

# Daniel Boateng (Dr. rer. nat.)

CLIMATE RISK EXPERT · CLIMATE DATA SCIENTIST|ENGINEER · NATCAT MODELLING ANALYST · STATISTICAL MODELLING

Rose-Senger-Strasse 6, Hannover, 30655, Germany

☎ (+49) 178-301-2296 | ✉ dannboateng@gmail.com | 🏠 dan-boat.github.io/ | 📱 Dan-Boat | 🌐 daniel-boateng-3892311b4 | 📧 Daniel Boateng

## Summary

Avid problem solver, climate data scientist|engineer with **5+ years** of experience and a certified Google Advanced Data Analytics Professional. Expert in **Machine Learning, Spatial Data Engineering, Climate Modelling, Time Series Analysis, Climate Risk Modelling, Parametric and Index-based Reinsurance solutions** and passionate in leveraging cutting-edge technologies in Cloud environment (AWS & GCP) to solve business challenges. Proven track record of applying advanced tech stacks-including **databases (SQL, NoSQL, MangoDB), programming (Python, R & Fortran), cloud engineering, data visualization, advanced statistics, and high performance computing**—to solve complex industry and scientific challenges.

## Skills

**Languages (Programming):** Python (Keras/TensorFlow/PyTorch, Sklearn, Numpy, Pandas, Statsmodel, Scipy, Xarray, Seaborn/Mathplotlib, Flask/FlaskAPI, Cartopy, Plotly, etc), C, Fortran, JavaScript, R, MATLAB

**Database:** SQL, PostSQL, NoSQL, MangoDB, SparkSQL, BigQuery, AWS S3

**Tools:** Git, Terraform, PySpark, Docker, Airflow, Kafka, Tableau, Mage, GitHub Actions, Datalog, Kibana, etc

**Cloud:** AWS, GCP, Oracle, Azure

**Frameworks and Methods:** Machine learning (Deep Learning (ConvLSTM, Computer Vision (NLP)), Classification (SVG, Random Forest, Logistic, Naive Bayes), Regression (LassoCV, MLP, XGBoost, Stacking, AdaBoost), Data transformation (EOF, PCA), Data Wrangling and Visualization, Statistical Models (ARIMA, Causal Inference, A/B testing, Distribution fitting, Stochastic simulation)

**NatCat Models:** CLIMADA, Verisk, KatRisk, Touchstone Re

**Languages:** English (professional | C1), German (conversational | B1)

## Work Experience

### Hannover Re

*Hannover, Germany*

CLIMATE RISK DATA SCIENTIST

*Feb. 2025 - present*

- Performs natural catastrophe modeling for main and side perils (Tropical Cyclone, Earthquakes, Flood, Winter Storms, Hail, etc) to price and validate reinsurance programs (proportional and non-proportional treaties of up to \$20 million capacity)
- Performs spatial data analytics for pricing of parametric and index-based insurance solutions for many lines of business (e.g., agriculture, energy, climate finance, business interruption).
- Developing spatial data infrastructure on AWS to automate data modeling pipeline for climate related reinsurance products
- Probabilistic adjustment of stochastic event sets from NatCat models using historical losses and exposure information gathered in-house
- Database management of stochastic events catalogs and various global climate and weather datasets from public APIs (e.g., ECMWF and NOAA)

### University of Tübingen

*Tübingen, Germany*

RESEARCH DATA SCIENTIST

*April. 2021 - present*

- Performed numerical modeling with complex climate models on HPCs to generate petabytes of climate data using tools like cdo, ECHAM5-wiso, MPI-ESM, NCL, Bash scripting, etc.
- Developed pyESD, a Python package for ML-based downscaling of climate model output to high resolution using sklearn, TensorFlow, xarray, pandas, PyPi, SciPy, etc.
- Created pyClimat, a Python package for automating spatial big data analysis pipelines with dask-array, xarray, cartopy, sklearn, numpy, etc.
- Designed spatial data analysis pipelines used by over 500 scientists, ensuring reproducibility of research results.
- Post-processed climate datasets: CMIP6, PMIP4, DWD weather stations data, ERA5, CRU.
- Developed a full-stack application for flight temperature calibration on Raspberry Pi, deployed on drones for temperature sensors.

### Equitech Futures, Chicago, USA

*Remote*

DATA SCIENTIST AND ENGINEER

*July. 2023 - March. 2024*

- Contributed to a platform estimating atmospheric CO<sub>2</sub> emissions for countries and regions using NASA satellite XCO<sub>2</sub> measurements.
- Processed extensive spatial satellite datasets of global carbon emissions since 1979.
- Managed Google Cloud Platform and data engineering using Terraform, BigQuery, pySpark, MongoDB, Cloud Run, and Docker.
- Engineered platform using matplotlib, pandas, xarray, dask, plotly, Cloud Run, Streamlit, and more.

# Education

---

## University of Tübingen

*Tübingen, Germany*

DOCTOR OF SCIENCE IN COMPUTATIONAL GEOSCIENCES (CLIMATE CHANGE MODELLING, CLIMATE DYNAMICS AND MACHINE LEARNING)

*April. 2021 - May. 2024*

- Spatial Data Analytics (Big Data), Software Development, (Paleo)Climate Modelling, Statistical Modelling, Climate Change Projection

## University of Tübingen

*Tübingen, Germany*

M.SC. IN APPLIED AND ENVIRONMENTAL GEOSCIENCE (FIRST CLASS)

*Oct. 2018 - Feb. 2021*

- Hydroclimatology, Environmental Modelling, Atmospheric Physics, Advanced Geophysics, Machine Learning

## Kwame Nkrumah University of Science and Technology

*Kumasi, Ghana*

B.SC. IN GEOLOGICAL ENGINEERING (FIRST CLASS)

*Oct. 2013 - June 2017*

- Hydrogeology, Geophysics, Geotechnical Engineering, Rock Mechanics, Computer Programming

# Projects

---

## Machine learning-based framework for downscaling climate change information

*[Github](#)*

PyESD | [DOCUMENTATION](#)

- An open-source framework for statistical downscaling of climate change information (e.g., precipitation, temperature, wind speed) using re-analysis products like ERA5 as predictors. The package encompasses all downscaling cycles, including data preprocessing, predictor selection, construction (e.g., using transformers like PCA, EOF), model selection, training, validation, evaluation, and future prediction.

## Python Package for Climate Data Analysis and Visualization

*[Github](#)*

PyCLIMAT | [CODE](#)

- An open-source framework for climate data analysis and visualization. Statistical analyses include correlation, A/B testing, causality testing, Principal Component Analysis, Empirical Orthogonal Analysis, and Z-score. Spatial data visualization features include Cartopy, time series animation, and interactive maps.

## Collection of Machine Learning Projects from Beginner to Advanced Level

*[Github](#)*

DATA SCIENCE PROJECTS | [CODE](#)

- Absenteeism prediction, RNN (LSTM) for Google stock price prediction, AudioBooks target ads prediction, etc

# Certifications

---

2024 **SQL for Data Science**, UC Davis (Coursera)

*[Certificate](#)*

2024 **Certificate of Excellence in Computational Tools for Climate Science**, Climatedata Academy

*[Certificate](#)*

2024 **Python for Everyone Specialisation (5 courses)**, University of Michigan (Coursera)

*[Certificate](#)*

2024 **Google Advanced Data Analytics Professional Certificate (7 courses)**, Google (Coursera)

*[Certificate](#)*

2023 **Machine Learning in Weather & Climate**, ECMWF

*[Certificate](#)*

2023 **Full Stack Software Engineering**, ALX Africa

*[Certificate](#)*

2023 **Complete Data Science Bootcamp**, Udemy

*[Certificate](#)*

2021 **From Machine Learning to Deep Learning: a concise introduction**, High-Performance Computing Center Stuttgart (HLRS)

*[Certificate](#)*

2021 **The Theory of Deep Learning-Deep Neural Networks**, Udemy

*[Certificate](#)*

# Publications

---

|      |                                                                                                                                                                                                                                                |                       |
|------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------|
| 2025 | <b>Boateng, D. et al.</b> , Refining paleoelevation estimates of the European Alps by simulating Middle Miocene climate and $\delta^{18}\text{O}$ responses to diachronous surface uplift scenarios. <i>Global and Planetary Change</i> , 2025 | <a href="#">Paper</a> |
| 2024 | <b>Boateng, D. et al.</b> , West African Monsoon Dynamics and Its Control on the Stable Oxygen Isotopic Composition of Precipitation in the Late Cenozoic. <i>JGR Atmospheres</i> , 2024                                                       | <a href="#">Paper</a> |
| 2023 | <b>Boateng, D. and Mutz, S. G.</b> , An open-source Python framework for empirical-statistical downscaling of climate information, <i>Geoscientific Model Development (GMD)</i> , 2023                                                         | <a href="#">Paper</a> |
| 2023 | <b>Boateng, D. et al.</b> , The effects of diachronous surface uplift of the European Alps on regional climate and the oxygen isotopic composition of precipitation, <i>Earth System Dynamics (ESD)</i> , 2023                                 | <a href="#">Paper</a> |
| 2022 | <b>Botsyun, S., et al (incl Boateng, D.)</b> , Middle Miocene climate and stable oxygen isotopes in Europe based on numerical modeling, <i>Paleoceanography and Paleoclimatology (JGR)</i> 2022                                                | <a href="#">Paper</a> |