Joint Extremes in Temperature and Mortality: A Bivariate POT Approach

Han Li[†] and Qihe Tang[*]

[†] Department of Actuarial Studies and Business Analytics
Macquarie University

[*] School of Risk and Actuarial Studies
UNSW Sydney
[*] Department of Statistics and Actuarial Science
University of Iowa

One World Actuarial Research Seminar

23rd September 2020

Acknowledgment

This research is sponsored by the Society of Actuaries through the 2019 Individual Grant Competition, project titled "*An EVT Approach to Modeling Joint Extremes in Climate and Mortality*".

SOCIETY OF ACTUARIES MACQUARIE University

Outline of the talk

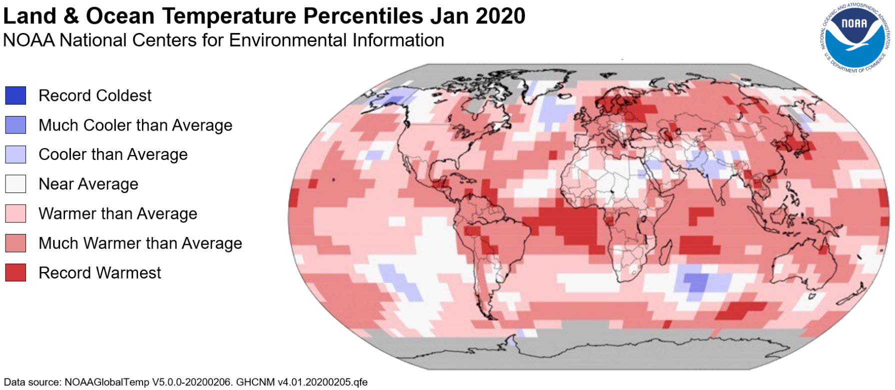
1. Motivation
2. Data Description and Modelling
   - Actuaries Climate Index
   - Mortality Data
3. Multivariate Extreme Value Theory
   - Fundamentals
   - Bivariate POT Theory
4. Empirical Results
5. Conclusions

Climate change quotes...

"*There's one issue that will define the contours of this century more dramatically than any other, and that is the urgent threat of a changing climate.*" - Barack Obama, speech to UN, 2014.

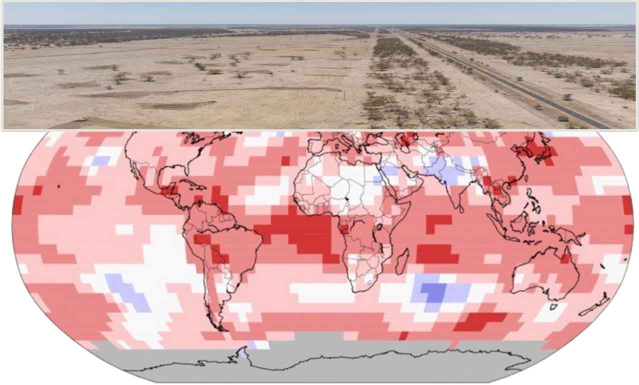
SOCIETY OF ACTUARIES   MACQUARIE University

Climate change = global warming?



Climate change = global warming?

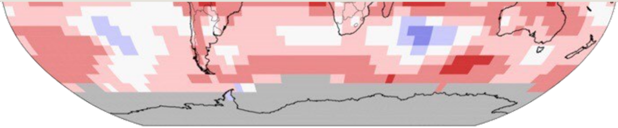
September 2019



Climate change = global warming?

September 2019

December 2019



Climate change or apocalypse?

September 2019

December 2019

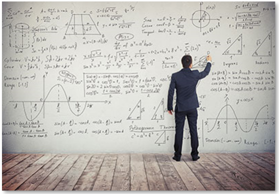
February 2020

SOCIETY OF ACTUARIES

MACQUARIE University

What can insurance do in a changing climate?

Unanticipated adverse claim experience due to climate change can lead to insolvency of insurance and reinsurance companies.



Enterprise Risk Management

Identify → Understand → Quantify → Mitigate

SOCIETY OF ACTUARIES

MACQUARIE University

How can climate change kill you?

According to the WHO:

Between **2030** and **2050**, climate change is expected to cause approximately **250,000** additional deaths **per year**.

Weather-related catastrophes (etc. floods, bushfires and earthquakes)

Extreme temperatures (etc. heat waves and cold spells)

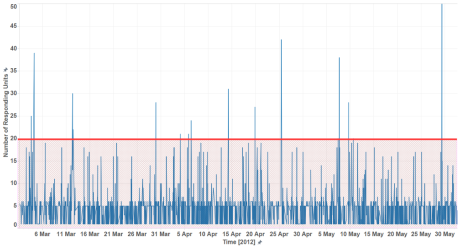
Climate-sensitive infectious diseases (etc. malaria disease and vector-borne diseases)

SOCIETY OF ACTUARIES

MACQUARIE University

Key challenge

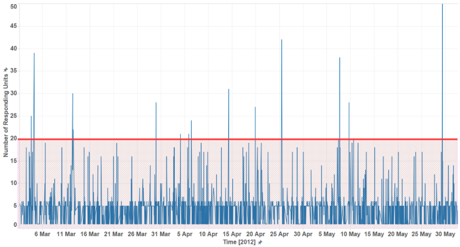
A key challenge in modeling extreme risks: scarcity of extreme observations.



Extreme value theory (EVT) tackles this problem by providing limiting results beyond observed values.

Key challenge

A key challenge in modeling extreme risks: scarcity of extreme observations.

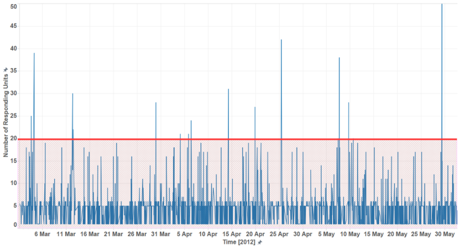


Extreme value theory (EVT) tackles this problem by providing limiting results beyond observed values.

- **Block Maxima**: distribution of the sample maximum

Key challenge

A key challenge in modeling extreme risks: scarcity of extreme observations.



Extreme value theory (EVT) tackles this problem by providing limiting results beyond observed values.

- **Block Maxima**: distribution of the sample maximum
- **Peaks Over Threshold**: distribution of values over a high threshold

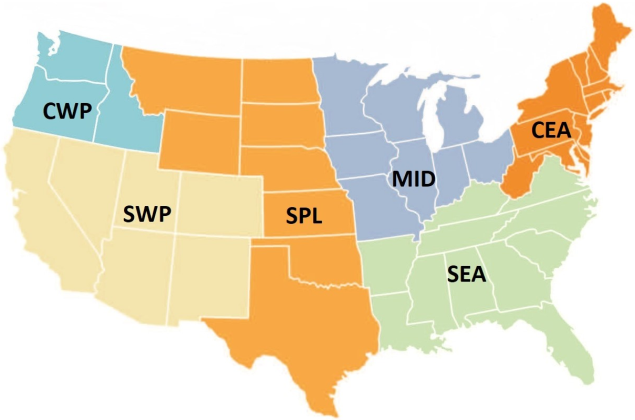
Data Description and Modeling

The ACI is developed jointly by the American Academy of Actuaries, the Casualty Actuarial Society, the Canadian Institute of Actuaries, and the Society of Actuaries. It consists of six components as follows:

- $T90$: Frequency of temperatures above the 90$^{\text{th}}$ percentile;
- $T10$: Frequency of temperatures below the 10$^{\text{th}}$ percentile;
- P: Maximum rainfall per month in five consecutive days;
- D: Annual maximum consecutive dry days;
- W: Frequency of wind speed above the 90$^{\text{th}}$ percentile;
- S: Sea level changes.

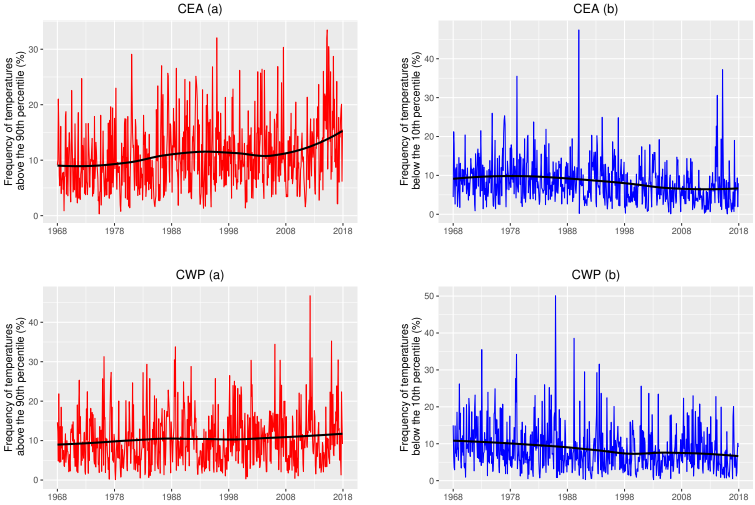
SOCIETY OF ACTUARIES MACQUARIE University

Data Description and Modeling

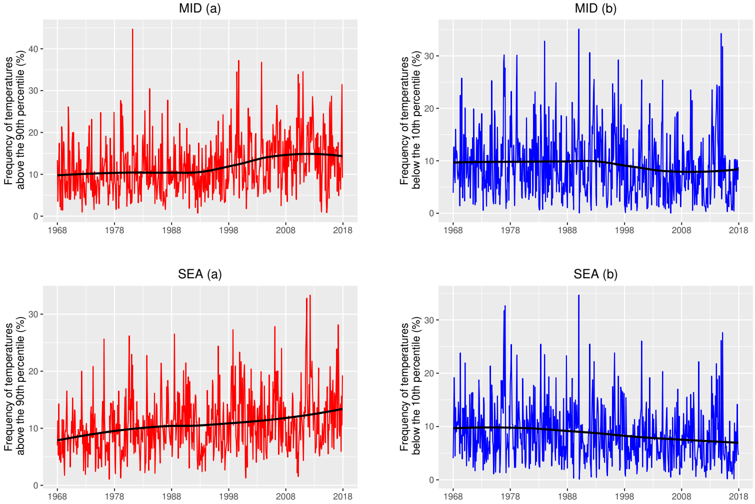


Source: Actuaries Climate Index Executive Summary (2018), Page 4, Figure 2

Data Description and Modeling

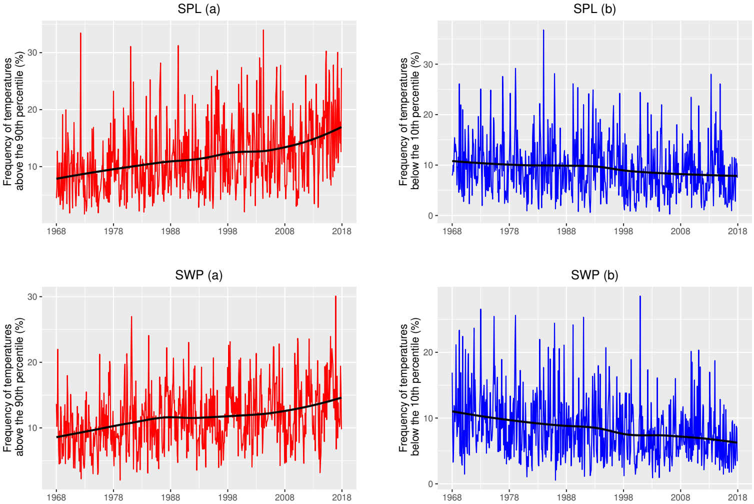


Data Description and Modeling



SOCIETY OF ACTUARIES MACQUARIE University

Data Description and Modeling



Data Description and Modeling

We adopt the seasonal ARIMA model which incorporates both non-seasonal and seasonal factors in a multiplicative model, which can be expressed as

$$\text{ARIMA}(p, d, q) \times (P, D, Q)_S, \tag{1}$$

where:

- p, d, and q denote the order of the AR model, the order of differencing, and the order of the MA model in the non-seasonal part, respectively,
- P, D, and Q denote the order of the AR model, the order of differencing, and the order of the MA model in the seasonal part, respectively, and
- S is the time span of repeating the seasonal pattern. Since we are modeling monthly $T90$ and $T10$ time series, S is set to be 12.

Data Description and Modeling

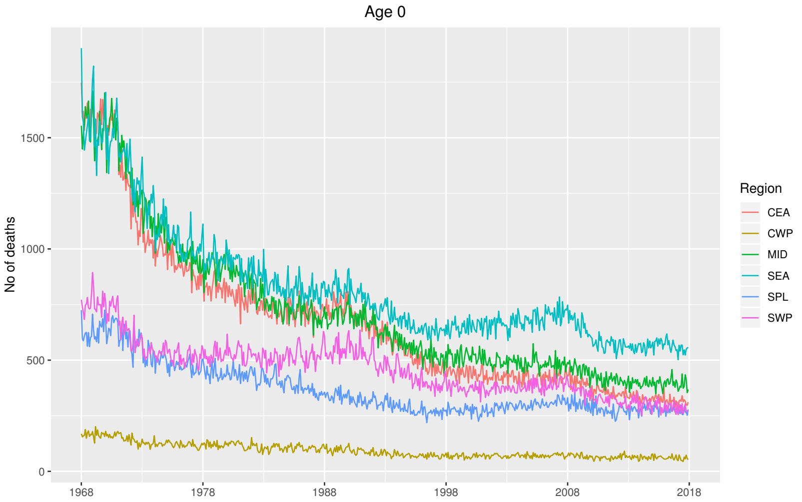
The U.S. regional-level death data for the time period 1968–2017 are obtained from two main sources listed as follows:

- *The National Center for Health Statistics (NCHS)*
- *Centers for Disease Control and Prevention (CDC) WONDER*

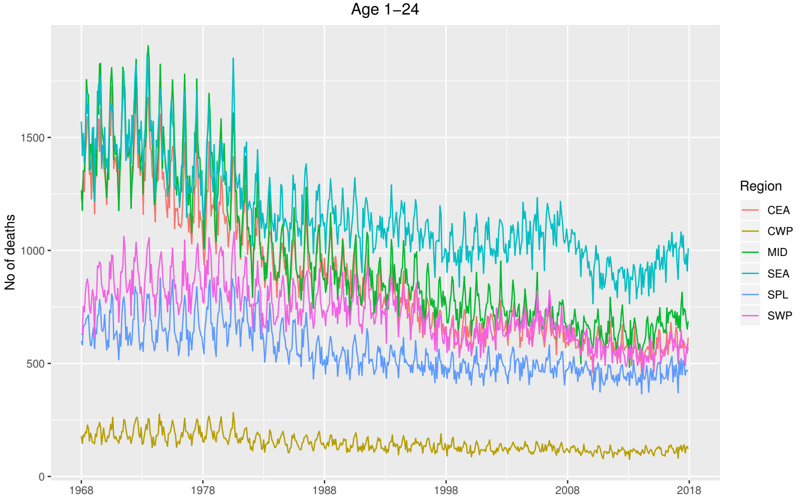
As there is no publicly available data on monthly age-specific population exposure, particularly at the state level, in this study we choose to directly model death counts rather than mortality rates.

SOCIETY OF ACTUARIES MACQUARIE University

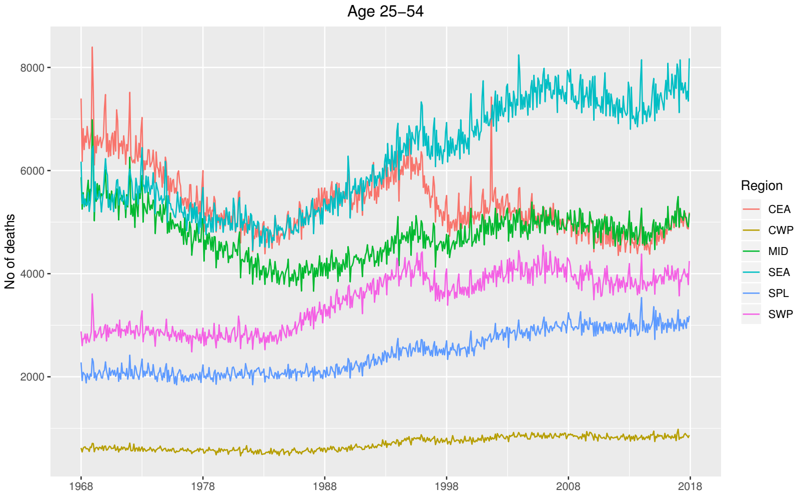
Data Description and Modeling



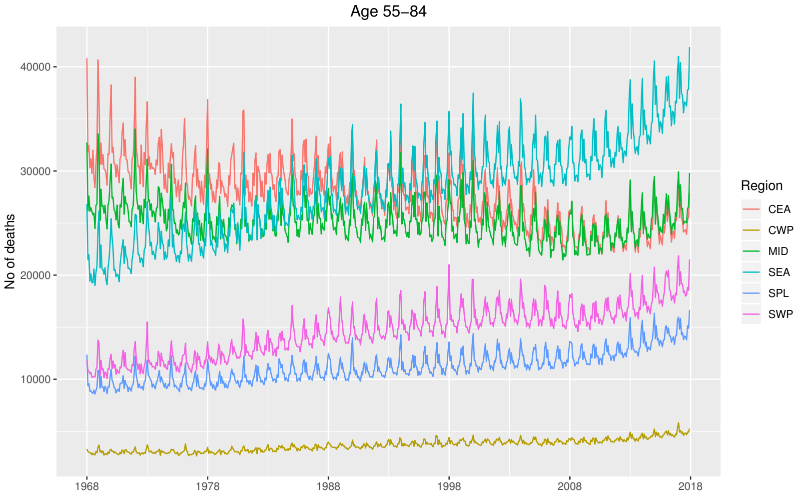
Data Description and Modeling



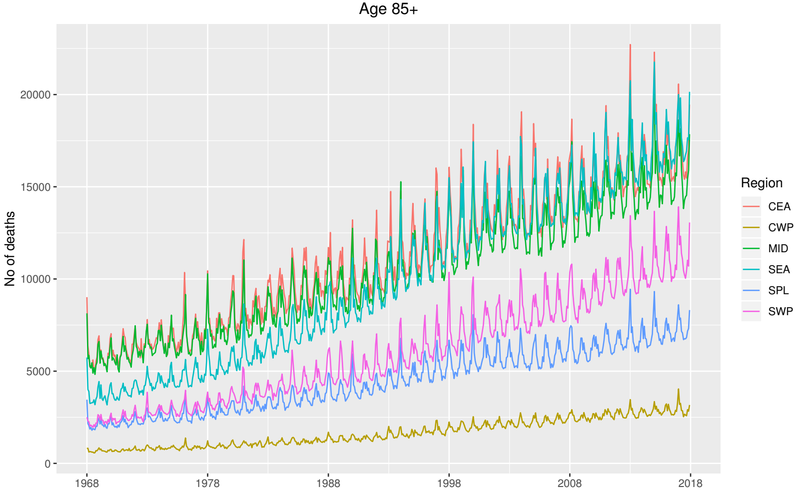
Data Description and Modeling



Data Description and Modeling



Data Description and Modeling



Data Description and Modeling

Similar to $T10$ and $T90$, we want to obtain the "noise" in death counts via time series models.

- We fit a seasonal ARIMA model first.
- We include the GARCH component if the resulting residuals fail the Ljung-Box test at 5% level of significance.
- The optimal model is selected based on AIC.

Multivariate Extreme Value Theory

Consider a random variable X with distribution F on $\mathbb{R}$ and denote by M_n the maximum of a sample of size n from F. If there exist norming constants $a_n > 0$ and $b_n \in \mathbb{R}$ such that

$$\lim_{n\to\infty} \Pr\left(\frac{M_n - b_n}{a_n} \leq y\right) = G(y), \qquad y \in \mathbb{R}, \tag{2}$$

then we say that F belongs to the max-domain of attraction of G, and call G a generalized extreme value (GEV) distribution. The GEV distribution function G must be of the same type as

$$G(y) = \exp\left\{-\left(1 + \gamma\frac{y-\mu}{\sigma}\right)_+^{-1/\gamma}\right\}, \tag{3}$$

where $\mu \in \mathbb{R}$, $\sigma > 0$, and $\gamma \in \mathbb{R}$ are the location, scale, and shape parameters, respectively, and $c_+ = \max(c, 0)$.

Multivariate Extreme Value Theory

Following the works of Balkema and de Haan (1974) and Pickands (1975), the conditional distribution of the normalized exceedance over a high threshold converges to a generalized Pareto distribution (GPD), that is,

$$\lim_{n\to\infty} \Pr\left(\frac{X-b_n}{a_n} \leq y \,\middle|\, X > b_n \right) = H(y), \qquad y > 0, \tag{4}$$

where H is of the same type as

$$H(y) = 1 - \left(1 + \gamma \frac{y-\mu}{\sigma}\right)_+^{-1/\gamma}, \tag{5}$$

with the location, scale, and shape parameters $\mu \in \mathbb{R}$, $\sigma > 0$, and $\gamma \in \mathbb{R}$. The GPD H above is supported on the region of y defined by $y > 0$ and $1 + \gamma \frac{y-\mu}{\sigma} > 0$.

Multivariate Extreme Value Theory

Consider a d-dimensional random vector $\boldsymbol{X}$ with distribution F on $\mathbb{R}^d$ and denote by $\boldsymbol{M}_n$ the component-wise maximum of a sample of size n from F. The limit distribution G, called a multivariate GEV distribution, has marginal distributions G_i for $1 \leq i \leq d$ identical to

$$\lim_{n\to\infty} \Pr\left(\frac{M_n^{(i)} - b_n^{(i)}}{a_n^{(i)}} \leq y\right) = G_i(y), \tag{6}$$

which therefore is of the same type as Equation (5).
In practice, it is common to first transform the marginal distributions to a particular distribution before fitting a multivariate GEV distribution. In this paper, we choose the unit Fréchet transformation

$$z = -\frac{1}{\log G_i(y)}. \tag{7}$$

SOCIETY OF ACTUARIES MACQUARIE University

Multivariate Extreme Value Theory

According to Propositions 5.10 and 5.11 in Resnick (1987), the representation of a multivariate GEV distribution with unit Fréchet margins can be written as

$$G(\boldsymbol{y}) = \exp\left\{-V(\boldsymbol{z})\right\}, \tag{8}$$

where $V(\cdot)$, the exponent measure, has a functional representation

$$V(\boldsymbol{z}) = \int_{S_d} \max_{1\leq i\leq d} \left(\frac{q_i}{z_i}\right) d\phi(\boldsymbol{q}), \tag{9}$$

with ϕ being a finite spectral measure on $S_d = \{\boldsymbol{q} \in \mathbb{R}^d : \|\boldsymbol{q}\| = 1\}$, and $\|\cdot\|$ representing an arbitrary norm in $\mathbb{R}^d$. We also impose a constraint such that, for $1 \leq i \leq d$,

$$\int_{S_d} q_i d\phi(q_i) = 1, \tag{10}$$

but beyond this the spectral measure ϕ is unknown.

Multivariate Extreme Value Theory

As in this study we focus on assessing the upper tail dependence between temperature and mortality, we adopt the symmetric logistic model for the function V, which is a natural candidate and a commonly used dependence model in bivariate POT studies (see *e.g.* Tawn, 1990; Coles *et al.*, 1999; Rootzén and Tajvidi, 2006). Under the symmetric logistic model,

$$V(z_1, z_2) = (z_1^{-r} + z_2^{-r})^{1/r}, \qquad r \geq 1, \tag{11}$$

which can be retrieved from Equation (9) with a suitably chosen spectral measure ϕ on S_2. The exponent measure $V(z_1, z_2)$ determines the strength of dependence between the two margins. In particular, independence is obtained when $r = 1$, and perfect dependence is obtained as $r \to \infty$.

Multivariate Extreme Value Theory

The multivariate POT theorem then states that, for a random vector $\mathbf{X}$ distributed by $F \in \mathrm{MDA}(G)$, assuming $0 < G(\mathbf{0}) < 1$ without loss of generality, the conditional distribution of $\boldsymbol{a}_n^{-1}(\boldsymbol{X} - \boldsymbol{b}_n)$ given $\boldsymbol{X} \not\leq \boldsymbol{b}_n$ converges to the multivariate GPD as

$$H(\boldsymbol{y}) = \frac{1}{-\log G(\mathbf{0})} \log \frac{G(\boldsymbol{y})}{G(\boldsymbol{y} \wedge \mathbf{0})}, \tag{12}$$

which is defined for all $\boldsymbol{y} \in \mathbb{R}^d$ such that $G(\boldsymbol{y}) > 0$. In particular, $H(\boldsymbol{y}) = 0$ for $\boldsymbol{y} < \mathbf{0}$ and $H(\boldsymbol{y}) = 1 - \frac{\log G(\boldsymbol{y})}{\log G(\mathbf{0})}$ for $\boldsymbol{y} > \mathbf{0}$ (Rootzén and Tajvidi, 2006; Rootzén *et al.*, 2018a,b).

SOCIETY OF ACTUARIES

MACQUARIE University

Empirical Results

The Pickands dependence function $A : [0, 1] \to [0, 1]$ is defined as

$$A(\omega) = \int_{S_2} \max\left(\omega q_1, (1-\omega) q_2\right) d\phi(\boldsymbol{q}), \qquad 0 \leq \omega \leq 1, \tag{13}$$

which links the function V through the relation

$$A(\omega) = \frac{V(z_1, z_2)}{z_1^{-1} + z_2^{-1}}, \tag{14}$$

with $\omega = \frac{z_2}{z_1 + z_2}$. By Equation (10), it is clear that $A(0) = A(1) = 1$. If two random variables with unit Fréchet margins are independent, then $A(\omega) = 1$ for all $0 \leq \omega \leq 1$, while if they are perfectly dependent, then $A(\omega) = \max(\omega, 1-\omega)$ for all $0 \leq \omega \leq 1$.

SOCIETY OF ACTUARIES MACQUARIE University

Empirical Results

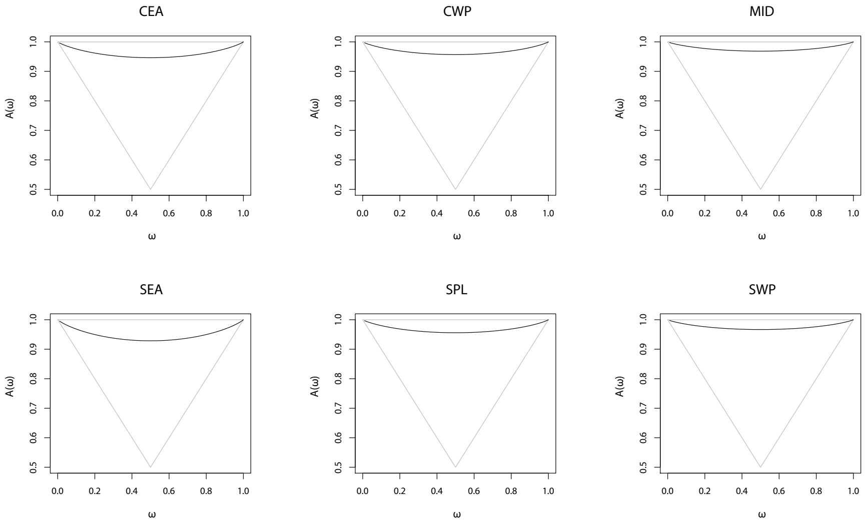
Coles *et al.* (1999) developed the index χ to measure extreme dependence for bivariate random variables. Assuming that random variables Z_1 and Z_2 have the same marginal distribution F, the index χ is defined as

$$\chi = \lim_{u \uparrow 1} \Pr\left(F(Z_2) > u \middle| F(Z_1) > u \right). \tag{15}$$

Thus, χ denotes the probability of one variable reaching the extreme value given that the other variable has already reached it. If $\chi = 0$, the two variables are said to be asymptotically independent. While for full tail dependence, we have $\chi = 1$. Furthermore, it can be verified (Coles *et al.*, 1999) that,

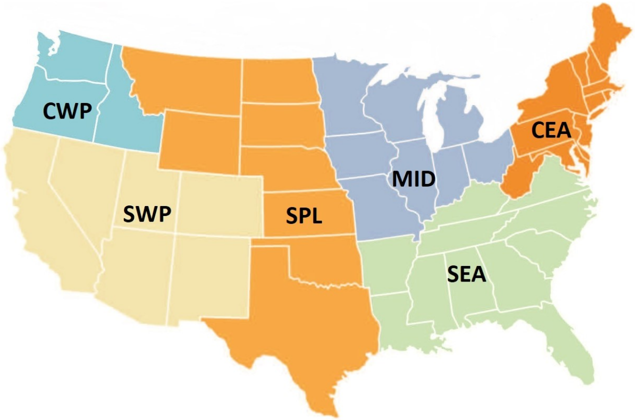
$$\chi = 2 - V(1,1) = 2 - 2A(0.5). \tag{16}$$

Empirical Results



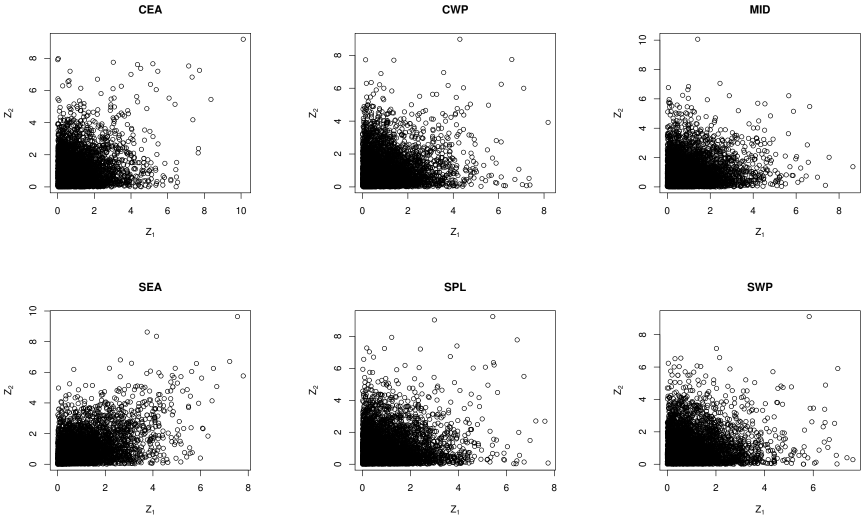
SOCIETY OF ACTUARIES
MACQUARIE University

Empirical Results



Source: Actuaries Climate Index Executive Summary (2018), Page 4, Figure 2

Empirical Results



Conclusions

Frequency of extreme hot temperatures
➔ weak impact on death counts

Heatwaves?? A heatwave is generally defined in terms of a consecutive period of excessively hot weather (4 days)

"Harvesting effect" or "Mortality displacement"

A simple measure of monthly hot temperature frequency may not be adequate

Frequency of extreme cold weather
➔ stronger impact on older people aged 55–84 and 85+

Cold weather can cause substantial short-term increase in mortality

Epidemics of influenza are likely to be associated with extreme cold weather

The increase in mortality following extreme cold is long lasting

The elderly are more fragile to extreme temperatures

SOCIETY OF ACTUARIES

MACQUARIE University

Another climate change quote...

“Global warming isn’t real because I was cold today! Also great news: world hunger is over because I just ate.”

– Stephen Colbert

SOCIETY OF ACTUARIES

MACQUARIE University

End of presentation

Thank you!

Any questions/ comments/ suggestions?

Contact email: han.li@mq.edu.au
Research Gate page: *http://www.researchgate.net/profile/Han_Li51*

References

Balkema, A. A. and de Haan, L. (1974). Residual life time at great age. *The Annals of Probability*, **2**(5), 792–804.

Coles, S., Heffernan, J., and Tawn, J. (1999). Dependence measures for extreme value analyses. *Extremes*, **2**(4), 339–365.

Pickands, J. (1975). Statistical inference using extreme order statistics. *The Annals of Statistics*, **3**(1), 119–131.

Resnick, S. I. (1987). *Extreme values, regular variation and point processes*. Springer.

Rootzén, H. and Tajvidi, N. (2006). Multivariate generalized pareto distributions. *Bernoulli*, **12**(5), 917–930.

Rootzén, H., Segers, J., and Wadsworth, J. L. (2018a). Multivariate generalized pareto distributions: Parametrizations, representations, and properties. *Journal of Multivariate Analysis*, **165**, 117–131.

Rootzén, H., Segers, J., and Wadsworth, J. L. (2018b). Multivariate peaks over thresholds models. *Extremes*, **21**(1), 115–145.

Tawn, J. A. (1990). Modelling multivariate extreme value distributions. *Biometrika*, **77**(2), 245–253.

SOCIETY OF ACTUARIES

MACQUARIE University