

## HighResLearn

High-resolution machine learning for the climate community in Austria

<b>Programm / Ausschreibung</b>	Austrian Climate Research Programme Ausschreibung 2022/01	<b>Status</b>	laufend
<b>Projektstart</b>	01.06.2024	<b>Projektende</b>	31.05.2027
<b>Zeitraum</b>	2024 - 2027	<b>Projektlaufzeit</b>	36 Monate
<b>Keywords</b>	Climate; machine learning; international collaboration		

### Projektbeschreibung

Hochaufgelöstes maschinelles Lernen für die Klimaforschung in Österreich - HighResLearn

Klimamodelle sind unser wichtigstes Werkzeug um das Klimasystem der Erde zu verstehen, zukünftige Änderungen zu simulieren und Entscheidungsträger\*innen die bestmögliche Information zur Verfügung zu stellen. Das gilt sowohl für großskalige Änderungen wie den globalen Temperaturanstieg, als auch für lokale Ereignisse wie Extremniederschläge in den Alpen. Für Ersteren werden typischerweise globale Klimamodelle mit einer eher groben horizontalen Auflösung von etwa 100km verwendet; für Letztere greift die Forschung oft auf hochaufgelöste, aber regional begrenzte Modelle zurück, die die Kilometer-Skala erreichen können.

Um Entscheidungsträger\*innen mit verlässlichen Informationen aus Klimamodellen versorgen zu können, benötigt es eine gründliche Qualitätssicherung. Dazu müssen Ergebnisse von Modellen kritisch gegeneinander und gegen andere Datensätze, zum Beispiel aus Beobachtungen, evaluiert werden. Für solche Evaluierungen werden die Daten typischerweise über mindestens 10 Jahre gemittelt, um strukturelle Probleme in der Simulation des Klimas von interner Variabilität, die durch das chaotische Verhalten des Klimasystems entsteht, zu separieren. Allerdings sind oft gerade die neuesten, am höchsten aufgelösten Klimamodelle (noch) nicht in der Lage, Daten für 10 Jahre oder mehr zu produzieren, da der Bedarf an Rechenleistung enorm ist, was die Qualitätskontrolle vor große Herausforderungen stellt. Das betrifft insbesondere die Entwicklung neuer Modelle auf der Kilometer-Skala, bei der typischerweise viele Modellversionen basierend auf wenigen Jahren evaluiert werden müssen. Beispiele dafür sind die geplanten neuen österreichischen Klimaszenarien (ÖKS26) und internationale Projekte wie NextGEMS, das globale Modelle auf der km-Skala entwickelt und damit an der Implementierung eines digitalen Zwillinges der Erde arbeitet.

Das HighResLearn Projekt hat zum Ziel neue Evaluierungsmethoden zu entwickeln, die speziell auf hochaufgelöste Klimadaten auf kurzen Zeitskalen abzielen, um die Etablierung von neuen regionalen Klimamodellen für Österreich zu unterstützen und zur internationalen Forschungstätigkeit im Feld der Modellevaluierung beizutragen. Das Projekt nutzt dafür Fortschritte in statistischem und maschinellem Lernen (ML) und wendet diese in der Klimaforschung an. Insbesondere zielt HighResLearn darauf ab, Modellfehler in der Simulation des Klimas im Alpenraum und in Österreich direkt aus jährlichen, monatlichen oder sogar täglichen hochaufgelösten Daten abzuschätzen. Dazu planen wir, ML-Methoden zu verwenden, die

wir zuerst auf Modelldaten trainieren, die für längere Zeiträume verfügbar sind und die wir dann für neue, im Training nicht verwendete Modelle generalisieren.

Dabei werden wir uns auf Daten mit Auflösungen von 5km oder feiner fokussieren, wie zum Beispiel von der letzten Generation der österreichischen Klimaszenarien (ÖKS15), von einer Pilotstudie des koordinieren regionalen Klima-Downscaling Experiments (CORDEX-FPS) und von NextGEMS. Mit der Verwendung dieser diversen Datensätze wird HighResLearn auch dazu beitragen, die Forschungsfelder von regionaler und globaler Klimamodellierung enger zu verknüpfen und das Forschungsnetzwerk in Österreich zu verstärken.

Ein weiteres Kernziel von HighResLearn ist es, Forschende in Österreich und international in die Lage zu versetzen, die globalen km-skalierten Modellergebnisse von NextGEMS einfach und effizient zu nutzen und in ihre Analysen zu integrieren. Dazu werden wir sogenannte Datenrezepte entwickeln, die Nutzer\*innen anleiten, wie sie auf die Daten zugreifen können und diese in ein Format äquivalent zu den ÖKS Daten prozessieren können. Die Rezepte werden sowohl Quellcode und Beispiele von prozessierten Daten wie auch eine ausführliche Dokumentation beinhalten. Sie werden außerdem, wie alle Ergebnisse von HighResLearn, ohne Copyright frei verfügbar und leicht an individuelle Anforderungen von Nutzer\*innen anpassbar sein.

In einem iterativen Prozess mit den Nutzer\*innen werden wir sicherstellen, dass unsere ML-Methoden und Datenrezepte ausreichend dokumentiert sind und alle relevanten Informationen beinhalten. Dazu werden wir mit Partnern wie dem Climate Change Center Austria (CCCA), der österreichischen Initiative Klimadashboard und möglicherweise mit dem use.AT-Projekt zusammenarbeiten, das ebenfalls bei dieser ACRP-Ausschreibung eingereicht wurde.

Wir erwarten, dass HighResLearn zu mehreren internationalen, peer-reviewed Publikationen über die Auswertung von hochaufgelösten Klimadaten mit Hilfe von ML führt und darüber hinaus zu mehreren Daten- und Quellcode-Produkten zur Verwendung in der Forschung an regionaler Klimamodellierung sowie zu einem gesteigerten öffentlichen Bewusstsein über die Auswirkungen des Klimawandels in Österreich durch unsere Aktivitäten im Bereich der Öffentlichkeitsarbeit.

## **Abstract**

Climate models are primary tools to understand the climate system, simulate future changes, and provide information to policy makers and the public. This holds for large-scale changes such as the rise in global mean temperatures as well as for local impacts such as extreme precipitation in the Alps. The former are typically assessed from simulations in global, coarse-resolution climate models with resolutions of about 100km, while for the latter often regional, high-resolution models on a km-scale are used.

To ensure data quality and provide the best possible information to decision makers and other users, it is imperative to evaluate model output against observations and other reference datasets. Such evaluations are typically performed on time-scales of 10 years or longer to isolate structural model biases in the representations of climate from internal variability that emerges from random fluctuations of daily weather on regional scales. However, high-resolution models in most cases do not provide multiple decades of output due to computational constraints. This limit in model output is a particular challenge for high-resolution model development efforts such as the next generation of the Austrian climate scenarios (ÖKS26) or the EU H2020 NextGEMS project, for which many versions of models need to be evaluated based on just a few years of simulations.

The proposed HighResLearn project aims at providing novel ways to evaluate high-resolution climate model output on short time scales, aiding regional climate modelling efforts in Austria as well as advancing ongoing international research.

HighResLearn will draw on advances in statistical and machine learning (ML) to separate climate information from noise. Specifically, our goal is to estimate climatological model biases and climate change signals for the greater Alpine domain

and Austria directly from high frequency data on annual, seasonal, monthly, or even daily time-scales using ML classifiers that we train on models available for longer periods.

We will, in particular, draw on data with resolutions of 5km and finer from the last generation ÖKS15, the coordinated regional climate downscaling experiment flagship pilot study (CORDEX-FPS), and global km-scale simulations provided by NextGEMS. This will also open ways to connect the global and regional climate modelling communities.

Another key goal of HighResLearn is to enable the regional modelling community in Austria and internationally to easily and efficiently use global km-scale model output from NextGEMS. To this end, we will develop and provide data recipes that showcase how data can be accessed and processed and that can be easily adapted to specific user needs. As all results of HighResLearn, the data recipes will be made freely available to the community and will combine download and preprocessing scripts, example data, and detailed documentation. This will allow regional climate modellers to integrate global high-resolution models into their workflows using a format similar to ÖKS15.

An iterative process with users will make sure that our ML classifiers and recipes are well-documented and provide useful and actionable information. For this we will collaborate with national partners like the Climate Change Centre Austria (CCCA), the Austrian initiative Klimadashboard, and, potentially, the use.AT project that has also been submitted to this ACRP call. We expect HighResLearn to result in several peer-reviewed publications on the evaluation of high-resolution climate information using ML, several data and code products for use by the climate modelling community, and increased public awareness on the impacts of climate change in Austria through our outreach activities.

### **Projektkoordinator**

- Universität Wien

### **Projektpartner**

- Universität Innsbruck
- GeoSphere Austria - Bundesanstalt für Geologie, Geophysik, Klimatologie und Meteorologie