Numerical Modelling Plan

Version 1 of the TEAMx Numerical Modelling Plan, dated **8 June 2023**. This document was presented and discussed at the 3rd TEAMx Workshop in Zürich on 15–16 June 2023, and is now undergoing review.

Document structure based on *The Legrand Orange Book*
LaTeX template downloaded from `http://www.LaTeXTemplates.com`

Image credits

Copyright

June 2023

Preface

TEAMx is an international research programme that aims at improving the understanding of processes in the atmosphere over mountains at multiple scales and at advancing the representation of these processes in numerical models for weather and climate prediction. Its acronyms stands for *Multi-scale transport and exchange processes in the atmosphere over mountains – Programme and experiment*.

TEAMx is a bottom-up initiative promoted by a number of universities, research institutions and national weather services, internationally integrated through a Memorandum of Understanding between interested parties. TEAMx is carried out by means of coordinated research projects initiated at the institutional but also multi-national level, and it is supported by a Programme Coordination Office at the Department of Atmospheric and Cryospheric Sciences of the University of Innsbruck, Austria. The present document describes all collaborative weather and climate modelling activities related to TEAMx.

Lead authors
Stefano Serafin (University of Vienna)
Juerg Schmidli (Goethe University Frankfurt)
Brigitta Goger (ETH Zurich)
Nikolina Ban (University of Innsbruck)
Sven Kotlarski (MeteoSwiss)
Stephanie Westerhuis (University of Innsbruck)
Marco Arpagaus (MeteoSwiss)

Contributors
Silvio Davolio (National Reasearch Council of Italy)
Lorenzo Giovannini (University of Trento)
Tina Katopodes Chow (UC Berkeley)
Daniel Kirshbaum (McGill University)
Jason Knievel (NCAR-RAL)
Martin Köhler (DWD)
Manuela Lehner (University of Innsbruck)
Claire Merker (MeteoSwiss)
Gudrun Nina Petersen (Icelandic Met Office)
Irina Sandu (ECMWF)
Yann Seity (Meteo France)
Ivana Stiperski (University of Innsbruck)
Annelize van Niekerk (Met Office)
Christoph Wittmann (GeoSphere Austria)


observation
processes
mountain
MoBL
climate
understanding
area
Alps
TEAMx
region
weather
exchange
atmosphere
research
experiments
coordinated
surface
atmospheric
modelling
simulations
complex
flow
impact
experimental
boundary
orography
campaign
programme
Meteorological
measurement
forecasts
MAP
COPS
T-REX
PYREX
MATERHORN
Perdig
ALPEX
Mediterranean
Europe
California

Contents

1. Introduction

The TEAMx initiative integrates both modeling and observational efforts. A large-scale observational campaign known as the TEAMx Observational Campaign (TOC) is scheduled to occur from autumn 2024 to autumn 2025. The TOC focuses on multiple target areas located along a transect through the Eastern Alps, from the German Alpine Foreland through the Inn and Adige Valley to the Po Valley. Two extended observation periods (EOPs), one during winter and another during summer, will provide valuable measurement data, particularly from aircraft operations. The main objective of this document is to provide guidance on effectively utilizing the abundant observational data and conducting modeling research within the TEAMx framework to enhance numerical weather prediction (NWP) and climate model development.

Modellers in the mountain meteorology and climatology community generally agree that that the accuracy of numerical models over mountainous regions is worse than over adjacent flat and homogeneous terrains. The diminished skill in mountainous areas is primarily attributed to the lack of parameterization schemes capable of adequately representing sub-grid-scale processes over complex terrain. Although these two statements enjoy broad consensus, scientific literature provides limited and inconclusive evidence in both cases.

In addition to deficiencies in parameterization schemes, other factors contribute to relatively poor weather forecasts and climate projections over mountains. These include discretization errors in terrain-following coordinate systems, inaccuracies in model terrain representation, and suboptimal data assimilation techniques, resulting in less accurate initial conditions. Determining the relative contributions of these factors to overall model skill remains an open question. Furthermore, the absence of reliable observations from high-elevation regions makes the verification and evaluation of weather and climate models over such areas particularly challenging. Chapter 2 provides a review of these topics, aiming to motivate coordinated modeling research within TEAMx.

Prior to the TOC, modeling activities will focus on conducting intercomparison studies (Chapter 3) to identify the limitations exhibited by current numerical models when applied to regions with complex terrain. The selection of processes and case studies aligns with the modeling challenges outlined in Chapter 2, and makes reference to the target areas of the TOC.

During the TOC, the focus will be on testing next-generation operational NWP systems and experimenting with near-real-time assimilation of campaign observations into operational-scale models (Chapter 4).

Following the TOC, the objective is to produce an unprecedented very-high resolution reanalysis (with a grid spacing of up to $\Delta x = 100\,\mathrm{m}$) of the EOPs (Chapter 5). As many as possible of the high-frequency and high-density observations collected during the campaign will be assimilted in high-resolution simulations, to obtain an optimal estimate of the atmospheric state in the TEAMx target areas. This local reanalysis, once available, can serve as the basis for defining initial and boundary conditions for subsequent high-resolution modeling case studies related to the TEAMx campaign. Additionally, it can serve as a benchmark for conducting free-running limited-area climate simulations.

The observations gathered during the TOC are expected to initiate a series of process-oriented modeling studies in NWP and climate research. These studies may involve detailed analyses of specific case studies, or idealized simulations designed to clarify the underlying dynamics of particular phenomena. While our current plan does not aim to coordinate these future studies, it aims to establish a common quality standard for all modeling research conducted within TEAMx. To achieve this, we summarize a few well-established best practices that we strongly encourage all TEAMx modelers to follow (Chapter 6).

Finally, we complete this document with a comprehensive review of published modeling studies that specifically address the TEAMx target areas or nearby regions (Chapter 7).

2. Known model errors over complex terrain

In recent decades, significant advancements were made in numerical atmospheric modeling and prediction techniques, largely driven by the increased computational power available (Bauer et al., 2015). Operational weather and climate models now employ horizontal grid spacing in the kilometric range, enabling reasonable resolution of a wide range of mesoscale and boundary-layer phenomena specific to mountainous terrain. However, some mountain weather processes remain under-resolved. For instance, while kilometric-scale resolution is generally sufficient for successfully simulating large-scale gravity waves, it is not optimal for capturing very localized slope flows. Additionally, high resolution in the numerical grid does not necessarily guarantee more accurate simulations. NWP and climate models continue to encounter various challenges in simulating atmospheric phenomena over mountainous terrain. In the following, we summarize some well-known model errors that arise when modeling atmospheric processes over mountains.

- **Numerical inaccuracies due to terrain-following grids**. Most models employ a vertical coordinate formulation in which the lowest model level follows the terrain surface, and the influence of the underlying orography diminishes as altitude increases towards a flat model top (Gal-Chen and Somerville, 1975). Consequently, irregularities in the lower boundary, such as orography, are reflected at all grid levels. The irregular geometry of terrain-following grids enhances discretization errors throughout the computational domain. To mitigate these issues, several approaches have been proposed. One approach is the use of hybrid coordinate systems, where the coordinate surfaces transition from terrain-following to pressure or geometric height with increasing altitude (Simmons and Burridge, 1981). Another approach involves designing a vertical coordinate that ensures a more rapid decay of small-scale disturbances with altitude, known as the Smooth-LEVEl vertical coordinate (SLEVE, Schär et al., 2002; Leuenberger et al., 2010). Alternatively, the application of a horizontal diffusion filter (Klemp, 2011; Westerhuis and Fuhrer, 2021; Westerhuis et al., 2021) has been considered. Immersed boundary methods (IBMs, Lundquist et al., 2010) and cut cells (Steppeler et al., 2002) have also been proposed as alternatives to represent orography on a regular Cartesian grid. These methods circumvent some of the challenges associated with terrain-following coordinates but introduce new complexities. For example, IBM requires interpolation across solid faces, which can be computationally demanding.

- **Computation of horizontal pressure gradients near sloping surfaces**. Large truncation errors arise when computing the horizontal pressure gradient term in the momentum equations in the presence of steep slopes, especially near the ground. These errors are particularly large when the difference in surface elevation between neighboring grid cells exceed the vertical grid spacing. Nonlinear amplification of these truncation errors, enhanced by aliasing, may even lead to numerical instability. To mitigate this issue, it is necessary to extrapolate the pressure vertically to a common level before calculating the horizontal gradient (Mahrer, 1984). Zängl (2012) has demonstrated that the implementation of a truly-horizontal pressure gradient discretization, combined with the use of the Exner function, allows for unprecedented representation of steep slopes, reaching up to 70%.
- **Computation of horizontal diffusion along horizontal surfaces**. When dealing with steep slopes and high model resolution, the diffusion of temperature, moisture and hydrometeor variables should be computed on truly horizontal surfaces rather than on the sloping coordinate surfaces (Zängl, 2002). When mixing is computed along coordinate surfaces, the vertical component of the gradients (which is typically much larger than the horizontal components) is erroneously projected onto the horizontal direction. This can have a substantial impact on precipitation forecasts, particularly for water vapor and hydrometeor mixing ratios (Zängl, 2004).
- **Too rapid fog dissipation**. In numerical simulations of the stable boundary layer over complex terrain, the dissipation of fog is often too rapid. This issue can be attributed not only to deficiencies in parameterization schemes but also to the intersecting of physically flat tops of stratus clouds by the sloping vertical coordinate surfaces. This intersection can promote excessive vertical mixing due to numerical diffusion associated with horizontal advection (Westerhuis et al., 2020, 2021). To address this problem, it is desirable to achieve a rapid decay of the orographic signal with altitude. However, in highly complex terrain such as the Alps, currently employed methods like the SLEVE coordinate fail to sufficiently flatten the coordinate surfaces over hilly terrain at the necessary (low) altitudes. A potential solution involves locally smoothing the coor-dinate surfaces in regions characterized by hilly terrain (Westerhuis and Fuhrer, 2021).
- **Elevation-dependent temperature bias**. Simulations at the kilometer-scale often feature a cold bias on mountain tops/slopes and a warm bias in valleys (e.g., Quéno et al., 2016; Vionnet et al., 2016). Many possible reasons have been identified, including: lack of realism in local breezes (and hence temperature advection) due to coarse model resolution; deficiencies in physical parameterisations (three-dimensional effects not represented in neither the radiation or turbulence parameterization); inadequate adaptation to complex terrain of the covariance modelling in the assimilation system. Models further struggle to simulate truly calm wind conditions, which are necessary for cold-air pool development resulting from nighttime radiative cooling in valleys and basins.
- **Elevation-dependent precipitation bias**. Observations indicate that the intensity of extreme short-duration precipitation decreases with increasing elevation, a phenomenon known as the "reverse orographic effect." However, current convection-permitting climate models tend to underestimate this effect (Dallan et al., 2023). To ensure accurate projections, it is necessary to develop bias-correction approaches specifically tailored to mountainous terrain (e.g., Velasquez et al., 2020).
- **Representation of turbulence in models**. All weather and climate models represent the diffusive effects of atmospheric turbulence with parameterization schemes that were generally developed for flat and homogeneous terrain. Parameterized turbulence is frequently distributed over the orography in a layer of approximately constant depth, which does not align

with observations of the mountain boundary layer structure (Rotach and Zardi, 2007). Additionally, parameterized turbulent kinetic energy is often underestimated in complex-terrain areas (Couvreux et al., 2016). Traditional turbulence parameterizations primarily consider vertical turbulent exchange, neglecting the significant impact of horizontal heterogeneities to the turbulence structure over mountainous terrain (Goger et al., 2018). This discrepancy between parameterizations and reality highlights the need for the development of hybrid or scale-aware turbulence parameterizations (Goger et al., 2019).

- **Wind speed overestimation in large-eddy simulations**. Overestimation of wind speed has been reported in most of the recent large-eddy simulation studies (dx<100 m) over mountainous terrain (Gerber et al., 2018; Goger et al., 2022). This systematic error in wind speed estimation can lead to an inaccurate representation of scale interactions, such as between dynamically-induced gravity waves and the boundary layer within valleys. It may also contribute to the models' inability to maintain realistic gradients in atmospheric variables and constituents over time. A major factor contributing to these issues is the smoothing of steep slopes, which can distort the flow patterns. Additionally, unrealistic representation of land-use characteristics could also play a significant role in the inaccuracies observed in wind speed estimation (Quimbayo-Duarte et al., 2022).
- **Precipitation spill-over**. Bias dipoles in precipitation forecasts, characterized by positive and negative biases occurring on opposite sides of a mountain range (windward/leeward), have been observed in several studies (Colle et al., 2005; Serafin and Ferretti, 2007). One potential reason for this phenomenon lies in microphysics parameterizations, which include inaccurate semi-empirical relationships for the fall speed of hydrometeors. Errors in fall speed, coupled with the horizontal advection of hydrometeors across the mountains, result in a horizontal shift of precipitation maxima from their optimal location. Specifically, an upstream shift is observed when the fall speed is overestimated, and a downstream shift occurs when it is underestimated. This issue is particularly relevant for hydrometeors with low fall speeds, such as snow, which can be horizontally advected over long distances.

3. Modelling before the TOC

3.1 NWP model intercomparison studies

Model intercomparison studies are an important component of the scientific process in NWP and climate modeling. Comparing multiple models against observational data and/or against each other sheds light on their strengths, weaknesses, and uncertainties. One of the best-known idealized model intercomparison studies in ABL research is the GEWEX Atmospheric Boundary Layer Study (GABLS), which consisted of several benchmark cases (GABLS1-4). The first two GABLS studies focused on the intercomparison of single-column models (SCMs) for idealized cases with prescribed surface temperatures and simplified wind profiles. GABLS3 examined the model comparability with observations and the interaction with the underlying surface. GABLS4 dealt with the arctic boundary layer and the performance of snow models. These benchmark cases highlighted inter-model variability in simulations of the same phenomenon and led to improvements of some parameterization schemes.

A single-column model approach such as that used in GABLS is not appropriate to make progress in atmospheric modelling over orographically complex terrain. In fact, the spatial heterogeneity of the lower boundary and the high degree of horizontal variability of the atmospheric state imply that one-dimensional (vertical) parameterizations are rigorously not applicable.

Previous model intercomparison studies in mountain meteorology research dealt with thermally- and dynamically-driven mesoscale processes. Schmidli et al. (2011) compared simulations of the daytime valley-wind system over idealized orography, and noticed that the largest differences between simulations depended on turbulence parameterization schemes. Doyle et al. (2011) compared simulations of mountain waves over both idealized and real orography. They demonstrated that differences in dynamical cores, lower boundary condition formulation, and surface-layer schemes caused marked discrepancies in the simulations by different models.

In the following, five new model intercomparison projects are presented (Table 1), focusing on phenomena commonly observed in mountainous regions. Three of these studies evaluate simulations against observations from campaigns that took place in the Inn Valley during autumn 2017 (Penetration and Interruption of Alpine Foehn project, PIANO Haid et al., 2020) and in summer/autumn 2019 (Cross-valley flow in the Inn Valley project, CROSSINN Adler et al., 2021).

Model	OD	CAP	TDF	CON	LES
ECMWF-IFS	✓				
ICON	✓	$*$	$*$	✓	
AROME			$*$	$*$	
MESO-NH		✓		✓	✓
UM		✓			
WRF		✓	$*$	$*$	✓
CM1					$*$
ARPS					✓
GRAMM-SCI			✓		

Table 1: Models represented in NWP intercomparison studies (as of 7 June 2023). An $*$ means that different model configurations are being tested. The acronyms CAP, TDF, CON and OD refer to the real-case intercomparison studies on cold-air pools, thermally-driven flows, moist convection and orographic drag, respectively. LES refers to the intercomparison of idealized large-eddy simulations of orographic convection initiation.

3.1.1 Orographic drag

Seamless modelling over orography from climate resolutions, to global and regional (sub-km) NWP is a recognised challenge. The amount of drag at resolutions from 10 km to 100 m should shift from parameterized to resolved. Simulations at resolutions from 5 km to 2.5 km are compared to high-resolution simulations (e.g. 500-m ICON) as well as to 10-40 km low-resolution control configurations. Sensitivity experiments target different settings of sub-grid scale orographic and turbulent orographic form drag parameterizations at low resolution.

3.1.2 Cold-air pools

This study is promoted the TEAMx Mountain Boundary Layer Working Group. It deals with the evolution of a cold-air pool in the Inn Valley over the city of Innsbruck, covering its entire life cycle from the formation to the breakup.

Cold-air pools are particularly challenging for NWP because of (i) the associated high stability, which means that traditional surface-layer parameterizations based on Monin-Obukhov similarity theory may not provide an adequate description of the turbulent transport and (ii) the oftentimes small scales and local processes, which necessitate very high horizontal and vertical resolution.

A case study was selected from an undisturbed period during the autumn 2017 PIANO field campaign in the Inn Valley, Austria.

The study compares simulations from four different models with a 1-km horizontal grid spacing in a domain covering the entire Alps. All models are run both in a configuration that matches as closely as possible some predetermined settings (vertical model levels, land cover properties, physics parameterizations, initialization time, boundary conditions), and in an optimal configuration determined on the basis of user experience.

The analysis focuses on the model representation of the strength, depth, and spatial extent of the cold-air pool in comparison to observations. The ensemble of simulations can also provide a type of benchmark for future cold-air pool simulations.

3.1.3 Thermally-driven flows

This study is promoted the TEAMx Mountain Boundary Layer Working Group. It focuses on the evaluation of model skill in reproducing thermally-driven winds and the associated thermodynamic fields in the Inn Valley, in real-case hindcast simulations. The initiative aims at updating and extending the findings by Schmidli et al. (2011), who considered idealized simulations (and therefore could not evaluate model skill) and focused on the daytime phase only. Here, the simulation of a full diurnal cycle permits evaluating model skill in the different phases of the diurnal evolution of the MoBL: daytime, nighttime, and transitions.

Simulations focus on IOP 8 of the CROSSINN field campaign (13 September 2019), characterized by weak synoptic forcing and a well-developed thermally-driven circulation in the Inn Valley. Model output is verified mainly in the CROSSINN target area with data from automatic weather stations, i-Box flux stations and CROSSINN observations (soundings, wind lidar profiles, coplanar retrievals).

Simulations are run at 1 km grid spacing, to evaluate model performance at the typical resolution of limited-area operational forecasts. A single computational domain is used, covering the entire Alpine region (Umek et al., 2021) and directly forced by IFS forecasts. Simulations at higher resolution may be performed in a second phase, to evaluate possible improvements at sub-kilometer resolution, which will probably become realistic for operational forecasts in the next few years. Models are configured with similar orography, land-use and vertical resolution. Multiple simulations with the same model in different configurations are also performed, mainly to evaluate the impact of different parameterization schemes.

The analysis considers the strength and timing of the thermally-driven circulation in the valley boundary layer in the Inn Valley, and the associated thermal field and horizontal pressure gradients. At a larger scale, measurements from temperature and humidity profilers and radiosoundings in the whole computational domain are used to evaluate to what extent model deficiencies in the Inn Valley are connected to larger-scale phenomena.

3.1.4 Convection: NWP

This study is promoted by the TEAMx Convection Working Group. It aims at evaluating the ability of current NWP kilometric models with explicit deep convection to forecast summer convection over the Alps in terms of location, timing and intensity. The study focuses on real cases featuring various conditions, from weakly forced summertime diurnal convection to synoptically triggered and organized convection.

The 23-29 July 2019 week has been chosen, as it represents a typical transition from stable conditions to days with localized/stationary convection, and finally to two days with widespread organized convection. The main focus area when choosing this period was the Inn valley, but for most of the days convection occurred and sometimes organized at a larger scale. All simulations cover the area from 43°N to 49 °N and from 5°E to 17°E. Simulation output from all models is remapped to a common grid with horizontal resolution of 0.01°. Model orography follows as closely as possible the operational COSMO model set-up at MeteoSwiss, while the vertical resolution is as close as possible to the 90 vertical levels used in AROME by Météo-France. The location of convective initiation in relation to the orography is examined in detail. Furthermore, the localation, intensity, and chronology of the precipitation events are of key interest.

3.1.5 Convection: LES

This study is promoted by the TEAMx Convection Working Group. Quasi-idealized LES are used to evaluate inter-model variability in the representation of boundary layer and cumulus development in a realistic but simplified summertime flow. In all simulations, the only parameterized processes are cloud microphysics and subgrid-scale turbulence, which helps to narrow and isolate the potential sources of error. Of particular interest is the determination of whether the different effective resolutions and physics schemes across the models meaningfully impacts the boundary-layer growth and turbulence, the mixing of clouds with their surroundings, and the vertical cloud development.

A section of the Italian Alps is extracted for the terrain, which is modified to isolate the ridge from the surrounding terrain and enforce periodicity along one axis. The initial thermodynamic sounding is a 10-year averaged ERA5 climatology over summer months at 06:00 UTC at the closest grid point to the city of Verona, Italy. The flow is initially quiescent, and sensible and latent heat fluxes are prescribed as diurnally varying sinusoidal functions. Initial experiments at coarse resolution indicate a multi-hour period of ABL deepening driven by surface heating, after which clouds begin to develop over lower ridges along the mountain flanks. These clouds progressively deepen and shift toward the higher terrain, ultimately giving rise to a mesoscale convective systems along the ridge top.

Analyses of the simulations focuses on time-evolving two-dimensional fields like surface precipitation, boundary-layer depth and turbulence, cloud cover, cloud-top altitude, and CAPE/CIN, along with surface fields (winds, water vapour, and potential temperature). Additional scientific insight is gained through analyses of cumulus dilution and detrainment, conditionally averaged buoyancy, cloud water, and updraft profiles, and power spectral densities over different portions of the domain.

3.2 Mountain climate

In parallel to the NWP group, the climate modeling group is working on understanding and modeling processes by which mountains are shaping regional climates and their spatial and temporal variability. Before the TEAMx FOC, the WG Mountain Climate group will exploit available high-resolution (km-scale) simulations available through several ongoing international projects with the main goal of recognizing the processes misrepresented in our current high-resolution models. Some of those data are the following:

- CORDEX FPS on convection over the Alps - This high-resolution multi-model ensemble of climate simulations is becoming available at a horizontal grid spacing of 3 km, integrated over 10-year long periods for present (2000-2009), historical (1991-2000) and future (2090-2099) climate - using RCP8.5 greenhouse gas and aerosol emission scenario (presented in Ban et al., 2021, and Pichelli et al., 2021).
- Austrian climate scenarios ÖKS15 - The ÖKS15 provides a standard ensemble of regional climate projections based on EURO-CORDEX simulations driven by CMIP5 global climate models (Jakob et al. 2014). The standard ensemble is provided with a horizontal grid spacing of 12 km. In addition, ÖKS15 provides (nominal) 1 km scenarios produced by bias-adjusting 12 km EURO-CORDEX simulations (Truhetz et al. 2016). This ensemble data set has served as an Austrian reference for climate change impact research since 2016.
- CH2018 Swiss Climate Scenarios - The CH2018 climate scenarios are based on the EURO-CORDEX RCM ensemble (EUR-11 and EUR-44) and involve comprehensive statistical post-processing and bias adjustment. A range of useful products is provided, including transient scenarios (1981-2099) at daily resolution for individual sites and on a regular 2 km grid. The CH2018 scenarios are planned to be extended/updated in 2025 by means of new approaches and user products. CORDEX simulations at 12 km grid spacing available through ESGF - new simulations are driven by CMIP6 GCMs.

In addition to the existing simulations, the groups will be running additional simulations most likely for shorter time periods and to address a specific research question.

The above-listed data is currently utilized in several ongoing projects listed below.

- *HighResMountains - Mountain weather in high-resolution climate data: How will the new generation of ÖKS benefit from new emerging datasets?*
  The main goal of HighResMountains is to gain a deeper understanding of extreme events and their processes and changes with further warming of the atmosphere over the Alps. The specific focus is on precipitation (rain and snow) and mountain wind systems (like foehn) which will be analyzed using different high-resolution datasets - more specifically CORDEX

FPS Alps and ÖKS15. The main results of the project will provide relevant information and guidelines on methods limitations for the development of new Austrian climate scenarios.

- *Orographic Convective precipitation*
  This PostDoc project aims to better understand convective phenomena over the eastern part of the Alps. The main goal of the project is to identify weather and climate conditions that influence most summer storms and their future evolution. In the first phase of the project, the focus is on the analysis of the CORDEX FPS Alps data, while in the second part sensitivity studies will be conducted to better understand the most impacting mechanisms.
- *Austrian Reanalysis - ARA*
  The main goal of the ARA project is to create first of its kind high resolution (2.5 km) re-analysis ensemble dataset for Austria by assimilating observations using the 3DVAR of the C-LAEF ensemble system based on the AROME model. This re-analysis will provide detailed spatially, temporally, and physically consistent 3D and 2D information on the state of the atmosphere in Austria from 2010 – 2020, with a potential extension to cover the FOC period. Successful completion of this prototype has the potential to further develop into a viable operational/commercial product. It will provide essential climate variables (ECVs) at spatial and temporal scales relevant for the NWP (numerical weather prediction)/ climate research community and can be further exploited by impact research to improve resilience in the community by strengthening mitigation and adaptation efforts.

- *Do kilometer-scale climate models really perform better over complex orography?*
  In this work, regional climate simulation with a grid spacing of 2 km and 12 km conducted with the COSMO model is evaluated against available observations over Europe. In contrast to previous studies which showed a blurred image of the model performance, here high versus low mountains and flatlands are distinguished. The preliminary results show that the increase in the resolution clearly improves the model's performance over flatland but the added value of using higher resolution is often smaller over complex mountainous terrain (i.e., higher mountains) than over flatland, especially for precipitation and clouds. The results suggest that the full potential of the kilometer-scale may not be reached in regions of complex orography and call for future research to improve those models, i.e., calls for the need for TEAMx research on climate scales as well. The work is presented in the manuscript that is currently in preparation (Poujol et al.). The work is however based on one model (COSMO), so there is a potential to extend it to the CORDEX FPS simulations.
- *Snow cover*
  A part of the research in WG Mountain Climate assesses the potential and limitations of km-scale climate models to represent past and future changes in snow conditions in the European Alps. In one of the earlier studies, Lüthi et al. (2019) showed how the representation of snow cover is better represented when going to higher resolution with the COSMO model. This work is further extended to simulations within CORDEX FPS with the aim of evaluation and estimation of the added value of high-resolution climate models versus Euro-CORDEX regional climate models, focusing on both past and future conditions over the Alpine region.
- *Urban climate in complex orography*
  The main question to be addressed here is how well the simulations perform in cities in mountainous terrain. As mentioned above, we can see a smaller added value in the application of km-scale climate models in complex topography than over flatland. However such analysis still needs to be done more specifically for the cities. In this work, existing CORDEX FPS simulations will be utilized in order to compare the model performance with the current settings, and then later additional sensitivity simulations could be conducted testing different urban parametrizations.

- *Application of machine learning*
  So far, the analysis of existing km-scale simulations was relying on conventional methods. However, the potential of machine learning has not been fully explored. For example, one could use it to identify the model biases different over different weather situations or to assess how much we can learn about the model performance from a very short time scale and how that applies to longer.

The above results and findings, together with previously conducted research will potentially be summarized in a Review paper on Mountain Climate (or Complex Orography) Modeling which should represent the main results and challenges of modeling climate over complex orography. The current idea is to capture all scales - from global to local and from global climate models down to regional high-resolution (km-scale or even LES scale) models.

4. NWP support during the TOC

The primary aim of running NWP modelling systems during the TEAMx Observational Campaign is to support the planning of the IOPs, especially for running observation platforms that do not operate on a 24/7 basis. This can partly be done with coarse-scale models such as ECMWF's IFS. However, in complex terrain, higher-resolution modelling systems provide more useful local information to guide intensive observations. Five different institutions - the national weather services from Austria (GeoSphere, GS), France (Météo-France, MF), Germany (DWD), Switzerland (MeteoSwiss, MCH), together with the Italian Institute of Atmospheric Sciences and Climate (ISAC) - will run various configurations of three different models (ICON, MOLOCH, AROME) at high resolutions over the Alps. Table 2 summarises some key aspects of each model setup and Figure 1 illustrates the model domains. Except for MOLOCH these limited-area modelling systems also run a data assimilation system at a high resolution, thereby capturing the state of the atmosphere over the Alps more in detail than global models can do.

A second motivation to run high resolution modelling systems for the EOPs and IOPs in real time is to get both immediate and retrospective feedback on their performance: scientists will be able to compare different modelling systems with different strengths and weaknesses with a wealth of observations in 4D. The direct intercomparison of different modelling systems with different approaches for data assimilation, discretizations of the dynamics, and physical parametrisations, provides an added value for the TEAMx research community.

Figure 1: Domains of models which will be available during the TOC. Météo-France (MF, blue) and GeoSphere Austria (GS, orange) run AROME, DWD (green) and MeteoSwiss (MCH, red) run ICON and ISAC (pink) provides MOLOCH. Subject to small changes.

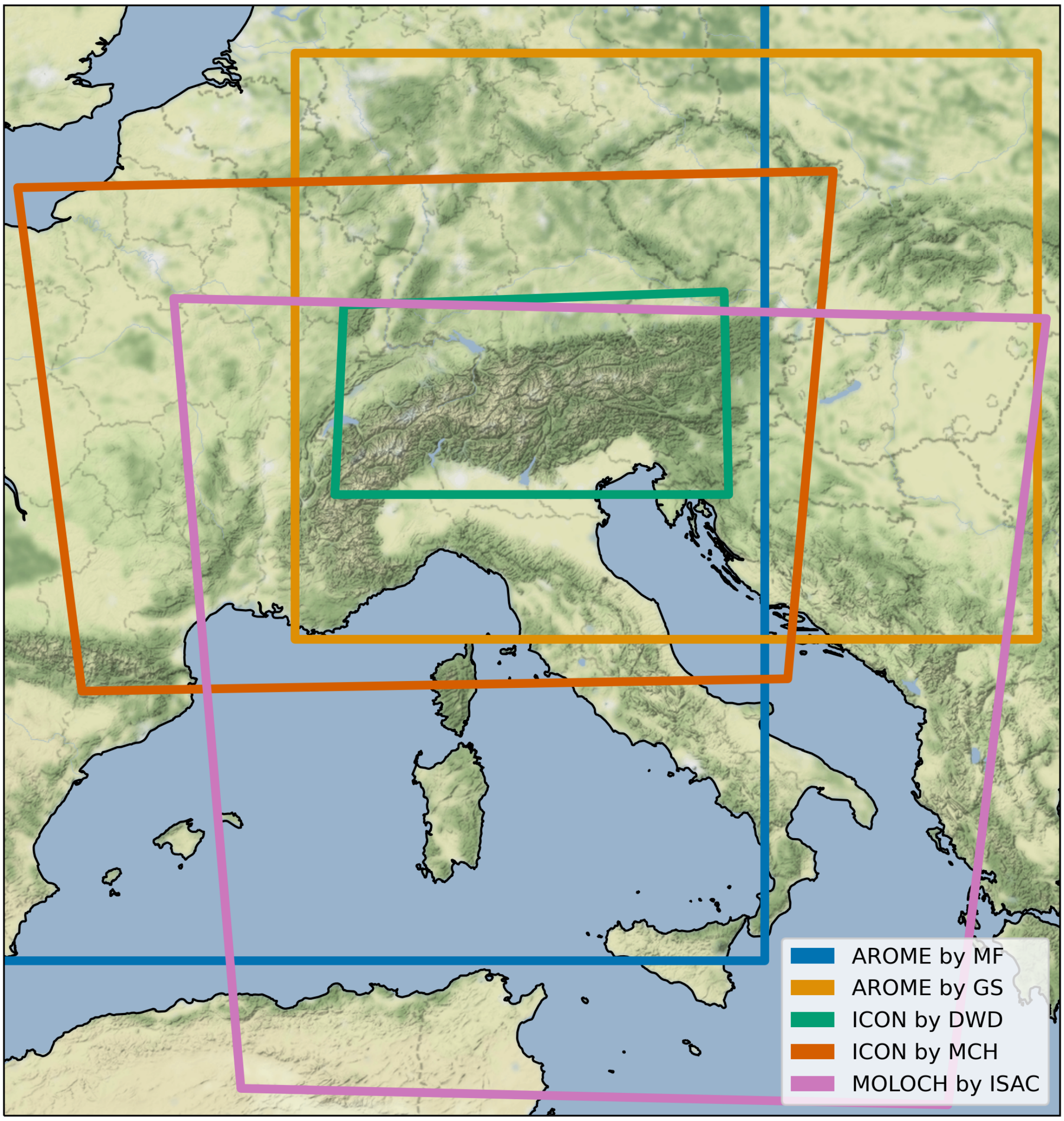

Table 2: Models and their configurations which will be available during the TOC. More details will follow at a later stage.

Model	Institution	Configuration	mesh-size (km)	N vert. levels	N ens. members	Lead time (h)
AROME	MF	AROME-FR	1.3	90	18	51
AROME	MF	AROME-FR-500	0.5	120	no ens.	36
AROME	GS	C-LAEF	2.5	90	16+1	tbd
AROME	GS	C-LAEF-1k	1	90	tbd	tbd
ICON	DWD	ICON-DWD-T	0.5	65	tbd	tbd
ICON	MCH	ICON-CH1-EPS	1.1	80	11	33
ICON	MCH	ICON-CHT-EPS	0.5	80	11	tbd
MOLOCH	ISAC	MOLOCH-ISAC	1.25	60	no ens.	45

5. Modelling after the TOC

5.1 High-resolution reanalysis of the TOC

Field campaigns such as TEAMx always spawn many modelling studies, where simulations in hindcast mode shed light on the meteorological processes that affect intensive observation periods, thereby aiding the interpretation of measurements. In the case of mesoscale NWP simulations ($\Delta x \sim$1 km), drawing their initial and boundary conditions from global analyses is often an adequate approach. However, tackling the TEAMx scientific objectives requires explicit resolution of advective transport and of coherent turbulent structures in the mountain boundary layer, which is feasible only with microscale simulations ($\Delta x \sim$100 m or smaller for convective boundary layers, $\Delta x \sim$10 m or smaller for stable boundary layers).

Large-eddy simulations of the mountain boundary layer are technically feasible, but their accuracy is severely limited by the lack of microscale detail in their initial and boundary conditions. The highest-resolution analysis products currently available, obtained from convective-scale data assimilation systems, have $\Delta x \sim$1 km. Over mountains, they suffer from non-negligible biases both because of their marginally adequate resolution and because the observations they ingest are too sparse to properly sample the spatial and temporal variability of the mountain boundary layer.

Therefore, we argue that an analysis with native grid resolution of about 100 m is necessary in order to achieve reasonably accurate modelling of the mountain boundary layer (MoBL), and we propose that the computation of such an analysis becomes a cornerstone of the modelling activities foreseen after the TEAMx Field Observation Campaign. In practice, the *TEAMx reanalysis* should be a four-dimensional gridded dataset, providing a statistically optimal estimate of the atmospheric state over the TEAMx target areas during the special observation periods of winter 2024-25 and summer 2025 (Inn Valley, Adige Valley, Northern and Southern Alpine Forelands in Bavaria and Po Valley), based on all conventional and special atmospheric measurements available during the campaign.

To align with the current state of the art on convective-scale data assimilation, we propose that the TEAMx reanalysis is produced with ensemble data assimilation (EnDA) methods. Data assimilation combines information from sparse observations and a preexisting simulation (the background) to compute an analysis. In EnDA, the background state and its flow-dependent uncertainty (error covariance) are estimated with an ensemble of numerical simulations. The observation and background error covariances are then used to determine the relative weight of the background and the observations in the final analysis. The process is iterative, meaning that each analysis provides the initial conditions for the subsequent ensemble run.

Because of the high computational requirements, a high-resolution ensemble analysis is only feasible in small nested domains. The TEAMx renalysis shall have a horizontal grid spacing that is similar or slightly smaller than that of operational NWP models (1 km or 500 m) in a broad domain covering the Alps, and of about 100 m in nested domains covering the TEAMx target areas. The parent coarse-resolution analysis provides initial and boundary conditions for the 100-m ones; in turn, synthetic observations from the 100-m model runs can be assimilated in the next cycle of the coarse-resolution one, thus accomplishing two-way feedback. The analyses at different scales may assimilate separate sets of observations (e.g., rain radar only at coarse resolution, doppler wind lidar only at high resolution). In the high-resolution nested domains, as many as possible of the TEAMx special observations shall be assimilated (including Doppler wind lidars; temperature, relative humidity and wind profilers; rain radars; radiosondes and dropsondes; surface and airborne in-situ measurements).

Based on the best of our knowledge, no high-resolution ensemble analysis product has ever been computed so far, although similar undertakings were recently made or are now underway:

- Deployments of the US Department of Energy Atmospheric Radiation Measurement Mobile Facilities are routinely complemented by simulations with the WRF-based LASSO environment (Gustafson et al., 2020). LASSO simulations assimilate boundary-layer observations in mesoscale runs using variational methods, and draw from these runs the boundary forcing for free-running large-eddy simulations. No direct assimilation of observation in LES runs is performed.
- Miyoshi et al. (2016) presented an experimental assimilation of high-frequency (1/30 s^{-1}) volumetric radar measurements in a 100-member ensemble of 100-m simulations. This ‘big data assimilation‘ exercise dealt with a single weather event and a single observation type, and lacks connection with field campaign measurements.
- There are plans to develop hectometric-scale on-demand simulations in the near future, for instance in the Destination Earth On-Demand Extremes Digital Twin DE_330_MF (Randriamampianina, 2023). The prevalent focus of this initiative is on prediction rather than analysis, and plans concerning the assimilation of non-conventional observations are unclear at the moment.

Computing a high-resolution campaign reanalysis with an ensemble filter poses significant conceptual and technical challenges. Among them:

- Ensemble runs at 100-m resolution are extremely demanding in terms of computational resources.
- The assimilation algorithm must be capable of handling massive amounts of observational data, possibly with correlated observation errors.
- Very dense observations sample scales of atmospheric motion that are smaller than the background model’s effective resolution, so observation thinning or averaging might be needed.
- For some of the special measurement platforms operated during TEAMx, observation operators may not be readily available.
- Background error statistics will likely evolve in a non-Gaussian and nonlinear manner, a challenging scenario for current operational ensemble filters.
- Frequent assimilation increments (still within the model spin-up phase) may introduce dynamical imbalances in the analyses, which need being mitigated.

Designing and testing sensible methods to overcome these challenges will require coordinated research, bringing together expertise in numerical weather prediction, high-resolution modelling and boundary-layer observation with in-situ and remote-sensing platforms. A joint effort to apply for funding is currently being made in a team comprising the Universities of Vienna and Innsbruck, the Karlsruhe Institute of Technology, Meteo Swiss and the Deutscher Wetterdienst.

The development of a methodology for high-resolution reanalyses has countless potential applications besides TEAMx (e.g., future field campaigns; parameterization development; energy meteorology).

5.2 Climate applications at convection permitting scale

Results and datasets obtained during the TEAMx FOC are expected to be of high value for subsequent climate applications. This climatological "digestion" of results will be coordinated by the TEAMx working group on Mountain Climate. A range of possible applications of the FOC results and achievements is envisaged:

- **Climate simulations at convection-permitting scale**: Most NWP modeling systems applied during the TEAMx FOC can also be run in climate mode without data assimilation and for long-term simulation periods, including future scenario simulations (e.g. Ban et al., 2014). These applications will profit from any further insight into model performance and bias structures obtained during the FOC and from any improvements in the representation of land-surface interactions in complex terrain. This includes future applications of limited-area model ensembles at convection-permitting scale (such as those planned during the upcoming phase of the CORDEX initiative) as well as kilometer-scale global model applications (such as those envisaged in the nextGEMS project).
- **Evaluation of *twin analogs***: Past and projected future global warming also strongly affects mountain climates, with the latter partly showing above-average temperature change signals and elevation-dependent patterns (Pepin and Lundquist, 2008; Pepin et al., 2015, 2022). While the short term TEAMx FOC does not cover the relevant time scales, "twin analogs" evaluated during the FOC could be of value for assessing the robustness of future climate change projections. For instance, quantifying the expected reduction in surface snow cover with climate warming and its possible feedback on near-surface air temperature change relies on the capability of climate models to correctly represent the snow-vs-no snow temperature contrast. Previous work indicates that this representation might be distorted in today's regional climate models (Winter et al., 2017). The TEAMx FOC offers the possibility to evaluate this relationship by transferring the temporal concept into a spatial one, i.e. by comparing model temperature and flux biases between snow-covered and snow-free sites. In addition to snow coverage, further systematic temporal trends expected as a consequence of global warming could be transferred into a spatial context and evaluated based on the FOC results (land cover change and upward moving tree line, modified aerosol burdens, intense precipitation events, etc.).
- **Storylines**: Storyline approaches have emerged as an important pillar of climate change communication and can provide physically-based insights into the nature of future climate extremes and their respective impacts (e.g. Shepherd et al., 2018; Sillmann et al., 2020). An often applied storyline technique consists of transferring a well-understood and well-simulated historical event into a future climate employing surrogate or pseudo global warming approaches. Depending on the evolvement of weather conditions during the TEAMx FOC and the ability to represent individual events in the applied modeling systems, a transfer of such events into a future climate along with an assessment of their respective impacts on depending natural and societal systems could be envisaged.

Note that these applications generally require models to be run in *climate mode*, i.e. in a free-running manner without data assimilation and possibly also without large-scale nudging.

6. Guidelines and best practices

6.1 Lower boundary conditions

Specification of the lower boundary condition for simulations of MoBL processes involves at least two aspects:

- Orography: Most idealized modelling studies so far adopt unrealistically smooth terrain profiles. To address research questions focusing on complex terrain, a realistic degree of variability in the properties of the lower boundary should be incorporated in the experiment design. This may be achieved by generating synthetic terrain whose variance follows a predetermined power spectrum, or by drawing information from real high-resolution terrain data (Kirshbaum et al., 2007). Commonly used orography datasets are digital elevation data from the Advanced Spaceborne Thermal Emission and Reflection radiometer(ASTER), the Multi-Error-Removed Improved-Terrain dataset (MERIT), and the Shuttle Radar Topography Mission (SRTM). Model intercomparison studies mimicking real cases require testing of various degrees of orography smoothing to find a setup that allows all models to run stably.
- Land use and soil type: The land use and soil type datasets should feature a resolution which is higher than the resolution of the simulations. Commonly used land use datasets include the COoRdination of INformation on the environment land cover (CORINE), the GLOBal land COVER (GLOBCOVER) map, and ECOlogical and CLImatic MAPping database ECOCLIMAP database. The soil type is often derived from the Harmonized World Soil Database.

6.2 Initial conditions and spin-up time

For (semi-)idealised simulations, ICs should be specified on the basis of available radiosoundings. For real-case simulations, high-resolution data-assimilation cycles can provide accurate analysis fields. However, if a model intercomparison project includes different models and the data-assimilation cycle associated to one of them serves as ICs for all this likely deteriorates comparability, especially in the early hours. As most models have functionalities to be driven by ECMWF's IFS this is a recommended choice despite the comparably lower resolution. In any case, a spin-up time of at least 6-12 hours should be considered.

6.3 Grid resolution and turbulence modelling

The properties of many phenomena over complex terrain (e.g. the strength and extent of the valley wind, the formation of gravity waves, or the three-dimensional structure of turbulence) are governed by the underlying orography. Therefore, the right choice of horizontal grid spacing depending on the phenomenon of interest is crucial for a successful weather simulation over mountainous terrain.

Wagner et al. (2014) suggest that at least ten grid points across a valley are necessary to simulate the relevant proceses for the formation of the thermally-induced circulation. The choice of grid spacing is also dependent on the diurnal cycle – during the daytime, coarser grid spacings at the hectometric range might be sufficient for a convective ABL to develop, but during the night-time, when a stable boundary layer persists, horizontal grid spacings well below 100 m are advisable (Cuxart, 2015; Muñoz Esparza et al., 2017). Adjustment of the grid spacing to the simulated phenomena is not feasible for operational NWP models, but it should be considered for tailored high-resolution case studies.

Another relevant decision for numerical modeling below the kilometric range is the choice of turbulence parameterization in the model. At grid spacings around and below 1 km, classic turbulence parameterization schemes, such as the Mellor-Yamada framework, are not enhtirely appropriate for complex orography, because three-dimensional effects, such as horizontal shear production of turbulence, become relevant (Goger et al., 2018). A solution to overcome this problem is the extension of 1D parameterizations to include a simplified treatment of 3D turbulence dynamics (Zhong and Chow, 2013; Goger et al., 2019; Juliano et al., 2022).

Even with these hybrid turbulence parameterizations, the grey zone of turbulence (Wyngaard, 2004; Honnert et al., 2020) needs special treatment in high-resolution mesoscale simulations. The turbulence grey zone lies at grid spacings between the "mesoscale limit" ($\approx$ 1 km, where turbulence is fully parameterized), and the LES range ($\approx$ 100 m, where the largest eddies are fully resolved). At grey zone resolution, turbulence is partly parameterized and partly resolved, which often leads to unrealistic flow structures in the simulations (Chow et al., 2019). Scale-aware turbulence schemes might bring a solution to this problem and are necessary for simulations in the hectometric range (e.g. Zhang et al., 2018).

6.4 Online diagnostics and post-processing

When simulations are run in LES mode, it is assumed that the largest eddies in the turbulent flow are resolved. Due to the turbulent nature of the flow, it is unlikely that instantaneous model output (e.g., every 30 minutes) is representative of the mean state of the flow. Therefore, frequent model output is required in order to accurately estimate the mean state of the flow. Then any turbulent model variable $\tilde{a}(\mathbf{x},t)$ can be split into a mean $A(\mathbf{x},t)$ and a fluctuating part $a(\mathbf{x},t)$ (Schmidli, 2013):

$$\tilde{a}(\mathbf{x},t) = A(\mathbf{x},t) + a(\mathbf{x},t).$$

Using such a decomposition, the Reynolds average of a product of two turbulent variables is given by:

$$\overline{\tilde{a}\tilde{b}} = AB + \overline{ab}.$$

If one of the variables is a velocity variable, the covariance $\overline{ab}$ represents a turbulent flux. In LES, a turbulent flux consists of a resolved and a subgrid part. When the turbulent motions are well-resolved in the LES (i.e., $\Delta x \ll l$, where l represents the length scale of the large turbulent eddies), the resolved turbulent flux is larger than the subgrid turbulent flux.

Due to the fact that turbulent motions are explicitly resolved, analysing LES output usually requires post-processing of the high-frequency model output with time- and space averaging (Schmidli, 2013; Göbel et al., 2022). Writing high-frequency output is demanding in terms of model runtime and storage space, so the recommended method to diagnose turbulence statistics is recursive averaging. Following the method of Schmidli (2013), Weinkaemmerer et al. (2022) and Wagner et al. (2014) implemented recursive averaging routines for turbulent flow quantities in idealized simulations for the CM1 model and the WRF model, respectively. The WRF implementation of recursive averaging was further improved by Umek et al. (2021, 2022) and by Göbel et al. (2022). The latter work also introduced a numerically consistent calculation of budget terms in Cartesian coordinates, which is otherwise hard to achieve (given the mass-based curvilinear grid adopted in WRF).

6.5 Model verification

When it comes to evaluating simulations against special observations (e.g., measurements available at a single site and for a short period), visual inspection of plots is the most common approach and it is acceptable in many circumstances. If a quantitative measure of the discrepancy between simulations and observations is desired, many verification scores for both continuous and binary predictands can be considered. See Wilks (2011) for a comprehensive review.

Some fundamental issues should be kept in mind:

- **Error magnitude relative to natural variability**. Comparing forecast error measures between sites with different degrees of natural variability can often be misleading. For instance, a mean absolute error of 2 m s^{-1} in wind speed forecasts may be exceedingly high at sites with near-zero mean wind speeds, but perfectly fine at sites with an aggressive wind climate. It is common practice to normalize error measures (e.g., the root mean square error) with the standard deviation of the verifying observations.
- **Systematic and random errors**. Forecast inaccuracy depends both on systematic deviation from the truth (bias) and on random errors due to forecast uncertainty. Bias is often large over mountainous terrain, but can easily be removed by post-processing. Some verification methods (e.g., Taylor diagrams; Taylor, 2001) are insensitive to mean bias by design, and their use is encouraged because they emphasize forecast errors that are inherently time-dependent or random.
- **Sampling error and small sample size**. Verification scores are always computed from a finite set of forecasts and verifying observations, which implies that they are subject to sampling error. Especially when sample sizes are small, it is important to evaluate the uncertainty margins of verification scores. For instance, when comparing two forecasts, a given difference between verification scores might be too small in relation to the their uncertainty, thus making it impossible to decide if a forecast is better than the other. The recommended approach is to estimate confidence intervals with non-parametric methods such as bootstrapping (resampling with replacement). In general it cannot be assumed that forecast-observation pairs in the sample are statistically independent, so their serial or spatial autocorrelation has to be taken into account with block bootstrapping (Wilks, 1997).
- **Testing differences between verification scores**. When comparing different forecasts, statistical hypothesis testing is the recommended method. In general, the null hypothesis to be tested is that the competing forecasts have equal skill. Because forecast errors are often correlated (if forecast A is wrong, most likely even forecast B will be similarly wrong), hypothesis tests should refer to the difference between scores. With this approach, the null hypothesis to be tested is that the score difference is zero.
- **Observation errors**. When verifying forecasts, observations are often taken at face value. Actually, they have their own uncertainty, which is determined both by instrumental error

and by representativeness error. Hacker et al. (2011) demonstrated that accounting for observation errors can change the conclusions drawn from verification: for instance, an ensemble forecast that is seemingly severely underdispersive could be judged reasonably reliable if observation errors were simulated in the verification process. This is possible by randomly perturbing the verifying observations according to a predetermined error variance. Rigorous estimates of observation error variance are hard to obtain, but can be achieved within a data assimilation framework (Desroziers et al., 2005). In general, the representativeness component of observation errors is model-dependent.

- **Double-penalty errors**. Highly resolved forecast fields contain spatial variability at small scales, which inherently enhances model errors. Rainfall forecasts are a typical example: a coarse-resolution model that totally misses a local precipitation maximum is penalized once; a high-resolution model that accurately captures the intensity of the maximum but misrepresents its location is penalized twice (once for missing the event in its right place, once for putting it in the wrong place). Verification scores that are specifically designed to circumvent double-penalty errors should be chosen. One example, in the context of spatial verification of binary events (e.g., accumulated rainfall exceeding a threshold), is the fractions skill score (Roberts and Lean, 2008).

6.6 Determination of the boundary-layer depth

The boundary-layer depth z_i is hard to evaluate over mountainous terrain, from both measurements and numerical simulations. The method followed to determine z_i should always be specified exactly. Comparison of multiple methods is encouraged. Most numerical models can be extended fairly easily to incorporate mass conservation equations for passive tracers emitted at the surface. In addition to common approaches (parcel method, gradient method, bulk Richardson number method), the determination of z_i on the basis of tracer mass fields is recommended.

6.7 Spatial interpolation

We recommend using standard, well-tested, computationally efficient interpolation tools instead of self-coded routines. Some interpolation functions (e.g., those available in `cdo`) require converting model output into CF-compliant format.

In addition, awareness of the exact georeferencing of the model grid (datum, projection) is necessary to correctly evaluate wind forecasts (wind directions on the model grid do not necessarily coincide with geographical wind directions).

6.8 Data availability and research reproducibility

All TEAMx modelling activities are expected to produce output which is reproducible and made openly available. NetCDF output should follow the CF-conventions.

7. Review of previous modelling studies

The following review summarizes published numerical modelling studies that refer to the three TEAMx target areas. Most of them were performed in the Inn Valley and surroundings (Sec. 7.1). Two large sets of simulations are connected with measurement campaigns on foehn winds (MAP, 1999 and PIANO, 2017), where a strong focus was laid on understanding windstorms over the city of Innsbruck. Another set of simulations refers to model evaluation with the i-Box turbulence observations, for purposes of turbulence parameterization evaluation and improvement.

Fewer numerical studies dealt with the other TEAMx target areas (Sec. 7.2-7.3). However, a few modelling studies on convection, gravity-wave breaking, cold-air pool dynamics and glacier-atmosphere interactions were conducted in neighboring regions in the Eastern Alps (Sec. 7.4).

7.1 Inn Valley Target Area

For clarity, we group weather modelling studies in the Inn Valley depending on the studied phenomenon.

- **Foehn winds.** The city of Innsbruck and surroundings are subject to frequent foehn wind episodes throughout the year, while the statistical maxima occur in spring and autumn. Field observations of foehn during the Mesoscale Alpine Programme (MAP) motivated several modelling studies. Gohm et al. (2004) and Zängl and Gohm (2006) investigated the mechanisms of foehn flow in the Wipp Valley (a side valley of the Inn Valley) in great detail. An additional set of simulations of the same location was conducted by Zängl et al. (2003), and the impact of vertical levels and the PBL scheme on simulations of foehn was also investigated Zängl et al. (2008). Gohm and Mayr (2004) discussed the hydraulic aspects of foehn flow as well. The major findings on foehn flow over Innsbruck and the adjacent side valleys are summarized by Mayr et al. (2007). The second large measurement campaign on foehn flow was conducted in autumn 2017 (PIANO, Haid et al., 2020), where the main focus was laid on foehn-cold air pool interactions. Accompanying simulations in the LES range were performed by Umek et al. (2021, 2022), and the simulations were utilized for process understanding and sensitivity experiments on horizontal grid spacing. A trajectory analysis for an unusual foehn event was performed by Saigger and Gohm (2022).

- **Thermally-induced circulations.** The valley wind system of the Inn Valley was first investigated numerically by Zängl (2004) with semi-idealized simulations, while a subsequent work studied the impact of synoptic flow on the valley wind system Zängl (2009). Measurements at i-Box flux measurement sites (Rotach et al., 2017) were used to evaluate boundary-layer parameterizations (Goger et al., 2016). Simulations of up-valley wind days with the COSMO model showed that 3D effects in the model's TKE progostic equation are essential for the correct simulation of TKE in the Inn Valley (Goger et al., 2018, 2019).
- Other relevant studies investigated wintertime smog episodes in the Inn Valley (Schicker and Seibert, 2009), and the impact of improved land-use datasets on weather forecasts over two Austrian regions, one of them being the Inn Valley (Schicker et al., 2015).

7.2 Adige Valley Target Area

In the Adige Valley, located south of the Alpine main crest, many applied modelling studies were conducted in the recent years. Topics included: the sensitivity of simulated wind speeds to horizontal grid spacings (Giovannini et al., 2014a); process studies on the Vaia storm over Northwestern Italy (Giovannini et al., 2021; Sioni et al., 2023); mountain boundary layer processes (Giovannini et al., 2014b); urban meteorology studies on building energy consumption (Pappaccogli et al., 2021); pollutant dispersion studies in connection with tracer release experiments (Zardi et al., 2021); improvement of turbulence parameterizations for dispersion modelling (Tomasi et al., 2019); evaluation and optimation of snowpack modelling in land-surface models (Tomasi et al., 2017).

7.3 Northern Pre-Alpine Target Area

The Northern Pre-Alpine Target Area is located in the Bavarian Alpine foreland, directly adjacent to the Alps. Weather in this region is heavily influenced by atmospheric processes related to mountainous terrain, for example Alpine pumping, which was studied in regional climate simulations by Graf et al. (2016). Siedersleben and Gohm (2016) studied the dynamics of potential vorticity banners leading to banded convection over the forelands in a wintertime episode of strong southerly synoptic flow. Hald et al. (2019) instead carried out LES of weather events during the ScaleX campaign and presented a qualitative comparison between simulations and observed turbulence statistics.

7.4 Vicinity of target areas

Zängl (2005a) investigated the interactions between cold-air pools and a valley wind system as well as cold-air pools in the Alpine foreland of Bavaria (Zängl, 2005b). Scheffknecht et al. (2017) investigated a long-lived supercell travelling along the Alpine main crest. Recently, the Hintereisferner glacier in the Ötztal was subject to a detailed study on glacier boundary layer processes (Goger et al., 2022).

7.5 Kilometer-scale climate applications

Climate modeling studies so far did not focus on the TEAMx targeted specific sub-regions in the Alps, but were analysing Alps as a whole and surrounding areas. Thus we here provide an overview of existing literature over the entire Alps.

Kilometer-scale applications of climate models over the Alpine domain now have an about 10-year long history. A review of the underlying techniques, assumptions, applications, and challenges is provided by Lucas-Picher et al. (2021) or Schär et al. (2020). The availability of model simulations was boosted by the CORDEX Flagship Pilot Study on Convective Systems in which the Alpine domain was selected as one of the focus areas (Ban et al., 2021; Pichelli et al., 2021; Coppola et al., 2020) as well as by the EUCP project `https://www.eucp-project.eu/`. Added value analyses indicate a clear benefit of the kilometer-scale resolution in many different aspects. Those are:

- The diurnal cycle of summer precipitation (see e.g., Ban et al., 2021; Knist et al., 2020; Lind et al., 2020; Leutwyler et al., 2017; Ban et al., 2015, 2014). In comparison to coarse resolution models which use parametrization of convection, km-scale climate models are able to reproduce the diurnal cycle of summer precipitation. These results are supported by many studies and are more recently confirmed by multi-model ensemble at the km-scale resolution (Ban et al., 2021). The better performance is visible in the timing of the diurnal cycle, precipitation intensity, and frequency.
- Precipitation extremes at daily and sub-daily scale (see e.g., Ban et al., 2021; Pichelli et al., 2021; Knist et al., 2020; Lind et al., 2020; Ban et al., 2020; Leutwyler et al., 2017; Ban et al., 2015, 2014). Extreme precipitation at short timescales with a potential to trigger flash floods, landslides, and debris flow, was for a long time misrepresented by regional climate models with grid spacing above 10 km. However, km-scale resolution improved the representation of these events, their relation with temperature, and can alter the climate change signal.
- Snow cover. A recent study using COSMO simulation at the km-scale resolution showed a much better representation of snow cover over the European Alps when using higher resolution (Lüthi et al., 2019). It was clearly shown that the 2 km model outperforms 12 and 50 km models, despite having a slight overestimation of snow cover in fall and too fast melt of it during springtime. It is also shown that only a high-resolution model can represent elevations higher than 2500 meters, which is important for glaciers and glacier modeling.
- Temperature (see e.g., Soares et al., 2022; Ban et al., 2014). Even though most of the km-scale models show warm bias, one can see the improvements in the simulation of diurnal temperature range (Ban et al., 2014).
- Winds (see e.g., Belušić Vozila et al., 2023; Belušić et al., 2018). Analysis of wind in high-resolution simulations is still very limited and in addition to the Alps, it includes the surrounding but also very complex areas like Adriatic. The existing literature shows that km-scale climate simulations are better in representing Bora and Sirocco along the Adriatic, as well as land-sea breezes.

In addition to those, future climate model applications in the Alpine domain, including future generations of national or Alpine scale climate scenarios, can be expected to more and more rely on ensembles of kilometer-scale climate model simulations.

Acronyms

AROME	Application of Research to Operations at MEsoscale model
ARPS	Advanced Regional Prediction System model
CM1	Cloud Model 1
CMIP	Coupled Model Intercomparison Project
CORDEX	COordinated Regional Climate Downscaling EXperiment
CORDEX FPS	CORDEX Flagship Pilot Study
COSMO	COnsortium for Small-scale MOdeling
CROSSINN	CROSS-valley flow in the Inn valley (project)
DWD	Deutscher WetterDienst
ECMWF	European Centre for Medium-range Weather Forecasts
EnDA	Ensemble Data Assimilation
EOP	Extended Observation Period
ERA5	ECMWF Reanalysis version 5
ESGF	Earth System Grid Federation
GABLS	GEWEX Atmospheric Boundary Layer Study
GEWEX	Global Energy and Water Exchanges
GRAMM-SCI	Graz Mesoscale Model–Scientific
HRES	High-RESolution
IBM	Immersed Boundary Method
i-Box	Innsbruck Box
ICON	ICOsahedral Nonhydrostatic weather and climate Model
ICs	Initial Conditions
IFS	Integrated Forecasting System
IOP	Intensive Observation Period
LASSO	LES ARM Symbiotic Simulation and Observation Workflow
LBCs	Lateral Boundary Conditions
LES	Large-Eddy Simulation
LIDAR	Laser Imaging Detection and Ranging

MAP	Mesoscale Alpine Programme
MesoNH	Mesoscale Non-Hydrostatic model
MoBL	Mountain Boundary Layer
MOLOCH	Local Model on Height coordinates
NWP	Numerical Weather Prediction
PIANO	Penetration and Interruption of Alpine Foehn (project)
SCM	Single-Column Model
SLEVE	Smooth-LevEVel VErtical (coordinate)
TEAMx	multi-scale Transport and Exchange processes in the Atmosphere over Mountains – programme and eXperiment
TKE	Turbulent Kinetic Energy
TOC	TEAMx Observational Campaign
UM	Unified Model
WRF	Weather Research and Forecasting model

(List is currently incomplete.)

Bibliography

Adler, B., et al., 2021: CROSSINN: A field experiment to study the three-dimensional flow structure in the Inn Valley, Austria. *Bull. Amer. Meteorol. Soc.*, **102 (1)**, E38–E60, doi:10.1175/BAMS-D-19-0283.1.

Ban, N., J. Rajczak, J. Schmidli, and C. Schär, 2020: Analysis of alpine precipitation extremes using generalized extreme value theory in convection-resolving climate simulations. *Climate Dynamics*, **55 (1-2)**, 61–75, doi:10.1007/s00382-018-4339-4.

Ban, N., J. Schmidli, and C. Schär, 2015: Heavy precipitation in a changing climate: Does short-term precipitation increase faster? *Geophysical Research Letters*, **42 (4)**, 1165–1172, doi:10.1002/2014GL062588.

Ban, N., J. Schmidli, and C. Schär, 2014: Evaluation of the convection-resolving regional climate modeling approach in decade-long simulations. *Journal of Geophysical Research: Atmospheres*, **119 (13)**, 7889–7907, doi:https://doi.org/10.1002/2014JD021478.

Ban, N., et al., 2021: The first multi-model ensemble of regional climate simulations at kilometer-scale resolution, part i: evaluation of precipitation. *Climate Dynamics*, **57 (1)**, 275–302, URL `https://doi.org/10.1007/s00382-021-05708-w`.

Bauer, P., A. Thorpe, and G. Brunet, 2015: The quiet revolution of numerical weather prediction. *Nature*, **525 (7567)**, 47–55, doi:10.1038/nature14956.

Belušić, A., M. T. Prtenjak, I. Güttler, N. Ban, D. Leutwyler, and C. Schär, 2018: Near-surface wind variability over the broader adriatic region: insights from an ensemble of regional climate models. *Climate Dynamics*, **50 (11)**, 4455–4480, doi:10.1007/s00382-017-3885-5, URL `https://doi.org/10.1007/s00382-017-3885-5`.

Belušić Vozila, A., D. Belušić, M. T. Prtenjak, I. Güttler, and et al., 2023: Evaluation of the near-surface wind field over the adriatic region: local wind characteristics in the convection-permitting model ensemble. *Climate Dynamics*, URL `https://doi.org/10.1007/s00382-023-06703-z`.

Chow, F. K., C. Schär, N. Ban, K. A. Lundquist, L. Schlemmer, and X. Shi, 2019: Crossing multiple gray zones in the transition from mesoscale to microscale simulation over complex terrain. *Atmosphere*, **10 (5)**, doi:10.3390/atmos10050274.

Colle, B. A., M. F. Garvert, J. B. Wolfe, C. F. Mass, and C. P. Woods, 2005: The 13–14 December 2001 IMPROVE-2 Event. Part III: Simulated Microphysical Budgets and Sensitivity Studies. *Journal of the Atmospheric Sciences*, **62 (10)**, 3535–3558, doi:10.1175/JAS3552.1.

Coppola, E., et al., 2020: A first-of-its-kind multi-model convection permitting ensemble for investigating convective phenomena over europe and the mediterranean. *Climate Dynamics*, **55 (1-2)**, 3–34, doi:10.1007/s00382-018-4521-8.

Couvreux, F., E. Bazile, G. Canut, Y. Seity, M. Lothon, F. Lohou, F. Guichard, and E. Nilsson, 2016: Boundary-layer turbulent processes and mesoscale variability represented by numerical weather prediction models during the BLLAST campaign. *Atmos. Chem. Phys.*, **16 (14)**, 8983–9002, doi:10.5194/acp-16-8983-2016.

Cuxart, J., 2015: When Can a High-Resolution Simulation Over Complex Terrain be Called LES? *Front. Earth Sci.*, **3 (87)**, 6, doi:10.3389/feart.2015.00087.

Dallan, E., F. Marra, G. Fosser, M. Marani, G. Formetta, C. Schär, and M. Borga, 2023: How well does a convection-permitting regional climate model represent the reverse orographic effect of extreme hourly precipitation? *Hydrology and Earth System Sciences*, **27 (5)**, 1133–1149, doi:10.5194/hess-27-1133-2023.

Desroziers, G., L. Berre, B. Chapnik, and P. Poli, 2005: Diagnosis of observation, background and analysis-error statistics in observation space. *Quarterly Journal of the Royal Meteorological Society*, **131 (613)**, 3385–3396, doi:10.1256/qj.05.108.

Doyle, J. D., et al., 2011: An Intercomparison of T-REX Mountain-Wave Simulations and Implications for Mesoscale Predictability. *Mon. Wea. Rev.*, **139 (9)**, 2811–2831, doi:10.1175/MWR-D-10-05042.1.

Gal-Chen, T. and R. C. Somerville, 1975: Numerical solution of the Navier-Stokes equations with topography. *Journal of Computational Physics*, **17 (3)**, 276–310, doi:10.1016/0021-9991(75)90054-6.

Gerber, F., et al., 2018: Spatial variability in snow precipitation and accumulation in COSMO–WRF simulations and radar estimations over complex terrain. *The Cryosphere*, **12 (10)**, 3137–3160, doi:10.5194/tc-12-3137-2018.

Giovannini, L., G. Antonacci, D. Zardi, L. Laiti, and L. Panziera, 2014a: Sensitivity of Simulated Wind Speed to Spatial Resolution over Complex Terrain. *Energy Procedia*, **59**, 323–329, doi:10.1016/j.egypro.2014.10.384.

Giovannini, L., S. Davolio, M. Zaramella, D. Zardi, and M. Borga, 2021: Multi-model convection-resolving simulations of the October 2018 Vaia storm over Northeastern Italy. *Atmos. Res.*, **253**, 105 455, doi:10.1016/j.atmosres.2021.105455.

Giovannini, L., D. Zardi, M. de Franceschi, and F. Chen, 2014b: Numerical simulations of boundary-layer processes and urban-induced alterations in an Alpine valley. *Int. J. Climatol.*, **34 (4)**, 1111–1131, doi:10.1002/joc.3750.

Göbel, M., S. Serafin, and M. W. Rotach, 2022: Numerically consistent budgets of potential temperature, momentum, and moisture in cartesian coordinates: application to the wrf model. *Geosci. Model Dev.*, **15 (2)**, 669–681, doi:10.5194/gmd-15-669-2022.

Goger, B., M. W. Rotach, A. Gohm, O. Fuhrer, I. Stiperski, and A. A. M. Holtslag, 2018: The impact of three-dimensional effects on the simulation of turbulence kinetic energy in a major alpine valley. *Boundary-Layer Meteorol*, **168 (1)**, 1–27, doi:10.1007/s10546-018-0341-y.

Goger, B., M. W. Rotach, A. Gohm, I. Stiperski, and O. Fuhrer, 2016: Current challenges for numerical weather prediction in complex terrain: Topography representation and parameterizations. *2016 International Conference on High Performance Computing Simulation (HPCS)*, 890–894, doi:10.1109/HPCSim.2016.7568428.

Goger, B., M. W. Rotach, A. Gohm, I. Stiperski, O. Fuhrer, and G. de Morsier, 2019: A new horizontal length scale for a three-dimensional turbulence parameterization in mesoscale atmospheric modeling over highly complex terrain. *J. Appl. Meteor. Climatol.*, **58 (9)**, 2087–2102, doi:10.1175/JAMC-D-18-0328.1.

Goger, B., I. Stiperski, L. Nicholson, and T. Sauter, 2022: Large-eddy simulations of the atmospheric boundary layer over an alpine glacier: Impact of synoptic flow direction and governing processes. *Q. J. R. Meteorol. Soc.*, **148 (744)**, 1319–1343, doi:10.1002/qj.4263.

Gohm, A. and G. J. Mayr, 2004: Hydraulic aspects of föhn winds in an Alpine valley. *Q. J. R. Meteorol. Soc.*, **130 (597)**, 449–480, doi:10.1256/qj.03.28.

Gohm, A., G. Zängl, and G. J. Mayr, 2004: South Foehn in the Wipp Valley on 24 October 1999 (MAP IOP 10): Verification of High-Resolution Numerical Simulations with Observations. *Mon. Wea. Rev.*, **132 (1)**, 78–102, doi:10.1175/1520-0493(2004)132<0078:SFITWV>2.0.CO;2.

Graf, M., M. Kossmann, K. Trusilova, and G. Mühlbacher, 2016: Identification and Climatology of Alpine Pumping from a Regional Climate Simulation. *Front. Earth Sci.*, **4 (5)**, 11, doi:10.3389/feart.2016.00005.

Gustafson, W. I., et al., 2020: The Large-Eddy Simulation (LES) Atmospheric Radiation Measurement (ARM) Symbiotic Simulation and Observation (LASSO) Activity for Continental Shallow Convection. *Bull. Amer. Meteor. Soc.*, **101 (4)**, E462–E479, doi:10.1175/BAMS-D-19-0065.1.

Hacker, J. P., et al., 2011: The U.S. Air Force Weather Agency's mesoscale ensemble: scientific description and performance results. *Tellus A: Dynamic Meteorology and Oceanography*, **63 (3)**, 625–641, doi:10.1111/j.1600-0870.2010.00497.x.

Haid, M., A. Gohm, L. Umek, H. C. Ward, T. Muschinski, L. Lehner, and M. W. Rotach, 2020: Foehn-cold pool interactions in the Inn Valley during PIANO IOP2. *Q. J. R. Meteorol. Soc.*, **146 (728)**, 1232–1263, doi:10.1002/qj.3735.

Hald, C., M. Zeeman, P. Laux, M. Mauder, and H. Kunstmann, 2019: Large-Eddy Simulations of Real-World Episodes in Complex Terrain Based on ERA-Reanalysis and Validated by Ground-Based Remote Sensing Data. *Mon. Wea. Rev.*, **147 (12)**, 4325–4343, doi:10.1175/MWR-D-19-0016.1.

Honnert, R., et al., 2020: The atmospheric boundary layer and the "gray zone" of turbulence: A critical review. *J. Geophys. Res. Atmos.*, **125 (n/a)**, e2019JD030 317, doi:10.1029/2019JD030317, e2019JD030317 2019JD030317.

Juliano, T. W., B. Kosović, P. A. Jiménez, M. Eghdami, S. E. Haupt, and A. Martilli, 2022: "Gray Zone" Simulations Using a Three-Dimensional Planetary Boundary Layer Parameterization in the Weather Research and Forecasting Model. *Mon Wea Rev*, **150** (**7**), 1585–1619, doi:10.1175/MWR-D-21-0164.1.

Kirshbaum, D. J., G. H. Bryan, R. Rotunno, and D. R. Durran, 2007: The triggering of orographic rainbands by small-scale topography. *Journal of the Atmospheric Sciences*, **64** (**5**), 1530–1549, doi:10.1175/JAS3924.1.

Klemp, J. B., 2011: A terrain-following coordinate with smoothed coordinate surfaces. *Mon. Wea. Rev.*, **139** (**7**), 2163–2169, doi:10.1175/MWR-D-10-05046.1.

Knist, S., K. Goergen, and C. Simmer, 2020: Evaluation and projected changes of precipitation statistics in convection-permitting WRF climate simulations over central europe. *Climate Dynamics*, **55** (**1-2**), 325–341, doi:10.1007/s00382-018-4147-x.

Leuenberger, D., M. Koller, O. Fuhrer, and C. Schär, 2010: A Generalization of the SLEVE Vertical Coordinate. *Mon. Wea. Rev.*, **138** (**9**), 3683–3689, doi:10.1175/2010MWR3307.1.

Leutwyler, D., D. Lüthi, N. Ban, O. Fuhrer, and C. Schär, 2017: Evaluation of the convection-resolving climate modeling approach on continental scales. *Journal of Geophysical Research: Atmospheres*, **122** (**10**), 5237–5258, doi:10.1002/2016JD026013, URL `http://dx.doi.org/10.1002/2016JD026013`, 2016JD026013.

Lind, P., et al., 2020: Benefits and added value of convection-permitting climate modeling over fenno-scandinavia. *Climate Dynamics*, **55** (**7**), 1893–1912, doi:10.1007/s00382-020-05359-3, URL `https://doi.org/10.1007/s00382-020-05359-3`.

Lucas-Picher, P., D. Argüeso, E. Brisson, Y. Tramblay, P. Berg, A. Lemonsu, S. Kotlarski, and C. Caillaud, 2021: Convection-permitting modeling with regional climate models: Latest developments and next steps. *WIREs Clim Change*, **n/a** (**n/a**), e731, URL `https://doi.org/10.1002/wcc.731`.

Lundquist, K. A., F. K. Chow, and J. K. Lundquist, 2010: An immersed boundary method for the weather research and forecasting model. *Mon. Wea. Rev.*, **138** (**3**), 796–817, doi:10.1175/2009MWR2990.1.

Lüthi, S., N. Ban, S. Kotlarski, C. R. Steger, T. Jonas, and C. Schär, 2019: Projections of alpine snow-cover in a high-resolution climate simulation. *Atmosphere*, **10** (**8**), 463, doi:10.3390/atmos10080463, URL `https://www.mdpi.com/2073-4433/10/8/463`.

Mahrer, Y., 1984: An improved numerical approximation of the horizontal gradients in a terrain-following coordinate system. *Mon. Wea. Rev.*, **112** (**5**), 918–922, doi:10.1175/1520-0493(1984)112<0918:AINAOT>2.0.CO;2.

Mayr, G. J., et al., 2007: Gap flows: Results from the Mesoscale Alpine Programme. *Q. J. R. Meteorol. Soc.*, **133** (**625**), 881–896, doi:10.1002/qj.66.

Miyoshi, T., et al., 2016: "Big Data Assimilation" Revolutionizing Severe Weather Prediction. *Bulletin of the American Meteorological Society*, **97** (**8**), 1347–1354, doi:10.1175/BAMS-D-15-00144.1.

Muñoz Esparza, D., J. K. Lundquist, J. A. Sauer, B. Kosović, and R. R. Linn, 2017: Coupled mesoscale-LES modeling of a diurnal cycle during the CWEX-13 field campaign:

From weather to boundary-layer eddies. *J. Adv. Model. Earth Syst.*, **9 (3)**, 1572–1594, doi: 10.1002/2017MS000960.

Pappaccogli, G., L. Giovannini, D. Zardi, and A. Martilli, 2021: Assessing the Ability of WRF-BEP+BEM in Reproducing the Wintertime Building Energy Consumption of an Italian Alpine City. *Journal of Geophysical Research: Atmospheres*, **126 (8)**, e2020JD033 652, doi:https://doi.org/10.1029/2020JD033652.

Pepin, N. and J. D. Lundquist, 2008: Temperature trends at high elevations: Patterns across the globe. *Geophysical Research Letters*, **35**, L14 701, doi:10.1029/2008GL034026.

Pepin, N., et al., 2015: Elevation-dependent warming in mountain regions of the world. *Nature climate change*, **5 (5)**, 424–430, doi:10.1038/nclimate2563.

Pepin, N. C., et al., 2022: Climate changes and their elevational patterns in the mountains of the world. *Rev Geophys*, **60**, e2020RG000 730, doi:10.1029/2020RG000730.

Pichelli, E., et al., 2021: The first multi-model ensemble of regional climate simulations at kilometer-scale resolution part 2: historical and future simulations of precipitation. *Climate Dynamics*, **56 (11)**, 3581–3602, URL `https://doi.org/10.1007/s00382-021-05657-4`.

Quéno, L., V. Vionnet, I. Dombrowski-Etchevers, M. Lafaysse, M. Dumont, and F. Karbou, 2016: Snowpack modelling in the pyrenees driven by kilometric-resolution meteorological forecasts. *The Cryosphere*, **10 (4)**, 1571–1589, doi:10.5194/tc-10-1571-2016.

Quimbayo-Duarte, J., J. Wagner, N. Wildmann, T. Gerz, and J. Schmidli, 2022: Evaluation of a forest parameterization to improve boundary layer flow simulations over complex terrain. *Geosci. Model Dev.*, **15**, 5195–5209, doi:10.5194/gmd-15-5195-2022.

Randriamampianina, R., 2023: Destination Earth On-Demand Extremes Digital Twin. Tech. rep., Copernicus Meetings. doi:10.5194/egusphere-egu23-6122.

Roberts, N. M. and H. W. Lean, 2008: Scale-Selective Verification of Rainfall Accumulations from High-Resolution Forecasts of Convective Events. *Monthly Weather Review*, **136 (1)**, 78–97, doi:10.1175/2007MWR2123.1.

Rotach, M. W. and D. Zardi, 2007: On the boundary-layer structure over highly complex terrain: Key findings from MAP. *Q. J. R. Meteorol. Soc.*, **133 (625)**, 937–948, doi:10.1002/qj.71.

Rotach, M. W., et al., 2017: Investigating Exchange Processes over Complex Topography: The Innsbruck Box (i-Box). *Bull. Amer. Meteor. Soc.*, **98 (4)**, 787–805, doi:10.1175/BAMS-D-15-00246.1.

Saigger, M. and A. Gohm, 2022: Is it north or west foehn? a lagrangian analysis of penetration and interruption of alpine foehn intensive observation period 1 (PIANO IOP 1). *Weather Clim. Dynam.*, **3 (1)**, 279–303, doi:10.5194/wcd-3-279-2022.

Schär, C., D. Leuenberger, O. Fuhrer, D. Lüthi, and C. Girard, 2002: A New Terrain-Following Vertical Coordinate Formulation for Atmospheric Prediction Models. *Mon. Wea. Rev.*, **130 (10)**, 2459–2480, doi:10.1175/1520-0493(2002)130<2459:ANTFVC>2.0.CO;2.

Scheffknecht, P., S. Serafin, and V. Grubišić, 2017: A long-lived supercell over mountainous terrain. *Q. J. R. Meteorol. Soc.*, **143 (709)**, 2973–2986, doi:10.1002/qj.3127.

Schicker, I., D. Arnold Arias, and P. Seibert, 2015: Influences of updated land-use datasets on WRF simulations for two Austrian regions. *Meteorol. Atmos. Phys.*, **128 (3)**, 279–301, doi: 10.1007/s00703-015-0416-y.

Schicker, I. and P. Seibert, 2009: Simulation of the meteorological conditions during a winter smog episode in the Inn Valley. *Meteorol. Atmos. Phys.*, **103 (1-4)**, 211–222, doi:10.1007/s00703-008-0346-z.

Schmidli, J., 2013: Daytime Heat Transfer Processes over Mountainous Terrain. *J. Atmos. Sci.*, **70 (12)**, 4041–4066, doi:10.1175/JAS-D-13-083.1.

Schmidli, J., et al., 2011: Intercomparison of mesoscale model simulations of the daytime valley wind system. *Monthly Weather Review*, **139 (5)**, 1389–1409, doi:10.1175/2010MWR3523.1.

Schär, C., et al., 2020: Kilometer-scale climate models: Prospects and challenges. *Bulletin of the American Meteorological Society*, **101 (5)**, E567–E587, doi:10.1175/bams-d-18-0167.1.

Serafin, S. and R. Ferretti, 2007: Sensitivity of a Mesoscale Model to Microphysical Parameterizations in the MAP SOP Events IOP2b and IOP8. *J. Appl. Meteor. Climatol.*, **46 (9)**, 1438–1454, doi:10.1175/JAM2545.1.

Shepherd, T. G., et al., 2018: Storylines: an alternative approach to representing uncertainty in physical aspects of climate change. *Climatic Change*, **151 (3-4)**, 555–571, doi:10.1007/s10584-018-2317-9.

Siedersleben, S. K. and A. Gohm, 2016: The missing link between terrain-induced potential vorticity banners and banded convection. *Mon. Wea. Rev.*, **144 (11)**, 4063–4080, doi:10.1175/MWR-D-16-0042.1.

Sillmann, J., T. G. Shepherd, B. van den Hurk, W. Hazeleger, O. Martius, J. Slingo, and J. Zscheischler, 2020: Event-based storylines to address climate risk. *Earths Future*, **n/a (n/a)**, e2020EF001 783, doi:10.1029/2020EF001783.

Simmons, A. J. and D. M. Burridge, 1981: An energy and angular-momentum conserving vertical finite-difference scheme and hybrid vertical coordinates. *Monthly Weather Review*, **109 (4)**, 758–766, doi:10.1175/1520-0493(1981)109<0758:AEAAMC>2.0.CO;2.

Sioni, F., S. Davolio, F. Grazzini, and L. Giovannini, 2023: Revisiting the atmospheric dynamics of the two century floods over north-eastern italy. *Atmospheric Research*, **286**, 106 662, doi: https://doi.org/10.1016/j.atmosres.2023.106662.

Soares, P. M. M., et al., 2022: The added value of km-scale simulations to describe temperature over complex orography: the cordex fps-convection multi-model ensemble runs over the alps. *Climate Dynamics*, doi:10.1007/s00382-022-06593-7, URL `https://doi.org/10.1007/s00382-022-06593-7`.

Steppeler, J., H.-W. Bitzer, M. Minotte, and L. Bonaventura, 2002: Nonhydrostatic atmospheric modeling using a z-coordinate representation. *Monthly Weather Review*, **130 (8)**, 2143–2149, doi:10.1175/1520-0493(2002)130<2143:NAMUAZ>2.0.CO;2.

Taylor, K. E., 2001: Summarizing multiple aspects of model performance in a single diagram. *Journal of Geophysical Research, D: Atmospheres*, **106 (D7)**, 7183–7192, doi:10.1029/2000JD900719.

Tomasi, E., L. Giovannini, D. Zardi, and M. de Franceschi, 2017: Optimization of Noah and Noah_MP WRF Land Surface Schemes in Snow-Melting Conditions over Complex Terrain. *Mon. Wea. Rev.*, **145 (12)**, 4727–4745, doi:10.1175/MWR-D-16-0408.1.

Tomasi, E., et al., 2019: Turbulence parameterizations for dispersion in sub-kilometer horizontally non-homogeneous flows. *Atmos. Res.*, **228**, 122–136, doi:10.1016/j.atmosres.2019.05.018.

Umek, L., A. Gohm, M. Haid, H. C. Ward, and M. W. Rotach, 2021: Large eddy simulation of foehn-cold pool interactions in the Inn Valley during PIANO IOP2. *Q. J. R. Meteor. Soc.*, **147 (735)**, 944–982, doi:10.1002/qj.3954.

Umek, L., A. Gohm, M. Haid, H. C. Ward, and M. W. Rotach, 2022: Influence of grid resolution of large-eddy simulations on foehn-cold pool interaction. *Q. J. R. Meteorol. Soc.*, **148 (745)**, 1840–1863, doi:10.1002/qj.4281.

Velasquez, P., M. Messmer, and C. C. Raible, 2020: A new bias-correction method for precipitation over complex terrain suitable for different climate states: a case study using WRF (version 3.8.1). *Geoscientific Model Development*, **13 (10)**, 5007–5027, doi:10.5194/gmd-13-5007-2020.

Vionnet, V., I. Dombrowski-Etchevers, M. Lafaysse, L. Quéno, Y. Seity, and E. Bazile, 2016: Numerical Weather Forecasts at Kilometer Scale in the French Alps: Evaluation and Application for Snowpack Modeling. *J. Hydrometeor.*, **17 (10)**, 2591–2614, doi:10.1175/JHM-D-15-0241.1.

Wagner, J. S., A. Gohm, and M. W. Rotach, 2014: The Impact of Horizontal Model Grid Resolution on the Boundary Layer Structure over an Idealized Valley. *Mon. Wea. Rev.*, **142 (9)**, 3446–3465, doi:10.1175/MWR-D-14-00002.1.

Weinkaemmerer, J., I. B. Ďurán, and J. Schmidli, 2022: The impact of large-scale winds on boundary layer structure, thermally driven flows, and exchange processes over mountainous terrain. *Journal of the Atmospheric Sciences*, **79 (10)**, 2685–2701, doi:10.1175/JAS-D-21-0195.1.

Westerhuis, S. and O. Fuhrer, 2021: A locally smoothed terrain-following vertical coordinate to improve the simulation of fog and low stratus in numerical weather prediction models. *J. Adv. Model. Earth Syst.*, **13 (8)**, e2020MS002 437, doi:10.1029/2020MS002437.

Westerhuis, S., O. Fuhrer, R. Bhattacharya, J. Schmidli, and C. Bretherton, 2021: Effects of terrain-following vertical coordinates on simulation of stratus clouds in numerical weather prediction models. *Q. J. R. Meteor. Soc.*, **147 (734)**, 94–105, doi:10.1002/qj.3907.

Westerhuis, S., O. Fuhrer, J. Cermak, and W. Eugster, 2020: Identifying the key challenges for fog and low stratus forecasting in complex terrain. *Q. J. R. Meteorol. Soc.*, **146 (732)**, 3347–3367, doi:10.1002/qj.3849.

Wilks, D., 1997: Resampling hypothesis tests for autocorrelated fields. *J. Climate*, **10**, 65–82, doi:10.1175/1520-0442(1997)010<0065:RHTFAF>2.0.CO;2.

Wilks, D., 2011: *Statistical Methods in the Atmospheric Sciences, Third Edition*. Academic Press, 704 pp.

Winter, K. J.-P. M., S. Kotlarski, S. Scherrer, and C. Schär, 2017: The alpine snow-albedo feedback in regional climate models. *Climate Dynamics*, **48**, 1109–1124, doi:10.1007/s00382-016-3130-7.

Wyngaard, J. C., 2004: Toward Numerical Modeling in the “Terra Incognita”. *J. Atmos. Sci.*, **61 (14)**, 1816–1826, doi:10.1175/1520-0469(2004)061<1816:TNMITT>2.0.CO;2.

Zängl, G., 2002: An improved method for computing horizontal diffusion in a sigma-coordinate model and its application to simulations over mountainous topography. *Monthly Weather Review*, **130 (5)**, 1423–1432, doi:10.1175/1520-0493(2002)130<1423:AIMFCH>2.0.CO;2.

Zängl, G., 2004: A reexamination of the valley wind system in the Alpine Inn Valley with numerical simulations. *Meteorol. Atmos. Phys.*, **87 (4)**, 241–256, doi:10.1007/s00703-003-0056-5.

Zängl, G., 2005a: Dynamical Aspects of Wintertime Cold-Air Pools in an Alpine Valley System. *Mon. Wea. Rev.*, **133 (9)**, 2721–2740, doi:10.1175/MWR2996.1.

Zängl, G., 2005b: Wintertime Cold-Air Pools in the Bavarian Danube Valley Basin: Data Analysis and Idealized Numerical Simulations. *J. Appl. Meteor.*, **44 (12)**, 1950–1971, doi: 10.1175/JAM2321.1.

Zängl, G., 2009: The impact of weak synoptic forcing on the valley-wind circulation in the Alpine Inn Valley. *Meteorol. Atmos. Phys.*, **105 (1-2)**, 37–53, doi:10.1007/s00703-009-0030-y.

Zängl, G., 2012: Extending the numerical stability limit of terrain-following coordinate models over steep slopes. *Monthly Weather Review*, **140 (11)**, 3722–3733, doi:10.1175/MWR-D-12-00049.1.

Zängl, G. and A. Gohm, 2006: Small-scale dynamics of the south foehn in the lower Wipp Valley. *Meteorol. Atmos. Phys.*, **93 (1-2)**, 79–95, doi:10.1007/s00703-005-0154-7.

Zängl, G., A. Gohm, and G. Geier, 2003: South foehn in the Wipp Valley – Innsbruck region: Numerical simulations of the 24 October 1999 case (MAP-IOP 10). *Meteorol. Atmos. Phys.*, **86 (3-4)**, 213–243, doi:10.1007/s00703-003-0029-8.

Zängl, G., A. Gohm, and F. Obleitner, 2008: The impact of the PBL scheme and the vertical distribution of model layers on simulations of Alpine foehn. *Meteorol. Atmos. Phys.*, **99 (1-2)**, 105–128, doi:10.1007/s00703-007-0276-1.

Zardi, D., et al., 2021: The bolzano tracer experiment (btex). *Bulletin of the American Meteorological Society*, **102 (5)**, E966–E989, doi:https://doi.org/10.1175/BAMS-D-19-0024.1.

Zhang, X., J.-W. Bao, B. Chen, and E. D. Grell, 2018: A Three-Dimensional Scale-Adaptive Turbulent Kinetic Energy Scheme in the WRF-ARW Model. *Mon. Wea. Rev.*, **146 (7)**, 2023–2045, doi:10.1175/MWR-D-17-0356.1.

Zhong, S. and F. K. Chow, 2013: Meso- and Fine-Scale Modeling over Complex Terrain: Parameterizations and Applications. *Mountain Weather Research and Forecasting*, F. K. Chow, S. F. J. De Wekker, and B. J. Snyder, Eds., Springer Netherlands, Springer Atmospheric Sciences, 591–653, doi:10.1007/978-94-007-4098-3.

Zängl, G., 2004: The sensitivity of simulated orographic precipitation to model components other than cloud microphysics. *Q. J. R. Meteorol. Soc.*, **130 (600)**, 1857–1875, doi:10.1256/qj.03.119.