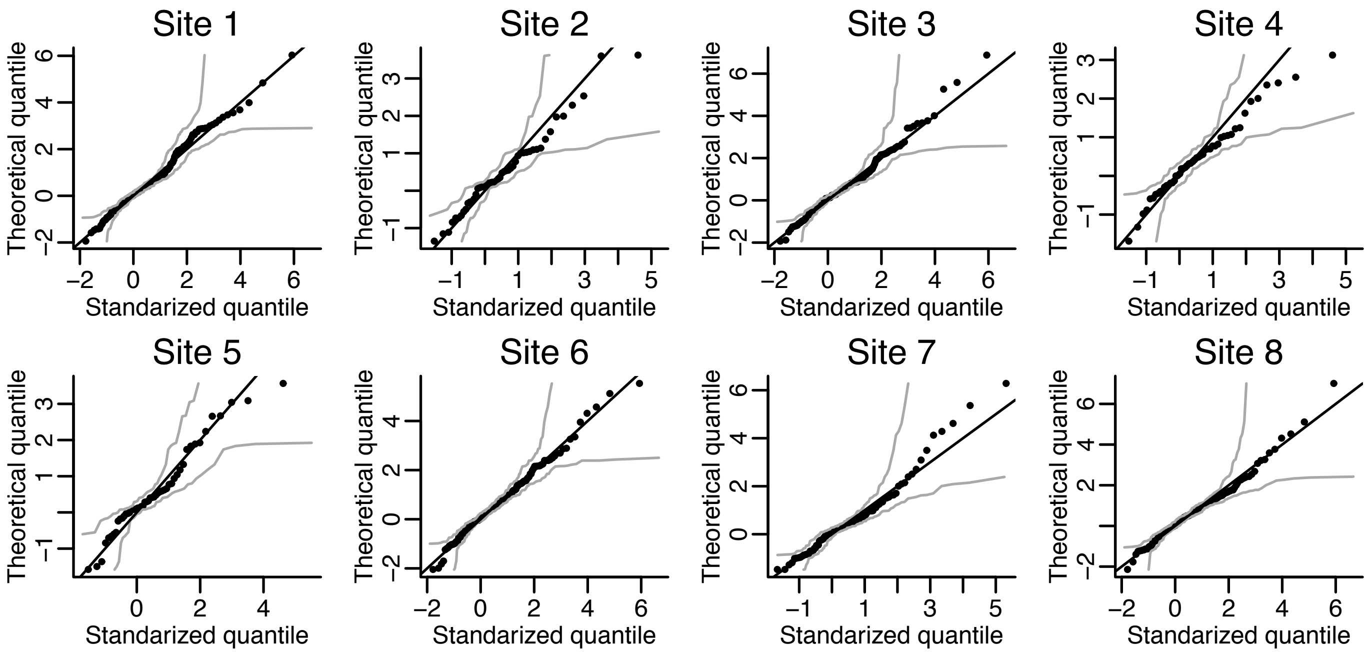
Because we have 1989 spatial locations, we made two approximations:

1. In updating the spatially varying GEV parameters, we calculated the acceptance ratios for Metropolis updates at each spatial location using the 15 **nearest neighbors**, rather than all the spatial locations.

2. In updating the hyperparameters in the spatial models, we broke the spatial field up into 4 **sub-regions** (NW, NE, SW and SE). This sped up matrix inversions.

Experiments demonstrated our results were robust to these approximations.

Model verification

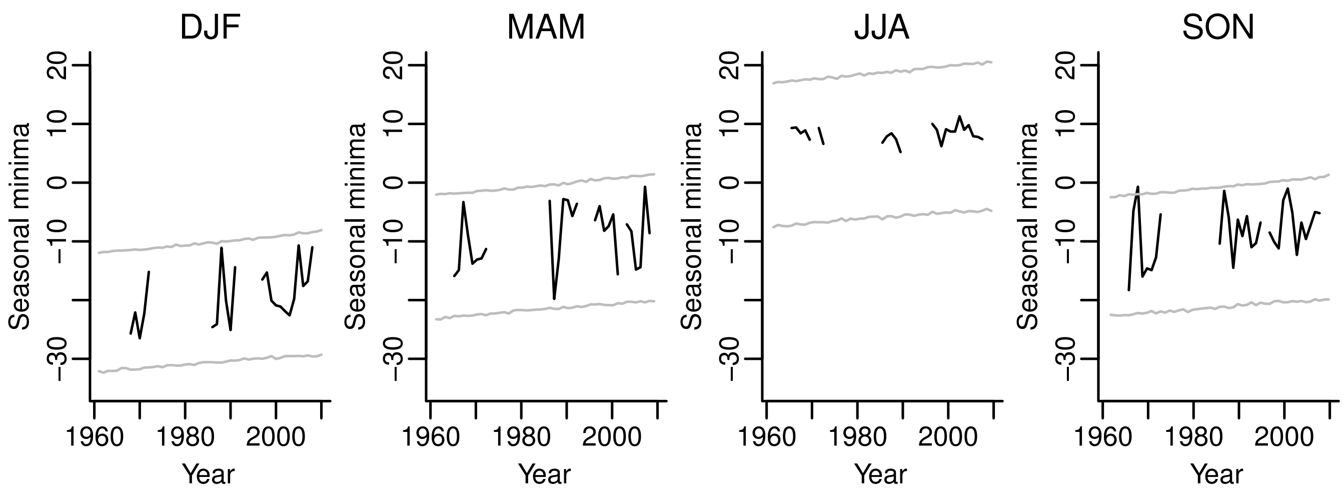


g

Quantile plots [see, e.g., Coles, 2001, Section 6.2.3] were used to assess the distributional assumptions made by the GEV models on a site-by-site basis.

- Excellent goodness of fit at all locations.
- Also indicates model fit improved over a model in which scale parameter was held fixed over locations.

Model verification, cont.



Also assess model using series at Borlänge — located at 15° 30' E and 60° 25' N, at an elevation of 152 meters. Station has been moved twice, and has long stretches of missing data, and was not included in the GEV model fit.

GEV model is **underestimating** the **variation** of seasonal minima in the summer (JJA).

GEV model results: the shape parameters

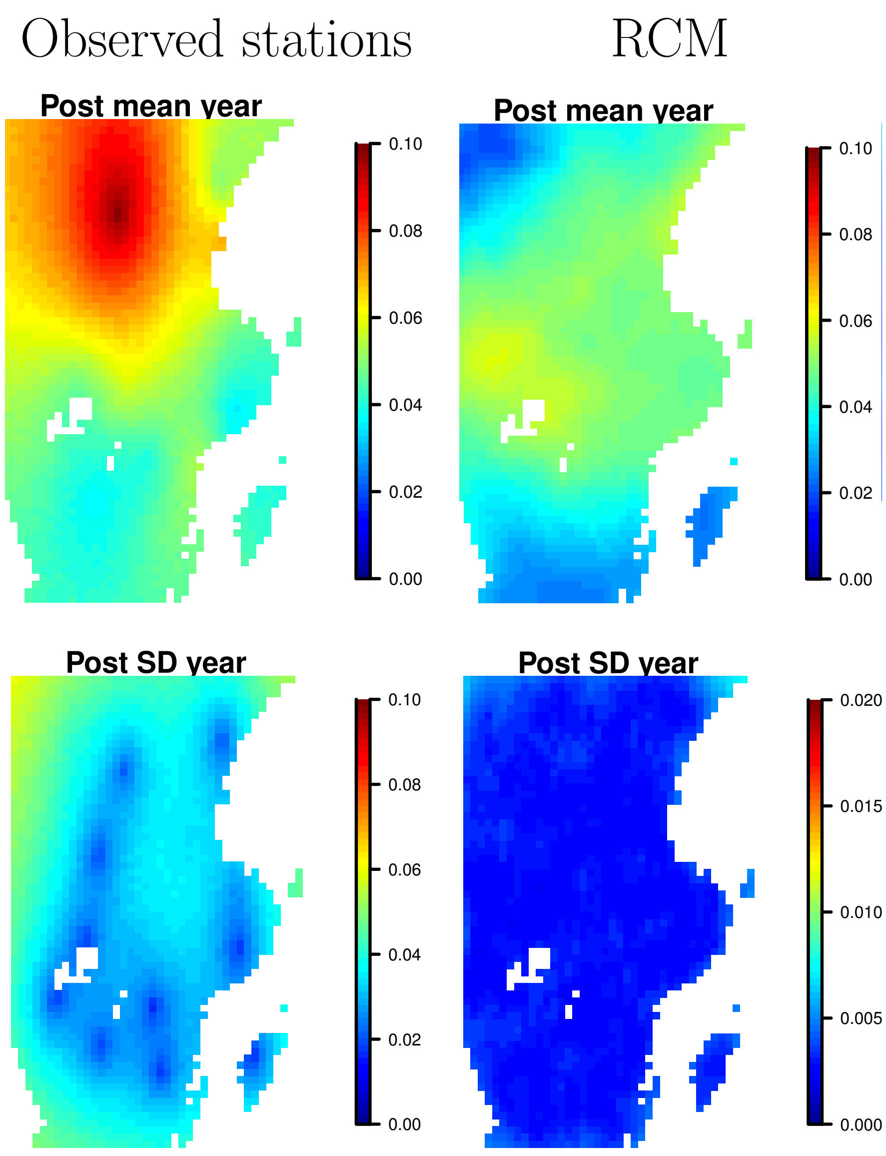
We fit our spatio-temporal GEV model to both the observational data and the RCM output.

	Observed stations		RCM	
Parameter	Post. mean	95% CI	Post. mean	95% CI
ξ	-0.18	(-0.21, -0.15)	-0.14	(-0.15, -0.14)

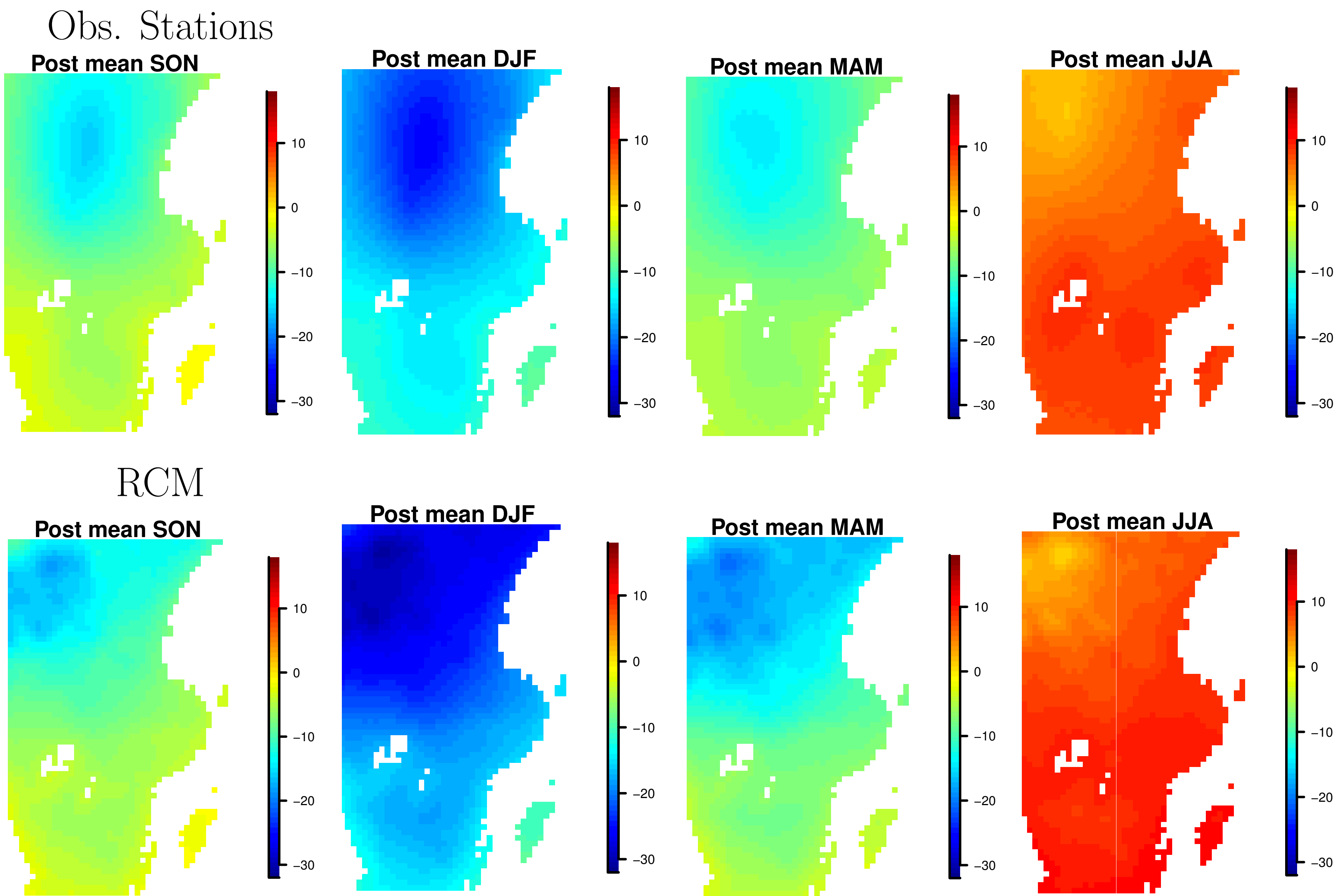
Shape parameters, ζ, are similar in both models.

Both parameters negative: hence distribution of seasonal minima is bounded.

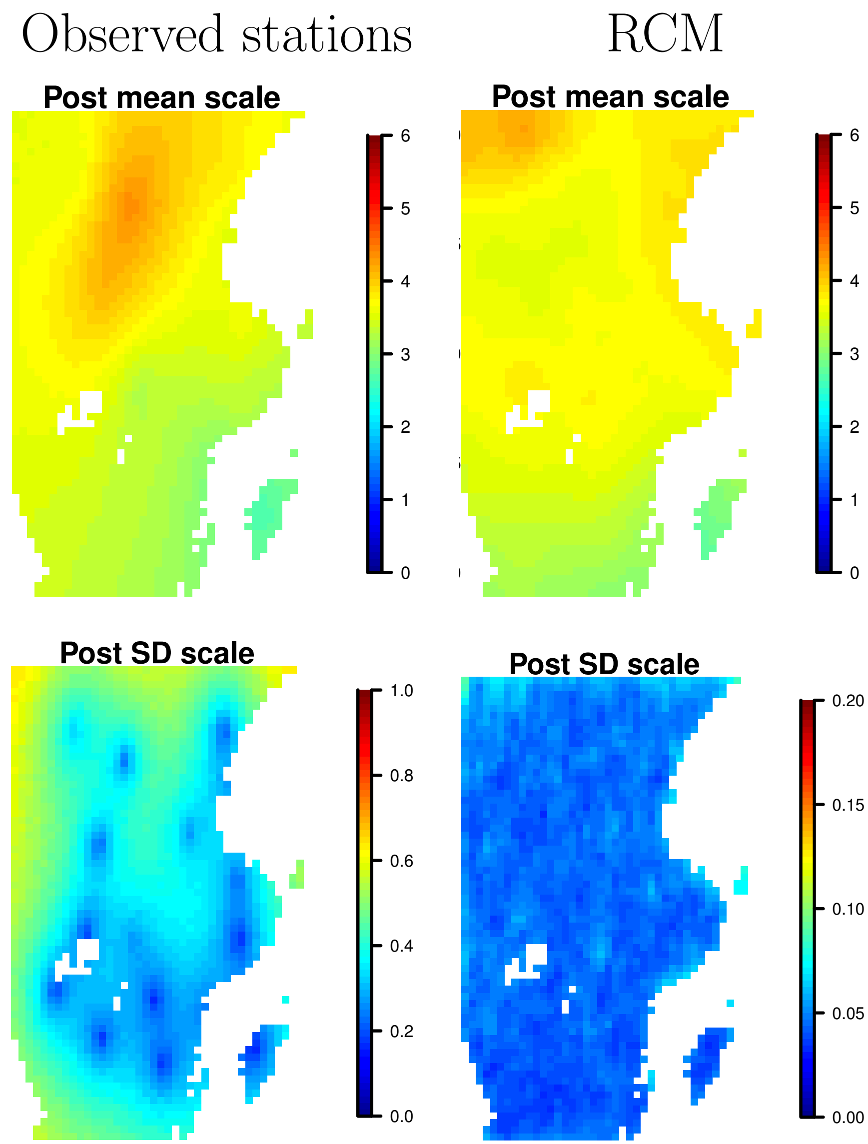
GEV model results: temporal trends



GEV model results: seasonal effects



GEV model results: scale parameters



Summary of GEV modeling results

1. Increasing trend in seasonal minima – warming underestimated in the RCM.

   For observations, change per year ranges from 0.04–0.10 ^{o}C.

2. Seasonal patterns in observations and RCM output are similar in direction.

   RCM too cold in DJF.

3. Differences in the spatial distribution of the scale parameter.

Thinking about change of support

In Berrocal et al. [2012] we spent considerable time building statistical models that **respect** the spatial and temporal scales of the observational data and climate model output.

Key idea: (I'll draw a doodle in the talk!)

Change of support for extremes

For RCM output, what does the seasonal minimum calculated for a **grid box** represent?

- Is it the minimum over the grid box region, or the minimum at the centroid?

We could answer these questions by **simulating minima** (based on the observation model) at a **finer** spatial scale than the regional climate model output, and then changing support by calculating minima at the grid box level.

But requires the use of **multivariate** models for extremes...

Joint modeling of extremes

There are many interesting approaches. Some include:

1. Conditional modeling [e.g., Ledford and Tawn, 1996, Heffernan and Tawn, 2004]
2. Max-stable approach [e.g. Davison et al., 2012, Wadsworth and Tawn, 2012, Huser and Genton, 2016]
3. Copula-based approaches [e.g., Ghosh and Mallick, 2011, Fuentes et al., 2013]

Multivariate inference is computationally demanding [e.g. Ribatet et al., 2012]. There is a question of whether it would be even be possible for this RCM.

Discussion: Why are we assessing RCMs?

Do we believe that observations are the best information about the truth?

- Are we using these observations (in some sense) to grade how well RCMs perform?

  See Heaton et al. [2014] for a non-trivial example of how one might grade a computer model.

What are the RCMs being used for?

[e.g. Jun, 2017]

- To learn about the spatio-temporal dynamics of the climate system
- To compare different climate models [e.g. Smith et al., 2009, Sain et al., 2010]
- To distinguish between different forcing scenarios [e.g. Tingley et al., 2015]
- To blend model outputs [e.g. Kang et al., 2012]
- For projections [e.g. Poppick et al., 2016]
- For detection and attribution [e.g. Hegerl et al., 2000]

How are we assessing regional climate model assessment

- Do we just compare the observations and RCMs directly? [e.g. Craigmile and Li., 2017]

- Do we build models independently on observations and RCMs and compare the parameters from each model?

- Do we regress one on the other, trying to learn about possible associations? [e.g. Sain et al., 2010, Braverman et al., 2016].

- What about joint modeling, rather than conditional modeling? [e.g. Philbin and Jun, 2015]

(The statistical modeling gets more complex as we move down the page.)

Challenges moving forward

- Building statistical models that capture the important features of climate.
- (Head-nod to Jim Zidek) how do we test the implementation of the statistical models?
- Handling massive datasets: model complexity versus computation
- Change of support
- How do we diagnose and comparing statistical models used for assessing RCMs?
- Uncertainty quantification (e.g., understanding the uncertainty of spatial features – for example, the peaks and troughs over space of temperature minima).

References

V. Berrocal, P. Craigmile, and P. Guttorp. Regional climate model assessment using statistical upscaling and downscaling techniques. *Environmetrics*, 23:482–492, 2012.

A. Braverman, S. Chatterjee, M. Heyman, and N. Cressie. Probabilistic evaluation of competing climate models. Technical report, National Institute for Applied Statistics Research, University of Wollongong, Australia, 2016. URL http://ro.uow.edu.au/niasrawp/42/.

S. Coles. *An Introduction to Statistical Modeling of Extreme Values*. Springer-Verlag, 2001.

D. Cooley, D. Nychka, and P. Naveau. Bayesian spatial modeling of extreme precipitation return levels. *Journal of the American Statistical Association*, 102: 824–840, 2007.

P. F. Craigmile and P. Guttorp. Space-time modelling of trends in temperature series. *Journal of Time Series Analysis*, 32:378–395, Apr. 2011.

P. F. Craigmile and P. Guttorp. Can a regional climate model reproduce observed extreme temperatures? *Statistica*, 73:103–122, 2013.

P. F. Craigmile and B. Li. Instruments, proxies, and simulations: exploring imperfect measures of climate. *CHANCE*, 2017. To appear.

A. C. Davison, S. A. Padoan, and M. Ribatet. Statistical modelling of spatial extremes. *Statistical Science*, 27:161–186, 2012.

K. A. Doksum. Empirical probability plots and statistical inference for nonlinear models in the two sample case. *Ann. Statist.*, 2:267–277, 1974.

K. A. Doksum and G. L. Sievers. Plotting with confidence: Graphical comparisons of two populations. *Biometrika*, 63:421–434, 1976.

M. Fuentes, J. Henry, and B. Reich. Nonparametric spatial models for extremes: Application to extreme temperature data. *Extremes*, 16:75–101, 2013.

S. Ghosh and B. K. Mallick. A hierarchical Bayesian spatio-temporal model for extreme precipitation events. *Enivronmetrics*, 22:192–204, 2011.

P. Guttorp and J. Xu. Climate change, trends in extremes, and model assessment for a long temperature time series from Sweden. *Environmetrics*, 22:456–463, 2011.

M. R. Haylock, N. Hofstra, A. M. G. Klein Tank, E. J. Klok, P. D. Jones, and co authors. European daily high-resolution gridded data set of surface temperature and precipitation for 1950–2006. *Journal of Geophysical Researh*, 113, 2008. doi: 10.1029/2008JD010201.

M. J. Heaton, W. Kleiber, S. R. Sain, and M. Wiltberger. Emulating and calibrating the multiple-fidelity Lyon-Fedder-Mobarry magnetosphere-ionosphere coupled computer model. *Journal of the Royal Statistical Society, Series C (Applied Statistics)*, 64:93–113, 2014.

J. E. Heffernan and J. A. Tawn. A conditional approach for multivariate extreme values. *Journal of the Royal Statistical Society, Series B*, 66:497–546, 2004.

G. C. Hegerl, P. A. Stott, M. R. Allen, J. Mitchell, S. Tett, and U. Cubasch. Optimal detection and attribution of climate change: Sensitivity of results to climate model differences. *Climate Dynamics*, 16:737–754, 2000.

R. Huser and M. G. Genton. Non-stationary dependence structures for spatial extremes. *Journal of Agricultural, Biological, and Environmental Statistics*, pages 1–22, 2016.

M. Jun. Climate model intercomparison. *CHANCE*, 2017. To appear.

E. L. Kang, N. Cressie, and S. R. Sain. Combining outputs from the North American Regional Climate Change Assessment Program by using a Bayesian hierarchical model. *Journal of the Royal Statistical Society, Series C (Applied Statistics)*, 61:291–313, 2012.

A. W. Ledford and J. A. Tawn. Statistics for near independence in multivariate extreme values. *Biometrika*, 83:169–187, 1996.

E. Mannshardt, P. F. Craigmile, and M. P. Tingley. Statistical modeling of extreme value behavior in North American tree-ring density series. *Climatic Change*, 117:843–858, 2013.

G. Nikulin, E. Kjellström, U. Hansson, G. Strandberg, and A. Ullerstig. Evaluation and future projections of temperature, precipitation and wind extremes over Europe in an ensemble of regional climate simulations. *Tellus A*, 63:41–55, 2011.

R. Philbin and M. Jun. Bivariate spatial analysis of temperature and precipitation from general circulation models and observation proxies. *Advances in Statistical Climatology, Meteorology and Oceanography*, 1:29, 2015.

A. Poppick, D. J. McInerney, E. J. Moyer, and M. L. Stein. Temperatures in transient climates: Improved methods for simulations with evolving temporal covariances. *The Annals of Applied Statistics*, 10:477–505, 2016.

M. Ribatet, D. Cooley, and A. C. Davison. Bayesian inference from composite likelihoods, with an application to spatial extremes. *Statistica Sinica*, 22:813–845, 2012.

W. B. Rossow, G. Tselioudis, A. Polak, and C. Jakob. Tropical climate described as a distribution of weather states indicated by distinct mesoscale cloud property mixtures. *Geophysical Research Letters*, 32, 2005. doi: 10.1029/2005GL024584.

S. R. Sain, D. Nychka, and L. Mearns. Functional anova and regional climate experiments: a statistical analysis of dynamic downscaling. *Environmetrics*, 22: 700–711, 2010.

P. Samuelsson, C. G. Jones, U. Willén, A. Ullerstig, S. Gollvik, U. Hansson, C. Jansson, E. Kjellström, G. Nikulin, and K. Wyser. The Rossby Centre regional climate model RCA3: model description and performance. *Tellus A*, 63:4–23, 2011.

H. Sang and A. E. Gelfand. Continuous Spatial Process Models for Spatial Extreme Values. *Journal of Agricultural, Biological and Environmental Statistics*, 15: 49–65, 2010.

R. L. Smith, C. Tebaldi, D. Nychka, and L. O. Mearns. Bayesian modeling of uncertainty in ensembles of climate models. *Journal of the American Statistical Association*, 104:97–116, 2009.

M. Tingley, P. F. Craigmile, M. Haran, B. Li, E. Mannshardt, and B. Rajaratnam. On discriminating between gcm forcing configurations using Bayesian reconstructions of late-holocene temperatures. *Journal of Climate*, 28:8264–8281, 2015.

S. Uppala, P. Kölberg, A. Simmons, U. Andrae, V. da Costa Bechtold, M. Fiorino, J. Gibson, J. Haseler, A. Hernandez, G. Kelly, X. Li, K. Onogi, S. Saarinen, N. Sokka, R. Allan, E. Andersson, K. Arpe, M. Balmaseda, A. Beljaars, L. van de Berg, J. Bidlot, N. Bormann, S. Caires, F. Chevallier, A. Dethof, M. Dragosavac, M. Fisher, M. Fuentes, S. Hagemann, E. Hólm, B. Hoskins, L. Isaksen, P. Janssen, R. Jenne, A. McNally, J.-F. Mahfouf, J.-J. Morcrette, N. Rayner, R. Saunders, P. Simon, A. Sterl, K. Trenberth, A. Untch, D. Vasiljevic, P. Viterbo, and J. Woollen. The ERA-40 re-analysis. *Quarterly Journal of the Royal Meteorological Society*, 131:2961–3012, 2005.

S. Uppala, D. Dee, S. Kobayashi, P. Berrisford, and A. Simmons. Towards a climate data assimilation system: status update of ERA-Interim. *ECMWF Newsletter*, pages 12–18, 2008.

J. L. Wadsworth and J. A. Tawn. Dependence modelling for spatial extremes. *Biometrika*, 99:253–272, 2012.

J. Wang, Y. Han, M. L. Stein, V. R. Kotamarthi, and W. K. Huang. Evaluation of dynamically downscaled extreme temperature using a spatially-aggregated generalized extreme value (GEV) model. *Climate Dynamics*, pages 1–17, 2016.