Applications of Tail Dependence II: Investigating the Pineapple Express

Dan Cooley
Grant Weller
Department of Statistics
Colorado State University

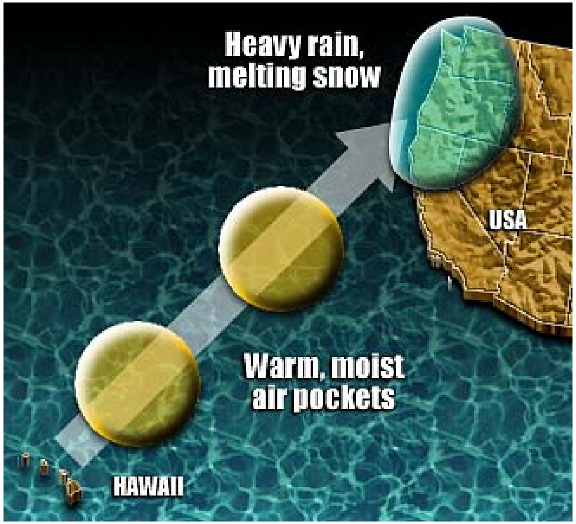
Joint work with:
Steve Sain, Melissa Bukovsky, Linda Mearns, NCAR

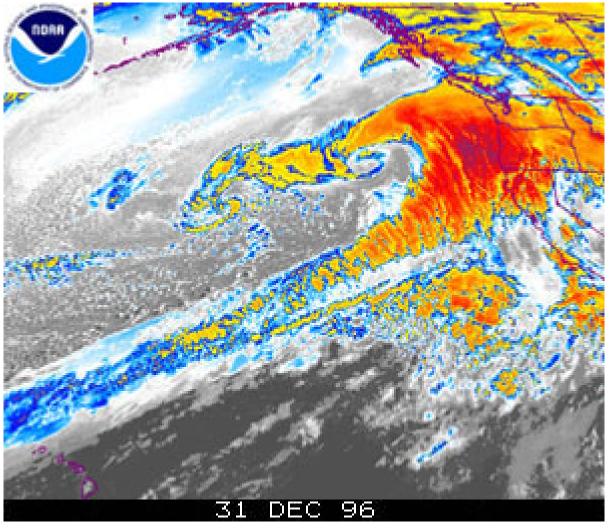
Acknowledgements:
Data provided by NARCCAP (NSF-ATM-0502977 and NSF-ATM-0534173)
GW supported by Weather & Climate Impact Assessment Program
GW & DC funded by NSF-DMS-0905315

What is the Pineapple Express?

PE storms: caused by atmospheric rivers hitting the west coast in winter

- Often bring heavy rain and warm temperatures
- Great impact on water resources of western US





This work aims to answer several questions related to this phenomenon:

Previous Work

- Dettinger (2004): List of PE events for 52 years.
- Leung and Qian (2009): Used an “index” approach to compare PE events in observations and model output.
- Other recent work looking at PE.

Overarching Question

What can be learned about the Pineapple Express by using extremes?

Previous Work

- Dettinger (2004): List of PE events for 52 years.
- Leung and Qian (2009): Used an "index" approach to compare PE events in observations and model output.
- Other recent work looking at PE.

Questions of Interest

1. Are regional climate models able to capture extreme PE precipitation events as seen in observational data?
2. Can we draw a connection between PE extreme precipitation events and short-lived synoptic-scale processes?
3. Given a future-scenario climate model run, what might extreme precipitation events look like in observations, and what is the uncertainty in these estimates?

Method: take an extreme-value approach for all questions.

Outline

1. Climate Models and Extremes Studies
2. Comparing RCM output extremes to observations
   - Study region and precipitation quantity
   - Modeling dependence in extremes
3. "Pineapple Express index"
   - North Pacific SLP fields
   - Toward a daily index of extremes
4. Examining future Pacific region precipitation extremes
   - Conditional simulation from dependence model
   - Future PE events & uncertainty
5. Summary and Future Work

Climate Models

Tools for simulating *weather* under different *climate* conditions. Discretized series of differential equations.

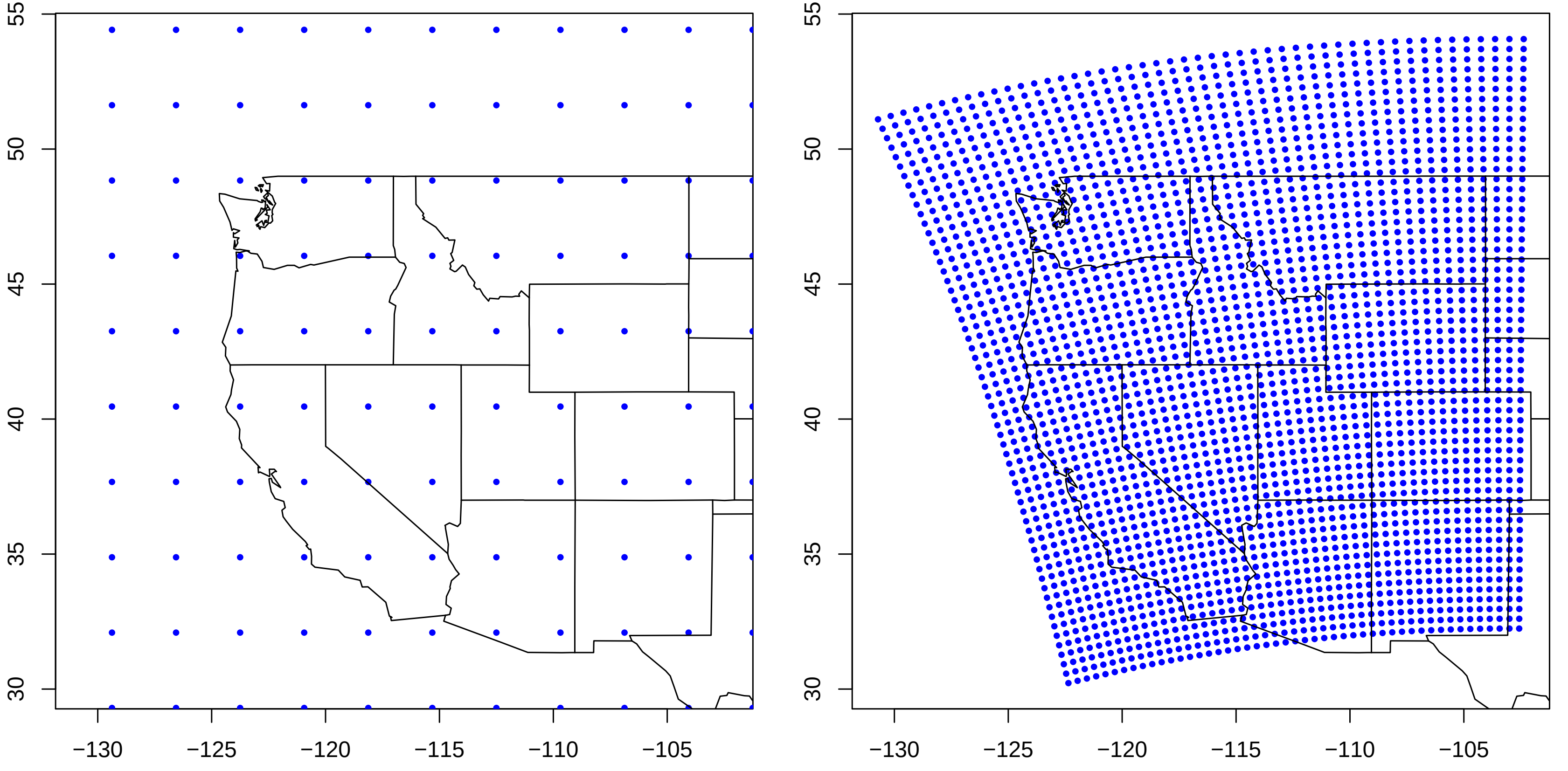
General Circulation Models or AOGCMs

- model large-scale processes over entire globe
- grid boxes on scale of 100's of km's

Regional Climate Models

- resolve smaller-scale processes over a region
- grid boxes on scale of 10's of km's
- driven by GCM's (or reanalysis)
- people would like to use them to assess local impacts—but *should* they be used?

Climate Models



Nature of Climate Model Predictions

The skeptics say:
"If models can't predict the weather more than 10 days in advance, how can models predict the climate 100 years from now?"

Nature of Climate Model Predictions

The skeptics say:
"If models can't predict the weather more than 10 days in advance, how can models predict the climate 100 years from now?"

The simulated weather generated by climate model runs is (usually) *not* supposed to correspond to the observed weather for a particular day.

Summary measures should be similar, although there is a downscaling issue.

Exception: reanalysis-driven RCM's *should* exhibit correspondence, as the reanalysis and the true state of the atmosphere should have similar synoptic-scale states.

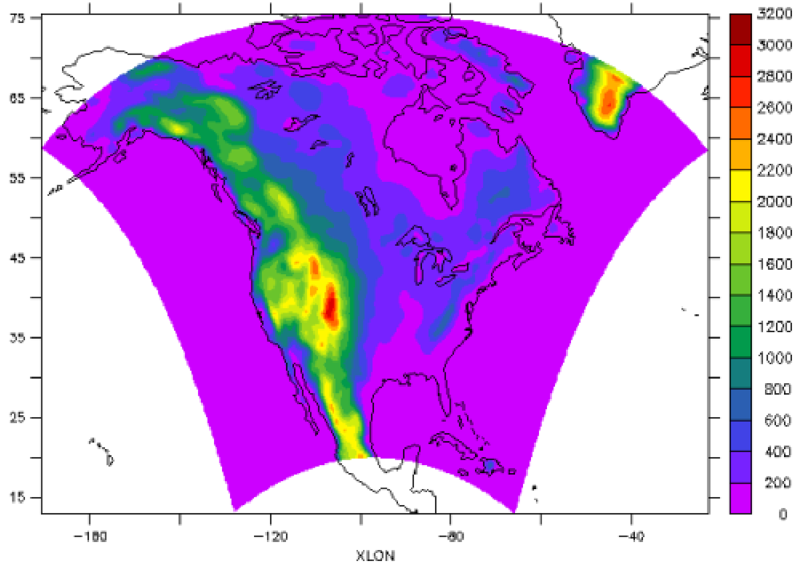
Our Studied RCMs: NARCCAP Project

Multi-model project to investigate uncertainties in regional climate model simulations.

	GCM			
RCM	CCSM	CGCM3	GFDL	NCEP
WRFG	X	X		X
ECP2			X	X
CRCM	X	X		X
MM5I	X			X
RCM3		X	X	X

- GCM-driven runs for current and future
- Reanalysis for current only
- Future runs (2041-2070): A2 emissions scenario.

Our Studied RCMs: NARCCAP Project



NARCCAP domain

Note that origin of PE moisture is outside the domain

Data and Model Output

We utilize several sources of climate model output and an observational product:

- Daily RCM precipitation output from NARCCAP - focus on WRF model
- NCEP/NCAR global reanalysis
- Daily gridded observational precipitation from University of Washington (*Maurer et al.*)
- Future run: WRF forced by CCSM global model

We study NDJF days from 1981-1999 (‘current’) and 2041-2070 (‘future’).

Question 1

Original question:
"Are regional climate models able to capture extreme PE precipitation events as seen in observational data?"

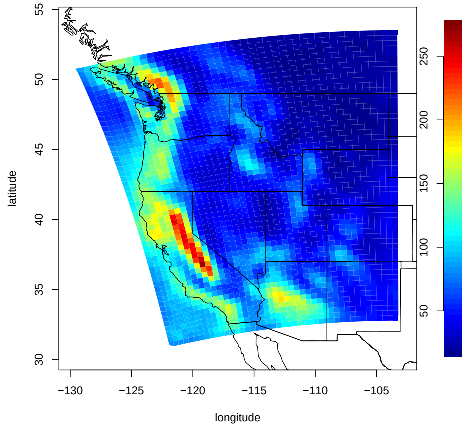
Reformulated:
"When large-scale conditions are such that PE events occur, does the RCM generate similar events?"

Aim: To assess the *correspondence* between extreme precipitation observed in the observational record and as produced by the reanalysis-driven climate model. Is there tail dependence? This is a "weather" study.

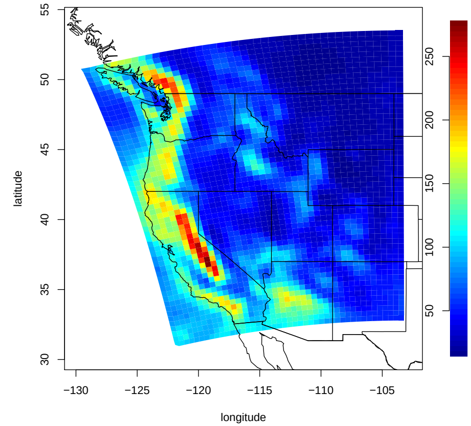
This differs from most previous extremes studies of climate model output which have been climatological.

Climatological Studies: Return Levels Estimates

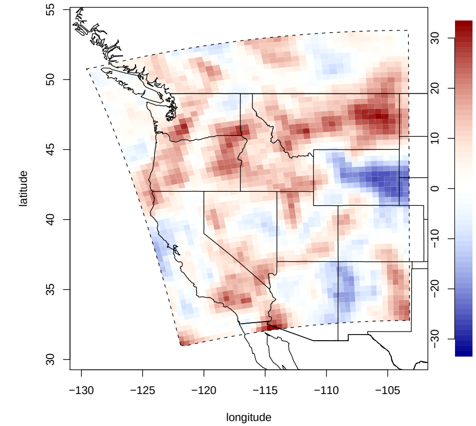
Control



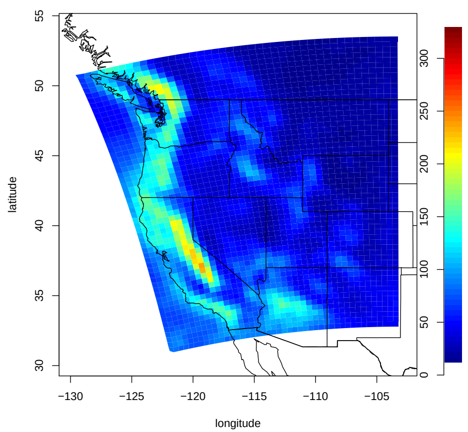
Future



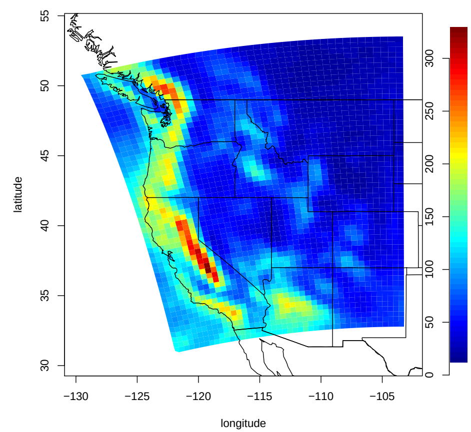
% Change



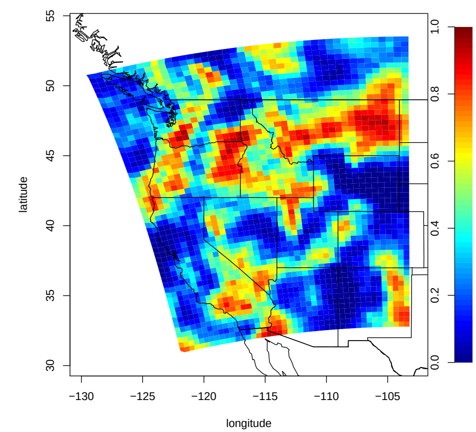
0.05 Quantile



0.95 Quantile

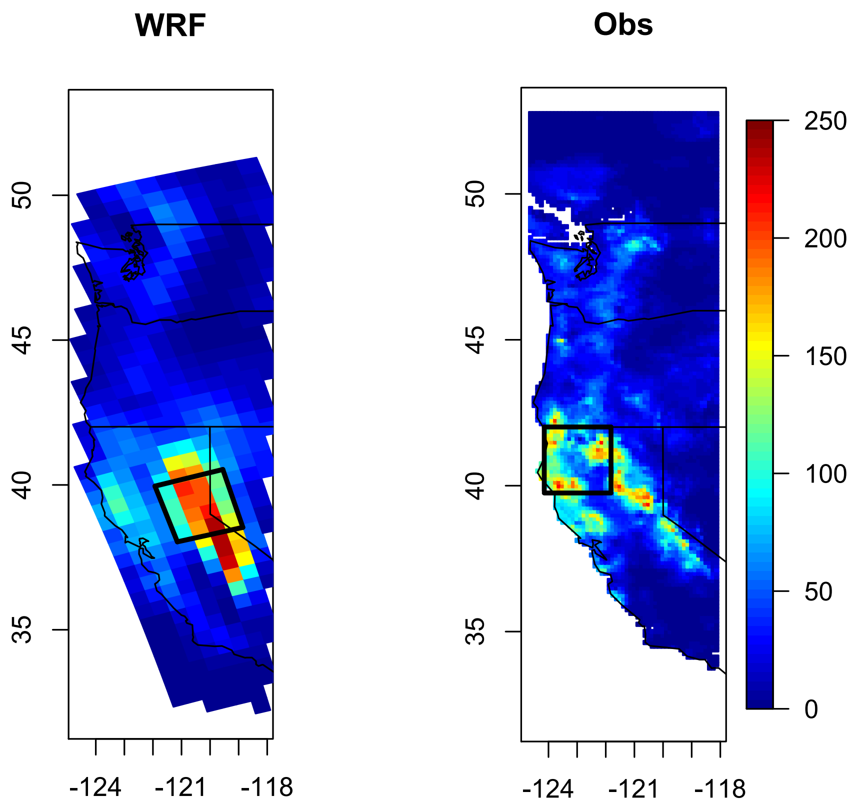


$\mathbb{P}$(10% increase)



Comparing WRF model output to observations

We define a study region and quantity with the purpose of capturing PE events identified by Dettinger et al. (2011).



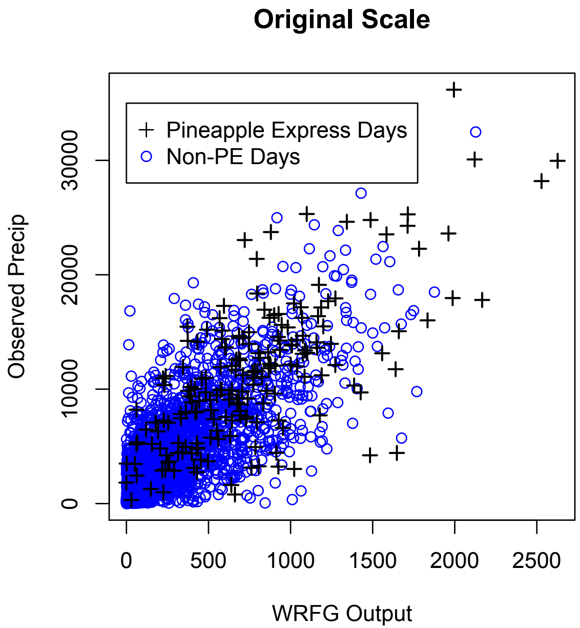
Precipitation from WRF-reanalysis output (left) and observational data product (right) on January 1, 1997.

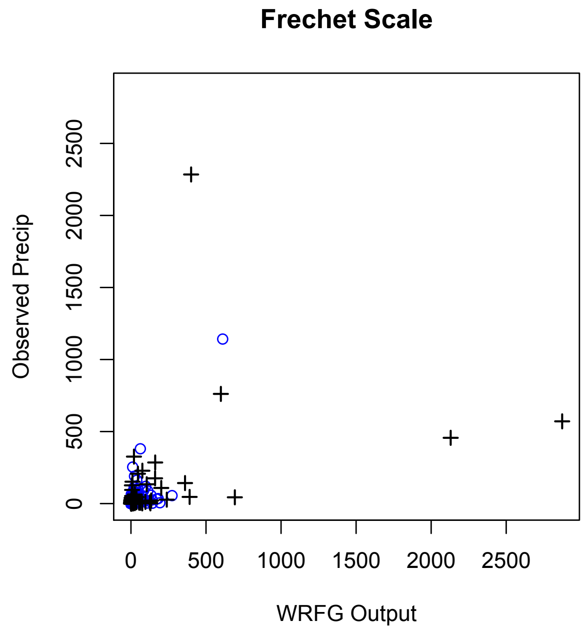
Estimation of marginal tails

GPDs are fit to the largest 5% of data in each margin:

Margin	$u_.$	$\hat{\psi}_.$ (se)	$\hat{\xi}_.$ (se)
X_t^{NC} (WRF)	1054	288.95(39.27)	0.0255(0.104)
Y_t^C (obs)	14240	3895.87(512.03)	0.0213(0.099)

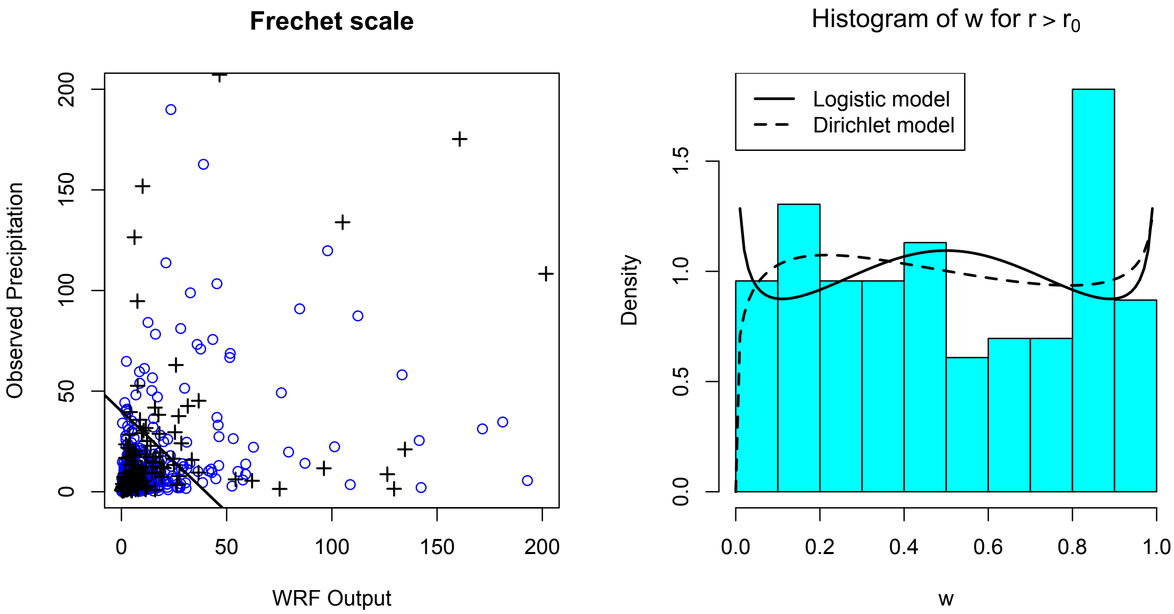
Each margin is transformed to unit Fréchet:





Examining tail dependence

We find tail dependence and fit a parametric model to the angular density of points with large 'radial' components.



+ WRF reproduces extreme events relatively well

– Not all 'extreme' events associated with Pineapple Express: aim to connect to synoptic-scale processes

Question 2

Original question: Can we draw a connection between PE extreme precipitation events and short-lived synoptic-scale processes?

Reformulated:

- Can we better understand the conditions which lead to a PE event? Can we capture those conditions in a number (index)?
- Can PE event conditions be seen on the GCM scale?
- If Dettinger hadn't made his list, could we identify PE events from GCM conditions?
- Will our index exhibit tail dependence?
- Asking *process* questions about extremes rather than *descriptive* ones.

The idea of tying extreme events to large scale processes is not new...

Regression approaches for extremes

An approach often used to understand the connection between extreme events and driving conditions is to allow the parameters of an extremes distribution to depend on covariates.

$\text{GEV}(\mu(\boldsymbol{x}), \sigma(\boldsymbol{x}), \xi(\boldsymbol{x}))$
$\text{GPD}(\psi(\boldsymbol{x}), \xi(\boldsymbol{x}))$

Atmospheric scientists *love* this approach.

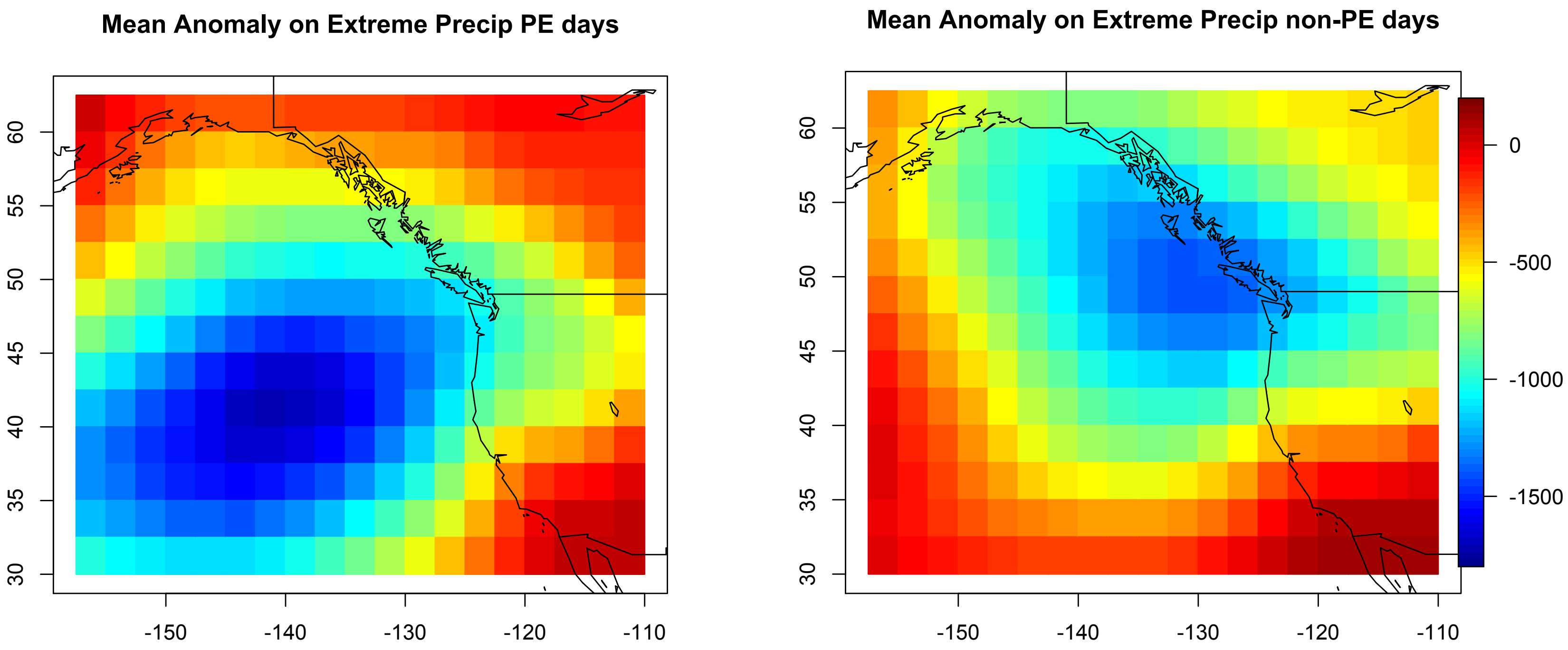
- Sillmann et al. (2011)
- Maraun et al. (2011)

They are in danger of loving it to death.

Because the large scale index is on a daily time scale, such an approach would be inappropriate for our study.

PE seen through pressure fields

Mean sea-level pressure fields are extracted from the NCEP reanalysis product

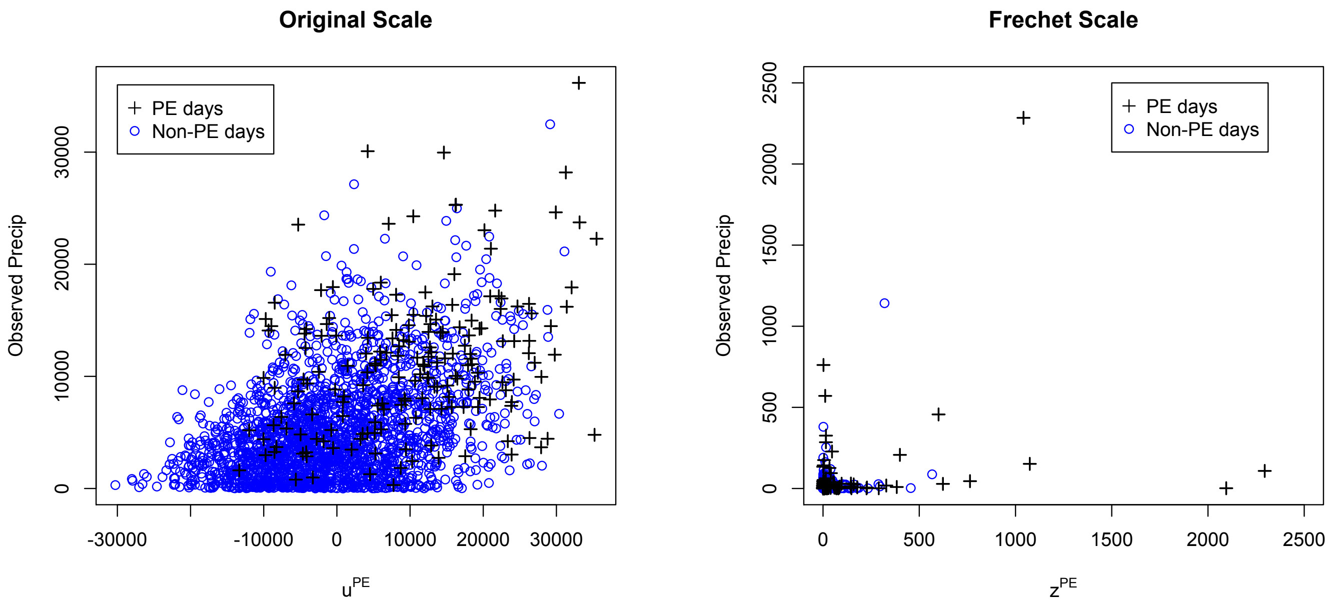


Composite anomaly fields for largest 130 observed precipitation days, partitioned into PE and non-PE

Pineapple Express Index

Define the index for day t as the projection of that day's SLP anomaly field onto the 'PE anomaly' field:

$$U_t^{PE} =: \mathbf{M}_t \cdot \boldsymbol{\mu}_{PE}$$



- Exhibits positive correlation *and* tail dependence with observed precipitation

Future PE Extremes

Question 3: Given a future-scenario climate model run, what might extreme precipitation events look like in observations, and what is the uncertainty in these estimates?

Reminder: Predicting the future is hard!

We analyze precipitation output from WRF driven by CCSM global model (2041-2070).

- Previous studies suggest increases in frequency and intensity of PE under A2.

Future PE Extremes

Aim: use fitted dependence model and PE index to simulate future observed precipitation extremes, given climate model output

Data Product	Current	Future
Observations	X	?
Reanalysis-Driven RCM	X	?
GCM-Driven RCM	X	X
PE Index	X	X

Challenge: we need to estimate

1. Marginal distribution of future reanalysis-driven precipitation
2. Marginal distribution of future observations

Extremes from the NARCCAP ensemble

Use other NARCCAP model combinations to infer the upper tail of future reanalysis-driven WRF precipitation:

	GCM			
RCM	CCSM	CGCM3	GFDL	NCEP
WRFG	X	X		X
ECP2			X	X
CRCM	X	X		X
MM5I	X			X
RCM3		X	X	X

For each RCM-GCM-time combination, obtain ML estimates and standard errors of GPD parameters.

Estimating future reanalysis-driven WRF

An 'ANOVA-like' model on the *parameters* of the GPD:

$$\begin{pmatrix}\psi_{ijr}\\ \xi_{ijr}\end{pmatrix} = \begin{pmatrix}\mu_{\psi}\\ \mu_{\xi}\end{pmatrix} + \begin{pmatrix}\alpha_{i\psi}\\ \alpha_{i\xi}\end{pmatrix} + \begin{pmatrix}\beta_{j\psi}\\ \beta_{j\xi}\end{pmatrix} + \begin{pmatrix}\gamma_{\psi}\\ \gamma_{\xi}\end{pmatrix} \mathbf{1}_{\{r=2\}}(r) + \epsilon_{ijr}$$

- $\alpha_i =$ effect of RCM i, $i = 1, ..., 5$
- $\beta_j =$ effect of GCM j, $j = 1, ..., 4$ ($4 =$ reanalysis)
- $\gamma =$ difference between current and future
- ϵ_{ijr} incorporates numerically estimated covariances

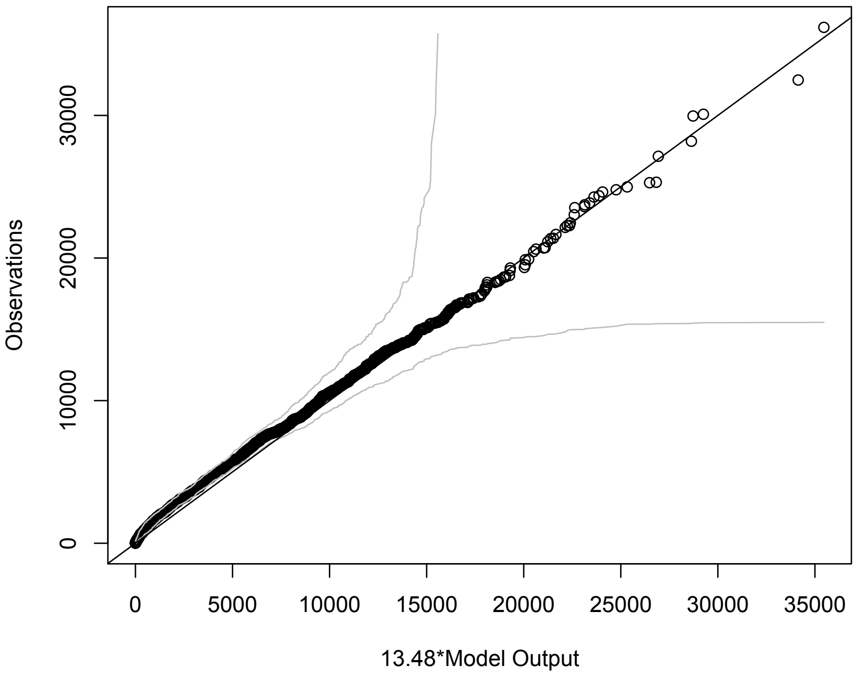
Estimates:

- $\widehat{\beta}_{4\xi} = 0.150 \Rightarrow$ NCEP-driven RCM runs produce *heavier tail* of precipitation than GCM-driven runs
- $\widehat{\gamma}_{\xi} = 0.057$: evidence for heavier-tailed precipitation in A2 scenario (WRF 100-year event becomes 36.3-year event)

What about future observations?

Exploit the relationship between reanalysis-driven WRF and observations:

NOT a scatterplot, but a qq plot



$$Y_t^F \stackrel{d}{=} aX_t^{NF}$$

Conditional Simulation of Extremes

Limiting Poisson process has density

$$\nu(dr \times dw) = r^{-2} h(w) dw$$

- Assumed to hold for large r
- Dirichlet model for $h(w)$ estimated by ML

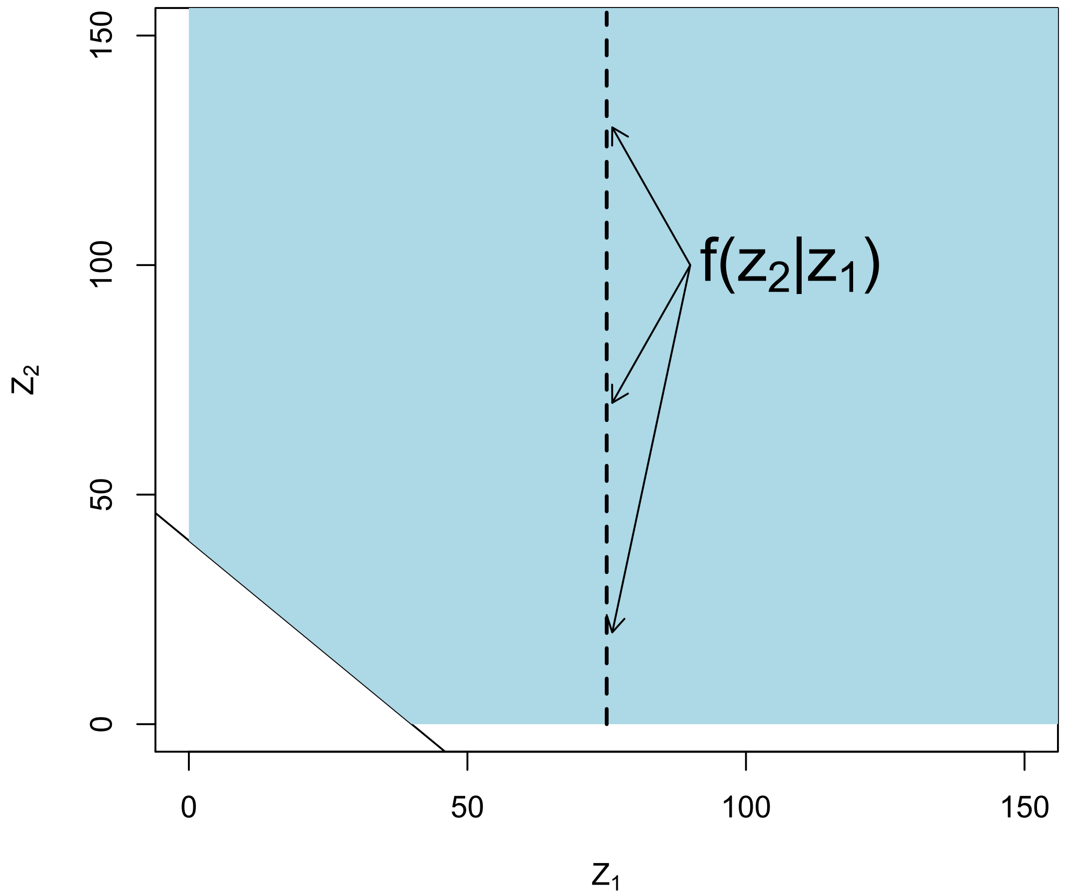
In Cartesian coordinates:

$$\nu(d\mathbf{z}) = \|\mathbf{z}\|^{-3} h(\mathbf{z}\|\mathbf{z}\|^{-1}) d\mathbf{z}$$

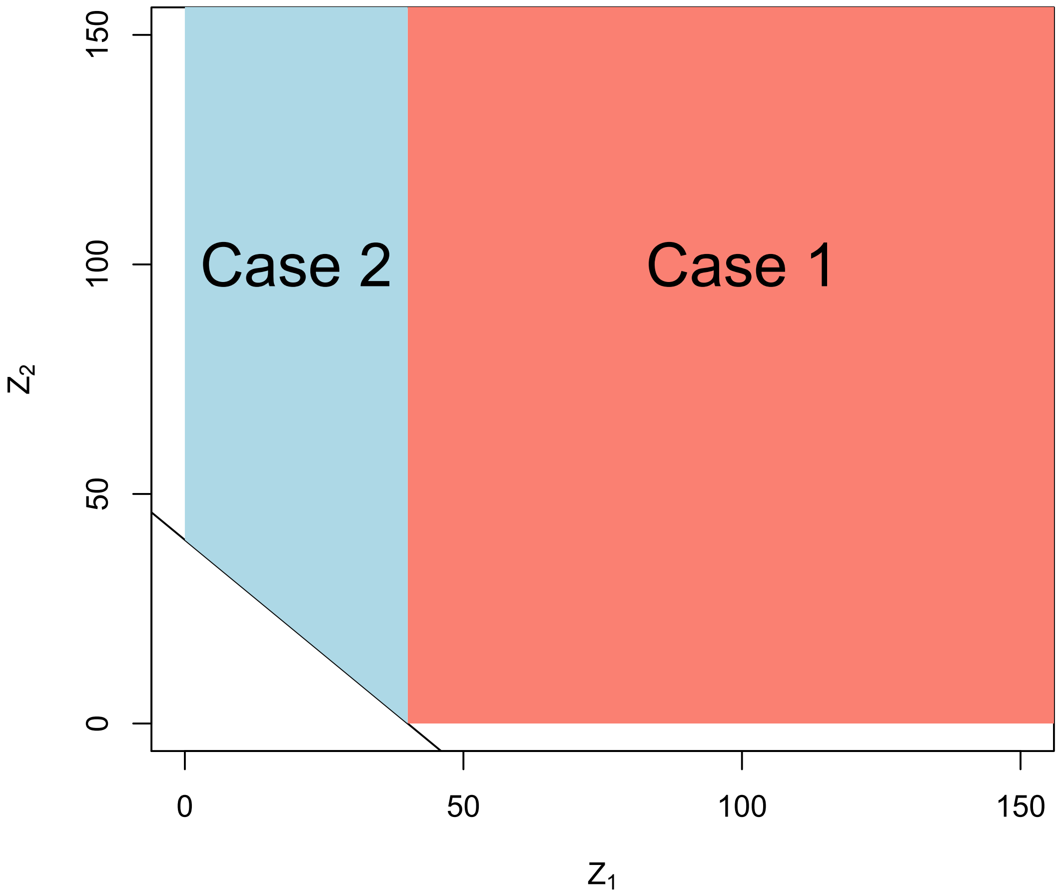
for $\|\mathbf{z}\|$ large.

Want to find the *conditional* density of Z_2 (observed precipitation), given Z_1 (climate model output).

Conditional Simulation of Extremes



Two cases



Two cases

Case 1: The full conditional density $f_{Z_2|Z_1=z_1}(z_2)$ can be approximated if z_1 is 'extreme'.

Case 2: The upper tail of the conditional density can be approximated even if z_1 is not 'extreme': for $z_2 > r^*$

$$\begin{aligned} f_{Z_2|Z_1=z_1}(z_2) \approx \mathbb{P}(Z_2 \in (z_2, z_2+dz)|Z_1 = z_1) \\ = \mathbb{P}(Z_2 \in (z_2, z_2+dz)|Z_1 = z_1, Z_2 > r^*) \\ \cdot \mathbb{P}(Z_2 > r^*|Z_1 = z_1) \end{aligned}$$

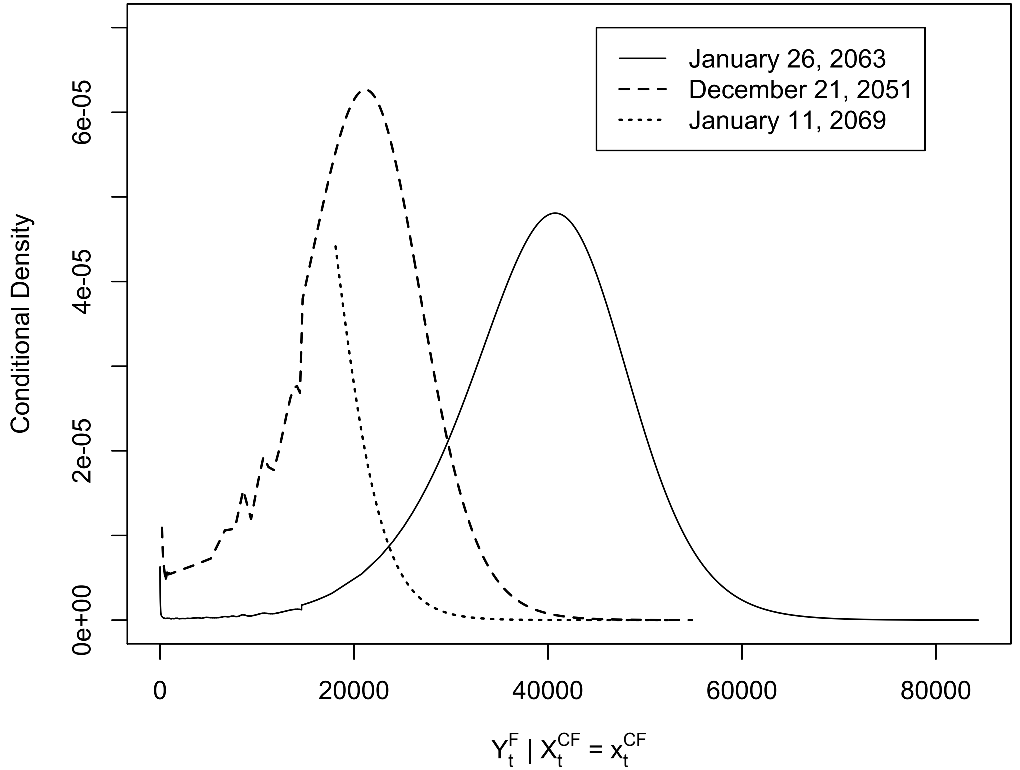
Bayes' rule:

$$\mathbb{P}(Z_2 > r^*|Z_1 = z_1) = \frac{\mathbb{P}(Z_1 = z_1|Z_2 > r^*) \cdot \mathbb{P}(Z_2 > r^*)}{\mathbb{P}(Z_1 = z_1)}$$

Works because marginal distributions are Fréchet (known).

Individual event studies

Conditional densities of future events given WRF-CCSM



We can then draw from these conditional densities.

Procedure review

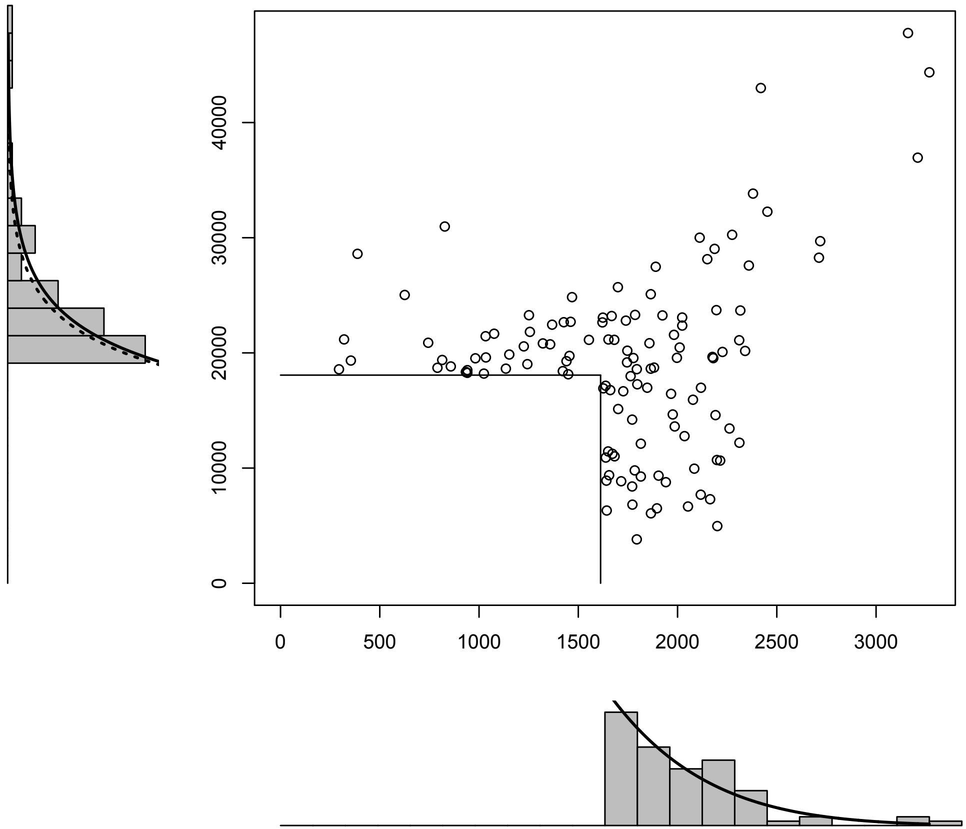
1. observe the CCSM-driven RCM output.
2. transform this to look like NCEP-driven RCM output.
3. use the tail dependence model to obtain the conditional density of the observations when the RCM output is large or the tail of the conditional density of the observations when RCM output is small.
4. could also draw from this density to "simulate" observations under the future climate.

---

5. Separately analyze the PE index produced by the CCSM model.

Simulation of observations

Repeated simulation gives uncertainty estimates based on how RCM represents extreme events.



x-axis: WRF-CCSM output. y-axis: simulated observations

PE Index in future scenario

Index is derived from Pacific SLP anomaly fields from driving GCM (CCSM)

- 8.9% of future period index values exceed 0.95 quantile of current period values - $\nearrow$ *frequency of PE events*
- Tail dependence of PE index with RCM output precipitation increases - $\nearrow$ *intensity of PE precipitation(?)*

Tail dependence with observed precipitation - studied through conditional simulation

Uncertainty through repeated simulation

We examine two quantities of interest through simulation:

- q_1: Proportion of 'extreme' observations that correspond to 'extreme' PE index values
- q_2: Proportion of 'extreme' observations occurring in years 2055-2070 (measure of nonstationarity)

Quantity	Estimate[1]	95% Interval[1]
q_1	0.203*	$(0.144, 0.257)$
q_2	0.571	$(0.477, 0.656)$

[1] Based on 500 conditional simulations

*Value from current period: 0.143

Evidence for increased correspondence of PE events and extreme precipitation - more intense PE events

Summary

This work is a novel application of bivariate EVT in a climate study.

- Tail dependence between RCM output and observations - modeled this parametrically
- PE Index - derived from SLP fields; tail dependent to observed precipitation
- Conditional simulation from parametric model given future RCM output - uncertainty estimates

Future work

Important to remember that we have studied one RCM, driven by one GCM, and compared it to one observational product.

- Improvement of the PE index - storms evolve over several days
- Applying methodology to other climate models
- Examining other regions/phenomena?

References

Cooley, D., Davis, R. A., and Naveau, P. (2012). Approximating the conditional density given large observed values via a multivariate extremes framework, with application to environmental data. *Annals of Applied Statistics*, 6:1406–1429.

Dettinger, M. (2004). Fifty-two years of pineapple-express storms across the west coast of North America. *US Geological Survey, Scripps Institution of Oceanography for the California Energy Commission, PIER Project Rep. CEC-500-2005-004*.

Dettinger, M., Ralph, F., Das, T., Neiman, P., and Cayan, D. (2011). Atmospheric rivers, floods and the water resources of California. *Water*, 3(2):445–478.

Leung, L. and Qian, Y. (2009). Atmospheric rivers induced heavy precipitation and flooding in the western US simulated by the WRF regional climate model. *Geophysical Research Letters*, 36(3):L03820.

Maraun, D., Osborn, T., and Rust, H. (2011). The influence of synoptic airflow on uk daily precipitation extremes. part i: Observed spatio-temporal relationships. *Climate Dynamics*, 36(1):261–275.

Sillmann, J., Croci-Maspoli, M., Kallache, M., and Katz, R. W. (2011). Extreme cold winter temperatures in europe under the influence of north atlantic atmospheric blocking. *Journal of Climate, to appear*, XX:XX–XX. doi: 10.1175/2011JCLI4075.1.

Weller, G., Cooley, D., and Sain, S. (2012). An investigation of the pineapple express phenomenon via bivariate extreme value theory. *Environmetrics*, 23:420–439.