

Die Bedeutung von Feature Engineering bei der KI-basierten räumlichen Modellierung am Beispiel einer DSM-Anwendung

Markus Möller¹, Simone Zepp², Martin Wiesmeier³, Heike Gerighausen¹ und Uta Heiden²

¹ Julius Kühn-Institut

² Deutsches Zentrum für Luft- und Raumfahrt e.V. (DLR)

³ Bayerische Landesanstalt für Landwirtschaft (LfL)

Exkursion zur DBG-Jahrestagung am 8.9.2023 (Eickendorf)



Bodenschätzung in Deutschland Rückblick – Würdigung – Ausblick

10

Von der Bodenschätzung höchstbewertete landwirtschaftlich nutzbare Böden



Sie befinden sich vor allem auf den Lössflächen der Magdeburger und Hildesheimer Börde sowie in der Kölner Bucht und sind mit L 1 Lö klassifiziert, z. B.

Eickendorf (FA Staßfurt)	Wertzahlen 100/104
Brumby (FA Staßfurt)	Wertzahlen 100/102
Börßum (FA Wolfenbüttel)	Wertzahlen 100/104
Wülknitz (FA Bitterfeld-Wolfen)	Wertzahlen 100/104
Nauendorf (FA Halle (S.)-Nord)	Wertzahlen 100/104
Leuththal (FA Jena)	Wertzahlen 100/98
Rannstedt (FA Jena)	Wertzahlen 100/100
Hönnersum (FA Hildesheim)	Wertzahlen 100/104

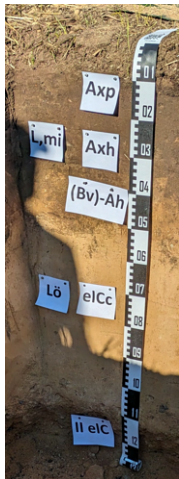
Beschreibung
des Musterstücks für Acker-Nr. 2

Landesfinanzamt Magdeburg Regierungsbezirk Magdeburg
Finanzamt Schönebeck "Elbe" Kreis Calbe
Katasteramt Schönebeck "Elbe" Gemeinde Eickendorf

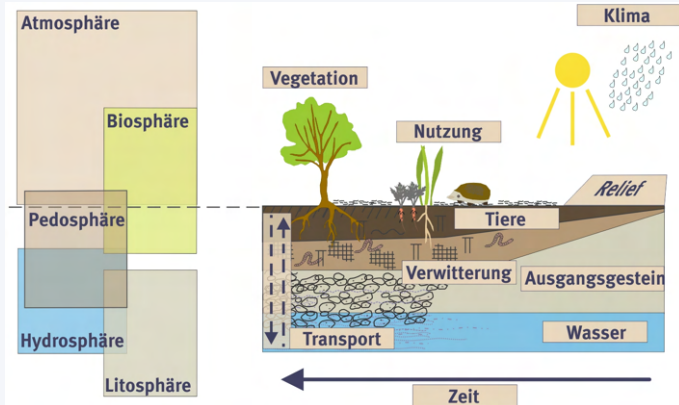
Tag der Beschätzung	Gemarkung	Flächennr.	Flächennr.	Eigentümer	Größe		Klima		Bodengefüge
					Ar	St	Jahr	Jahr	
18.5.1934	Eickendorf	2	277/24	Haberhauffe, Wilh. Elze geb. Wiegand	100	100	100	100	5. mi. 1.2 5. mi. 1.2 5. mi. 1.2

Klasse	Zus. zur Klasse	Bodenwert	Besondereheiten		Anmerkungen
			Einzelheiten	Einzelheiten	
L	1	100			

Motivation



Bodenkartierung



UBA, 2010. Die Böden Deutschlands – Sehen, Erkunden, Verstehen. Umweltbundesamt.

Motivation

B



2410182

Aufnahme: D. Kühn 2000

TT: u/leu
Tschermosom aus Schluff über tiefem Carbonatschluff

TTn: u-u(b)/f-eu(b)
Normtschermosom aus Hangschluff über tiefem Fluvicarbonatschluff (Beckenbildung)

vkTTn: uk-su(b)/pky-su(b)/fgl-clu(b)
kolluvialer (Acker)Normtschermosom aus Kolluvialsandschluff über Sandschluff über tiefem Beckenkalklehm-schluff



2410184

Aufnahme: D. Kühn 2000

TT: l
Tschermosom aus Lehm

BB-TT: p-l(Lp)/p-l(Mg)
Braunerde-Tschermosom aus Lehm (Decklehm) über Lehm (Geschiebemergel)

v.sBB-TTp: pky-sl(Lp)/pfl-l(Mg)/gm-cl(Mg)
pseudovergleyte (Acker)Braunerde-Tschermosom aus Sandlehm (Decklehm) über Fließnormallehm über tiefem Grundmoränenkalknormallehm



2410185

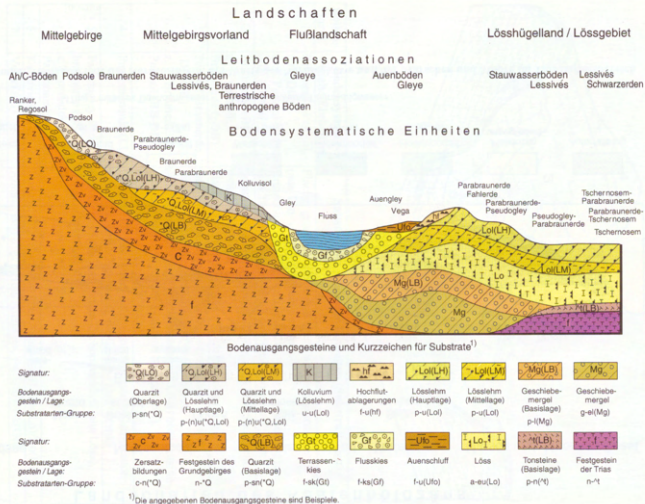
Aufnahme: D. Kühn 2000

TT: l/lel
Tschermosom aus Lehm über tiefem Carbonatlehm

SS-TT: u-l(Lp)/g-el(Mg)
Pseudogley-Tschermosom aus Hanglehm (Decklehm) über Moränencarbonatlehm

vkBB-SS-TT: uk-l(Lp)/pky-l(Lp)/gm-cs-l(Mg)
kolluviale (Acker)Braunerde-Pseudogley-Tschermosom aus Kolluvialnormallehm über Normallehm (Decklehm) über tiefem Grundmoränenkalksandlehm

Bodenkartierung



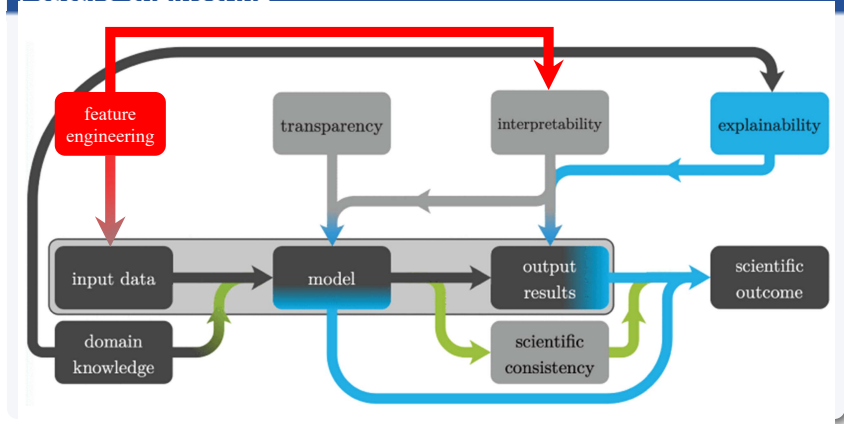
Ad-hoc-Arbeitsgruppe
Boden, 2005.
Bodenkundliche
Kartieranleitung,
5. Auflage.
Schweizerbart'sche
Verlagsbuchhandlung
(Nägele und Obermiller),
Stuttgart.

Bodenkartierung

	Expert/knowledge-driven soil mapping	Data/technology-driven soil mapping
<i>Target variables:</i>	Soil types (soil series)	Analytical soil properties
<i>Spatial data model:</i>	Discrete (soil bodies)	Continuous/hybrid (quantities / probabilities)
<i>Major inputs:</i>	Expert knowledge / soil profile description	Laboratory data / proximal soil sensing
<i>Important covariates:</i>	Soil delineations (photo-interpretation)	Remote sensing images, DEM-derivatives
<i>Spatial prediction model:</i>	Averaging per polygon	Automated (geo)statistics
<i>Accuracy assessment:</i>	Validation of soil mapping units (kappa)	Cross-validation (RMSE)
<i>Data representation:</i>	Polygon maps + attribute tables	Gridded maps + prediction error map
<i>Major technical aspect:</i>	Cartographic scale	Grid cell size
<i>Soil sampling strategies:</i>	Free survey (surveyor selects sampling)	Statistical (design/model-based) sampling

Hengl, T., MacMillan, R.A., 2019. Predictive Soil Mapping with R. OpenGeoHub foundation, Wageningen.
<https://www.soilmapper.org>, ISBN: 978-0-359-30635-0

Feature engineering



Roscher, R., Bohn, B., Duarte, M.F., Garcke, J., 2020. Explainable Machine Learning for Scientific Insights and Discoveries. IEEE Access 8, 42200–42216. <https://doi.org/10.1109/ACCESS.2020.2976199>

Rentschler, T., 2021. Explainable machine learning in soil mapping: Peeking into the black box. <https://doi.org/10.15496/PUBLIKATION-57960>

Transparency considers the ML approach, **interpretability** considers the ML model together with data, and **explainability** considers the model, the data, and human involvement.

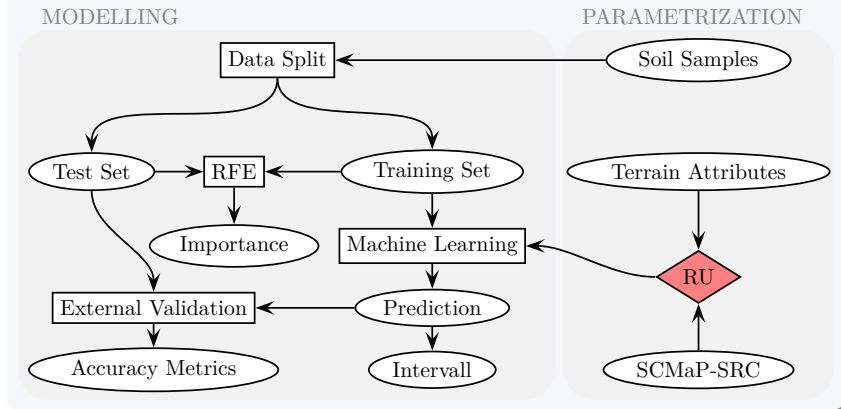
Feature engineering

- adjusting and reworking the predictors to enable models to better uncover predictor-response relationships
- the process of creating representations of data that increase the effectiveness of a model

Roscher, R., Bohn, B., Duarte, M.F., Garcke, J., 2020. Explainable Machine Learning for Scientific Insights and Discoveries. IEEE Access 8, 42200–42216. <https://doi.org/10.1109/ACCESS.2020.2976199>

Kuhn, M., Johnson, K., 2019. Feature Engineering and Selection: A Practical Approach for Predictive Models, 1st ed. Chapman and Hall/CRC. <https://doi.org/10.1201/9781315108230>

Feature engineering



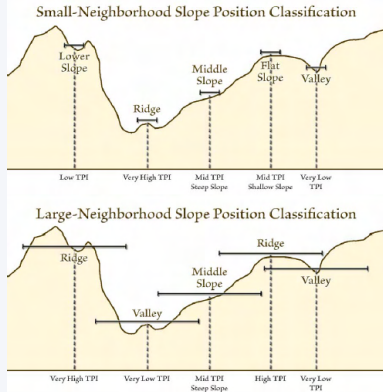
Möller, M., Zepp, S., Wiesmeier, M., Gerighausen, H., Heiden, U., 2022.
Scale-Specific Prediction of Topsoil Organic Carbon Contents Using Terrain
Attributes and SCMaP Soil Reflectance Composites. Remote Sensing 14, 2295.
<https://doi.org/10.3390/rs14102295>



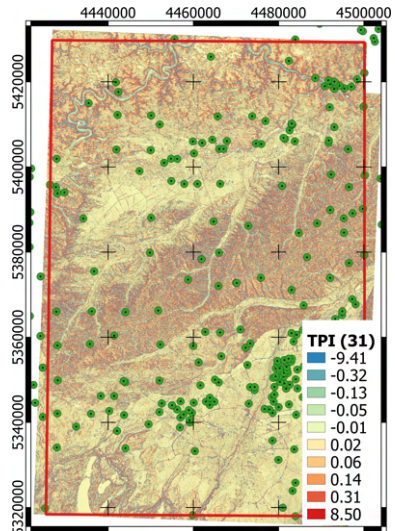
DOI [10.5281/zenodo.7895529](https://doi.org/10.5281/zenodo.7895529)

Feature engineering: Scale

Topographic Position Index

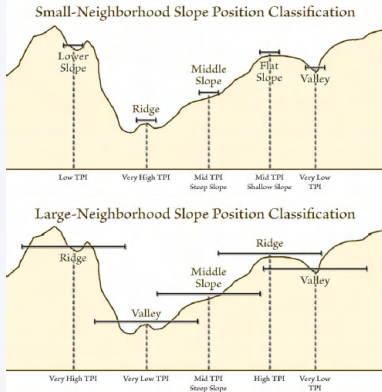


Guisan, A., Weiss, S.B., Weiss, A.D., 1999. GLM versus CCA spatial modeling of plant species distribution. *Plant Ecology* 143, 107–122.
<https://doi.org/10.1023/A:1009841519580>

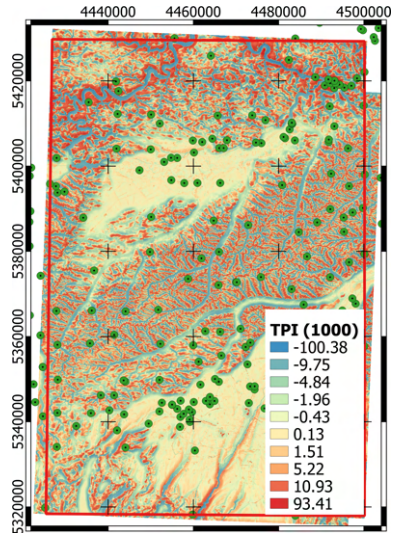


Feature engineering: Scale

Topographic Position Index



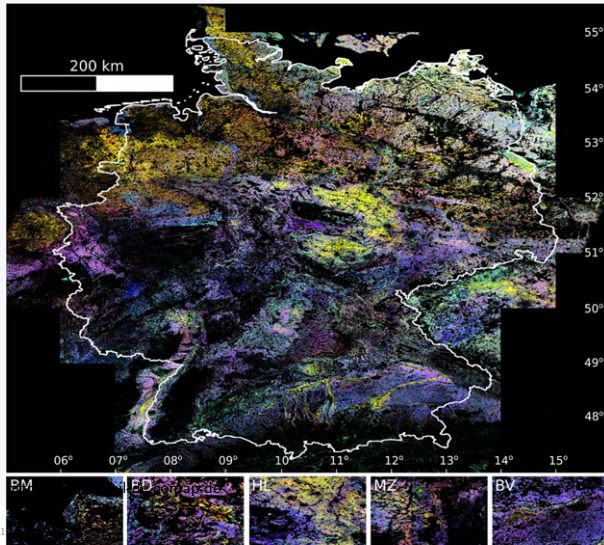
Guisan, A., Weiss, S.B., Weiss, A.D., 1999. GLM versus CCA spatial modeling of plant species distribution. *Plant Ecology* 143, 107–122.
<https://doi.org/10.1023/A:1009841519580>



Feature engineering: Time

Bodenreflektanzkomposite (SCMaP-SRC)

Remote Sensing of Environment 205 (2018) 1–17



Rogge, D., Bauer, A., Zeidler, J., Mueller, A., Esch, T., Heiden, U., 2018. Building an exposed soil composite processor (SCMaP) for mapping spatial and temporal characteristics of soils with Landsat imagery (1984–2014). *Remote Sensing of Environment* 205, 1–17. <https://doi.org/10.1016/j.rse.2017.11.004>

Zepp, S., Heiden, U., Bachmann, M., Möller, M., Wiesmeier, M., van Wesemael, B., 2023. Optimized Bare Soil Compositing for Soil Organic Carbon Prediction of Topsoil Croplands in Bavaria using Landsat. *ISPRS Journal of Photogrammetry and Remote Sensing* 202, 287–302. <https://doi.org/10.1016/j.isprsjprs.2023.06.003>

Feature engineering ...

- formalisiert Experten- und Domainwissen,
- erhöht die Modellgüte und -akzeptanz.

Feature engineering ...

- formalisiert Experten- und Domainwissen,
- erhöht die Modellgüte und -akzeptanz.

KoBoS

Deutschlandweite raum-zeitliche Modellierung von **K**ohlenstoffgehalten landwirtschaftlicher **B**öden durch eine integrative Auswertung von **S**atellitenbildzeitreihen und Geodaten (**KoBoS**)

- Bereitstellung von Webdiensten deutschlandweit erklärender erklärenden Variablen nach FAIR-Prinzipien,
- Entwicklung eines erweiterbaren, dynamischen und webbasierten Open-Source-Modells, das für beliebige Gebiete in Deutschland anwendbar ist.

gefördert durch das BMEL-Klimaschutz-Sofortprogramm 2022 | #RessortForschtKlima | <https://wissen.julius-kuehn.de/klimaschutz/projekte/erhoehung-kohlenstoffspeicherpotentiale/kobos>

Fragen?

