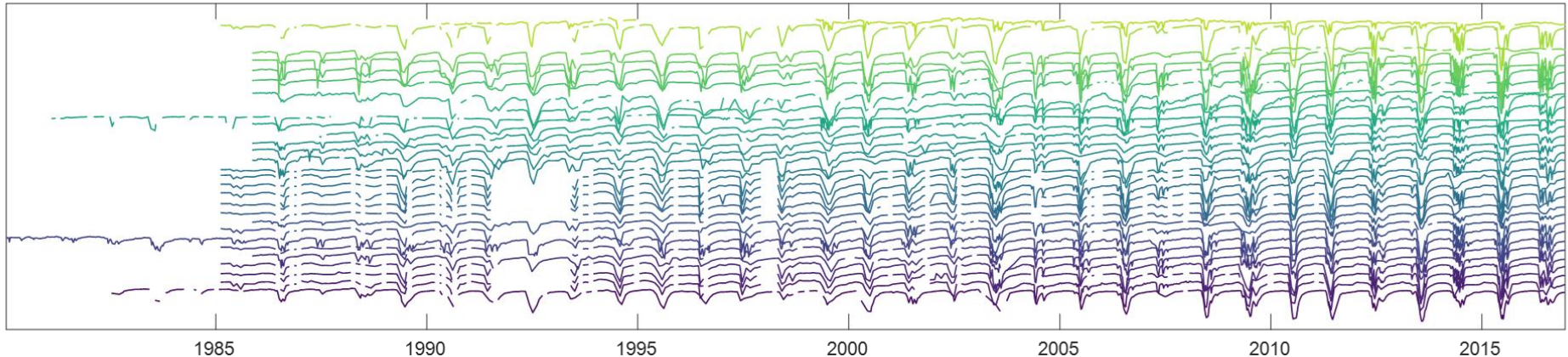


Feature-basiertes Clustering von Umweltzeitreihen mit Self-Organizing-Map-Ensembles

am Beispiel von Grundwasserstandszeitreihen des Oberrheingrabens

M.Sc. Andreas Wunsch, Dr. Tanja Liesch, Dr. Stefan Broda

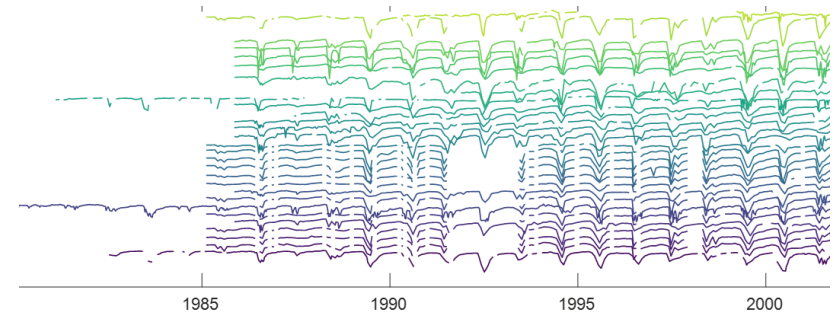


Umweltzeitreihen

Idee: Gruppierung von Zeitreihen gleicher Dynamik + Identifikation räumlicher Muster

Hauptproblem realer Messdaten: **Datenqualität**

- Datenlücken
- Ausreißer
- Inhomogenitäten
- variable Beobachtungszeiträume



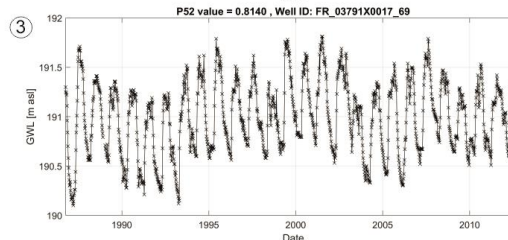
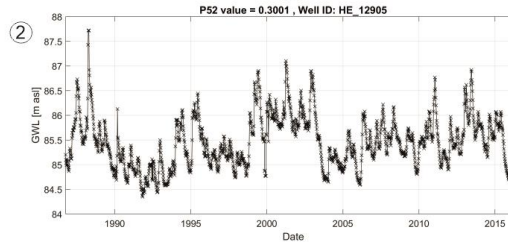
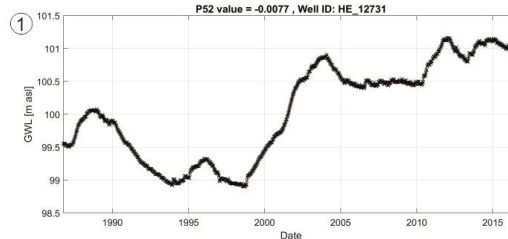
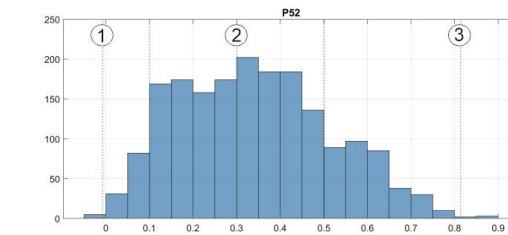
→ Feature-basierte Beschreibung der
Zeitreihendynamik

Features

→ beschreibende statistische Indizes

z.B.

- Periodizität
- Verteilungsschiefe
- Position des Maximums im Kalenderjahr
- [...]



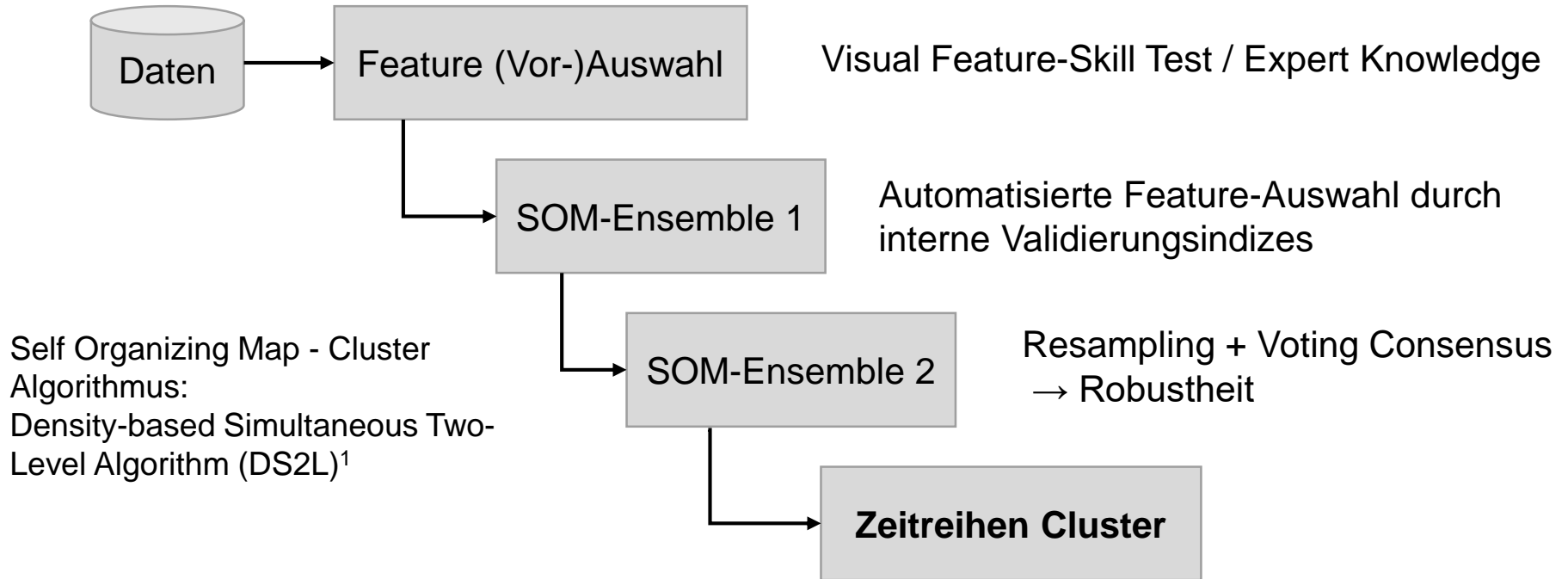
niedrig



Periodizität

hoch

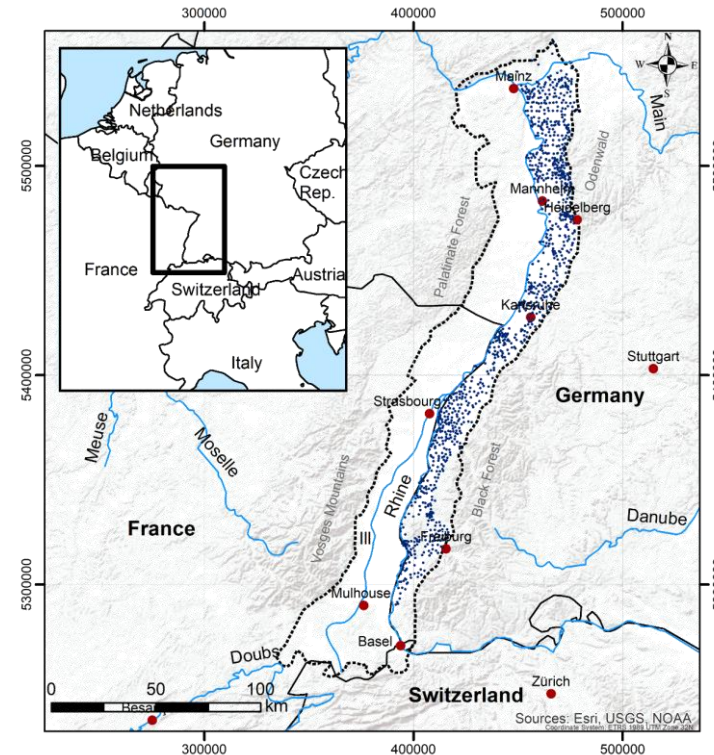
Vereinfachter Workflow



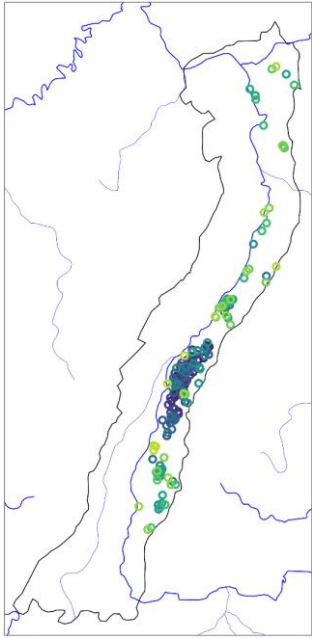
¹Cabanes et al. (2012): Enriched topological learning for cluster detection and visualization. *Neural Networks*, 32, 186–195

Beispielanwendung: Grundwasserdaten

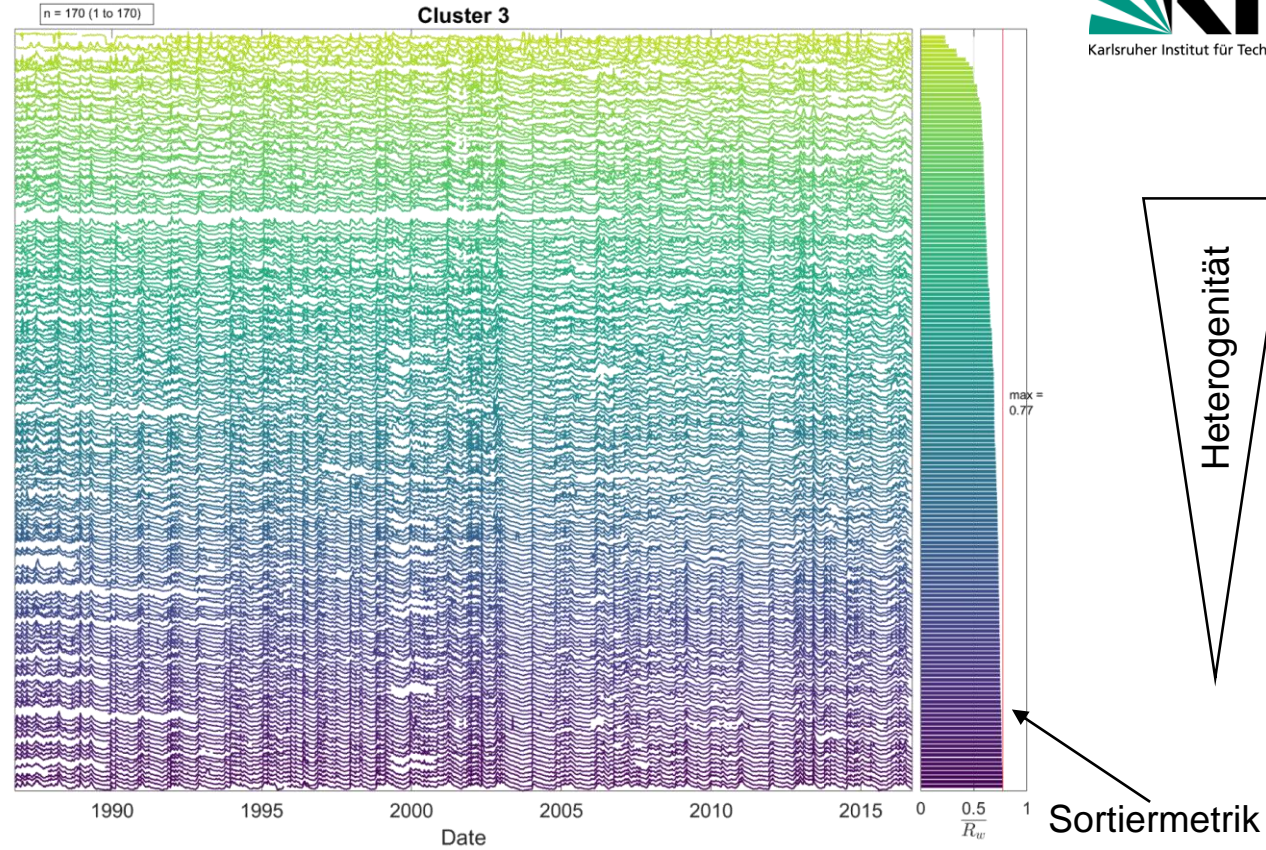
- Oberrheingraben (BW + HE)
- 1195 Zeitreihen (Wochendaten)
- Zeitraum: 1986-2016
- Heterogene Daten: Datenlücken und variable Zeitreihenlängen



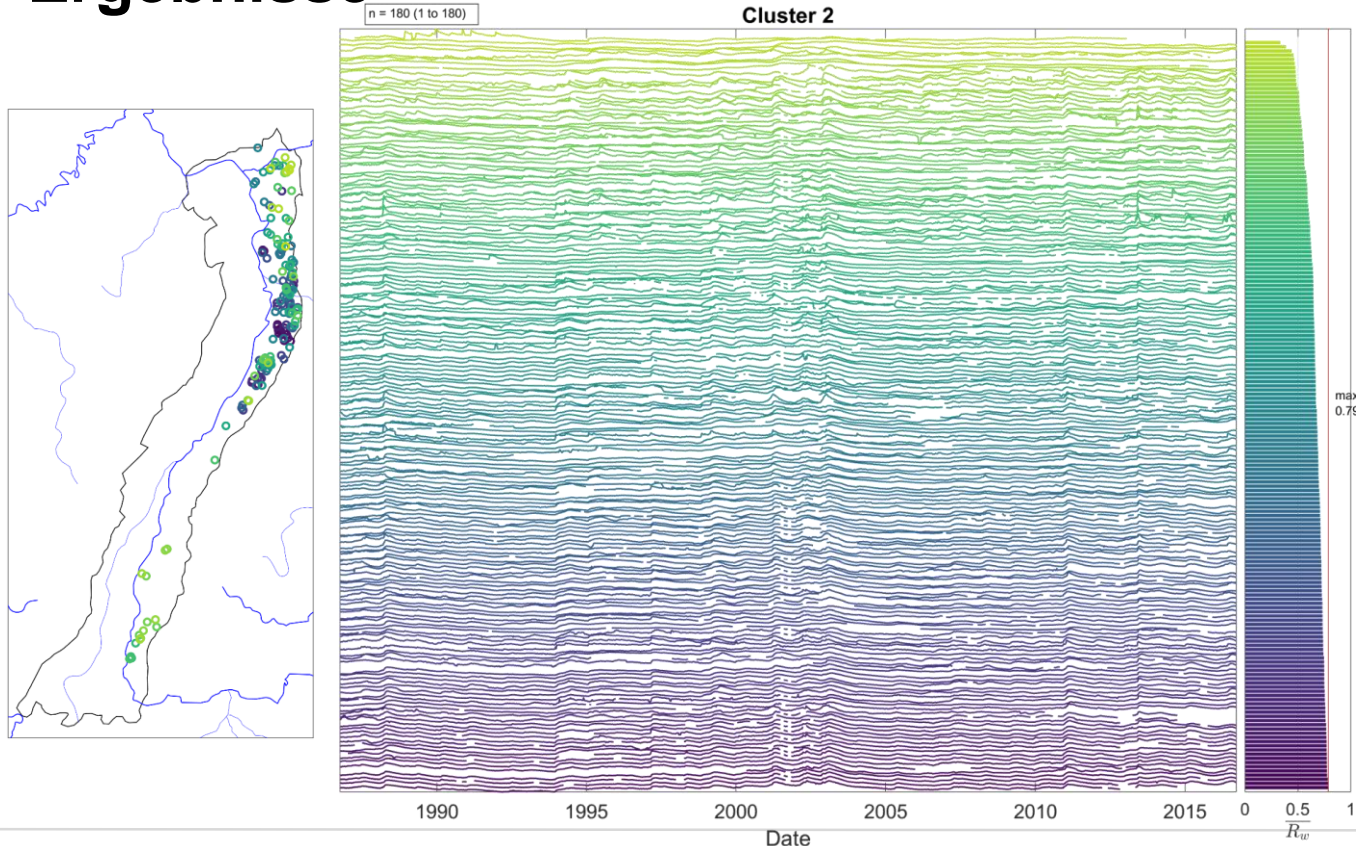
Ergebnisse



(insg. 22 Cluster)



Ergebnisse

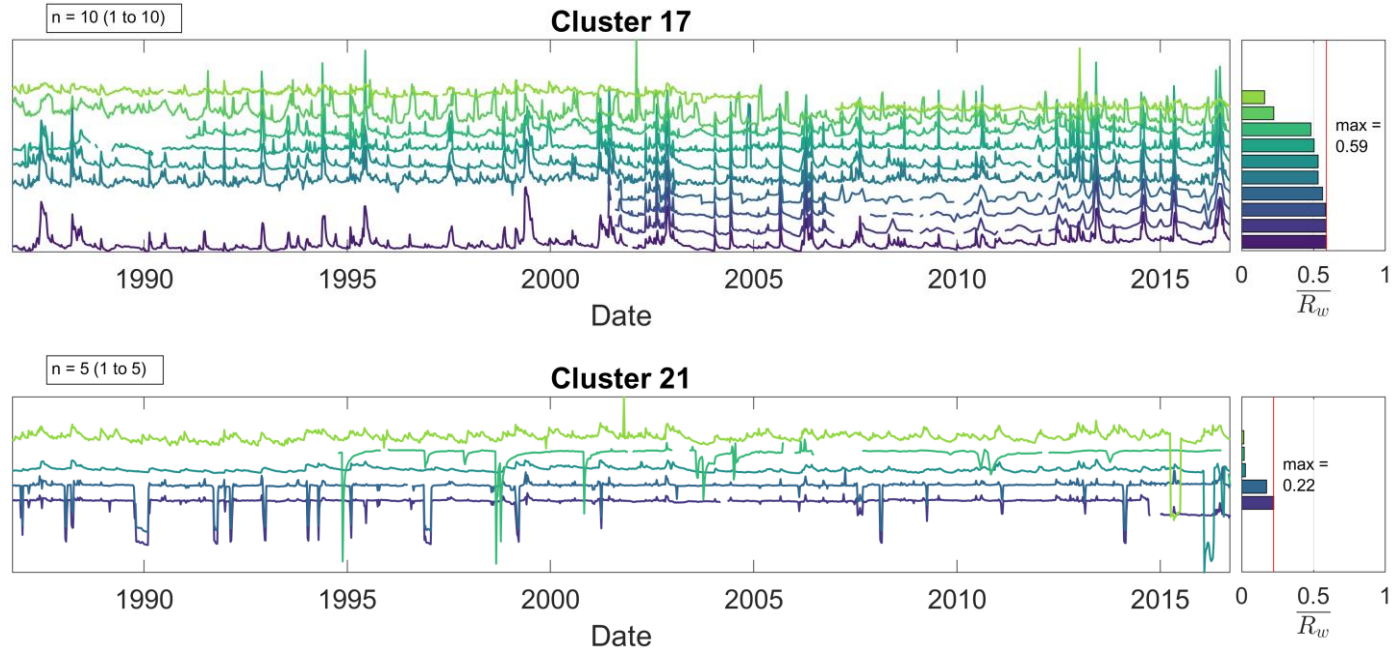


Geringere Variabilität
wegen starker GW-
Entnahmen und
geringerer Neubildung

→ Fachliche
Interpretierbarkeit der
Cluster

Ergebnisse

- Variable Clustergrößen
- zuverlässige Erkennung von Inhomogenitäten und Ausreißern



Zusammenfassung

- Features + ML → Handhabung großer und heterogener Datensätze
- Ensembles → robuste und praxisorientierte Ergebnisse
→ u.U. hoher Rechenaufwand
- Features + Cluster-Alg. → Flexibilität in Bezug auf Clusteranzahl, Clustergröße
und Ausreißerererkennung
→ Benutzerfreundliche Modifikationsmöglichkeiten
- Hydrogeologische Interpretierbarkeit der Ergebnisse

Vielen Dank!

Kontakt:

andreas.wunsch@kit.edu



