

# HENRY

Hydraulic Engineering Repository

Ein Service der Bundesanstalt für Wasserbau

---

Conference Poster, Published Version

**Kittner, Jonas; Demuzere, Matthias; Bechtel, Benjamin**  
**Crowdsourcing von Lufttemperaturmessdaten zur  
kleinräumigen Modellierung von städtischen  
Temperaturverteilungen**

Zur Verfügung gestellt in Kooperation mit/Provided in Cooperation with:  
**Deutsche Meteorologische Gesellschaft, KlimaCampus Hamburg**

---

Verfügbar unter/Available at: <https://hdl.handle.net/20.500.11970/107514>

Vorgeschlagene Zitierweise/Suggested citation:

Kittner, Jonas; Demuzere, Matthias; Bechtel, Benjamin (2021): Crowdsourcing von Lufttemperaturmessdaten zur kleinräumigen Modellierung von städtischen Temperaturverteilungen. Poster präsentiert bei: 12. Deutsche Klimatagung, Online-Tagung, 15. bis 18. März 2021. <https://doi.org/10.5194/dkt-12-30>.

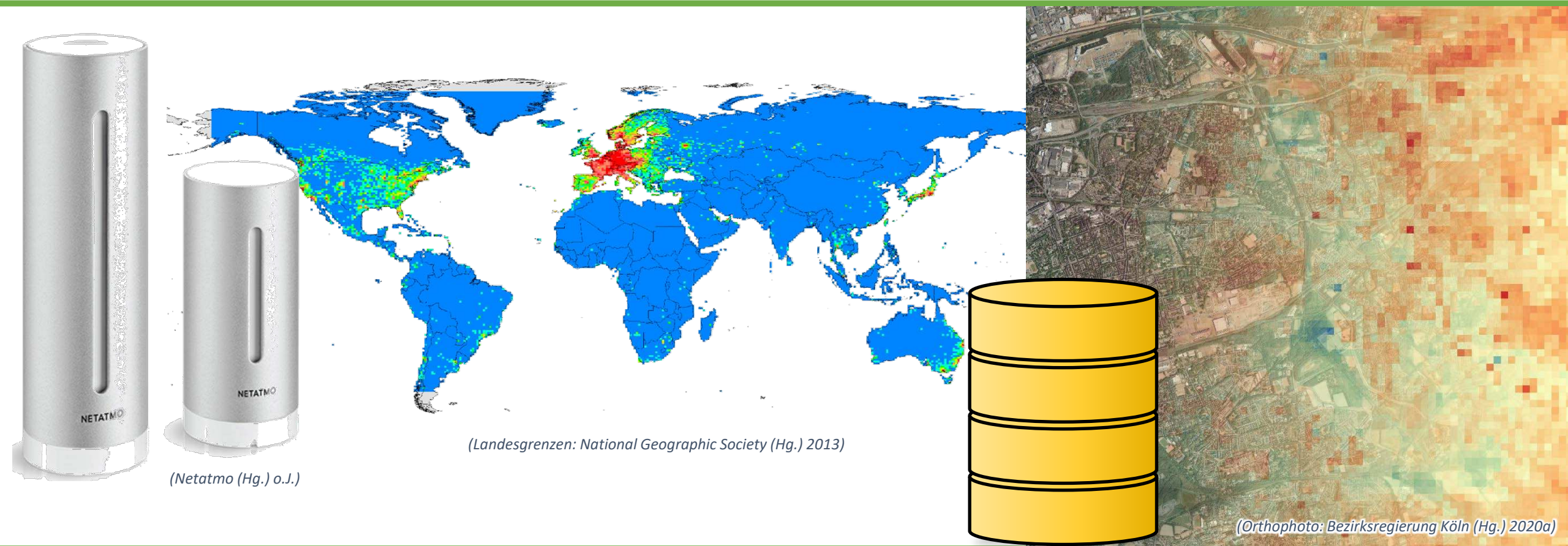
**Standardnutzungsbedingungen/Terms of Use:**

Die Dokumente in HENRY stehen unter der Creative Commons Lizenz CC BY 4.0, sofern keine abweichenden Nutzungsbedingungen getroffen wurden. Damit ist sowohl die kommerzielle Nutzung als auch das Teilen, die Weiterbearbeitung und Speicherung erlaubt. Das Verwenden und das Bearbeiten stehen unter der Bedingung der Namensnennung. Im Einzelfall kann eine restriktivere Lizenz gelten; dann gelten abweichend von den obigen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Documents in HENRY are made available under the Creative Commons License CC BY 4.0, if no other license is applicable. Under CC BY 4.0 commercial use and sharing, remixing, transforming, and building upon the material of the work is permitted. In some cases a different, more restrictive license may apply; if applicable the terms of the restrictive license will be binding.



# B.Sc. Arbeit: Crowdsourcing von Lufttemperaturmessdaten zur kleinräumigen Modellierung von städtischen Temperaturverteilungen





# Getting started ...

Kleinräumige Temperaturkarten wozu?

Welche Daten?

Klimamessstationen?

Crowdsourcing?

Modellierung?

Fernerkundung?



*“Crowdsourcing is traditionally defined as obtaining data or information by enlisting the services of a (potentially large) number of people. [...] this definition can now be expanded to include [...] ‘from a range of public sensors, typically connected via the Internet’ ”*

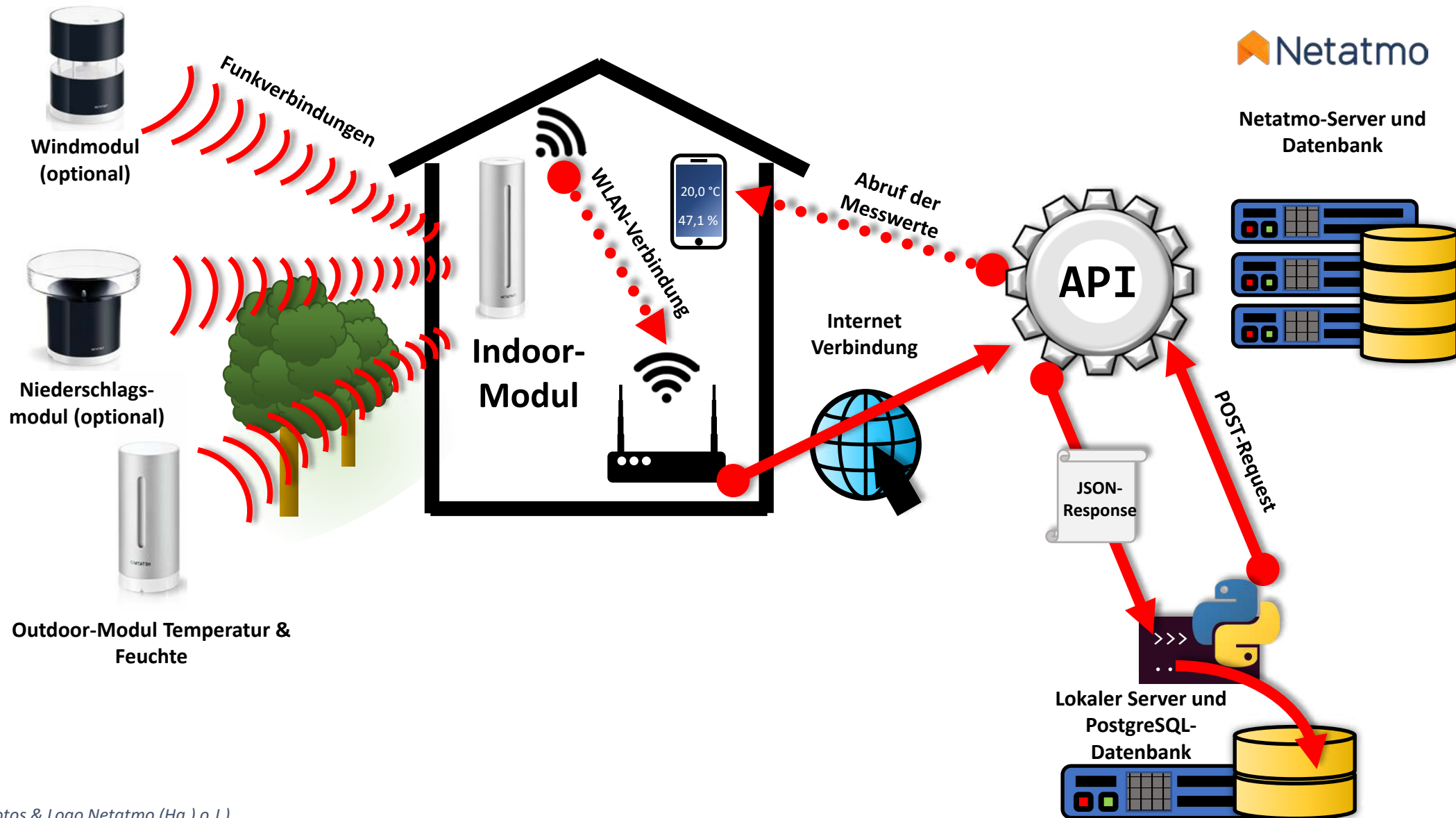
*(Muller et al. 2015)*

- CWS (Citizen Weather Station)
  - + Hohe räumliche & zeitliche Dichte
  - + In Städten vorhanden
  - Messdatenqualität *(Muller et al. 2015)*
  - Begrenzte Metadaten(-qualität) *(Bell et al. 2015)*
  - Heterogene, vom Stationstyp abhängige Fehler *(Bell et al. 2015)*
- Netatmo
  - Einheitlicher Sensortyp
  - Sensorqualität geprüft (0 °C bis 30 °C ± 0,3 K) *(Meier et al. 2017)*
  - Öffentliche *API* (Application Programming Interface) verfügbar



*(Netatmo (Hg.) o.J.)*

# Netatmo CWS-Netzwerk



# Datenbankaufbau – Sicherung der Metadatenqualität

## → Aufbau einer Datenbank

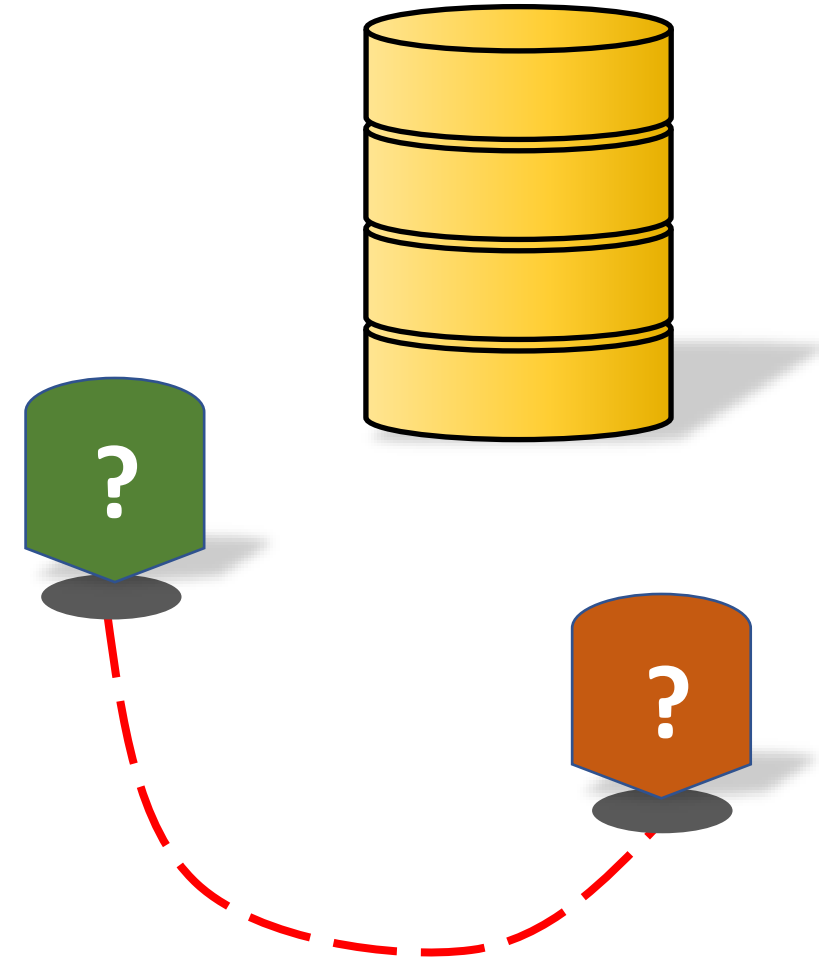
- Stationen mit Metadaten
- Schnelle Verfügbarkeit großer Mengen von Messdaten

## → Schwierigkeit: Metadatenqualität

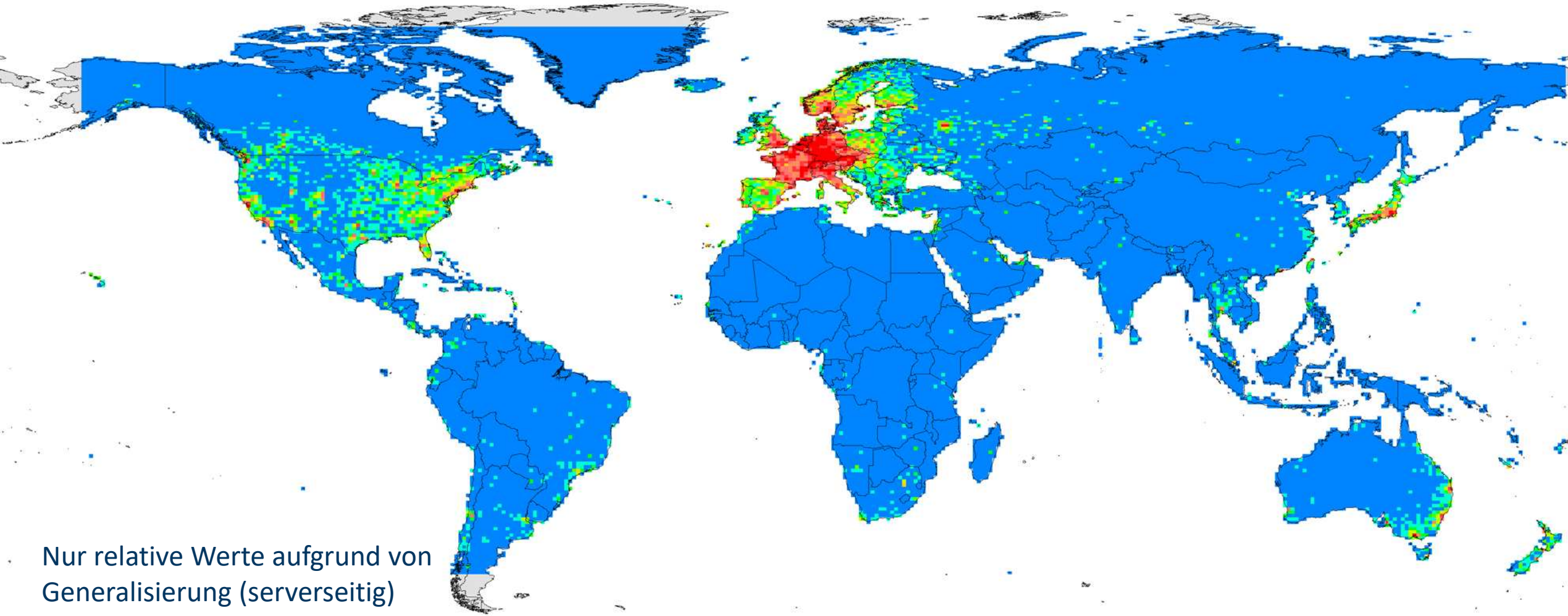
- Metadatenabruf: nur Daten für den aktuellen Zeitpunkt
- Retrospektiver Datendownload: vergangene Positionsänderungen unbekannt
  - Fehlerhafte Messdaten bzw. falsche Position für diese Daten

## → Metadaten von Stationen müssen regelmäßig geprüft werden!

- Umzug/Verkauf → Positionsänderung
- Neue Sensormodule
- Neue Stationen



# Weltweite Netatmo Stationsdichte (08/2019)

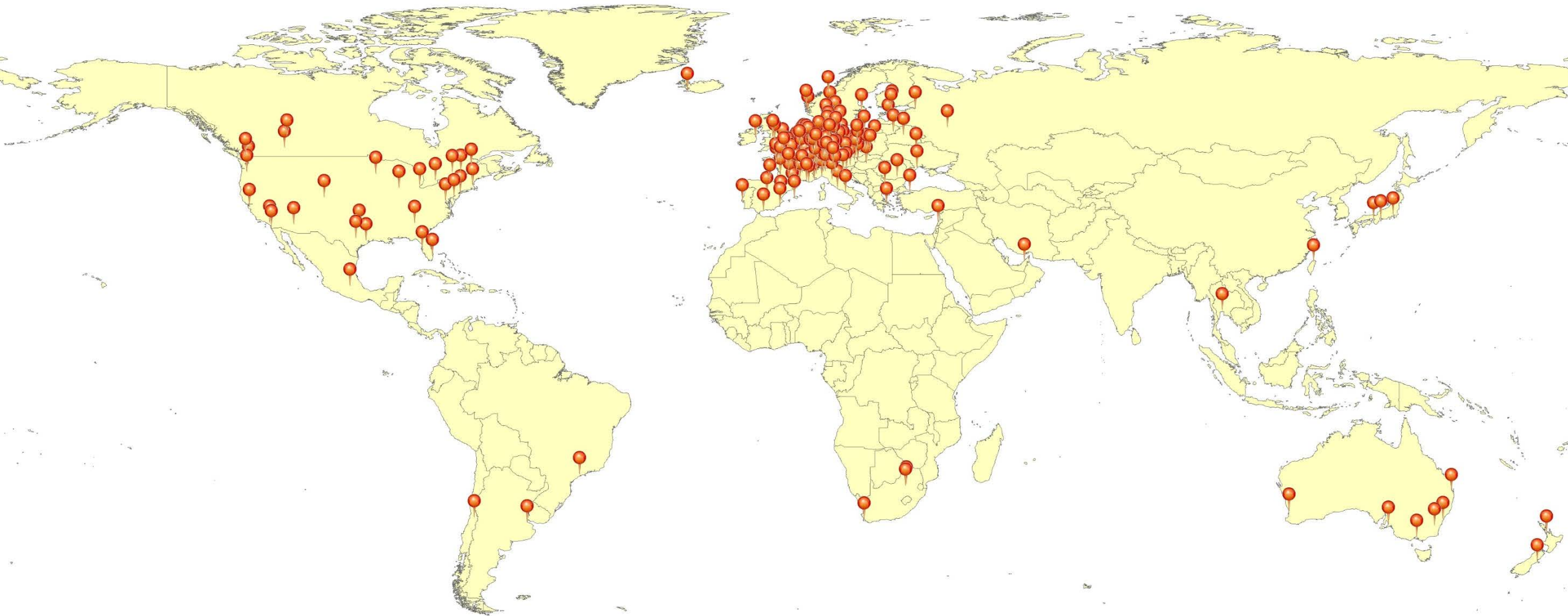


Nur relative Werte aufgrund von Generalisierung (serverseitig)





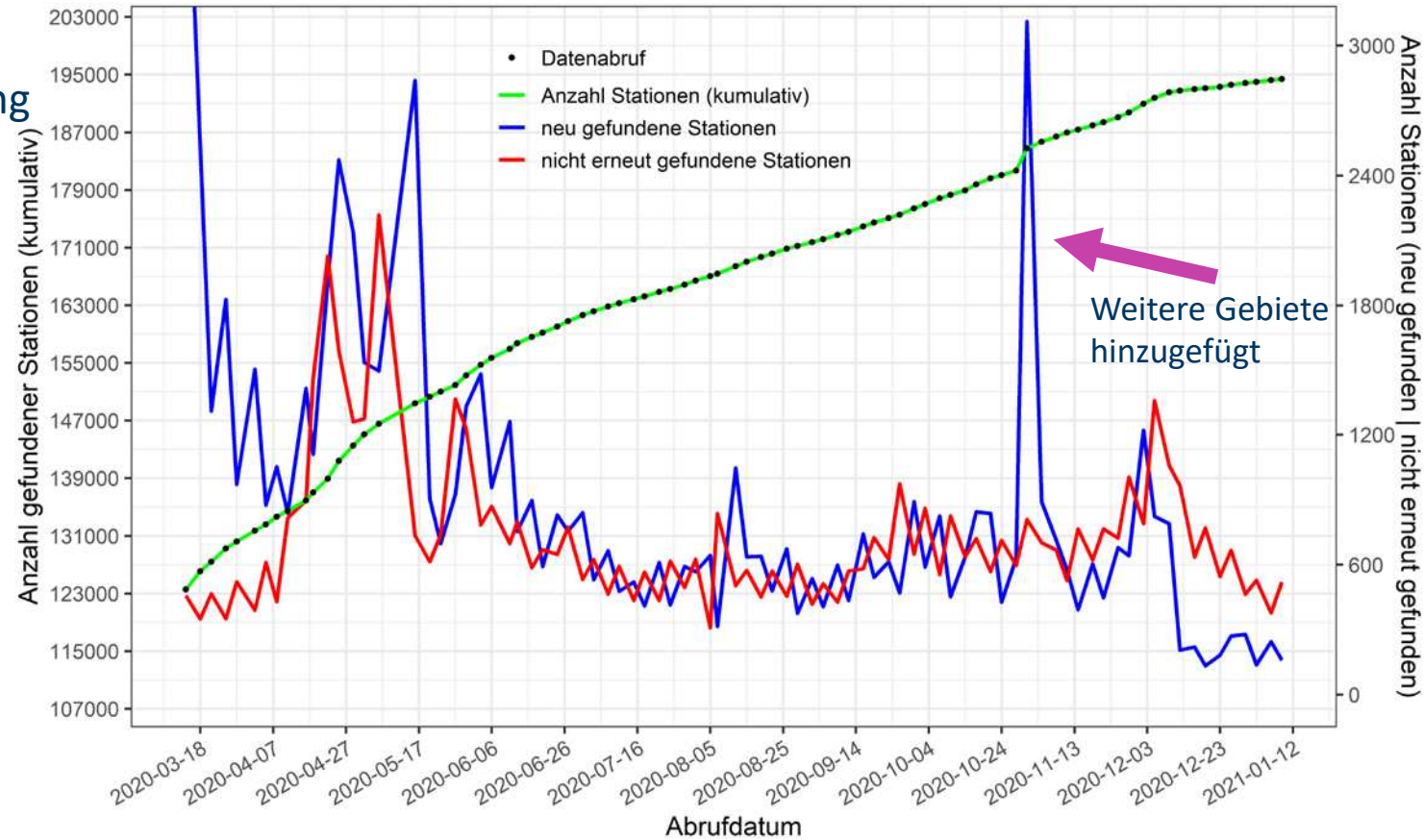
# 200 Untersuchungsgebiete





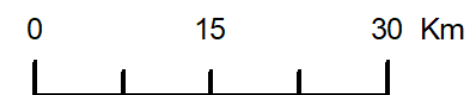
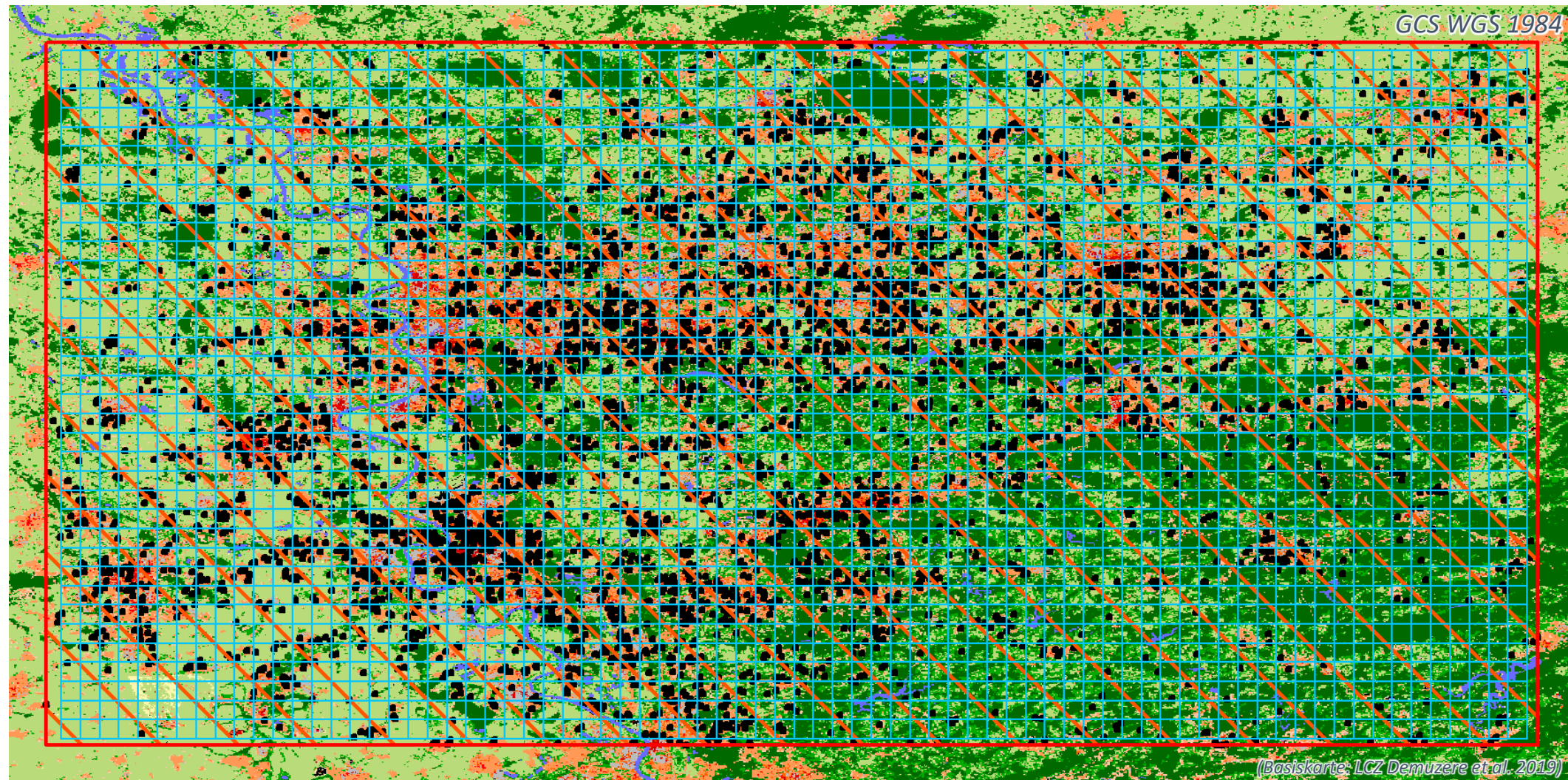
# Entwicklung seit 03/2020

- ≈ 20 000 (11 %) Stationen mit Positionsänderung
  - > 0.002 °
- ≈ 59 000 (29 %) Stationen mit Updates
  - Positionsupdate (< 0.002 °)
  - Neue Sensormodule
- ≈ 175 000 individuelle Messtationen
- Hohe Fluktuation
  - ca. 600-700 neue Stationen pro Woche
  - Aber auch ähnlich viele nicht erneut gefunden





# Anwendung: Untersuchungsgebiet Rhein-Ruhr



Abgefragte Kachel



Tatsächliches Gebiet

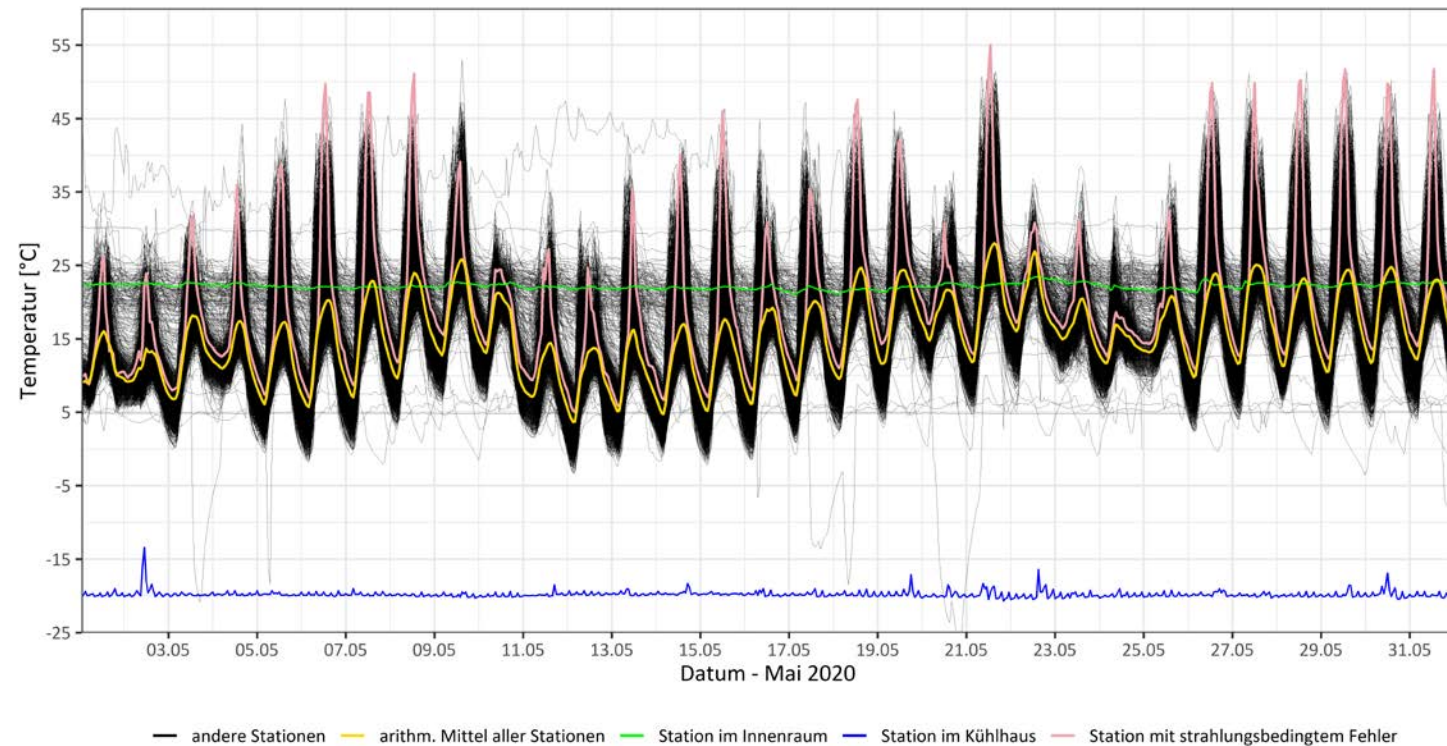


Netatmo Station



# Qualitätskontrolle – State of the Art

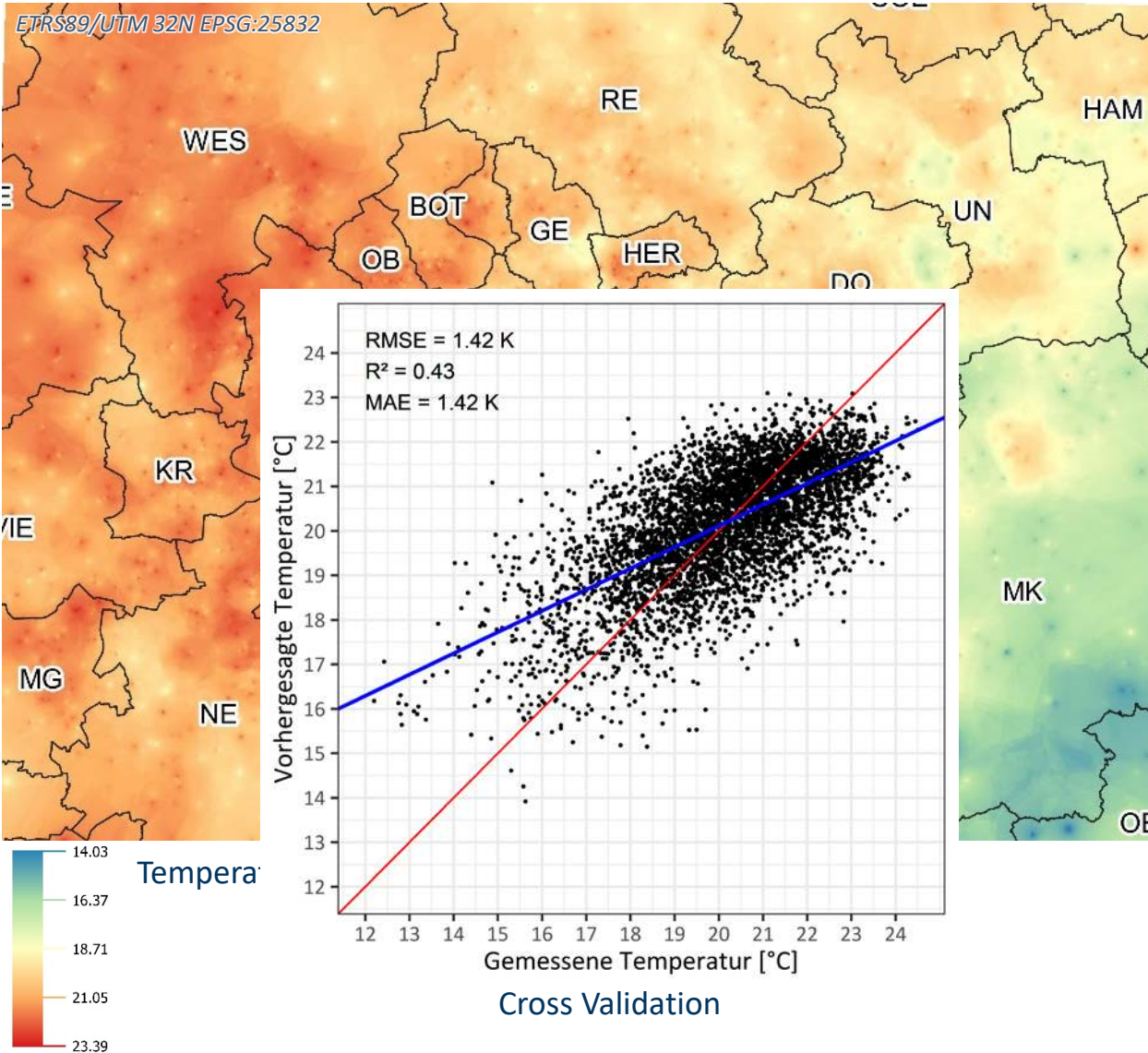
- Mehrschrittige Kontrolle für typische Fehler
  - Strahlungsbedingt
  - Ort (Innenraum/Kühlhaus)
  - Metadaten (z.B. Position per IP-Adresse)
  - Zeitliche Verfügbarkeit
- Mithilfe Daten Dritter *(Meier et al. 2017)*
  - Vergleich mit offiziellen Messdaten
  - Temperatur & Globalstrahlung
- Rein Statistisch „Wisdom of the Crowd“ *(Napoly et al. 2018)*
  - U.a. Korrelation aller Stationen mit Median
  - Frei als R-Package verfügbar *(Grassmann et al. 2018)*



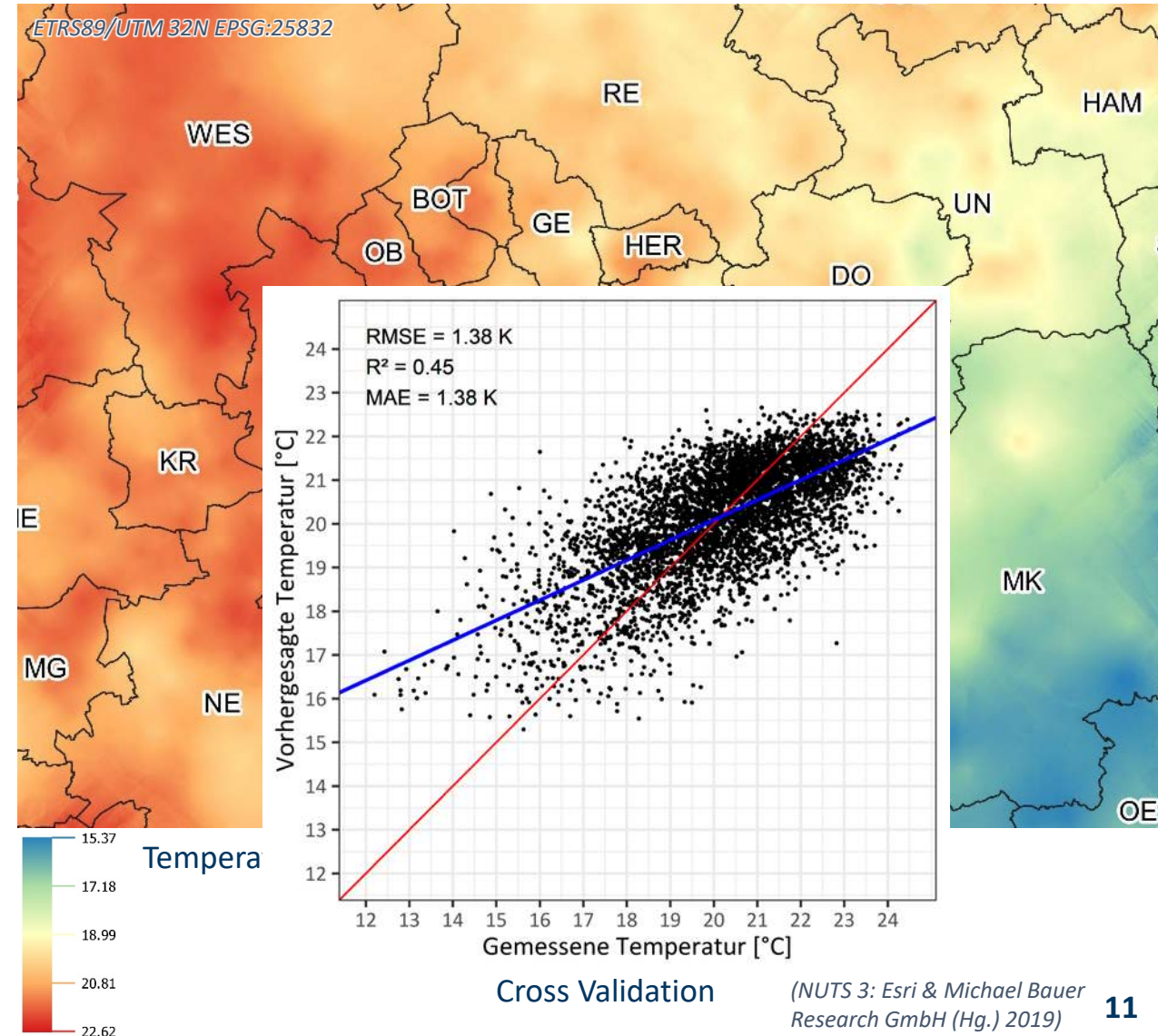
*Ungefilterte Netatmo-Daten aus stündlichen Mitteln für Mai 2020 im Bereich des Ruhrgebiets. Typische Messfehler sind farblich hervorgehoben.*

# Räumliche Interpolation: 25.06. 22:00 bis 26.06.2020 02:00 UTC

## IDW (Inverse Distance Weighting)

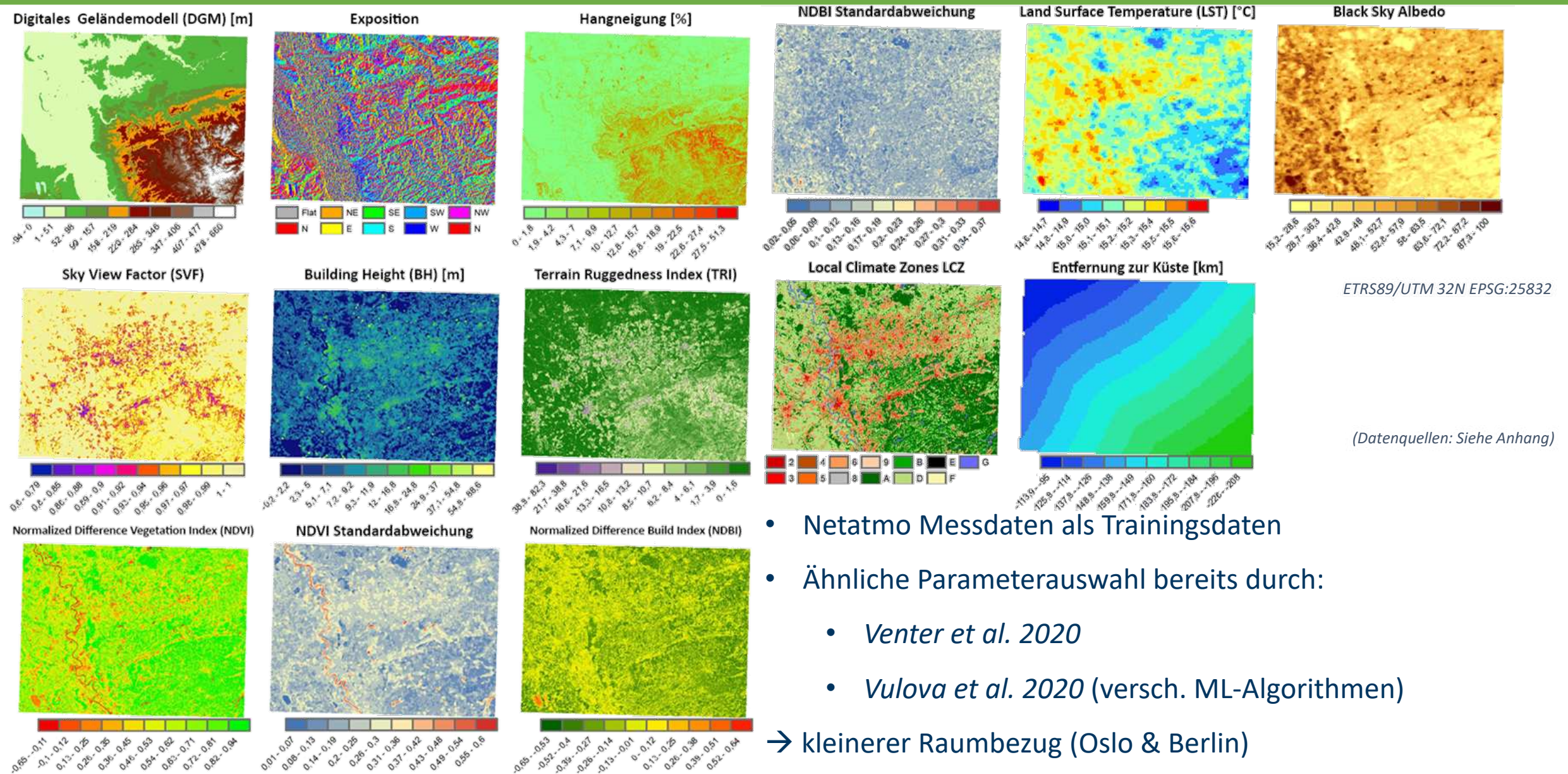


## Kriging (mit DGM (Digitales Geländemodell))





# Prädiktoren für Random Forest (RF) Modell

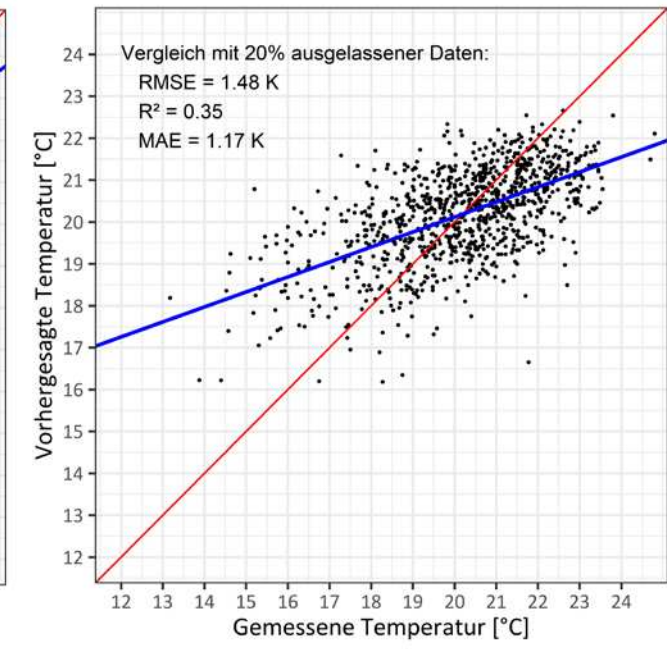
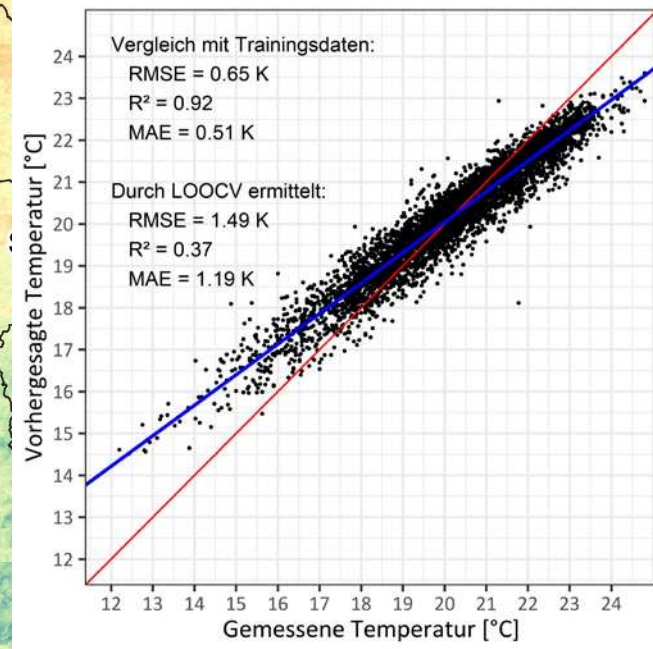
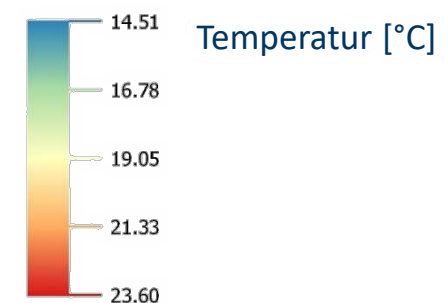
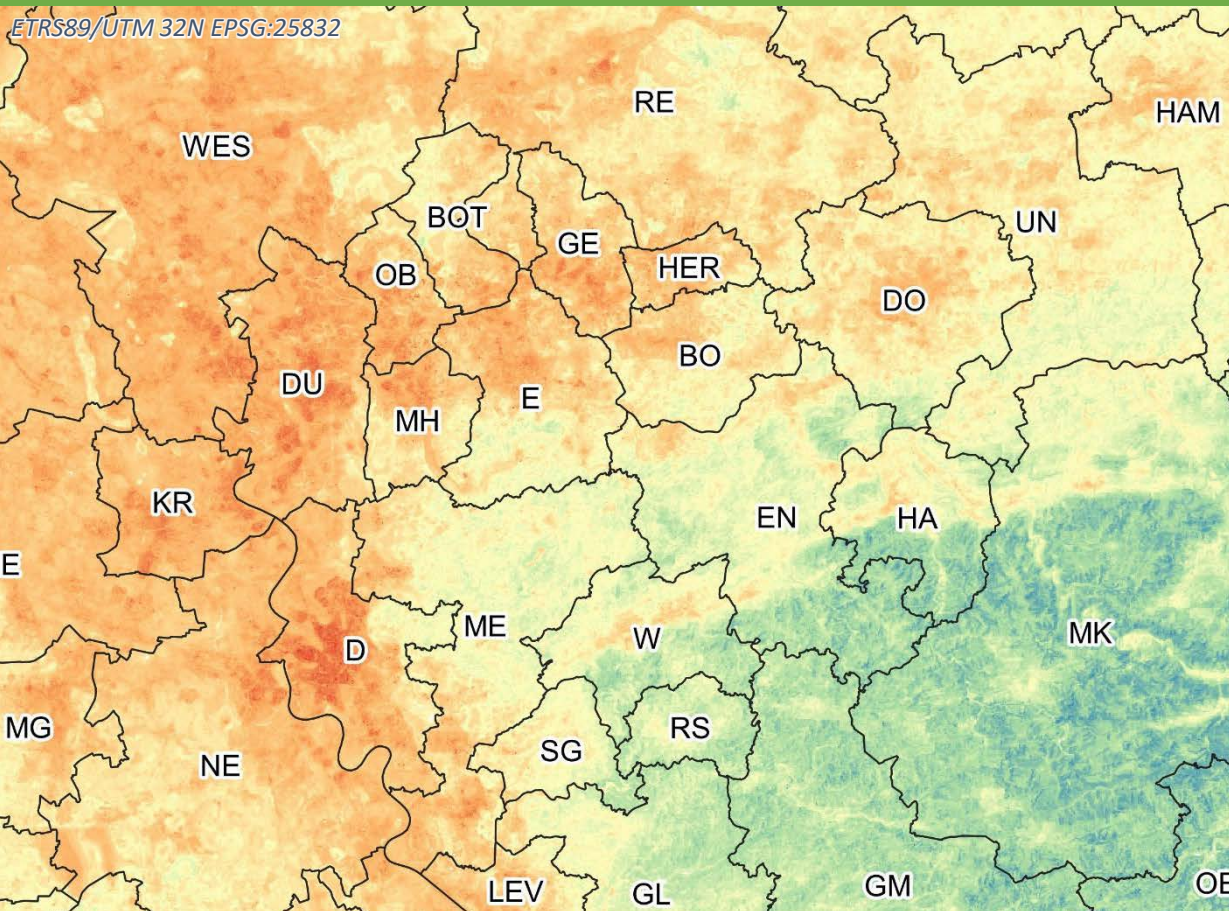


- Netatmo Messdaten als Trainingsdaten
- Ähnliche Parameterauswahl bereits durch:
  - *Venter et al. 2020*
  - *Vulova et al. 2020* (versch. ML-Algorithmen)

→ kleinerer Raumbezug (Oslo & Berlin)



# Random Forest Modellergebnis



- LOOCV (Leave One Out Cross Validation)
- 20% Bootstrap → Ähnlichkeit mit Interpolationen
- Overfitting?



# Kleinräumige Auflösung (100 m)

25.06. 22:00 bis 26.06.2020 02:00 UTC

Kleingartenanlage  
Emschertal

Gelsenkirchen  
Innenstadt

Halde und Brachfläche  
ehem. Zeche Alma

Temperatur [°C]



- Crowdsourcing-Daten bieten Grundlage für flächenhafte Temperaturkarten
- Datenbank: Systematische, kontinuierliche Datensammlung mit kontrollierten Metadaten
- Flächenhafte Modellierung
  - Interpolation vs. RF-Modell
- **Ausblick:**
  - Mehr Trainingsdaten Ergänzung durch neu entstehende Stadtklimanetzwerke
  - Anwendung in der Stadt- und Regionalplanung (Klimaschutz- und Klimaanpassungsmaßnahmen)



**Vielen Dank!**

**Bell, S.; Cornford, D.; Bastin, L. (2015):** How good are citizen weather stations? Addressing a biased opinion. In: Weather 70 (3): 75–84. DOI:10.1002/wea.2316.

**Bezirksregierung Köln (Hg.) (2020a):** Digitale Orthophotos (DOP). [https://www.bezreg-koeln.nrw.de/brk\\_internet/geobasis/webdienste/geodatendienste/](https://www.bezreg-koeln.nrw.de/brk_internet/geobasis/webdienste/geodatendienste/) [07.07.2020].

**Bezirksregierung Köln (Hg.) (2020b):** 3D-Gebäudemodelle. LoD1. [https://www.bezregkoeln.nrw.de/brk\\_internet/geobasis/3d\\_gebaeudemodelle/index.html](https://www.bezregkoeln.nrw.de/brk_internet/geobasis/3d_gebaeudemodelle/index.html) [07.07.2020].

**Bezirksregierung Köln (Hg.) (2020c):** Digitales Geländemodell. [https://www.bezregkoeln.nrw.de/brk\\_internet/geobasis/hoehenmodelle/gelaendemodell/index.html](https://www.bezregkoeln.nrw.de/brk_internet/geobasis/hoehenmodelle/gelaendemodell/index.html) [07.07.2020].

**Bezirksregierung Köln (Hg.) (2020d):** Hausumringe. [https://www.bezregkoeln.nrw.de/brk\\_internet/geobasis/liegenschaftskataster/hausumringe/index.html](https://www.bezregkoeln.nrw.de/brk_internet/geobasis/liegenschaftskataster/hausumringe/index.html) [07.07.2020].

**Demuzere, M.; Bechtel, B.; Middel, A.; Mills, G. (2019):** Mapping Europe into local climate zones. In: PloS one 14 (4): e0214474. DOI:10.1371/journal.pone.0214474.

**Florczyk, A.; Corbane, C.; Schiavina, M.; Pesaresi, M.; Maffenini, L.; Melchiorri, M.; Politis, P.; Sabo, F.; Freire, S.; Ehrlich, D.; Kemper, T.; Tommasi, P.; Airaghi, D.; Zanchetta, L. (2019a):** GHS-UCDB R2019A - GHS Urban Centre Database 2015, multitemporal and multidimensional attributes. <http://data.europa.eu/89h/53473144-b88c-44bc-b4a3-4583ed1f547e> [15.06.2020].

**Grassmann, T.; Napoly, A.; Meier, F.; Fenner, D. (2018):** Quality control for crowdsourced data from CWS. <https://depositonce.tuberlin.de//handle/11303/7520.3>. DOI:10.14279/DEPOSITONCE-6740.3 [20.06.2020].



# Literaturverzeichnis II



**Lindberg, F.; Grimmond, C.S.B.; Gabey, A.; Huang, B.; Kent, C.; Sun, T.; Theeuwes, N.; Järvi, L.; Ward, H.; Capel-Timms, I.; Chang, Y.; Jonsson, P.; Krave, N.; Liu, D.; Meyer, D.; Olofson, K.; Tan, J.; Wästberg, D.; Xue, L.; Zhang, Z. (2018):** Urban Multi-scale Environmental Predictor (UMEP): An integrated tool for city-based climate services. In: Environmental Modelling & Software 99: 70–87. DOI:10.1016/j.envsoft.2017.09.020.

**Meier, F.; Fenner, D.; Grassmann, T.; Otto, M.; Scherer, D. (2017):** Crowdsourcing air temperature from citizen weather stations for urban climate research. In: Urban Climate 19: 170–191. DOI:10.1016/j.uclim.2017.01.006.

**Muller, C.; Chapman, L.; Johnston, S.; Kidd, C.; Illingworth, S.; Foody, G.; Overeem, A.; Leigh, R. (2015):** Crowdsourcing for climate and atmospheric sciences: current status and future potential. In: International Journal of Climatology 35 (11): 3185–3203. DOI:10.1002/joc.4210.

**Napoly, A.; Grassmann, T.; Meier, F.; Fenner, D. (2018):** Development and Application of a Statistically-Based Quality Control for Crowdsourced Air Temperature Data. In: Frontiers in Earth Science 6. DOI:10.3389/feart.2018.00118.

**National Geographic Society (Hg.) (2013):** World Countries (Generalized).  
<https://www.arcgis.com/home/item.html?id=2b93b06dc0dc4e809d3c8db5cb96ba69> [15.06.2020].

**Netatmo (Hg.) (o.J.):** All products. <https://shop.netatmo.com/> [04.06.2020].

**NASA Ocean Biology Processing Group (OBPG) (Hg.) (2009):** Distance to the Nearest Coast.  
[https://oceancolor.gsfc.nasa.gov/docs/distfromcoast/GMT\\_intermediate\\_coast\\_distance\\_01d.zip](https://oceancolor.gsfc.nasa.gov/docs/distfromcoast/GMT_intermediate_coast_distance_01d.zip) [22.07.2020].

# Literaturverzeichnis III



**Riley, S.; Gloria, S. de; Elliot, R. (1999):** A Terrain Ruggedness that Quantifies Topographic Heterogeneity. In: Intermountain Journal of Sciences 5 (1-4): 23–27.

[https://www.researchgate.net/publication/259011943\\_A\\_Terrain\\_Ruggedness\\_Index\\_that\\_Quantifies\\_Topographic\\_Heterogeneity](https://www.researchgate.net/publication/259011943_A_Terrain_Ruggedness_Index_that_Quantifies_Topographic_Heterogeneity).

**Schaaf, C.; Wang, Z. (2015):** MCD43A3 MODIS/Terra+Aqua BRDF/Albedo Daily L3 Global - 500m V006. DOI:10.5067/MODIS/MCD43A3.006.

**Venter, Z.; Brousse, O.; Esau, I.; Meier, F. (2020):** Hyperlocal mapping of urban air temperature using remote sensing and crowdsourced weather data. In: Remote Sensing of Environment 242 (111791): 1–14. DOI:10.1016/j.rse.2020.111791.

**Vulova, S.; Meier, F.; Fenner, D.; Nouri, H.; Kleinschmit, B. (2020):** Summer Nights in Berlin, Germany: Modeling Air Temperature Spatially With Remote Sensing, Crowdsourcing Weather Data, and Machine Learning. In: IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing PP(99):1-1. DOI: 10.1109/JSTARS.2020.3019696

**Wan, Z.; Hook, S. (2015):** MOD11A1 MODIS/Terra Land Surface Temperature/Emissivity Daily L3 Global 1km SIN Grid V006. DOI:10.5067/MODIS/MOD11A1.006.



# Anhang I - Datenquellen



DATENSATZ	DATENQUELLE/-BERECHNUNG	AUFLÖSUNG
DGM	(Bezirksregierung Köln (Hg.) 2020c)	1 m, resampling auf 100 m
Exposition	Berechnung aus DGM	100 m
Hangneigung	Berechnung aus DGM	100 m
SVF	Berechnung aus DGM, LoD1 3D-Gebäudemodell (Bezirksregierung Köln (Hg.) 2020b) und Hausumringen (Bezirksregierung Köln (Hg.) 2020d) mithilfe von UMEP (Lindberg et al. 2018).	10 m, resampling auf 100 m
BH	Berechnung aus DGM, LoD1 und Hausumringen	10 m, 300 m Focal Buffer und resampling auf 100 m
TRI	Berechnung aus DOM (DGM, LoD1, Hausumringen) nach (Riley et al. 1999)	10 m, resampling auf 100 m
NDVI	Sentinel 2 (MSI Level-2A)	100 m
NDVI dev	Sentinel 2 (MSI Level-2A)	100 m, 500 m Focal Buffer
NDBI	Sentinel 2 (MSI Level-2A)	100 m
NDBI dev	Sentinel 2 (MSI Level-2A)	100 m, 500 m Focal Buffer
LST	MODIS Terra (MOD11A1) (Wan & Hook 2015)	1000 m
Black Sky Albedo	MODIS (MCD43A3) (Schaaf & Wang 2015)	500 m
LCZ	(Demuzere et al. 2019)	100 m
Entfernung zu Küste	(OBPG (Hg.) 2009)	1000 m