

# Hydroinformatik - SoSe 2026

## UW-BHW-414-09: Maschinelles Lernen

Prof. Dr.-Ing. habil. Olaf Kolditz

<sup>1</sup>Helmholtz Centre for Environmental Research – UFZ, Leipzig

<sup>2</sup>Technische Universität Dresden – TUD, Dresden

<sup>3</sup>Center for Advanced Water Research – CAWR

<sup>4</sup>TUBAF-UFZ Center for Environmental Geosciences – C-EGS, Freiberg / Leipzig

Dresden, 05.06.2026

# Zeitplan: Hydroinformatik I+II

Sommersemester 2026: Stand: 06.04.2026

Nr.	KW	Datum	ID	Thema
01+02	16	17.04.2026	UW-BHW-414-01/02	Einführung in die Vorlesung, Umweltinformatik
03	16	17.04.2026	UW-BHW-414-03	Werkzeuge, Hello World (in C++)
05	17	24.04.2026	UW-BHW-414-04	Selbststudium: Software-Installationen
07	19	08.05.2026	UW-BHW-414-05	Objekt-Orientierte Programmierung: C++, Klassen
09	20	15.05.2026	UW-BHW-414-06	Programmiersprache Python
11	21	22.05.2026	UW-BHW-414-07/08	Modellierung, Digitalisierung - Wasser 4.0
00	22	29.05.2026		Vorlesungsfreie Woche
13	23	05.06.2026	UW-BHW-414-09	KI, Maschinelles Lernen, Neuronale Netzwerke
15	24	12.06.2026	UW-BHW-414-H	Kontinuumsmechanik, Hydromechanik
17	25	19.06.2026	UW-BHW-414-I	Differentialgleichungen, Näherungsverfahren
19	26	26.06.2026	UW-BHW-414-J	Finite-Differenzen, explizite Verfahren
21	27	03.07.2026	UW-BHW-414-K	Finite-Differenzen, implizite Verfahren
23	28	10.07.2026	UW-BHW-414-L	Gerinnehydraulik, Grundwasserhydraulik
25	29	17.07.2026	UW-BHW-414-M	Grundwasserhydraulik
27	30	24.07.2026	UW-BHW-414-N	Zusammenfassung, Klausurvorbereitung

- 1 UW-BHW-414-09: Maschinelles Lernen
  - Semesterplan

**0** Rückblick letzte Veranstaltung: Digitalisierung

---

**1** ML: Definition

**2** ML: Anwendungsmöglichkeiten

**3** ML: Methodik

**4** ML: Algorithmen

**5** ML in der Forschung

**6** ML in der Hydrologie

---

**7** Ausblick auf die nächste Veranstaltung: Mechanik

# Maschinelles Lernen

# Differences Between Machine Learning vs Neural Networks

”Machine Learning is an application or the **subfield of artificial intelligence** (AI). Machine Learning enables a system to automatically learn and progress from experience without being explicitly programmed. Machine Learning is a continuously developing practice. The goal of Machine learning is to understand the structure of data and fit that data into models, these models can be understood and used by people. In Machine Learning generally, the tasks are classified into broad categories. These categories explain how learning is received, two of the most widely used machine learning methods are supervised learning and unsupervised learning.

The neural network is inspired by the structure of the brain. The neural network contains highly interconnected entities, called units or nodes. Neural networks are deep learning technologies. It generally focuses on solving complex processes. A typical neural network is a group of algorithms, these algorithms model the data using neurons for machine learning.”

<https://www.educba.com/machine-learning-vs-neural-network/>

”**Maschinelles Lernen** ist ein Oberbegriff für die „künstliche“ Generierung von Wissen aus Erfahrung: Ein künstliches System lernt aus Beispielen und kann diese nach Beendigung der Lernphase verallgemeinern. Das heißt, es werden nicht einfach die Beispiele auswendig gelernt, sondern es „erkennt“ Muster und Gesetzmäßigkeiten in den Lerndaten. So kann das System auch unbekannte Daten beurteilen (Lerntransfer) oder aber am Lernen unbekannter Daten scheitern (Überanpassung).”

Source: [https://de.wikipedia.org/wiki/Maschinelles\\_Lernen](https://de.wikipedia.org/wiki/Maschinelles_Lernen)

”Maschinelles Lernen ist ein Oberbegriff für die „künstliche“ **Generierung von Wissen aus Erfahrung**: Ein künstliches System lernt aus Beispielen und kann diese nach Beendigung der Lernphase **verallgemeinern**. Das heißt, es werden nicht einfach die Beispiele auswendig gelernt, sondern es „erkennt“ **Muster und Gesetzmäßigkeiten in den Lerndaten**. So kann das System auch unbekannte Daten beurteilen (**Lerntransfer**) oder aber am Lernen unbekannter Daten scheitern (Überanpassung).”

Source: [https://de.wikipedia.org/wiki/Maschinelles\\_Lernen](https://de.wikipedia.org/wiki/Maschinelles_Lernen)

”Aus dem weiten Spektrum möglicher Anwendungen seien hier genannt: automatisierte Diagnoseverfahren, Erkennung von Kreditkartenbetrug, Aktienmarktanalysen, Klassifikation von Nukleotidsequenzen, Sprach- und Texterkennung sowie autonome Systeme.” Naturwissenschaften ?

Source: [https://de.wikipedia.org/wiki/Maschinelles\\_Lernen](https://de.wikipedia.org/wiki/Maschinelles_Lernen)

”Das Thema ist eng verwandt mit „Knowledge Discovery in Databases“ und „Data-Mining“, bei dem es jedoch vorwiegend um das Finden von neuen **Mustern** und **Gesetzmäßigkeiten** geht. Viele Algorithmen können für beide Zwecke verwendet werden. Außerdem können Methoden der „Knowledge Discovery in Databases“ genutzt werden, um Lerndaten für „maschinelles Lernen“ zu produzieren oder vorzuverarbeiten, und Algorithmen aus dem maschinellen Lernen finden beim Data-Mining Anwendung.“

Source: [https://de.wikipedia.org/wiki/Maschinelles\\_Lernen](https://de.wikipedia.org/wiki/Maschinelles_Lernen)

”Die praktische Umsetzung geschieht mittels Algorithmen. Verschiedene Algorithmen aus dem Bereich des maschinellen Lernens lassen sich grob nach diesem Schema einteilen:”

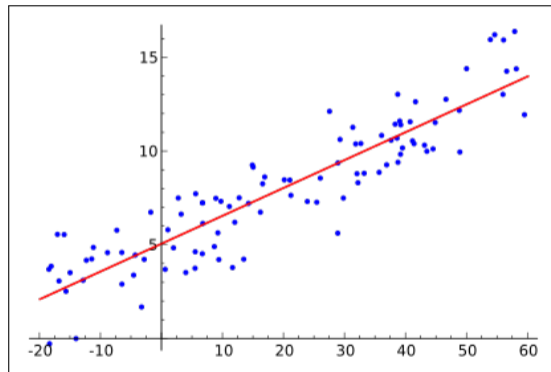
- ▶ Überwachtes Lernen (engl. supervised learning)
- ▶ Unüberwachtes Lernen (engl. unsupervised learning)

Source: [https://de.wikipedia.org/wiki/Maschinelles\\_Lernen](https://de.wikipedia.org/wiki/Maschinelles_Lernen)

”Der Algorithmus lernt eine Funktion aus gegebenen Paaren von Ein- und Ausgaben. Dabei stellt während des Lernens ein „Lehrer“ den korrekten Funktionswert zu einer Eingabe bereit. Ziel beim überwachten Lernen ist, dass dem Netz nach mehreren Rechengängen mit unterschiedlichen Ein- und Ausgaben die Fähigkeit antrainiert wird, **Assoziationen** herzustellen. Ein Teilgebiet des überwachten Lernens ist die automatische Klassifizierung. Ein Anwendungsbeispiel wäre die Handschrifterkennung.”

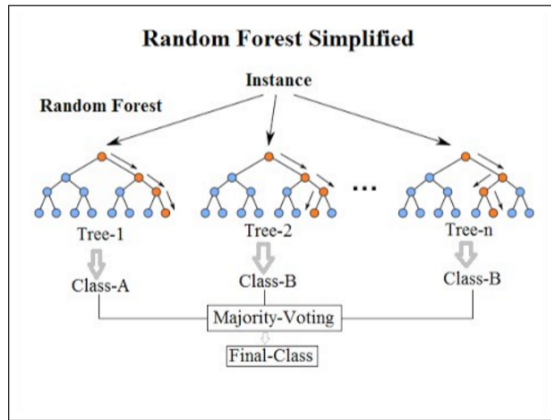
Source: [https://de.wikipedia.org/wiki/Maschinelles\\_Lernen](https://de.wikipedia.org/wiki/Maschinelles_Lernen)

”Die lineare Regression ist ein Spezialfall der Regressionsanalyse, also ein statistisches Verfahren, mit dem versucht wird, eine beobachtete abhängige Variable durch eine oder mehrere unabhängige Variablen zu erklären. Bei der linearen Regression wird dabei ein lineares Modell angenommen. Es werden also nur solche Zusammenhänge herangezogen, bei denen die abhängige Variable eine Linearkombination der Regressionskoeffizienten (aber nicht notwendigerweise der unabhängigen Variablen) ist.”



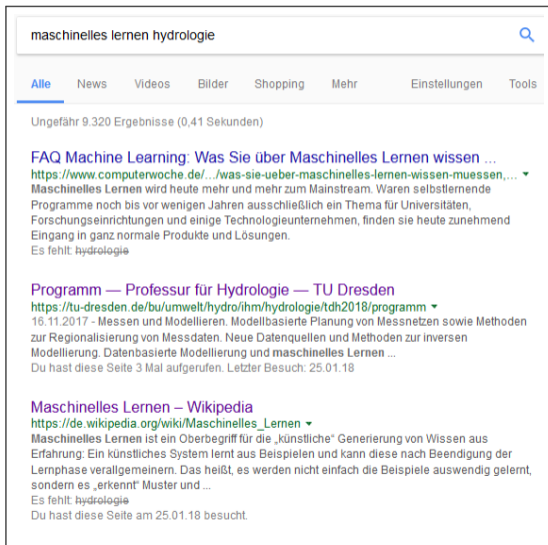
Quelle: [https://de.wikipedia.org/wiki/Lineare\\_Regression](https://de.wikipedia.org/wiki/Lineare_Regression)

”Ein Random Forest ist ein Klassifikations- und Regressionsverfahren, das aus mehreren unkorrelierten Entscheidungsbäumen besteht. Alle Entscheidungsbäume sind unter einer bestimmten Art von Randomisierung während des Lernprozesses gewachsen. ... Random Forests können auch zur Regression eingesetzt werden.”



Quelle: [https://de.wikipedia.org/wiki/Random\\_Forest](https://de.wikipedia.org/wiki/Random_Forest)

# ML Beispiele aus der Hydrologie



The screenshot shows a Google search interface with the query 'maschinelles lernen hydrologie'. The search results are filtered to 'Alle' (All). The first result is a FAQ titled 'FAQ Machine Learning: Was Sie über Maschinelles Lernen wissen ...' from computerwoche.de. The second result is a program page from TU Dresden titled 'Programm — Professur für Hydrologie — TU Dresden'. The third result is the Wikipedia page for 'Maschinelles Lernen'.

maschinelles lernen hydrologie

Alle News Videos Bilder Shopping Mehr Einstellungen Tools

Ungefähr 9.320 Ergebnisse (0,41 Sekunden)

**FAQ Machine Learning: Was Sie über Maschinelles Lernen wissen ...**  
<https://www.computerwoche.de/.../was-sie-ueber-maschinelles-lernen-wissen-muessen,...>  
Maschinelles Lernen wird heute mehr und mehr zum Mainstream. Waren selbstlernende Programme noch bis vor wenigen Jahren ausschließlich ein Thema für Universitäten, Forschungseinrichtungen und einige Technologieunternehmen, finden sie heute zunehmend Eingang in ganz normale Produkte und Lösungen.  
Es fehlt: hydrologie

**Programm — Professur für Hydrologie — TU Dresden**  
<https://tu-dresden.de/bu/umwelt/hydro/ihm/hydrologie/tdh2018/programm>  
16.11.2017 - Messen und Modellieren. Modellbasierte Planung von Messnetzen sowie Methoden zur Regionalisierung von Messdaten. Neue Datenquellen und Methoden zur inversen Modellierung. Datenbasierte Modellierung und maschinelles Lernen ...  
Du hast diese Seite 3 Mal aufgerufen. Letzter Besuch: 25.01.18

**Maschinelles Lernen – Wikipedia**  
[https://de.wikipedia.org/wiki/Maschinelles\\_Lernen](https://de.wikipedia.org/wiki/Maschinelles_Lernen)  
Maschinelles Lernen ist ein Oberbegriff für die „künstliche“ Generierung von Wissen aus Erfahrung: Ein künstliches System lernt aus Beispielen und kann diese nach Beendigung der Lernphase verallgemeinern. Das heißt, es werden nicht einfach die Beispiele auswendig gelernt, sondern es „erkennt“ Muster und ...  
Es fehlt: hydrologie  
Du hast diese Seite am 25.01.18 besucht.

## Suche: maschinelles lernen hydrologie

- ▶ Google
- ▶ ChatGPT
- ▶ DeepSeek

## Water Resources Research



### RESEARCH ARTICLE

10.1002/2016WR019933

#### Key Points:

- Groundwater level change can be modeled with high accuracy using machine learning methods
- Model framework does not require subsurface parameters and simulates comparable groundwater levels to numerical models of physical flow
- Seasonal irrigation demand has the highest relevance to groundwater level change compared to climate and streamflow inputs for most wells

#### Supporting Information:

- Supporting Information S1

#### Correspondence to:

T. A. Russo,

## Machine learning algorithms for modeling groundwater level changes in agricultural regions of the U.S.

S. Sahoo<sup>1</sup> , T. A. Russo<sup>1</sup> , J. Elliott<sup>2,3</sup> , and I. Foster<sup>2,3,4</sup>

<sup>1</sup>Department of Geosciences, Pennsylvania State University, University Park, Pennsylvania, USA, <sup>2</sup>Computation Institute, University of Chicago, Chicago, Illinois, USA, <sup>3</sup>Mathematics and Computer Science Division, Argonne National Laboratory, Lemont, Illinois, USA, <sup>4</sup>Department of Computer Science, University of Chicago, Chicago, Illinois, USA

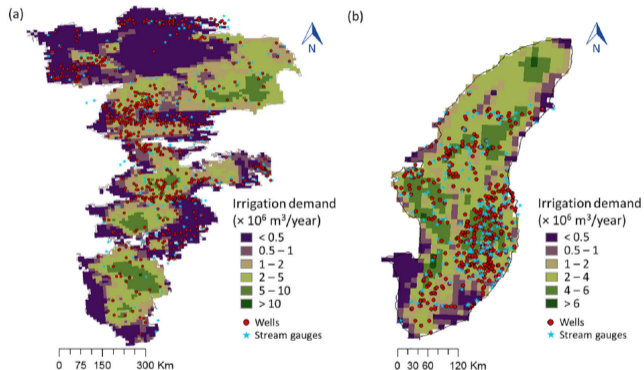
**Abstract** Climate, groundwater extraction, and surface water flows have complex nonlinear relationships with groundwater level in agricultural regions. To better understand the relative importance of each driver and predict groundwater level change, we develop a new ensemble modeling framework based on spectral analysis, machine learning, and uncertainty analysis, as an alternative to complex and computationally expensive physical models. We apply and evaluate this new approach in the context of two aquifer systems supporting agricultural production in the United States: the High Plains aquifer (HPA) and the Mississippi River Valley alluvial aquifer (MRVA). We select input data sets by using a combination of mutual information, genetic algorithms, and lag analysis, and then use the selected data sets in a Multilayer Perceptron network architecture to simulate seasonal groundwater level change. As expected, model results suggest that irriga-

Sahoo, S., T. A. Russo, J. Elliott, and I. Foster (2017), Machine learning algorithms for modeling groundwater level changes in agricultural regions of the U.S., *Water Resour. Res.*, 53, 3878–3895, doi:10.1002/2016WR019933.

**HPA** High Plains Aquifer

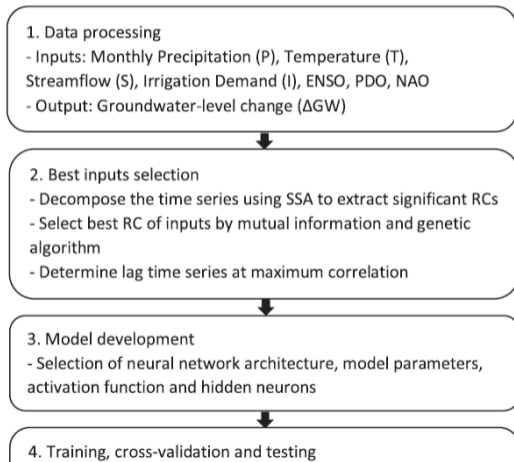
**MRVA** Mississippi River Valley alluvial Aquifer

Source: Sahoo et al. (2017)

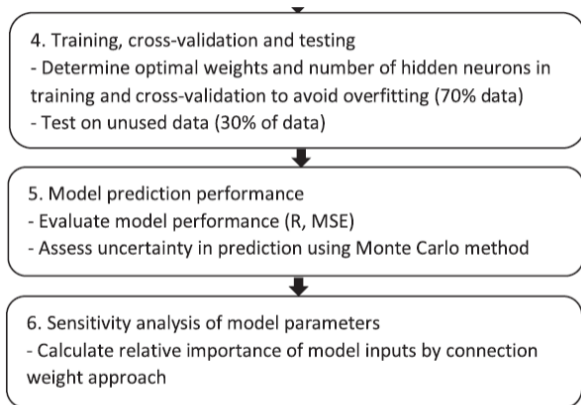


**Figure 1.** (a) HPA and (b) MRVA showing observation wells (red circles), stream gauges (blue stars), and DSSAT-simulated annual average irrigation demand (m<sup>3</sup>/yr) at 5 arc min resolution. The irrigation demand in the HPA is shown for seven crops: corn, sorghum, soybean, cotton, barley, spring wheat, and winter wheat, and accounts for 83% of the total irrigated cropland. The irrigation demand in the MRVA is shown for five crops: cotton, sorghum, soybean, corn, and winter wheat, and accounts for 93% of the total irrigated cropland.

Source: Sahoo et al. (2017)

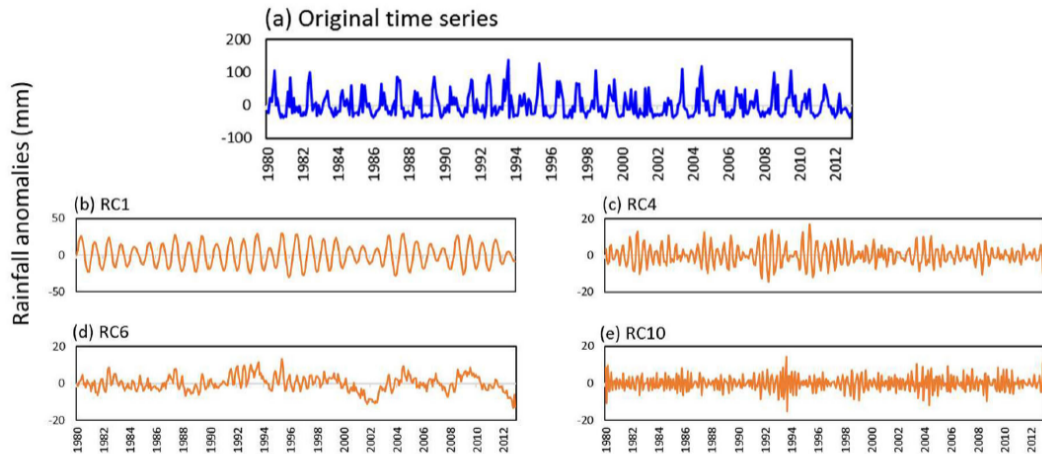


Source: Sahoo et al. (2017)



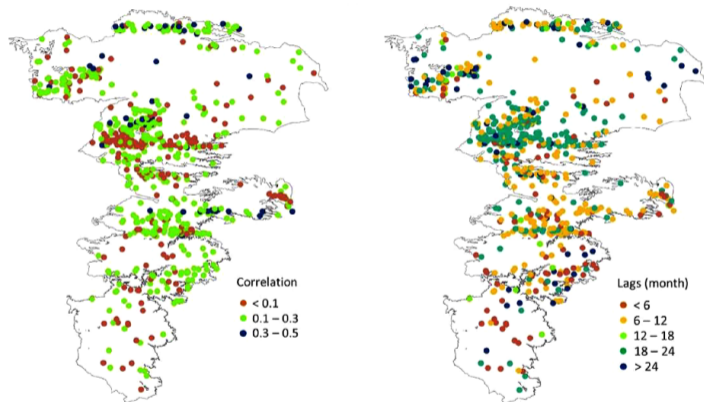
**Figure 2.** Methodological framework of the proposed groundwater model.

Source: Sahoo et al. (2017)



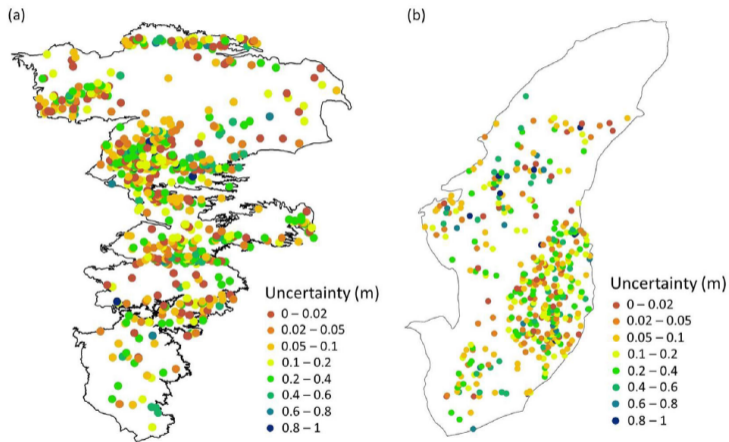
**Figure 3.** (a) Original time series of the rainfall anomalies. (b-e) Decomposed time series (4 RCs shown here out of 10 RCs) using SSA.

Source: Sahoo et al. (2017)



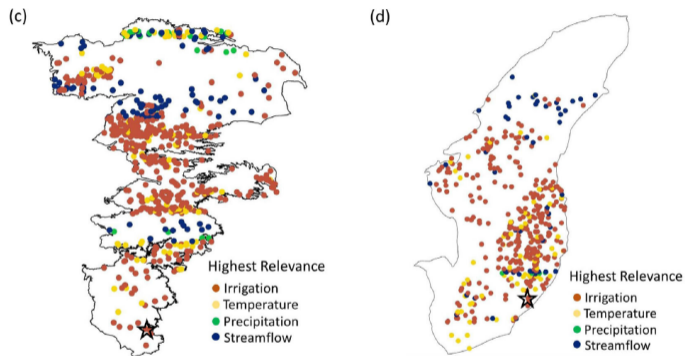
**Figure 4.** (a) Correlations and (b) respective time lags between monthly groundwater level and precipitation over the HPA for 687 well locations.

Source: Sahoo et al. (2017)



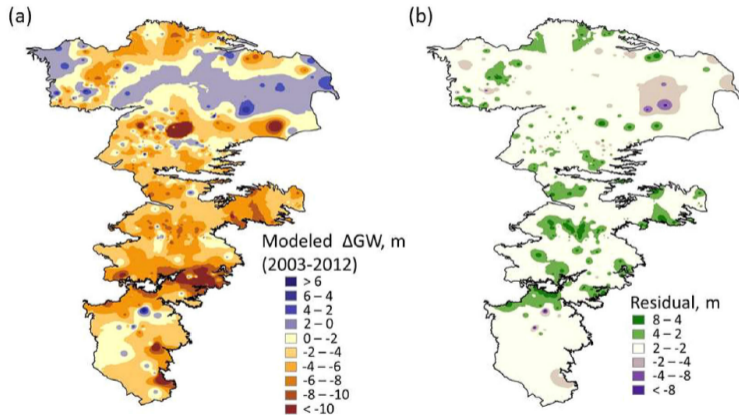
**Figure 8.** Average seasonal groundwater level uncertainty in model prediction for (a) 687 wells in the HPA and (b) 437 wells in MRVA calculated using HANN model and Monte Carlo method.

Source: Sahoo et al. (2017)



**Figure 9.** Relative importance of model inputs to groundwater level change prediction by connection weight approach for the example wells in the (a) HPA and (b) MRVA (I: irrigation demand, P: precipitation, T: temperature, S: stream discharge), and the input parameter with the highest relevance at each well location using the HANN model for (c) 687 wells in the HPA and (d) 437 wells in the MRVA. The black stars in the HPA and MRVA indicate the locations of two example sites. The ranking for all of the input parameters and percentage of wells that fall under each category is included in Table 2.

Source: Sahoo et al. (2017)



Source: Sahoo et al. (2017)

## Geophysical Research Letters

### RESEARCH LETTER

10.1002/2017GL075661

#### Key Points:

- A new geothermal heat flux map of Greenland is obtained within ~15% accuracy using machine learning techniques
- The new map honors regional geology, tectonic settings, and ice core measurements
- Pockets of high heat flux are predicted in central-north Greenland and upstream of several fast-flowing outlet glaciers






#### Supporting Information:

- Supporting Information S1

#### Correspondence to:

S. Rezvanbehbahani,  
soroushr@ku.edu

## Predicting the Geothermal Heat Flux in Greenland: A Machine Learning Approach

Soroush Rezvanbehbahani<sup>1,2</sup> , Leigh A. Stearns<sup>1,2</sup> , Amir Kadivar<sup>3</sup> , J. Doug Walker<sup>1</sup> ,  
and C. J. van der Veen<sup>4</sup> 

<sup>1</sup>Department of Geology, University of Kansas, Lawrence, KS, USA, <sup>2</sup>Center for Remote Sensing of Ice Sheets, University of Kansas, Lawrence, KS, USA, <sup>3</sup>Department of Mathematics and Statistics, McGill University, Montréal, Québec, Canada, <sup>4</sup>Department of Geography and Atmospheric Sciences, University of Kansas, Lawrence, KS, USA

**Abstract** Geothermal heat flux (GHF) is a crucial boundary condition for making accurate predictions of ice sheet mass loss, yet it is poorly known in Greenland due to inaccessibility of the bedrock. Here we use a machine learning algorithm on a large collection of relevant geologic features and global GHF measurements and produce a GHF map of Greenland that we argue is within ~15% accuracy. The main features of our predicted GHF map include a large region with high GHF in central-north Greenland surrounding the NorthGRIP ice core site, and hot spots in the Jakobshavn Isbræ catchment, upstream of Petermann Gletscher, and near the terminus of Nioghalvfjærdsfjorden glacier. Our model also captures the trajectory of Greenland movement over the Icelandic plume by predicting a stripe of elevated GHF in central-east Greenland. Finally, we show that our model can produce substantially more accurate predictions if additional measurements of GHF in Greenland are provided.

Source: AGU

**HELMHOLTZ**  
RESEARCH FOR GRAND CHALLENGES

**DIGITAL**

Applicability of machine learning-based approaches to predict CRNS roving-derived soil moisture estimates on the catchment scale

Dirk Nixdorf\*, Moritz Schirler\*, Marco Heesmann\* and Thomas Kalbacher\*

\* Helmholtz Institute for Environmental Research, Leipzig, 04215

**Motivation**

- Soil moisture is an essential critical state variable in soil surface hydrology and a key component of bioclimatology
- Measuring principles exist on different scales (TDR, CRNS and satellite observations)
- We integrate CRNS-derived soil moisture estimates with machine learning algorithms towards an improvement of soil moisture predictions on the catchment scale

**Study Site**

- Müglitz River Basin south of Dresden (~200 km<sup>2</sup>)
- The basin is heavily impacted by heavy rain events and flash floods
- Ongoing water-borne erosion watershed-scale research and monitoring efforts (soil/science)

**Methodology**

**Preprocessing workflow**

- 3 learners selected for regression problem
- Trained with subset from entire CRNS dataset and tested with Müglitz CRNS data
- 3 target variables, 2 different weather datasets, multiple hydrological models, or different resolutions

**Learner Application**

- 3 learners selected for regression problem
- Trained with subset from entire CRNS dataset and tested with Müglitz CRNS data
- 3 target variables, 2 different weather datasets, multiple hydrological models, or different resolutions

**Hydrological Modeling**

- Hydrological model of Müglitz Basin using SWAT 2.2.5 (Nash et al. 2012)
- Model calibration with gauge data
- Provision of daily soil moisture estimates

**ML Model Performance**

- RF outperforms LR and ANN for each target variable
- Large moisture estimates are underestimated and vice versa
- Important features are altitude, bulk density and aspect, but also moisture target

**Basin-scale Prediction**

- Regenerated data-driven soil moisture estimates are compared with daily output of state hydrological model on a 200m grid
- Although a systematic bias between the results of both methods exists, both temporal trends and spatial patterns resemble each other

**Conclusions**

- Random Forest outperforms other approaches to predict CRNS products using exclusively features from open data sources
- Most relevant features for soil moisture prediction are weather input measured at gauging stations (bulk density, altitude, slope, topographic aspect and meteorological features). This reveals that existing correction functions on CRNS data require adaptation
- Applying the trained random forest on the entire Müglitz catchment, obtained daily moisture estimates are in the range of the CRNS cover records
- Comparing with the soil moisture product of mHM, soil moisture predicted by random forest is underestimated. Though, temporal trends as well as spatial distribution patterns of both approaches resemble each other

**References**

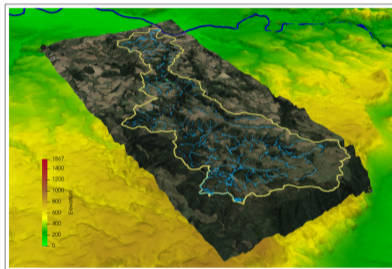
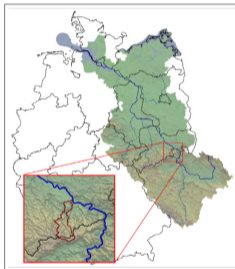
1. Nixdorf, D., Schirler, M., Heesmann, M., Kalbacher, T. (2023) Machine Learning-Based Prediction of Catchment-Scale Soil Moisture from CRNS Data. *Water*, 15(12), 2188. <https://doi.org/10.3390/w15122188>

2. Nixdorf, D., Schirler, M., Heesmann, M., Kalbacher, T. (2023) Machine Learning-Based Prediction of Catchment-Scale Soil Moisture from CRNS Data. *Water*, 15(12), 2188. <https://doi.org/10.3390/w15122188>

3. Nixdorf, D., Schirler, M., Heesmann, M., Kalbacher, T. (2023) Machine Learning-Based Prediction of Catchment-Scale Soil Moisture from CRNS Data. *Water*, 15(12), 2188. <https://doi.org/10.3390/w15122188>

**HELMHOLTZ**  
RESEARCH FOR GRAND CHALLENGES

## Müglitz-Studie



## Video

Simulationen: Thomas Kalbacher, Thomas Fischer, Johannes Boog, Erik Nixdorf  
Visualisierung: Karsten Rink, Ozan Sen (Umweltinformatik)

- 1** Tobias Reitmaier: Aktives Lernen für Klassifikationsprobleme unter der Nutzung von Strukturinformationen. kassel university press, Kassel 2015, ISBN 978-3-86219-999-0, S. 1 (Google books).
- 2** Lillian Pierson: Data Science für Dummies. 1. Auflage. Wiley-VCH Verlag, Weinheim 2016, ISBN 978-3-527-80675-1, S. 105 f. (Google books).
- 3** Ralf Mikut: Data Mining in der Medizin und Medizintechnik. KIT Scientific Publishing, 2008, ISBN 978-3-86644-253-5, S. 34 (Google books).
- 4** Paul Fischer: Algorithmisches Lernen. Springer-Verlag, 2013, ISBN 978-3-663-11956-2, S. 6–7 (Google books).
- 5** Andreas C. Müller / Sarah Guido: Einführung in Machine Learning mit Python. Hrsg.: O'Reilly-Verlag. ISBN 978-3-96009-049-6 (oreilly.de [abgerufen am 28. November 2017]).
- 6** Sebastian Raschka: Machine Learning mit Python | ISBN 978-3-95845-422-4. mitp-Verlag, abgerufen am 28. November 2017

# Testfragen

- 1 Beschreiben sie die Methode "Maschinelles Lernen" mit eigenen Worten (3 Sätze).
- 2 Was ist der Unterschied zwischen "supervised" und "un-supervised" Maschinellern Lernen ?
- 3 Welche Methoden des Maschinellen Lernens kennen Sie ?
- 4 ...