Räumliche Modellierung extremer Scheehöhen in Österreich

Harald Schellander [1] Naomi Auer [3] Tobias Hell [2] Stefan Rainer [2] Claudia Schmuck [3] and Christoph Zingerle [1]

[1] ZAMG

[2] Institut für Mathematik, Univ. Innsbruck

[3] Institut für Mathematik, Univ. Innsbruck/ZAMG FEMtech

8.11.2013

Überblick

- Einleitung
- Räumliche Modellierung
- Ergebnisse
- Zusammenfassung und Ausblick

Motivation

Gefragt: 50-jährliche return levels der Schneelast

- Schneelast wird praktisch nicht gemessen
- return levels der Schneehöhe einfach berechenbar

Ziel: Return level map

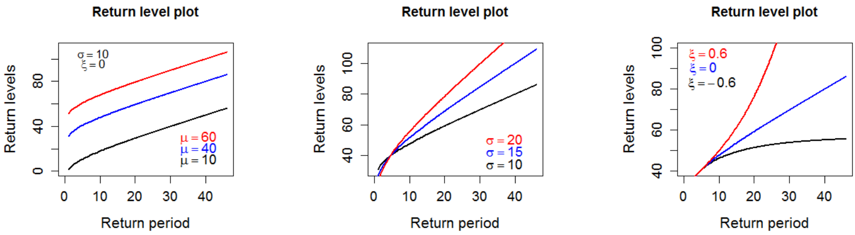
Wie erhält man Extremwerte für bestimmte Jährlichkeit an beliebigem Punkt?

Es sei $Z_n = max\{X_1, ..., X_n\}$, X_i unabhängig und identisch verteilt. Dann gilt für sehr große n, dass $\mathbb{P}(Z_n \leq z)$ durch die folgende Funktion approximiert werden kann:

Generalized Extreme Value Distribution (GEV)

$$G(z; \mu, \sigma, \xi) = exp\{-[1 + \xi(\frac{z-\mu}{\sigma})]^{-\frac{1}{\xi}}\}$$

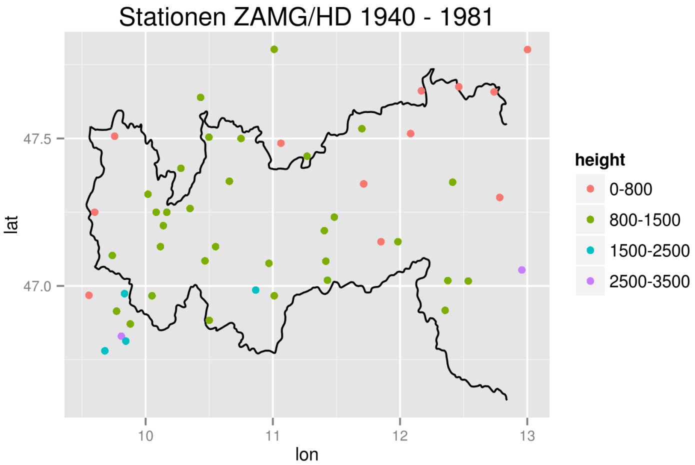
Verteilung der Extremwerte wird durch drei Parameter bestimmt:



- Für flächige Berechnung können die drei Parameter interpoliert werden. Ergebnisse sind aber nicht zufriedenstellend[1]:
  - Es werden bereits geschätzte Daten interpoliert
  - Interpolation erfolgt ohne Information über Extremwertverteilung zwischen Stationen
  - Keine Qualitätsinformation (confidence intervals)
- Besser funktioniert “smooth GEV modeling”: Parameter werden direkt über lineares Modell mit bestimmten Kovariablen aus den Messreihen berechnet.

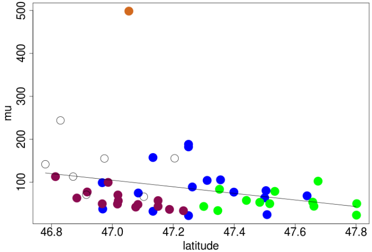
[1] Blanchet J. and M. Lehning, 2010: Mapping snow depth return levels: smooth spatial modeling versus station interpolation. Hydrol. Earth Syst. Sci., 14, 2527 - 2544

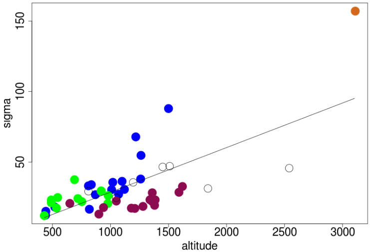
Schneehöhendaten der ZAMG und des HD Österreich.
50 Stationen mit vollständigen Messreihen von 1940 bis 1981.



Bestimmung der Parameter μ, σ und ξ

Die Parameter hängen z.B. von geogr. Länge, Breite und Höhe ab.





Es ist daher naheliegend, die Parameter als lineares Modell mit den Kovariablen Länge, Breite, Höhe und mittlere maximale Schneehöhe zu modellieren.

zum Beispiel: $\mu(x) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4$

Maximum Likelihood Methode

- Parameter werden nun so optimiert, dass

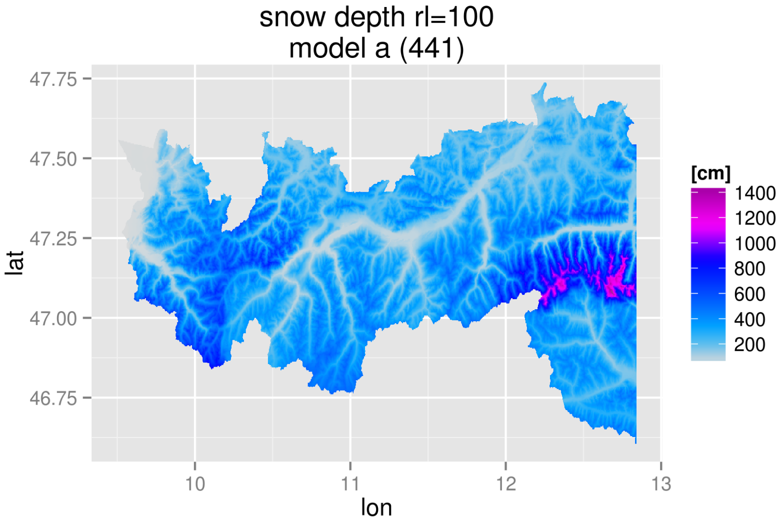
$$\mathbb{P}\left\{G(\mu,\sigma,\xi)=G(\hat{\mu},\hat{\sigma},\hat{\xi})\right\}\rightarrow max$$

- Für 1 Station $l(\mu,\sigma,\xi)\rightarrow max$
- Für N Stationen suchen wir gemeinsame Verteilung der Jahreshöchstwerte der Schneehöhe. Wenn Jahreshöchstwerte zeitlich und räumlich unabhängig, dann gilt

$$L(\mu,\sigma,\xi)=\sum_{i=1}^{N} l(\mu_i,\sigma_i,\xi_i) \text{ und } L(\mu,\sigma,\xi)\rightarrow max$$

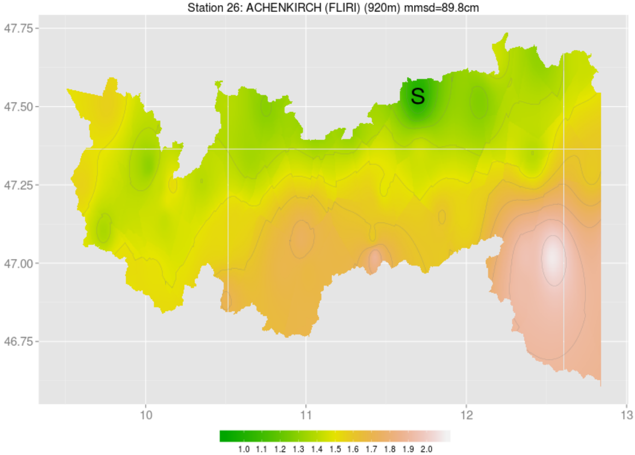
- Jahresmaxima der Schneehöhe sind
  - zeitlich unabhängig? ja
  - räumlich unabhängig? nein ($\rightarrow$ später)
- 4 Modelle für μ, 6 Modelle für σ, 2 Modelle für ξ
- Maximiere $L(\mu, \sigma, \xi)$ für alle Modellkombinationen $(4 * 6 * 2 = 48)$
- Finde am besten angepasstes Modell mit Informationskriterium TIC

$$\mu(spline(lon, lat), alt, z), \sigma(lat, alt, \mu), \xi(lon, lat, alt)$$



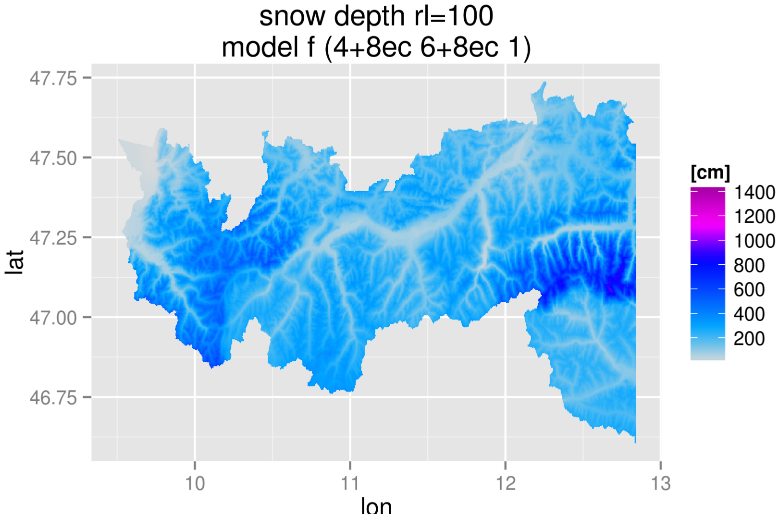
Extremal coefficients als Kovariable

Räumliche Abhängigkeit über extremal coefficient mit Schlather max-stable Prozess: $1 \leq \theta_{xx'} \leq 2$



- $\theta_{xx'}$ für alle Stationen: $L(\mu, \sigma, \xi) \rightarrow max$ nicht mehr durchführbar
- daher clustering der Stationen: k-Means mit 4 bzw. 8 clustern. Jeder cluster liefert eine Kovariable: $\theta_{xx'}$ jener Station, die den signifikantesten Anteil zur Modellierung der Schneehöhe beiträgt $\rightarrow$ 4 bzw. 8 statt 50 Kovariablen zusätzlich
- Vorgangsweise wie ohne räumliche Abhängigkeit: Verschiedene Modelle für μ, σ und ξ, Kombinationen aus 4 “alten” und 4 bzw. 8 “neuen” Kovariablen

$$\mu(spline(lon, lat), alt, z, 8ec), \sigma(spline(lon, lat), alt, z, 8ec), \xi(lon, lat, alt)$$

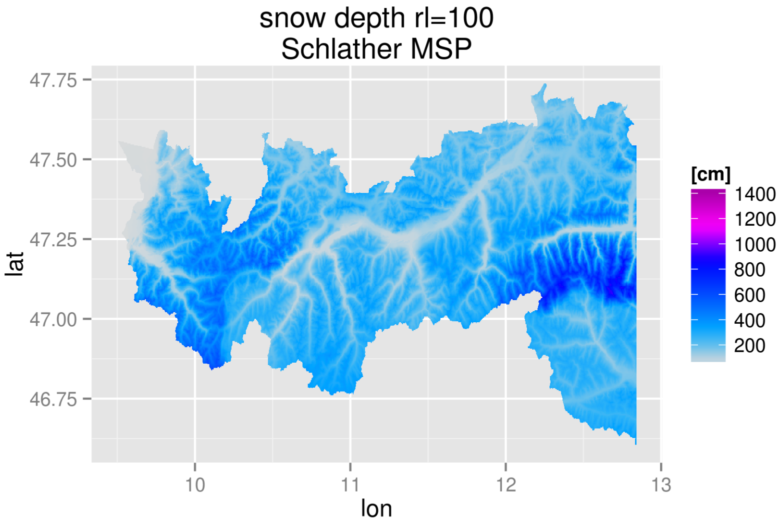


Max-stable process (MSP)

- MSP-basierten Methode zur räumlichen Modellierung extremer Schneehöhen in der Schweiz[2].
- Betrachte Jahreshöchstwerte der Schneehöhe paarweise, d.h. an zwei Orten gleichzeitig $\rightarrow$ bivariates Problem
- MSP sind eine Erweiterung der multivariaten Extremwerttheorie und können Verteilung und räumliche Abhängigkeit der Schneehöhe gleichzeitig beschreiben
- Wahl des charakterisierenden Prozesses: Schlather
- pairwise likelihood und angepasster TIC
- derzeit nur einfachstes Modell: alle Parameter abhängig von Länge, Breite, Höhe, mittlere maximale Schneehöhe

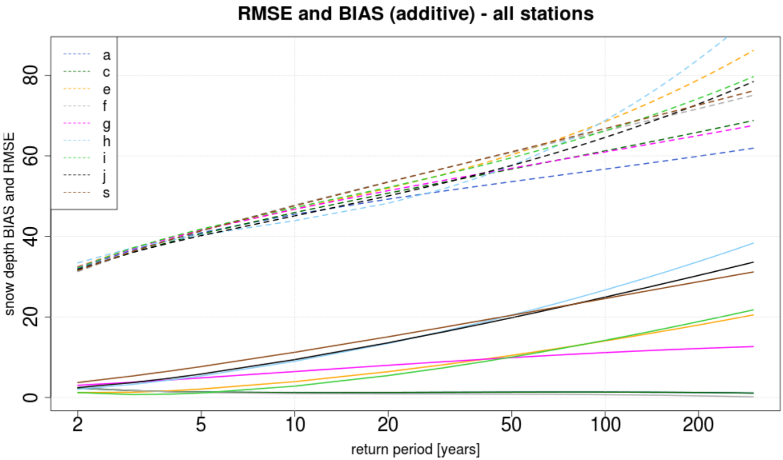
[2] J. Blanchet and A. Davison, 2011: Spatial modeling of extreme snow depth. The Annals of Applied Statistics, Vol. 5, Nr. 3, 1699 - 1725

$$\mu(lon, lat, alt, z), \sigma(lon, lat, alt, z), \xi(lon, lat, alt, z)$$

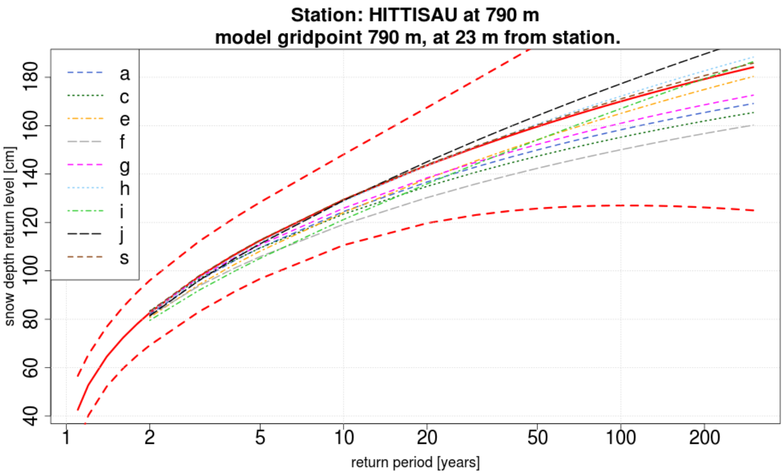


Verifikation mit Stationen an denen Modelle geeicht wurden

RMSE 30 bis etwa 70 cm; BIAS meist positiv



Model a: ohne räumliche Abhängigkeit.
Models mit räumlicher Abhängigkeit bringen Verbesserung



Zusammenfassung

- Fehler nehmen mit return period zu
- Schneehöhen werden von fast allen Modellen eher überschätzt
- Grössenordnung des mittleren Fehlers
  - bis etwa return period 50 bei allen Modellen gleich
  - zwischen 30 und etwa 70 cm
- Modelle mit räumlicher Abhängigkeit besser
- Modelle mit spline besser
- MSP mit einfachstem Ansatz vielversprechend

Ausblick

- verschiedene Modelle für MSP
- erweiterte Verifikation: cross validation, unbeteiligte Stationen
- Erweiterung auf Österreich
- vorgestellte Methoden liefern Framework für räumliche Modellierung anderer Extremwerte (zum Beispiel 3-stündiger Niederschlag, 3-Tages Neuschneesummen, etc.)
- Räumliche Modellierung der Schneelast (siehe nächster Vortrag von Steffi Gruber)