

Künstliche Intelligenz in der N–A Modellierung

N. Schütze, S.Pahner, T. Morgenstern

HoWa-innovativ-Verbundtreffen

Offenbach, 30.09.2021

Übersicht

Blick zurück

Methodik der datengetriebenen Hochwasservorhersage

KI der ersten Generation

KI der zweiten Generation

Neuentwicklungen

Beispiel: Wavelet-Transformation und Singular Spektrum

Analyse

Beispiel: Encoder-Decoder LSTM

Übersicht

Blick zurück

Methodik der datengetriebenen Hochwasservorhersage

KI der ersten Generation

KI der zweiten Generation

Neuentwicklungen

Beispiel: Wavelet-Transformation und Singular Spektrum

Analyse

Beispiel: Encoder-Decoder LSTM

KI in der Hochwasserfrühwarnung?

Nutzer

- pro-aktive Versorgung der Nutzer (SMS, push, ...)
- Frühwarnung als Teil des HW-Nachrichtendienstes
- Von Landkreisen zu Flussgebieten
- Benötigte VHS-Zeiträume/Vorwarnzeiten
→ 3h? 6h? 12h? 24h?
→ update-Zyklus
- Rückmeldung (i.S.v. Verifikation der Abschätzung) durch Nutzer; crowdsourcing
- Kulisse der HW-Entstehungsgebiete

Daten

- Probabilistische Produkte
→ Bewarnung/Kommunikation?
- Generalisierte Produkte/Schwellenwert-Ansätze (2.8km → 28km)
- Grenzen akzeptieren ($FAR > 0.5$, Grenzen der Vorhersagbarkeit $A_E < ?? km^2$)
- recht generelle Quantil-VHS schneidet ebenfalls gut ab (+Experteneinschätzung)
→ räumliche Grenzen der Methoden?
- Frühwarnung für kleine EZG/flächenhaften Abfluss (lt. Jeschke-Bericht) → nicht möglich

Modelle

- Wenn deterministische Modellierung
→ Synthese mit operationellen, pegelbezogenem VHS-System ("Frühwarnung und VHS aus einem System")
- Pegelbezogene Aussagen
→ HQ(T)s
→ konkrete Auswirkungen "am Ort" (z.B. Gemeinde)
- es braucht den operationellen Hydrologen (kein Automatismus)
- Warnphilosophie
- Multi-Modell/Modell-Mix
→ Kombination?

Whiteboard vom LfULG Workshop 2016 »Hochwasserfrühwarnung«

Hochwasserentstehung in sächsischen Einzugsgebieten (2)



06/2013
Reichstädter Bach
Foto: Ralf Kämmerer

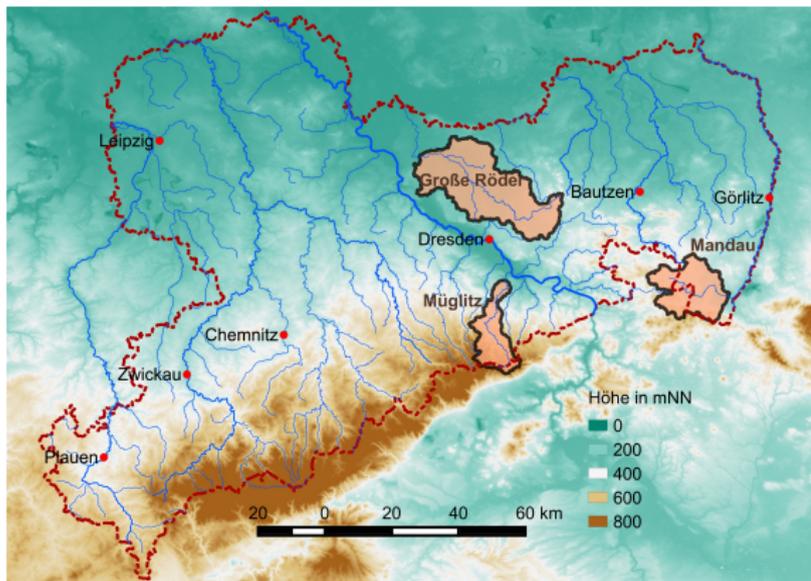
08/2010

Bertsdorfer Wasser
Foto: Olaf Menges



Kleinräumige Ereignisse in den Jahren 2010 und 2013

Hochwasserentstehung in sächsischen Einzugsgebieten (3)



Pilotgebiete im Projekt »Hochwasserwarnung und -vorhersage in kleinen EZG«

Übersicht

Blick zurück

Methodik der datengetriebenen Hochwasservorhersage

KI der ersten Generation

KI der zweiten Generation

Neuentwicklungen

Beispiel: Wavelet-Transformation und Singular Spektrum

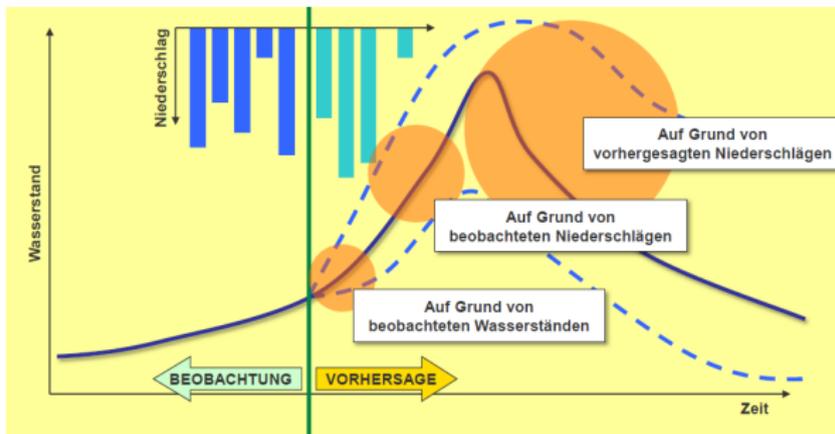
Analyse

Beispiel: Encoder-Decoder LSTM

Methodik der datengetriebenen Hochwasservorhersage

Gegebene Daten, u.a.:

- ▶ gemessene Niederschläge $\mathbf{p}_M = (p_1, \dots, p_t)$
- ▶ vorhergesagte Niederschläge $\mathbf{p}_V = (p_{t+1}, \dots, p_{t+l})$
- ▶ gemessene Durchflüsse $\mathbf{q} = (q_1, \dots, q_t)$



Problemstellung Hochwasservorhersage (Quelle: Andy Philipp)

Methodik der datengetriebenen Hochwasservorhersage

Gegebene Daten, u.a.:

- ▶ gemessene Niederschläge $\mathbf{p}_M = (p_1, \dots, p_t)$
- ▶ vorhergesagte Niederschläge $\mathbf{p}_V = (p_{t+1}, \dots, p_{t+l})$
- ▶ gemessene Durchflüsse $\mathbf{q} = (q_1, \dots, q_t)$

Problem:

Finde

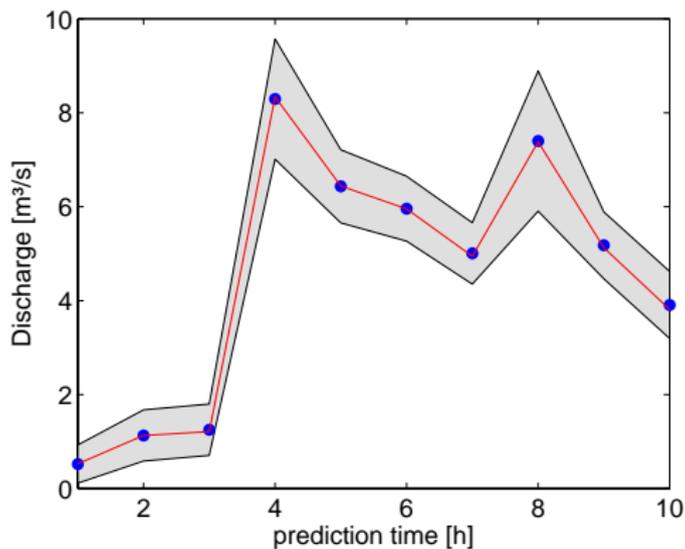
- ▶ geeignete Eingangs- und Trainingsdaten,
- ▶ eine mathematische Funktion oder einen Algorithmus \mathbf{f} , der q_{t+1} und die damit verbundene Unsicherheit ε berechnet:

$$(q_{t+1}, \varepsilon) = \mathbf{f}(\mathbf{p}_M, \mathbf{p}_V, \mathbf{q}, \theta).$$

- ▶ sowie einen geeigneten (Lern-) Algorithmus (einschl. Gütekriterium), der den optimalen Parametervektor θ berechnet.

Methodik der datengetriebenen Hochwasservorhersage (2)

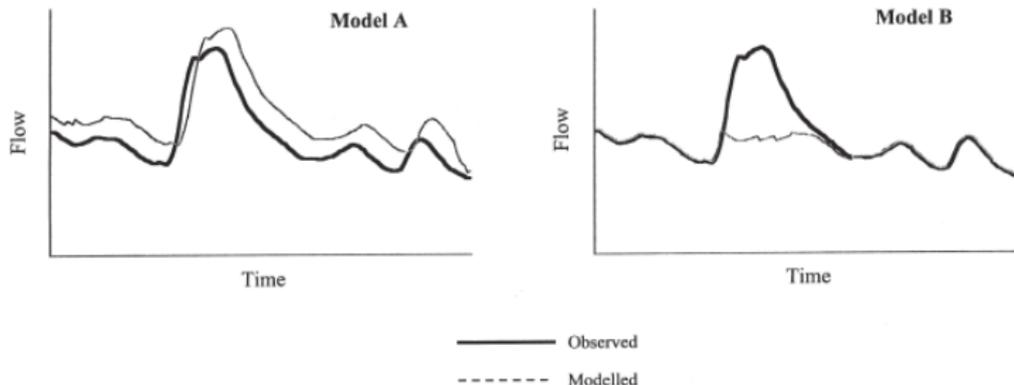
Ergebnisprodukt eines stochastischen Modells (GP-Modell)



Vorhersage mit Unsicherheitsbereich

Methodik der datengetriebenen Hochwasservorhersage (3)

Mögliche Probleme datengetriebener Modelle



Minns und Hall (1996)

a) Phasenfehler durch Autokorrelation

b) Extrapolationsfehler

Übersicht

Blick zurück

Methodik der datengetriebenen Hochwasservorhersage

KI der ersten Generation

KI der zweiten Generation

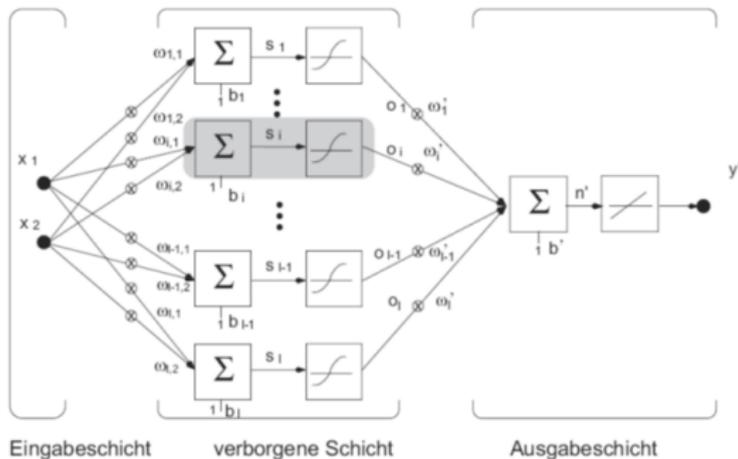
Neuentwicklungen

Beispiel: Wavelet-Transformation und Singular Spektrum

Analyse

Beispiel: Encoder-Decoder LSTM

Standardmethoden und Workflow



$$y' = \sum_{i=1}^l \omega_i' \frac{e^{\sum_{j=1}^2 \omega_{ij} x_j + b_i} - e^{-\left(\sum_{j=1}^2 \omega_{ij} x_j + b_i\right)}}{e^{\sum_{j=1}^2 \omega_{ij} x_j + b_i} + e^{-\left(\sum_{j=1}^2 \omega_{ij} x_j + b_i\right)}} + b'$$

Die Referenz – Multilayer Perceptron (Rummelhardt 1989)

Standardmethoden und Workflow

Multilayer Perceptron

Architektur: MLP zusammen mit
Backpropagation-Trainingsalgorithmus
(Gradientenverfahren) + u.a. regularization and
early stopping

Datengrundlage: Zeitreihen (Stationsdaten) beobachteter
Niederschläge & Durchflüsse (täglich)

Datenteilung: ad hoc → training|validation|test

Datentransformation: Standardisierung (0,1)

Zielfunktion (loss function): RMSE

Standardmethoden und Workflow (2)

Beispiel: geeignete Eingangs- und Trainingsdaten

Eingangsdaten (Messdaten)

- ▶ interpolierte Niederschlags-Stationsdaten

Eingangsdaten (Vorhersagen)

- ▶ Quantilvorhersage – DWD
- ▶ deterministische Vorhersage – COSMO-DE
- ▶ Ensemble-Vorhersage – COSMO-DE-EPS

Trainingsdaten

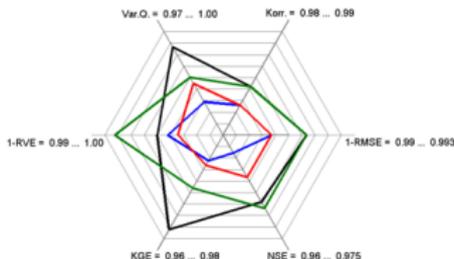
- ▶ Durchflussflussdaten (aus Wasserständen an Pegeln)
- ▶ Hochwasserwarnstufen der Pegel

Standardmethoden und Workflow (3)

Beispiel: geeignete Eingangs- und Trainingsdaten

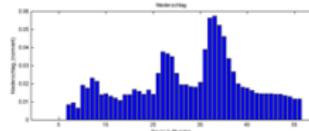
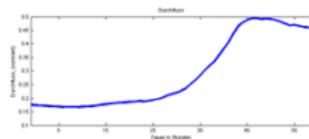
Eingangsdaten, z.B.

- ▶ Durchfluss zu Zeitpunkten $t=i, i-1, \dots, i-3$
- ▶ Niederschlag zu Zeitpunkten $t=i, i-1, \dots, i-3$
- ▶ (Vorhergesagter) summierter Niederschlag $t=i+1$ bis $t=i+12$



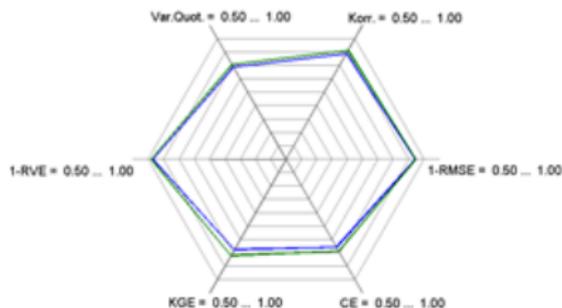
Trainingsdaten

- ▶ Zeiträume (kontinuierliche Trainingsdaten) versus
- ▶ Ereignisse (ereignisbasierte Trainingsdaten)

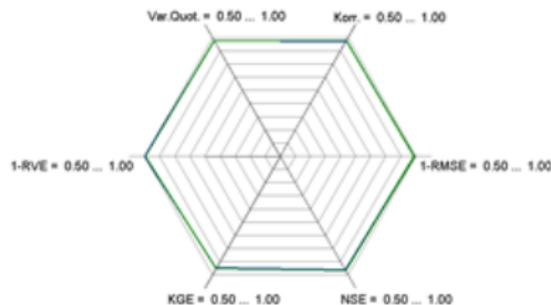


Standardmethoden und Workflow (4)

Beispiel: Datenvorverarbeitung



normierte Daten



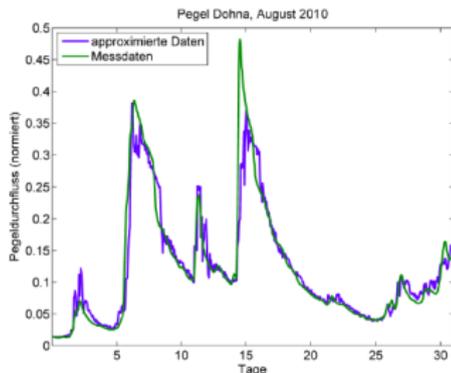
nicht normierte Daten

Bibliothek »datengetriebener Modellbausteine« (5)

Beispiel: untersuchte KI-Methoden

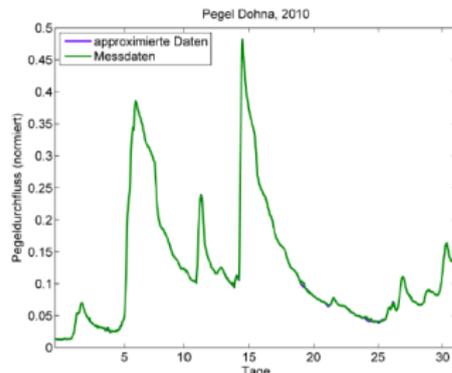
Parametrische Modelle

- ▶ Lineare Regression
- ▶ Regressionsbäume
- ▶ Neuronale Netze (MLP)



Nichtparametrische Modelle

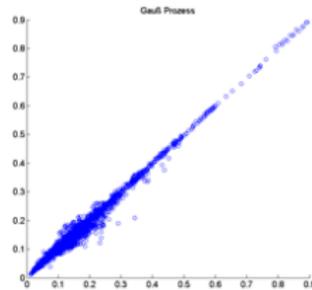
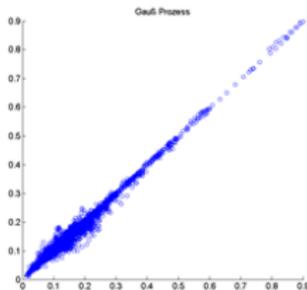
- ▶ Gauß-Prozesse (GP)
- ▶ Support Vektor Regression



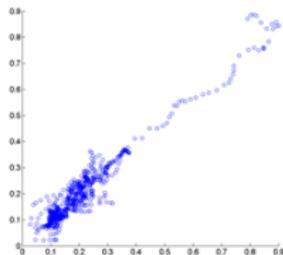
Standardmethoden und Workflow (6)

Ausgewählte Ergebnisse (Müglitz):

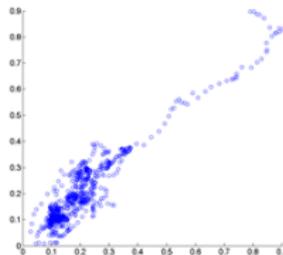
Approximation



Vorhersage



6 h

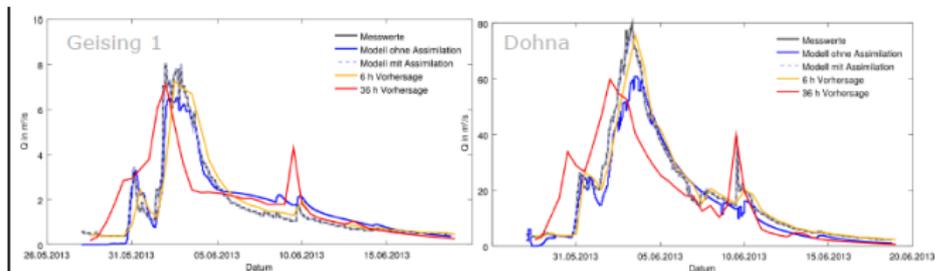


12 h

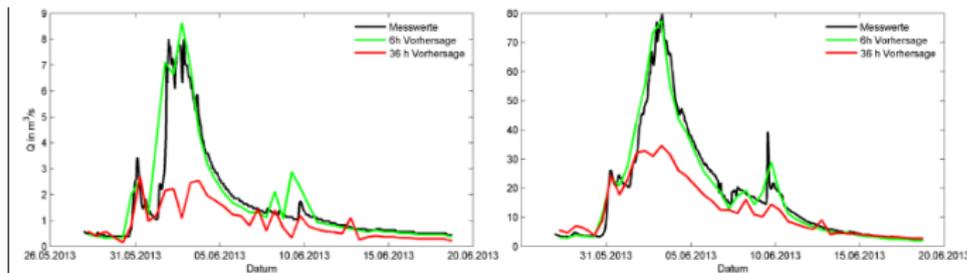
Standardmethoden und Workflow (7)

Ausgewählte Ergebnisse (Müglitz):

DeHM



DaHM



Quell-EZG

Mündungs-EZG

Standardmethoden und Workflow (8)

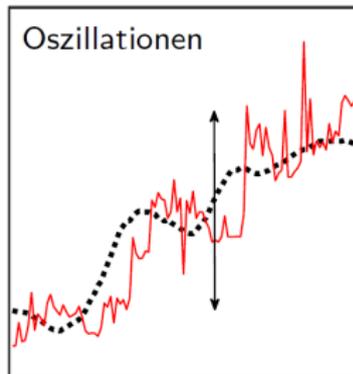
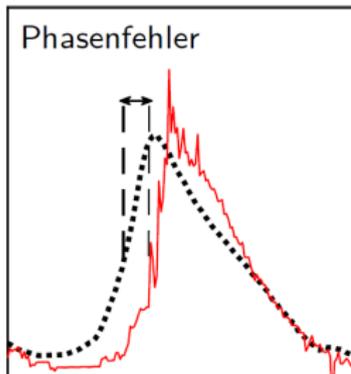
Validierung: datengetriebene Vorhersagemodelle

Gütekriterium	MLR	MLP	GP	SVR	RT
NSE	0,87	0,90	0,94	0,88	0,73
PPD (h)	-5	-4	-4	-3	-11

Pegel Dohna (Müglitz), 12 h - Vorhersage, Ereignisbasiertes Training

Die schlechteren Nachrichten

Beispiel: typische Fehler, insbesondere bei kleinen EZG



..... Qbeo ——— Qsim

Übersicht

Blick zurück

Methodik der datengetriebenen Hochwasservorhersage

KI der ersten Generation

KI der zweiten Generation

Neuentwicklungen

Beispiel: Wavelet-Transformation und Singular Spektrum

Analyse

Beispiel: Encoder-Decoder LSTM

Standardmethoden und Workflow

Multilayer Perceptron

Architektur: MLP zusammen mit
Backpropagation-Trainingsalgorithmus
(Gradientenverfahren) + u.a. regularization and
early stopping

Datengrundlage: Zeitreihen (Stationsdaten) beobachteter
Niederschläge & Durchflüsse (täglich)

Datenteilung: ad hoc → training|validation|test

Datentransformation: Standardisierung (0,1)

Zielfunktion (loss function): RMSE

Neuentwicklungen

Architektur:

MLP zusammen mit Backpropagation-Trainingsalgorithmus (Gradientenverfahren) + u.a. regularization and early stopping

- ▶ deep MLP mit mehr als 2 »hidden layers«
- ▶ Convolutional Neural Networks zur Bilderkennung
- ▶ Long Short-Term Memory Networks (LSTMs) als Weiterentwicklung von rekurrenten Netzen
- ▶ Frame et al.: Mass Conserving LSTM
<https://doi.org/10.5194/hess-2021-423>
- ▶ neue Trainingsalgorithmen: stochastische Gradientenverfahren, Bayes'sche Optimierungsalgorithmen

Neuentwicklungen

Datengrundlage: Zeitreihen (Stationsdaten) beobachteter Niederschläge & Durchflüsse (täglich)

- ▶ Niederschlagsfelder (RADOLAN) statt Stationsdaten oder Gebietsmittel

Datenteilung: ad hoc → training|validation|test

- ▶ umfangreiche cross-validation: Leave One Out, Stratified Cross Validation

Neuentwicklungen

Datentransformation: Standardisierung (0,1)

- ▶ z. B. Wavelet-Transformation und Singular Spektrum Analyse

Zuverlässige Deep learning Infrastrukturen und Bibliotheken:

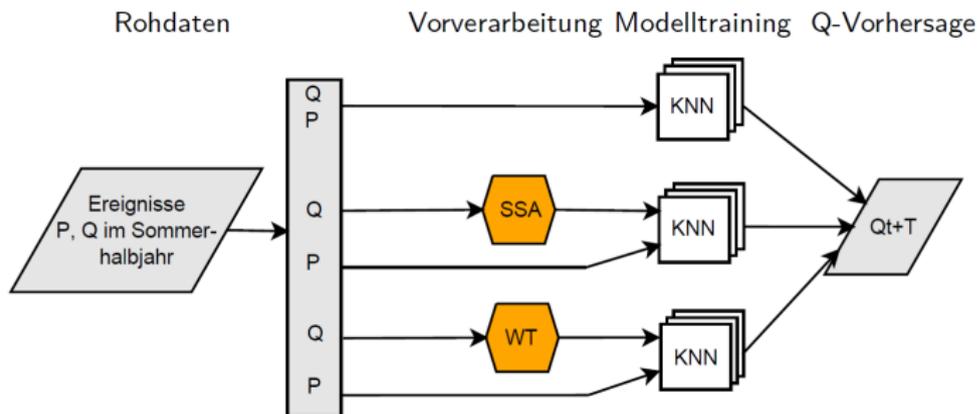
- ▶ z. B. Keras und TensorFlow

Neuentwicklungen

hybride Modelle:

- ▶ Verbindung von hydrologischen Modell und KNN zur Fehlerkorrektur (Training auf Basis der Residuen)
- ▶ modulare Netze und Modell-Ensemble (z.B. Regressionsbäume oder gradient boosting)

Beispiel: Wavelet-Transformation und Singular Spektrum Analyse Vorgehen

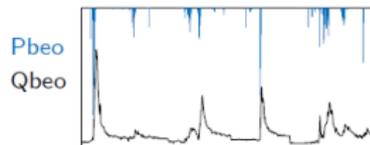


Beispiel: Wavelet-Transformation und Singular Spektrum Analyse

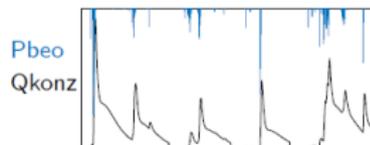
Vorgehen (2)

- stündliche KNN-Eingangsdaten:

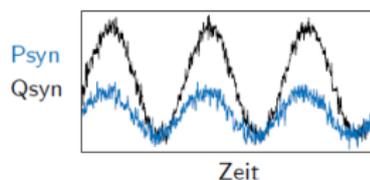
→ reale Ebene



→ konzeptionelle Ebene



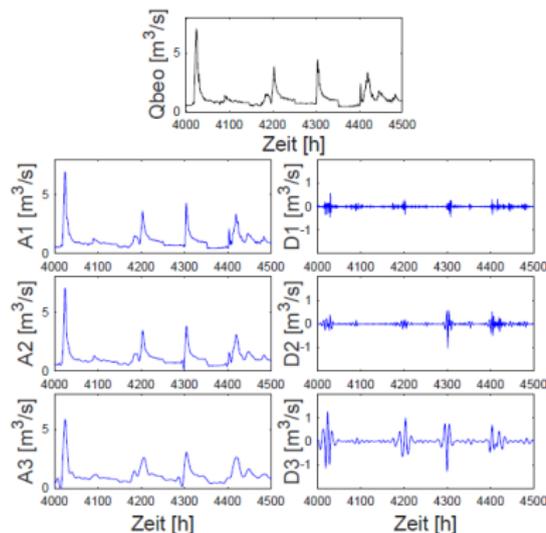
→ analytische Ebene



Beispiel: Wavelet-Transformation und Singular Spektrum Analyse Methoden

Wavelet-Transformation

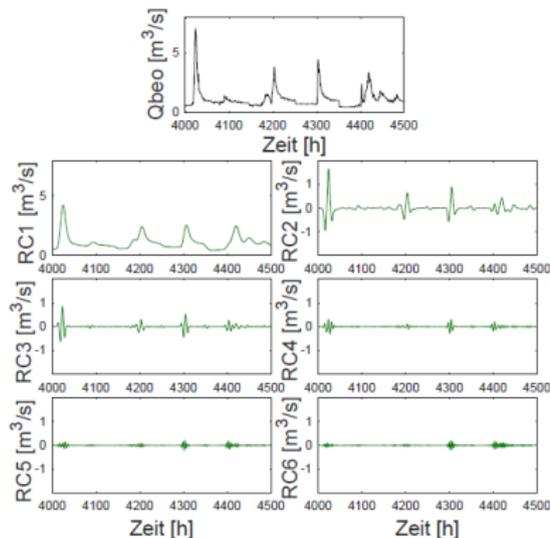
- 8 Mutterwavelets (Daubechies, Symlet)
- Auswahlalgorithmus
Filterkomponenten (A, D)
als KNN-Input



Beispiel: Wavelet-Transformation und Singular Spektrum Analyse Methoden (2)

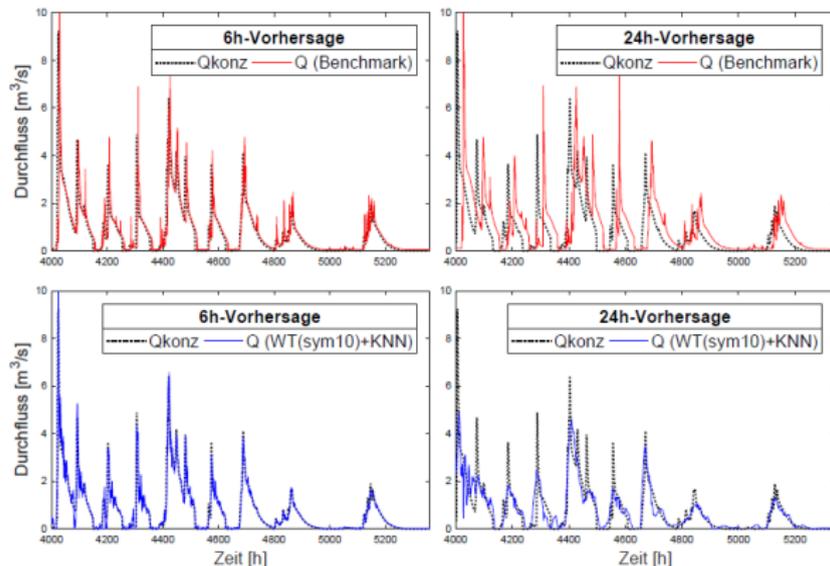
Singular Spektrum Analyse

- Fensterlänge M
- Auswahlalgorithmus
Filterkomponenten (RC)
als KNN-Input



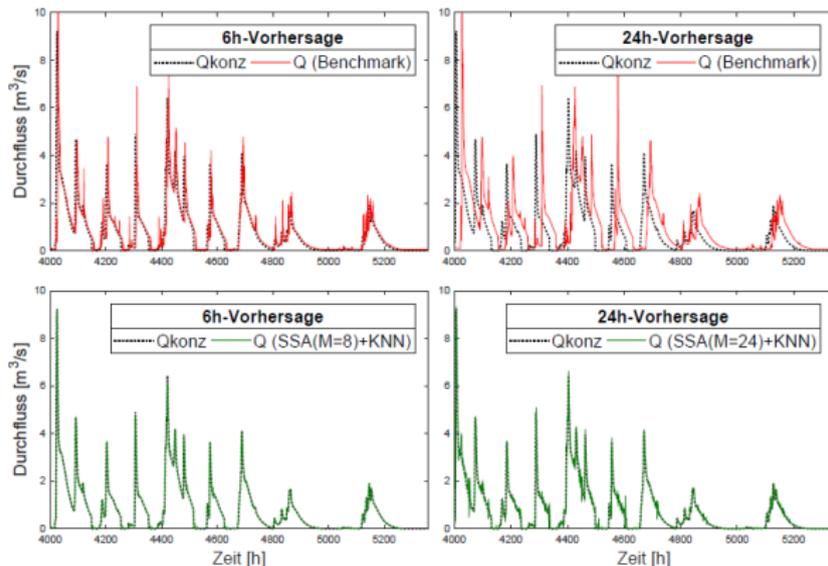
Beispiel: Wavelet-Transformation und Singular Spektrum Analyse

Ergebnisse Modell (wavelet)



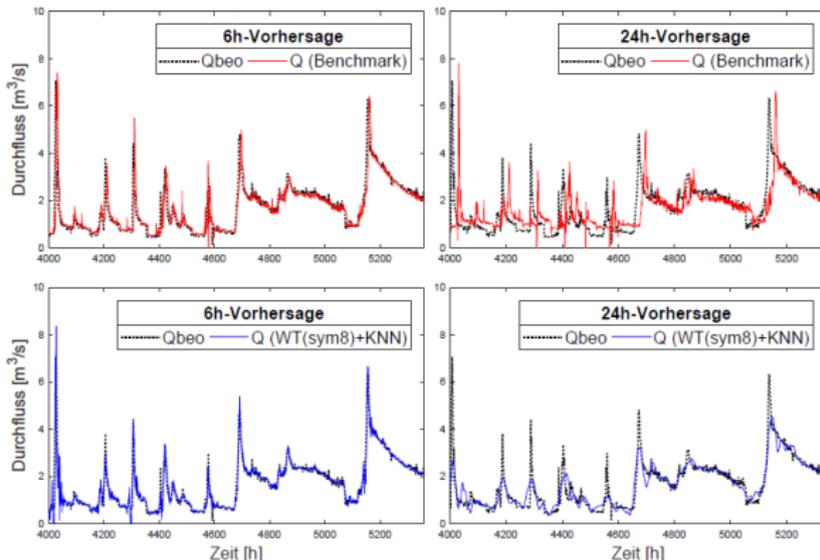
Beispiel: Wavelet-Transformation und Singular Spektrum Analyse

Ergebnisse Modell (SSA)



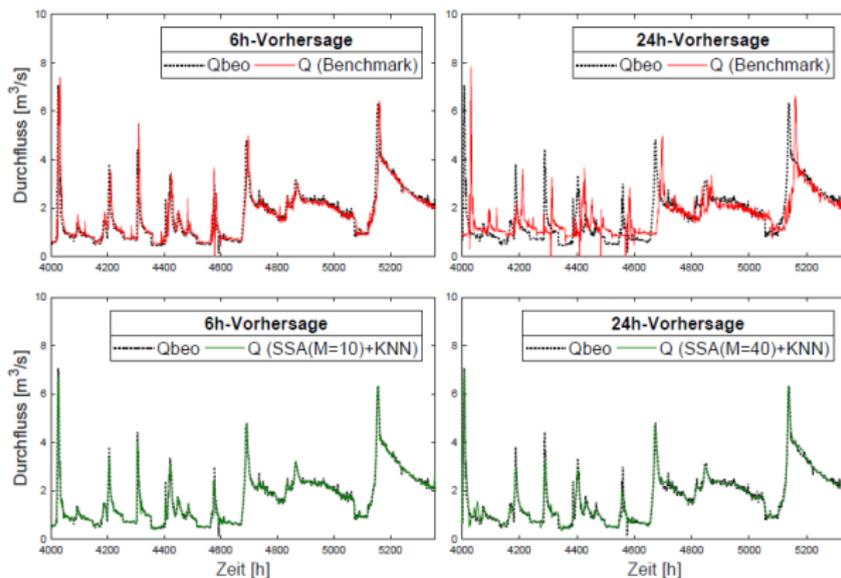
Beispiel: Wavelet-Transformation und Singular Spektrum Analyse

Ergebnisse Real (wavelet)



Beispiel: Wavelet-Transformation und Singular Spektrum Analyse

Ergebnisse Real (SSA)



Beispiel: Wavelet-Transformation und Singular Spektrum Analyse

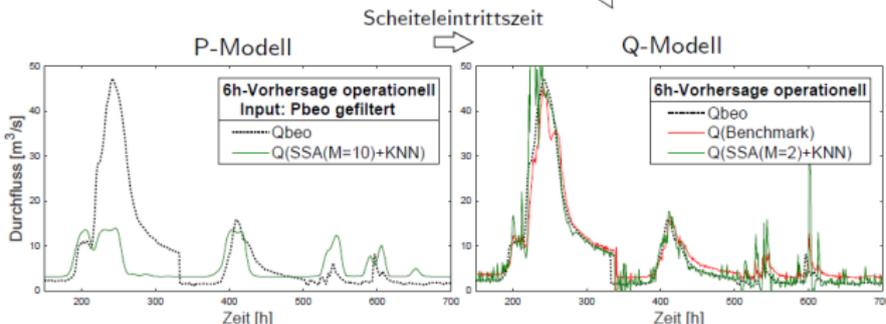
Problem Real (SSA)

Bsp. Elbersdorf-Einzugsgebiet (SSA-KNN)

operationell
6 x 1h-Vorhersage

$$Q(t-1) \cdots Q(t) \cdots Q(t+1)$$

nichtlineare
Extrapolation

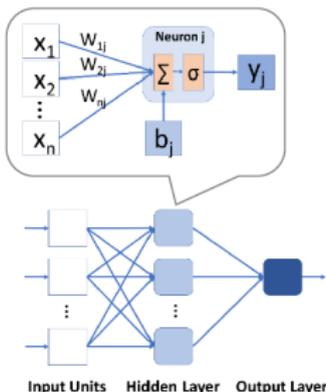


Beispiel: Encoder-Decoder LSTM

Methode

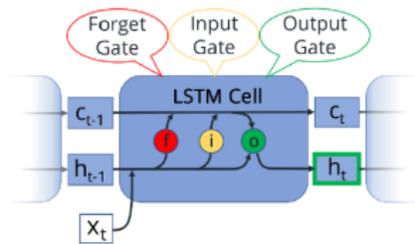
Multi Layer Perceptrons (MLP)

„Klassische“ geschichtete KNN



Long Short-Term Memory Netze (LSTM)

KNN mit effizienter, sequentieller Verarbeitung des Inputs



- x_t ... Eingangssignal (zum Zeitschritt t)
- h_t ... Hidden State (zum Zeitschritt t)
- c_t ... Cell State (zum Zeitschritt t)
- w_{nj} ... Gewichte der Eingangssignale
- b_j ... Bias
- σ ... sigmoide Aktivierungsfunktion

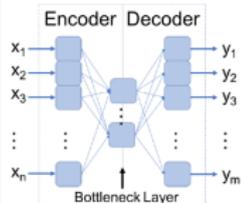
Beispiel: Encoder-Decoder LSTM

Vorgehen

4 Ansätze

Architektur-basierte Variationen

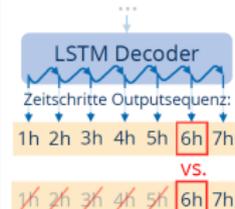
Encoder-Decoder Architektur



(Abbildung: originales Konzept für Multi Layer Perceptron)

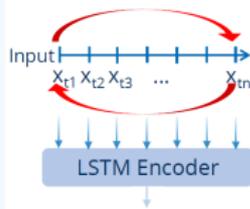


Startzeitpunkt der Vorhersage

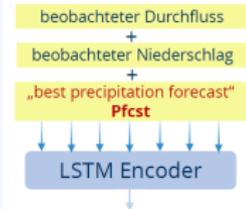


Input-basierte Variationen

Rückläufige Inputzeitreihe



„Best Precipitation Forecast“



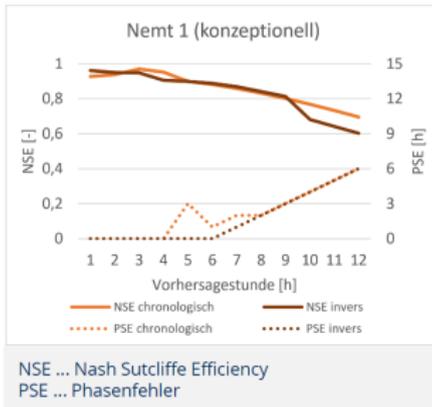
(Pfcst = nachfolgender beobachteter Niederschlag)



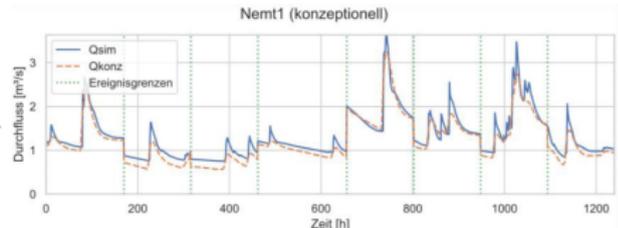
Beispiel: Encoder-Decoder LSTM

Ergebnisse

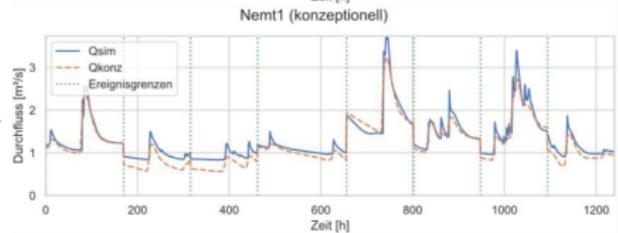
Rüchläufige Inputzeitreihe



chronologischer
Input



rüchläufiger
Input



Beispiel: Encoder-Decoder LSTM

Ergebnisse (2)

„Best Precipitation Forecast“

