

PAPER • OPEN ACCESS

Dynamic deep learning based super-resolution for the shallow water equations

To cite this article: Maximilian Witte *et al* 2025 *Mach. Learn.: Sci. Technol.* **6** 015060

View the [article online](#) for updates and enhancements.

You may also like

- [Comparing AI versus optimization workflows for simulation-based inference of spatial-stochastic systems](#)
Michael Alexander Ramirez Sierra and Thomas R Sokolowski
- [A novel dynamic machine learning-based explainable fusion monitoring: application to industrial and chemical processes](#)
Husnain Ali, Rizwan Safdar, Yuanqiang Zhou et al.
- [Asymptotically stable data-driven koopman operator approximation with inputs using total extended DMD](#)
Louis Lortie and James Richard Forbes



PAPER

OPEN ACCESS

RECEIVED
6 August 2024REVISED
20 November 2024ACCEPTED FOR PUBLICATION
18 December 2024PUBLISHED
11 March 2025

Original Content from
this work may be used
under the terms of the
[Creative Commons
Attribution 4.0 licence](#).

Any further distribution
of this work must
maintain attribution to
the author(s) and the title
of the work, journal
citation and DOI.



Dynamic deep learning based super-resolution for the shallow water equations

Maximilian Witte^{1,*} , Fabrício R Lapolli² , Philip Freese³ , Sebastian Götschel³ , Daniel Ruprecht³ , Peter Korn²  and Christopher Kadow¹ 

¹ Data Analysis, German Climate Computing Center (DKRZ), 20146 Hamburg, Germany

² Climate Variability, Max-Planck Institute for Meteorology, 20146 Hamburg, Germany

³ Chair Computational Mathematics, Institute of Mathematics, Hamburg University of Technology, 21073 Hamburg, Germany

* Author to whom any correspondence should be addressed.

E-mail: witte@dkrz.de, fabricao.lapolli@mpimet.mpg.de, philip.freese@tuhh.de, sebastian.goetschel@tuhh.de, ruprecht@tuhh.de, peter.korn@mpimet.mpg.de and kadow@dkrz.de

Keywords: super-resolution, deep learning, convolutional neural network, shallow water equation, galewesky test case, hybrid modeling, numerical ocean model ICON

Abstract

Correctly capturing the transition to turbulence in a barotropic instability requires fine spatial resolution. To reduce computational cost, we propose a dynamic super-resolution approach where a transient simulation on a coarse mesh is frequently corrected using a U-net-type neural network. For the nonlinear shallow water equations, we demonstrate that a simulation with the Icosahedral Nonhydrostatic ocean model with a 20 km resolution plus dynamic super-resolution trained on a 2.5 km resolution achieves discretization errors comparable to a simulation with 10 km resolution. The neural network, originally developed for image-based super-resolution in post-processing, is trained to compute the difference between solutions on both meshes and is used to correct the coarse mesh solution every 12 h. We show that the ML-corrected coarse solution correctly maintains a balanced flow and captures the transition to turbulence in line with the higher resolution simulation. After an 8 d simulation, the L_2 -error of the corrected run is similar to a simulation run on a finer mesh. While mass is conserved in the corrected runs, we observe some spurious generation of kinetic energy.

1. Introduction

Numerical models of the atmospheric and oceanic circulation greatly aid our understanding of climate [34] and are instrumental for numerical weather prediction (NWP). While there are fundamental differences regarding scales and driving forces between weather prediction and climate simulations, there are many similarities in the underlying equations and numerical techniques. Several earth system models, like the Icosahedral Nonhydrostatic Model (ICON), are designed to be used for both. These models comprise two important components, the dynamical core and sub-grid models, so-called parameterizations. While the former is responsible for solving partial differential equations (PDEs) on the discrete grid, the latter is responsible for approximating the effect of processes such as diffusion, mixing, or turbulence that can not be resolved explicitly at a given spatio-temporal resolution [38]. To improve the accuracy of weather predictions and climate projections, there is a strong push towards ever finer grid resolutions [30]. However, finer mesh resolutions incur increasing computational cost and runtimes of climate models and energy consumption of high-performance computing (HPC) machines.

Machine learning (ML) has been gaining popularity in many areas of simulation based science. ML techniques are applied in NWP and climate sciences in different ways. Pure data-driven ML approaches, for example, are trained by reanalysis data that contain the optimal combination of model and observational information. These systems have started to challenge short and medium term weather predictions in terms of accuracy while being much more efficient [23, 33].

Physics-informed ML aims to include physical relationships such as dynamical equations as a constraint within the optimization process [18]. ML techniques have also been applied to emulate unresolved physical processes such as convection, closures for eddy parameterizations of the ocean, or replacing specific subcomponents of a climate model [43, 44].

In recent years, the advent of super-resolution has significantly advanced the frontiers of image and video processing. Super-resolution attempts to enhance the resolution of an image or video, aspiring to construct a high-resolution output from one or more low-resolution inputs [40]. It is used across a spectrum of applications, ranging from medical imaging and security surveillance to entertainment [19]. Central to super-resolution in ML is the deployment of deep convolutional neural networks [5] (CNNs), generative adversarial networks [25] or diffusion models [12]. These networks, through training on datasets comprising low-resolution and high-resolution image pairs, have demonstrated remarkable proficiency in generating high-resolution images from low-resolution inputs. A blend of super-resolution and image inpainting techniques has shown potential in enhancing outcomes, illustrating the synergy between these approaches. An example of this approach within climate science is the work of Kadow *et al* [17], who apply image inpainting techniques, leveraging deep learning algorithms, to mitigate missing data issues in the HadCRUT4 temperature dataset. However, the utility of ML extends beyond the realms of image resolution enhancement and data gap mitigation and can help with bias correction in numerical climate models [29] or in other related [39] and non-related disciplines [41].

The transfer of information to higher resolutions or smaller scales is referred to as downscaling in climate science and can be divided into two main approaches: statistical and dynamical downscaling [7]. Statistical downscaling relies on empirical relationships between large-scale atmospheric variables and local climate features, producing fine-resolution data by linking observed local data to coarser model outputs [27]. Dynamical downscaling, on the other hand, uses physically based regional climate models to simulate local climate processes based on boundary conditions from global climate models, capturing complex physical interactions on a finer spatial scale [42]. Super-resolution ML models for fluid flow downscaling have been extensively studied and summarized by Fukami *et al* [10]. The use of convolutional neural networks in a multi-scale design, such as their downsampled skip-connection/multi-scale model and U-nets, has been shown to provide superior results compared to simple CNNs [9, 10]. Due to their efficient spatial dimension reduction, U-net architectures have been successfully applied to the reconstruction of turbulent flows, showing accurate results with excellent convergence properties and computational performance [16, 31].

In our work, we integrate high-resolution information obtained by a neural network into a low-resolution configuration of the same model. Our approach augments a coarse, mesh-based algorithm by frequent modifications of the solution by the deep neural net, designed to improve the ‘effective resolution’, that is, correct the coarse solution towards the restriction of a solution computed on a finer mesh. Since this happens at runtime in the time-stepping loop, we call this approach ‘dynamic super-resolution’. We show that this hybrid system is capable of maintaining a balanced flow state and transition to geostrophic turbulence in a similar way as the high-resolution model. This is different from so-called ‘nudging-to-fine’ approaches, that use ML to learn a state-dependent bias correction which is included in the source term of the governing PDE [3, 4].

A related hybrid ML-PDE solver was introduced by Pathak *et al* for a two-dimensional turbulent flow simulation [32], but was based on the fine scale features of a quasi-steady solution, and thus cannot capture transient behavior like the onset of instabilities. Recently, Barthelemy *et al* investigated a similar approach using a convolutional super-resolution ML model to downscale low-resolution data of a quasi-geostrophic test case [2]. In contrast to Barthelemy *et al* we omit the use of an Ensemble Kalman filter to assimilate the output of the ML model. This allows us to focus on the impact of the ML algorithm without interference of the Kalman filter’s error covariance matrices. Our approach shares some similarities with the recent work of Margenberg *et al* [28]. They use a DNN to provide some fine scale information to a multi-grid solver to provide a solution that is more accurate compared to using the coarse grid alone. This avoids the need to store and process fine mesh data after training is finished but, of course, does not produce the actual fine mesh solution. Furthermore, in contrast to their architecture, our neural network is purely convolutional, thus independent of the spatial domain, and can be applied localized as well. As we do not apply the neural network in the context of a geometric multigrid method, the network does not need access to residuals or cell geometries. While Margenberg *et al* focus on reproducing the statistics of turbulence, we study the subtle interplay between balance and loss-of-balance for a geostrophic jet in a deterministic setting.

2. Methodology

The aim of our super-resolution approach is to run a simulation on a coarse mesh while frequently correcting it using a trained ML-model to resemble the restriction of a simulation that was run on a finer

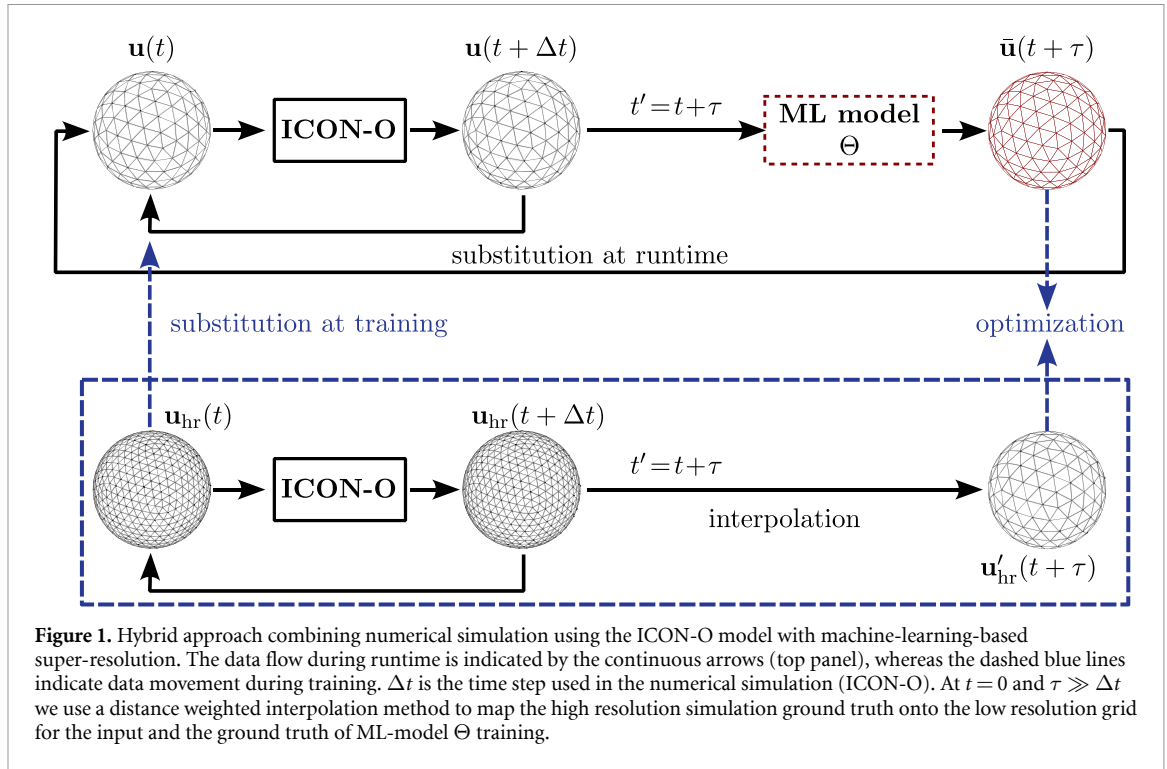


Figure 1. Hybrid approach combining numerical simulation using the ICON-O model with machine-learning-based super-resolution. The data flow during runtime is indicated by the continuous arrows (top panel), whereas the dashed blue lines indicate data movement during training. Δt is the time step used in the numerical simulation (ICON-O). At $t = 0$ and $\tau \gg \Delta t$ we use a distance weighted interpolation method to map the high resolution simulation ground truth onto the low resolution grid for the input and the ground truth of ML-model Θ training.

mesh. This approach does integrate numerical subgrid-scale information into the coarse resolution simulation by correcting the whole state vector. This differs from classical parameterization approaches where specific physical processes are modeled. Our proposed method is visually summarized in figure 1, which shows the interaction between the ICON Ocean model (ICON-O) and the ML model Θ during simulation runtime and ML model training.

During runtime, we use a trained ML model to periodically correct the velocity field \mathbf{u} at time steps $t' = t + \tau$ via

$$\bar{\mathbf{u}}(t + \tau) = \Theta(\mathbf{u}(t + \tau)). \quad (1)$$

The ML model was previously trained according to the scheme shown in the blue box in figure 1. We define the initial condition on a high resolution grid $\mathbf{u}_{\text{hr}}(t)$ and interpolate it to a low resolution grid $\mathbf{u}'_{\text{hr}}(t)$. Then we integrate both states using the model time step Δt until $t' = t + \tau$. We use the output velocity field of the low-resolution simulation $\mathbf{u}(t + \tau)$ to train the ML model using the high-resolution, interpolated velocity field $\mathbf{u}'_{\text{hr}}(t + \tau)$ as ground truth

$$\min_{\theta} \mathcal{L}(\Theta(\mathbf{u}(t + \tau)), \mathbf{u}'_{\text{hr}}(t + \tau)), \quad (2)$$

where θ are the parameters of the ML model Θ and \mathcal{L} the loss function which is minimized during training. In the next sections, we elaborate on the details of our method, the ICON-O model, the ML architecture, its training and the experiment to show the potential of our hybrid approach.

2.1. ICON-O numerical scheme

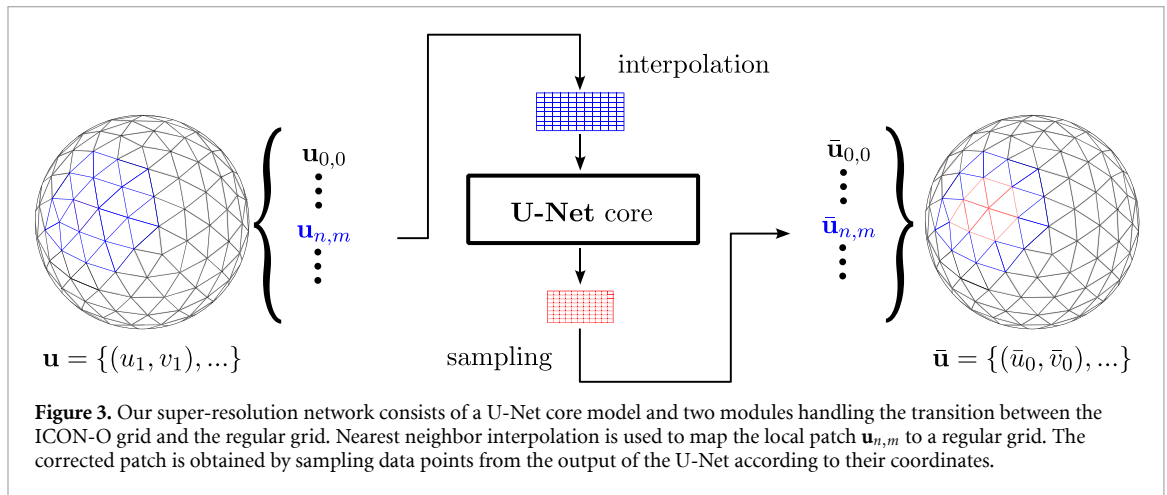
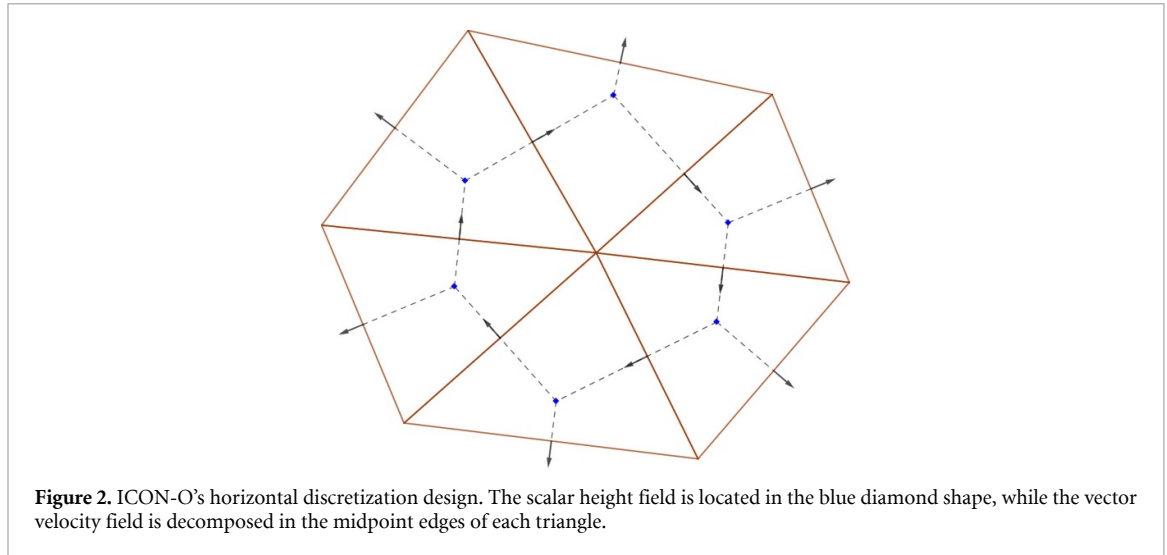
ICON-O is the oceanic component of the Earth system model ICON [14]. It uses the hydrostatic Boussinesq equations on a sphere with a free surface boundary condition and solves these equations on a triangular grid with Arakawa C-staggering as shown in figure 2 [20, 21].

The ocean model equations contain the two-dimensional shallow water equations (SWEs) used in this paper as a special case. The SWE are

$$\partial_t h = \nabla \cdot \mathbf{u}h \quad (3a)$$

$$\partial_t \mathbf{u} = \nabla [g(h + b) + E_k] + (f + \omega) \mathbf{u}^\perp, \quad (3b)$$

where h and $\mathbf{u} = (u, v)$ are scalar height and vector velocity field consisting of zonal (u) and meridional (v) velocity components, respectively. Here, g is the Earth's gravity acceleration, b is the topography, $E_k = |\mathbf{u}|^2/2$



is the kinetic energy, f is the Coriolis parameter, ω is the relative vorticity, and $\mathbf{u}^\perp = \mathbf{k} \times \mathbf{u}$ is the perpendicular velocity, where \mathbf{k} is the unitary vertical vector.

The numerical scheme of ICON-O in its shallow-water configuration satisfies discrete conservation laws for mass, total energy, potential vorticity and potential enstrophy [21]. A study of its accuracy was provided by Lapolli *et al* [24].

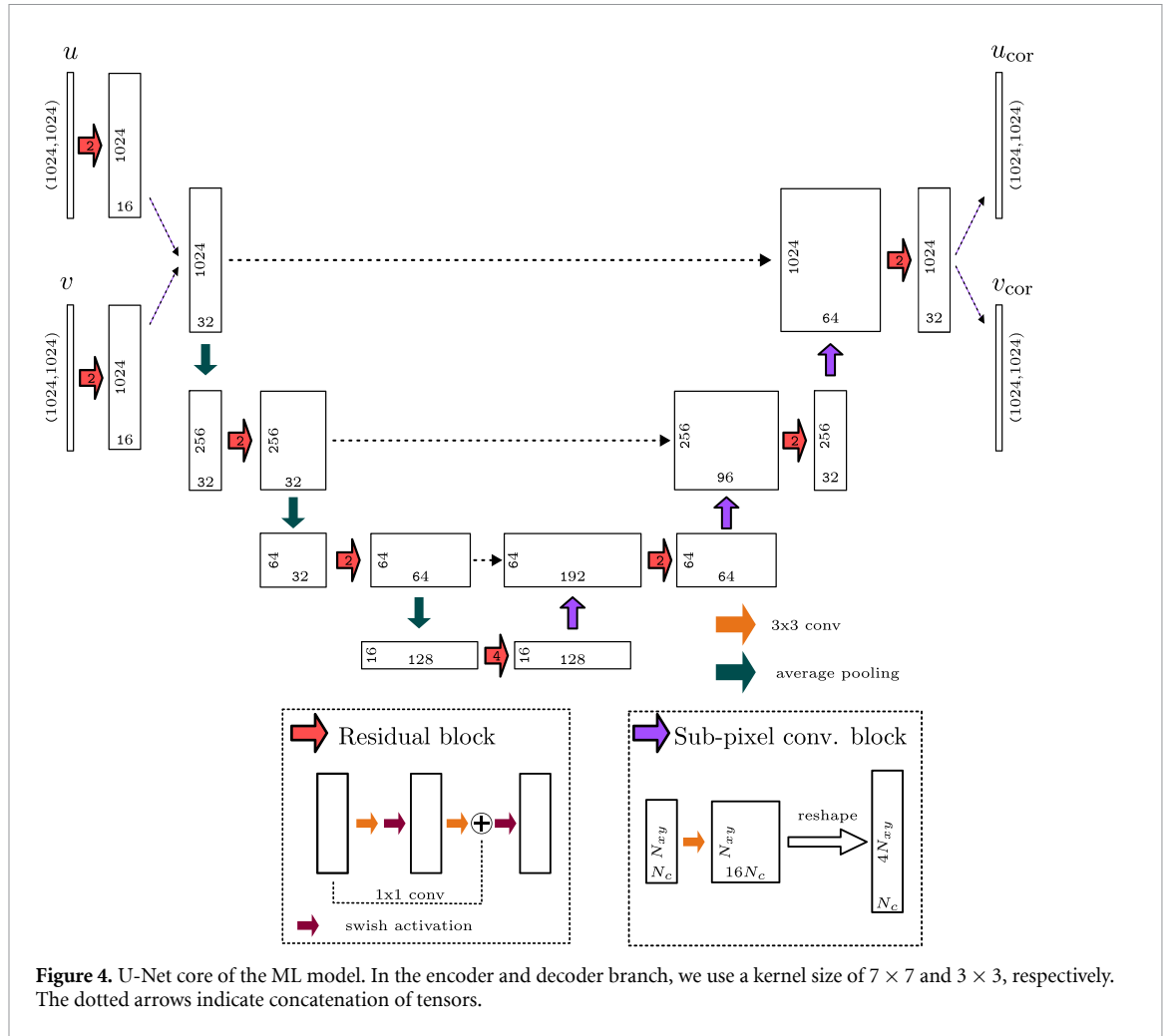
2.2. Network architecture

We assume that our choice of τ is sufficiently small so that the difference between the integrated high-resolution and low-resolution velocity fields at a given location depends only on a local neighborhood. We therefore define the ML model on local domains as shown in figure 3. The advantage of this approach is that the model otherwise tends to overfit to global structures. In addition, model prediction of local patches can be done in parallel, which is particularly useful when the model is running on high performance computing systems.

We define a local patch by $\mathbf{u}_{n,m}$ with $n, m \in \{0, \dots, 7\}$ as a subset of the unstructured grid of ICON-O. We use nearest neighbor interpolation to interpolate the subset to a high-resolution regular grid of 1024×1024 pixels. We use a 10% overlap of the patches during training. The output data $\bar{\mathbf{u}}_{n,m}$ is sampled from the output of the core model based on their relative coordinates.

The core model has a U-net structure [35] with skip connections to take advantage of the computational efficiency of convolutions in image to image tasks. We define the convolutional layers of the U-net without biases. As a consequence, the trivial field correction is itself a trivial field $\Theta(\mathbf{0}) = \mathbf{0}$.

The detailed structure of the U-net is shown in figure 4. The basic element of our U-Net consists of two consecutive residual blocks, or ResNet blocks [13], with a swish activation function [8] and a 1×1 convolutional layer for the residual connection. We use average pooling layers with a kernel size of 4 to effectively decrease the spatial dimension.



Upsampling, i.e. increasing the spatial dimension in the decoder part of the U-Net, is realized by sub-pixel convolutional blocks [37]. These consist of a convolutional layer, which increases the channel dimension by a factor of r^2 , where r is the factor by which the resolution is to be increased in both spatial dimensions. A consecutive pixel-shuffle operation rearranges the tensor with a shape of $(r^2 N_c, N_x, N_y)$ such that the channel dimension is decreased by a factor of $1/r^2$, whereas both spatial dimensions are increased by a factor of r . We use the kernel weight initialization proposed by Aitken [1] to ensure the absence of checkerboard artifacts.

The output of the encoder branch is concatenated with the respective upsampled output of the decoder layer and fed into the next layer. Note that for the input and output, we are using grouped convolutions to efficiently learn input and output features for both velocity components.

3. Experiment design

We analyze the performance of our approach for the Galewsky testcase [11]. This testcase describes the transition from a geostrophically balanced steady state to a barotropic instability. This transition is triggered by a small perturbation that is added to the balanced initial condition and that needs several days to build up. The instability of barotropic jets is a classical and well studied process in geophysical fluid dynamics that occurs in both the atmosphere and the ocean.

The initial condition is defined as follows

$$u(\phi) = \begin{cases} 0 & \phi \leq \phi_0 \\ u_{\max}/e_n \cdot \exp(1/[(\phi - \phi_0)(\phi - \phi_1)]) & \phi_0 \leq \phi \leq \phi_1 \\ 0 & \phi \geq \phi_1, \end{cases} \quad (4)$$

where $u_{\max} = 80 \text{ ms}^{-1}$, $e_n = \exp(-4/(\phi_1 - \phi_0)^2)$, $\phi_0 = \pi/7 - n_\phi\pi/36$, $n_\phi \in \{-1, 0, 1\}$, and $\phi_1 = \pi/2 - \phi_0$. This is balanced by the height field

$$gh(\phi) = gh_0 - \int^{\phi} au(\phi') \left[f + \frac{\tan \phi'}{a} u(\phi') \right] d\phi', \quad (5)$$

where $a = 6.371 \times 10^6 \text{ m}$ is the radius of the Earth and the resultant integral is a function of the latitude ϕ . Here, the integral is unfeasible to calculate analytically, therefore, we chose to compute it using Romberg's quadrature, from the bounds of 90° S , where $u = 0$ to ϕ .

To trigger the instability and development of turbulence we add the perturbation

$$h'(\lambda, \phi) = \hat{h} \exp\left(-[(\lambda_2 - \lambda)/\alpha]^2\right) \exp\left(-[(\phi_2 - \phi)/\beta]^2\right) \cos \phi, \quad (6)$$

where λ is the longitude, and $\lambda_2 = n_\lambda\pi/4 - \pi$, $n_\lambda \in \{0, \dots, 7\}$, $\phi_2 = \pi/4 - n_\phi\pi/36$, $\alpha = 1/3$, $\beta = 1/15$, and $\hat{h} = 120 \text{ m}$.

3.1. Training of the ML model

We employ the training scheme shown in figure 1 (blue box). Starting from the same initial conditions, we integrate the SWEs (3a) and (3b) on both, the high-resolution (2.5 km) and the low-resolution (20 km) grid. At $t = \tau = 12 \text{ h}$ we generate outputs from both grids $\mathbf{u}_{\text{hr}}(t + \tau)$ and $\mathbf{u}_{\text{r}}(t + \tau)$. We obtain the ground truth $\mathbf{u}'(t + \tau)$ by interpolating the 2.5 km output to the 20 km grid using distance weighted interpolation *Climate Data Operators* (CDO) [36]. We normalize both velocity components by the 95% quantile of their magnitudes calculated on the training data distribution. Note that the sign of both components is preserved. After forward feeding the ML model with the low resolution output $\bar{\mathbf{u}} = \Theta(\mathbf{u}_{\text{r}}(t + \tau))$, we calculate gradients based on the loss

$$\mathcal{L} = \mathcal{L}_{\text{abs}} + \gamma \mathcal{L}_{\text{rel}} \quad (7)$$

where

$$\mathcal{L}_{\text{abs}} = \frac{1}{N} \sum_i^N (|\bar{u}_i - u'_i|)^2 + (|\bar{v}_i - v'_i|)^2 \quad (8)$$

and

$$\mathcal{L}_{\text{rel}} = \frac{1}{N} \sum_i^N \left(\min \left[1, \frac{|\bar{u}_i - u'_i|}{u'_i + \epsilon} \right] + \min \left[1, \frac{|\bar{v}_i - v'_i|}{v'_i + \epsilon} \right] \right). \quad (9)$$

Here, u'_i, v'_i and \bar{u}_i, \bar{v}_i denote the i th element of the zonal and the meridional velocity components and $\epsilon = 10^{-12}$ an offset to avoid division by zero. We chose a squared error loss in the calculation of equation (8) as an absolute error measure to penalize large valued errors. In addition, we add a non-squared relative error loss to equation (7) to account for the imbalance between both velocity components. We chose a scaling factor of $\gamma = 0.1$ such that the initial magnitudes of both loss terms are approximately the same. Especially at early stages of the experiment, v is orders of magnitude lower than u but plays a crucial role in disturbing the geostrophic flow. To stabilize the training process, we restrict maximum relative errors to 1 (equation (9)). Note that we use normalized dimensionless quantities in the calculations of equations (8) and (9).

The training data set of the ML model is provided by running variations of the original Galewsky test by modifying the location of the jet equation (4) by setting $n_\phi \in \{-1, 1\}$, and the location of the perturbation equation (6) by using $n_\lambda \in \{0, \dots, 7\}$. For validation, we use the original Galewsky test case with $n_\phi = n_\lambda = 0$. The choice for training and validation data is summarized in table 1. We run the simulations over 20 d with a timestep of 10 s and an output frequency of 12 h. The total number of unique patches, which are randomly sampled, amounts to $N = 32(\text{patches}) \cdot 1184(\text{snapshots}) = 37888$ and 2368 for training and validation, respectively. We train the model for 50 000 mini-batch iterations with a batch size of 32 on $4 \times \text{A100 GPUs}$ using a decaying learning rate with an initial learning rate of 10^{-4} (figure 5).

3.2. ML correction during runtime

Given the trained ML model, we are now running ICON-O coupled with the ML model (figure 1, top panel only). We again use the original Galewsky test case with $n_\lambda = n_\phi = 0$, a model time step of 10 s and a correction frequency of 12 h.

Table 1. Simulation setups for generating the training and validation data. We use a 12 h output frequency and a maximum simulation time of 20 d.

	Jet location n_ϕ	Perturbation location n_λ	Total number of snapshots
Training	$\{-1, 1\}$	$\{0, \dots, 7\}$	1184
Validation	0	0	37

For evaluating the performance of the ML coupling with ICON, we compute the accuracy error norms for the simulated fields, i.e. velocity and height. We compute errors in both L_2 - and L_{\max} -norm, as they will provide information about the integrated error as well as local errors. These norms are defined as:

$$L_2 = \sqrt{\frac{S \left[\left(f^{[n]} - f^{[r]} \right)^2 \right]}{S \left[\left(f^{[r]} \right)^2 \right]}}, \quad (10a)$$

$$L_{\max} = \frac{\max_j \left| f_j^{[n]} - f_j^{[r]} \right|}{\max_j \left| f_j^{[r]} \right|}, \quad (10b)$$

where $f_j^{[n]}$ and $f_j^{[r]}$ is the reference (high resolution) and coarse fields at cell j , respectively. S is defined as

$$S[f] = \frac{\sum_j A_j f_j}{\sum_j A_j}$$

where A_j is the area of the j th cell. The simulated field is the low resolution (LR 20 km) simulated outputs, while the reference solution is the interpolated high resolution (HR 2.5 km) to LR simulated outputs. For a proper comparison, we interpolate the HR outputs with a distance weighted interpolation from the CDO software.

Additionally, to evaluate the conservation properties of our schemes we track the energy over time. The calculations of the kinetic and potential energy are done by the following formulae

$$\bar{E}_k = \sum_j \frac{|\mathbf{u}_j|^2}{2} A_j, \quad \text{and} \quad \bar{E}_p = g \sum_j h_j A_j.$$

Another important metric to evaluate is the transfer of energy between different scales within the simulation, which for the 2D fluid simulation follows well known scaling laws (see, e.g. [22]). We evaluate this metric analyzing the spectra of kinetic energy and enstrophy. This is done by decomposing the velocity in its divergence and vorticity components, and, for each component, expanding them in their respective spherical harmonic components

$$\psi = \sum_m \sum_n \psi_n^m P_n^m(\phi) e^{im\lambda}, \quad (11)$$

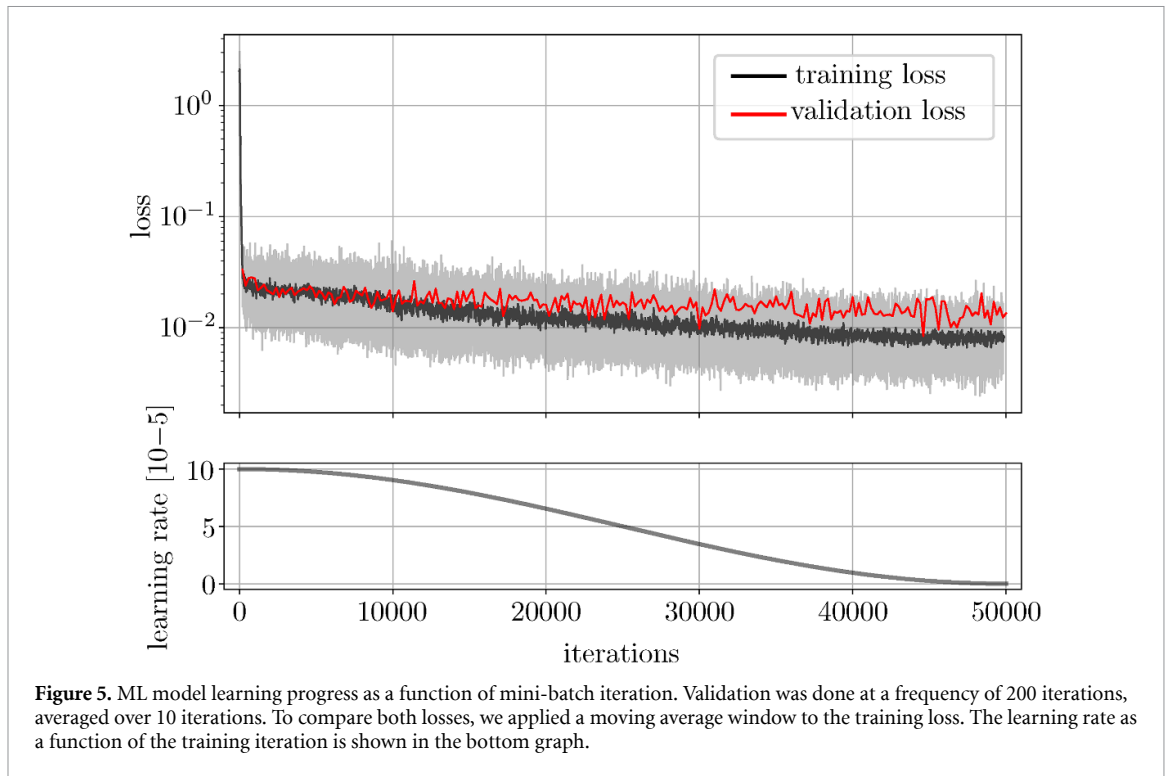
where ψ is either divergence or curl of the velocity, ψ_n^m are the coefficients, and P_n^m are the associated Legendre functions. The spectra are given by the coefficients ψ_n^m , which can be obtained, analogously to the Fourier transform, as

$$\psi_n^m = \int_{-1}^1 \int_0^{2\pi} \psi e^{-im\lambda} d\lambda P_n^m(\phi) d\phi. \quad (12)$$

In order to calculate the coefficients, we use nearest neighbor interpolation to a Gaussian Grid [15] and Fourier transform in the zonal component. We then compute the meridional component by exploiting the zeroes of the Legendre polynomial using the Gaussian quadrature method. Finally, the spectra are obtained by the following equations

$$\text{KE}_n = \frac{a^2}{n(n+1)} \left[|\zeta_n^0|^2 + \sum_m |\zeta_n^m|^2 + |\delta_n^0|^2 + \sum_m |\delta_n^m|^2 \right], \quad (13)$$

$$\text{EN}_n = |\zeta_n^0|^2 + \sum_m |\zeta_n^m|^2, \quad (14)$$



where KE_n and EN_n are the kinetic energy and enstrophy spectra; ζ_n^m and δ_n^m are the spectra coefficients of the vorticity and divergence obtained from equation (12) and a is the radius of the earth.

4. Numerical accuracy

We first analyze the performance of the ML model correcting the low resolution Galewsky test case that periodically has been initialized by a high resolution run as shown in figure 1. This serves as a best case for the ML correction performance, as the re-initialization from high resolution avoids error propagation. We then compare these results to the coupled run, where the output of the ML is used for initialization figure 1(top panel only).

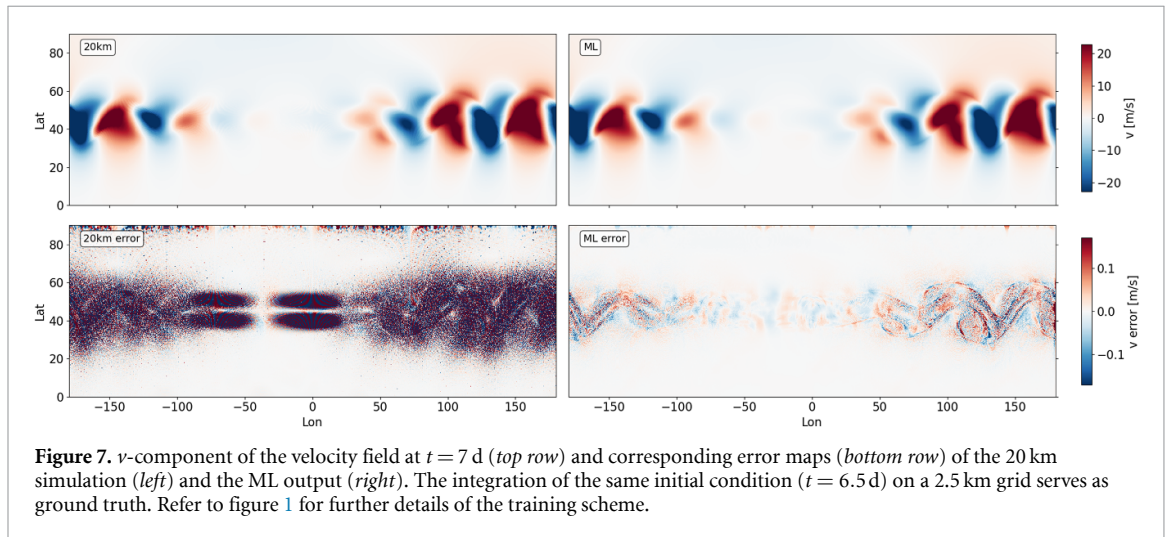
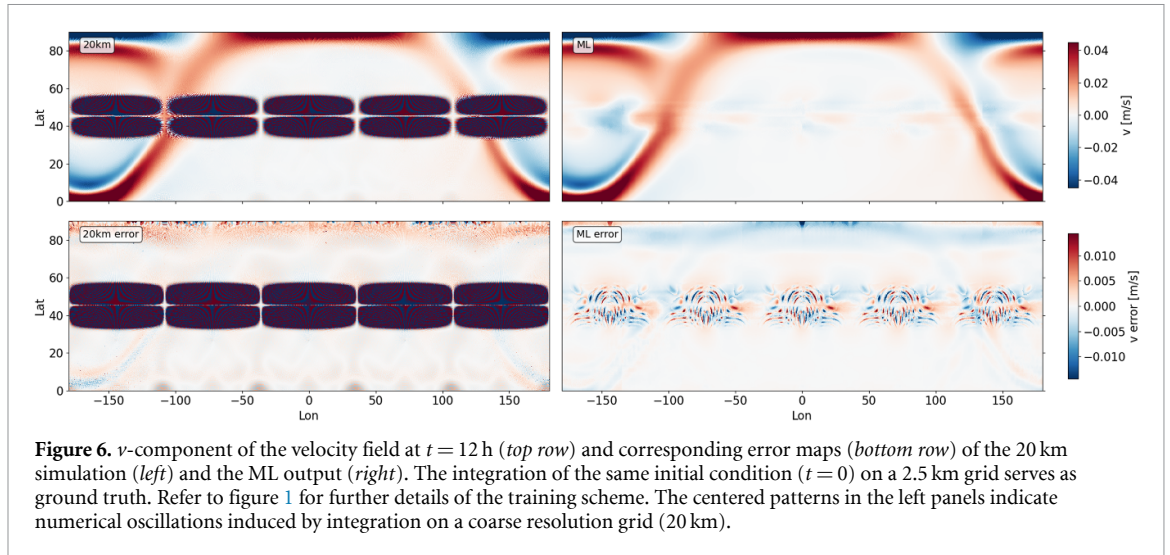
4.1. ML training

The training progress is shown in figure 5. The presence of large fluctuations in the training loss is because of the random sampling of local patches. Even though we restrict maximum losses to values of 1, low amplitudes of v result in a high contribution to the relative error equation (9). To avoid overfitting of the data, we stopped the optimization after 50 000 mini-batch iterations.

Figures 6 and 7 show the global snapshots of the uncorrected ML input and the ML output of the v -component of the velocity field \mathbf{u} at $t = 12\text{h}$ and $t = 7\text{d}$. The ML model is able to remove the numerical oscillations induced by the coarse resolution at both time points, with v ranging over several orders of magnitude. At early time steps, these numerical oscillations are particularly pronounced around the center of the jet, where large, high-frequency patterns appear in the left panels of figure 6. The global output of the ML model is smooth and shows no pattern to indicate local patches. Thus, our ML model is robust to small variations of the input data in neighboring patches.

4.2. Numerical accuracy

The uncoupled ICON simulation shows a small error in both norms for all fields, quickly increasing from day 4 to day 8, where it then saturates until the end of the computational simulation, see figure 8. For both L_2 and L_{\max} norms, the errors of the uncoupled simulations for the last day of integration are found in table 2. Note that, since the meridional component of the velocity is zero at the initial state, the v -component is not normalized as the other fields. Regardless, we observe a near first order accuracy of these fields on both norms, which is expected from the model. Additionally, it can be observed that the order of accuracy is slightly lower for the maximum norm. It may be the case that the testcase we are evaluating is prone to numerical oscillations or that the reference simulation of 2.5 km is not fine enough to achieve first order accuracy in the maximum norm.



In contrast, the coupled ICON/ML run has its first correction, displaying a significantly improved error in comparison to the uncoupled 20 km. The error of the velocity components, although displaying substantially lower values than the uncoupled 10 km, quickly increases as the field is integrated in time. This increase is also reflected in the height field, showing an increase in the error slightly larger than the uncoupled 20 km simulation. Despite this, as the model is integrated in time, the overall fields' accuracies tend to the order of the uncoupled 10 km simulation. It is also observed that for some days in the later stages of the simulation, the accuracy of the coupled run overtakes even the uncoupled 10 km, especially in the L_2 norm of the u and v velocity components. Additionally, despite the height field not being directly affected by the ML, it is impacted by the corrections performed in the velocity field. The observed deterioration of the error in the early stages of the simulation is driven by the velocity field corrections. However, as the velocity accuracy improves over time, it adjusts the height field to a lower error, improving its accuracy.

The improvement in the domain becomes particularly evident when examining the vorticity field in the 7th day of simulation. Here, we observe the early stages of the triggering of the fluid instability. We see the vorticity field generating vortices as a consequence of the instability, with these vortices being transported eastward (see figure 9). These vortical formations generate tight gradients which, besides ICON's own numerical noise, are also a source of spurious oscillations [11]. For coarser simulations, these oscillations are greater and the instability is shown to be more developed away from the initial region of perturbation, visible in the uncoupled 20 km simulation.

This illustrates the effects of the grid resolution in the fluid simulation. From this perspective, it is noticeable that the coupled model accurately portrays the onset of instability and delays local instabilities generated by the interaction between the spurious oscillations of both the unstructured nature of the grid and tight gradients by the testcase. However, the ML correction add its own set of artifacts, see figure 10, which is likely playing a role in the accuracy of the model. Therefore, we may interpret that the difficulty to

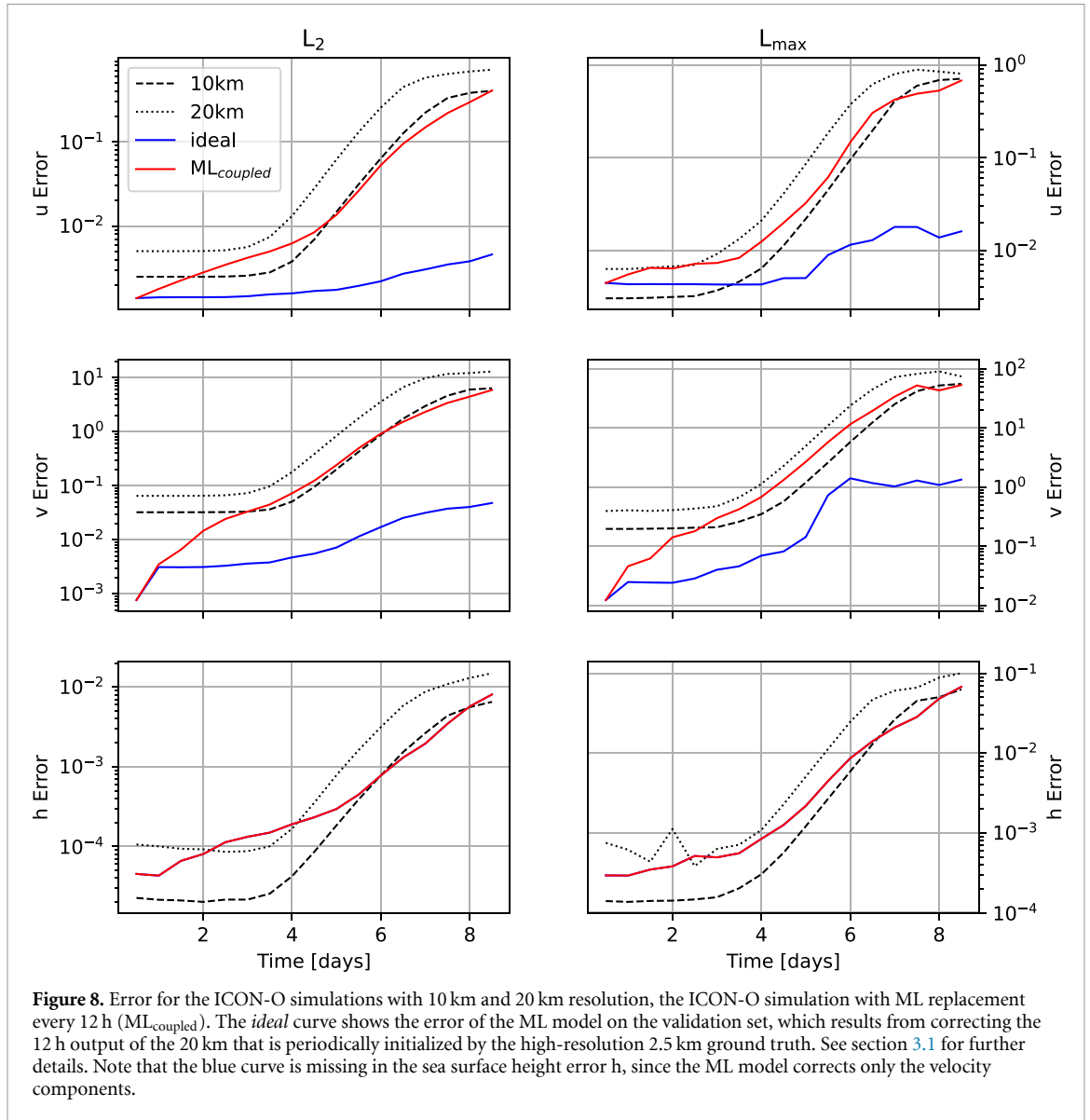


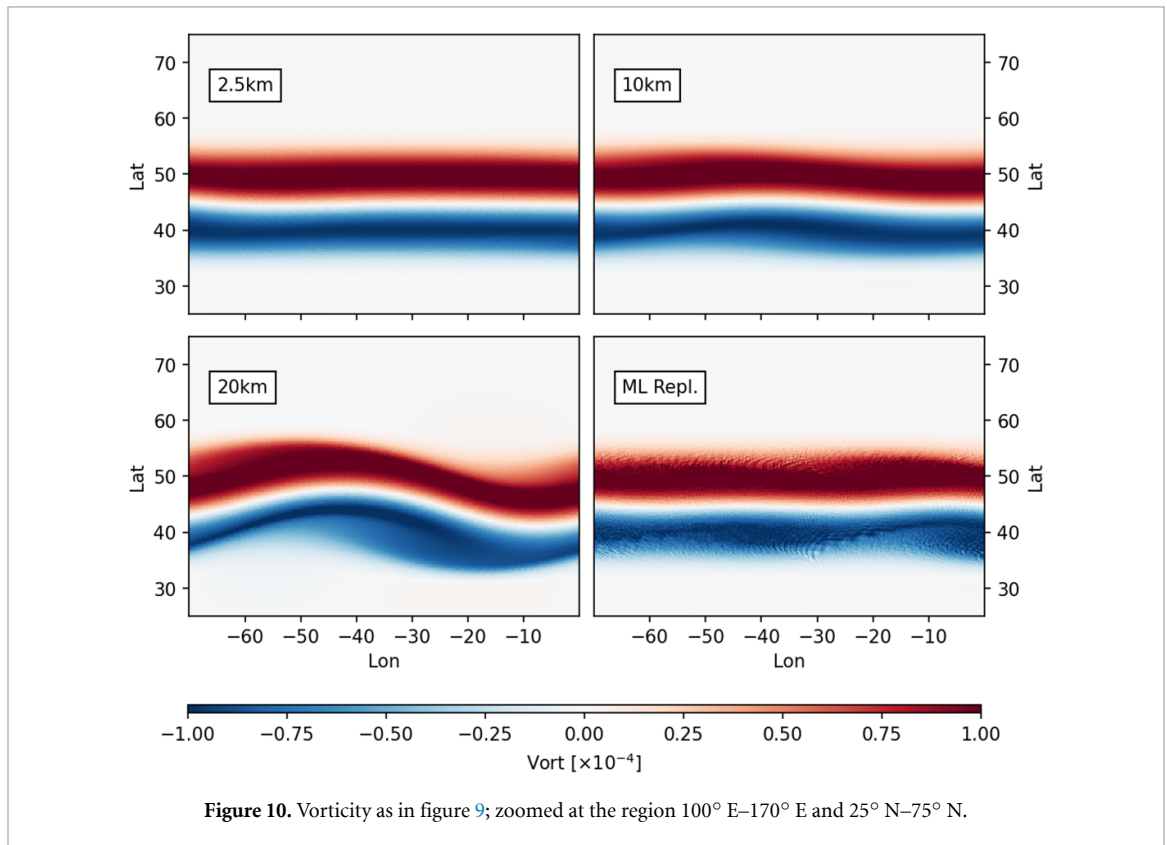
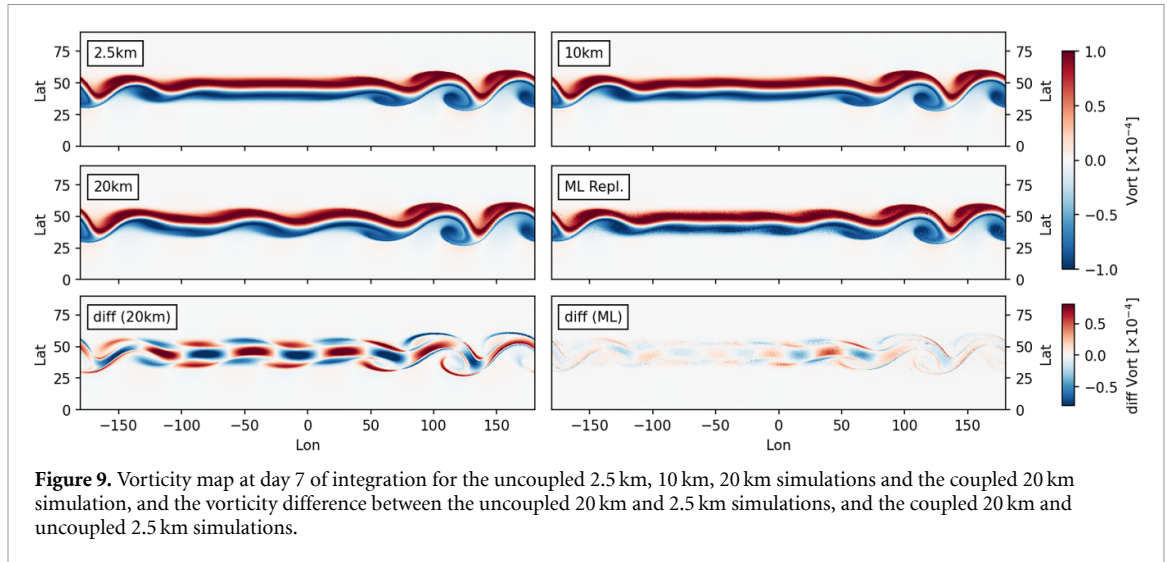
Table 2. Errors of the uncoupled simulation between 2.5 km and 10 km/20 km resolutions for the different fields at the last day of integration.

	L_2			L_{max}		
	10 km	$ML_{coupled}$	20 km	10 km	$ML_{coupled}$	20 km
u	4.40×10^{-1}	4.43×10^{-1}	7.20×10^{-1}	0.78	0.76	1.25
v	7.47	7.58	13.99	58.79	73.19	76.59
h	0.81×10^{-2}	0.84×10^{-2}	1.01×10^{-2}	6.78×10^{-2}	7.10×10^{-2}	9.23×10^{-2}

maintain low error in the first few days is a consequence of these artifacts, especially in the early stages of the simulation. As a consequence, due to the lack of diffusivity in the model, these noises do not disappear, and they are still present in the solution and integrated in time. These oscillations, although being visible, do not amplify nor seem to meaningfully disrupt the solution of the integrated fields, since we can achieve a 10 km-order-error simulation.

4.3. Energy and enstrophy

Despite the improvement, it is important to understand if this artificial impact of the coupled model, especially concerning its numerical oscillations, may disrupt the energy pathway of the system. For this purpose we consider the energy and enstrophy spectra. From the seminal work by Kraichnan [22] the fluid can contain two main inertial ranges with different slopes in the shallow water regime: inverse energy and forward enstrophy cascades with a dissipation rate of $-5/3$ and -3 , respectively. We investigate these spectra



although they apply to developed turbulence only while the barotropic instability covers the transition to turbulence, because a potential artificial injection of energy or enstrophy due to the coupled model should show up in the spectra. In this case the spectra would also reveal if this injection occurs at small or large scales. Due to the local nature of the ML correction we expect changes to appear in the small scales. For a qualitative assessment whether ICON-O reproduces the theoretical benchmark of the energy and enstrophy spectra the experiment is not appropriate. The results of the Galewsky experiment at day 20 display the -3 slope through most of the wavenumber space, while the $-5/3$ slope is likely restricted to the low wavenumbers only [26].

In our experiments, in day 7, prior to the full turbulence development, we observe a spectrum for the energy of a slope steeper than k^{-3} (figure 11 upper panel). In our analysis, the slope of the spectra is well represented in all models down to the 300 km wavelength. Towards finer scales, the coupled ML run starts to deviate, displaying an increase in energy, while the uncoupled 20 km simulation still maintains the slope along with reference solution. The deviation of the uncoupled 20 km simulation is only detected at index

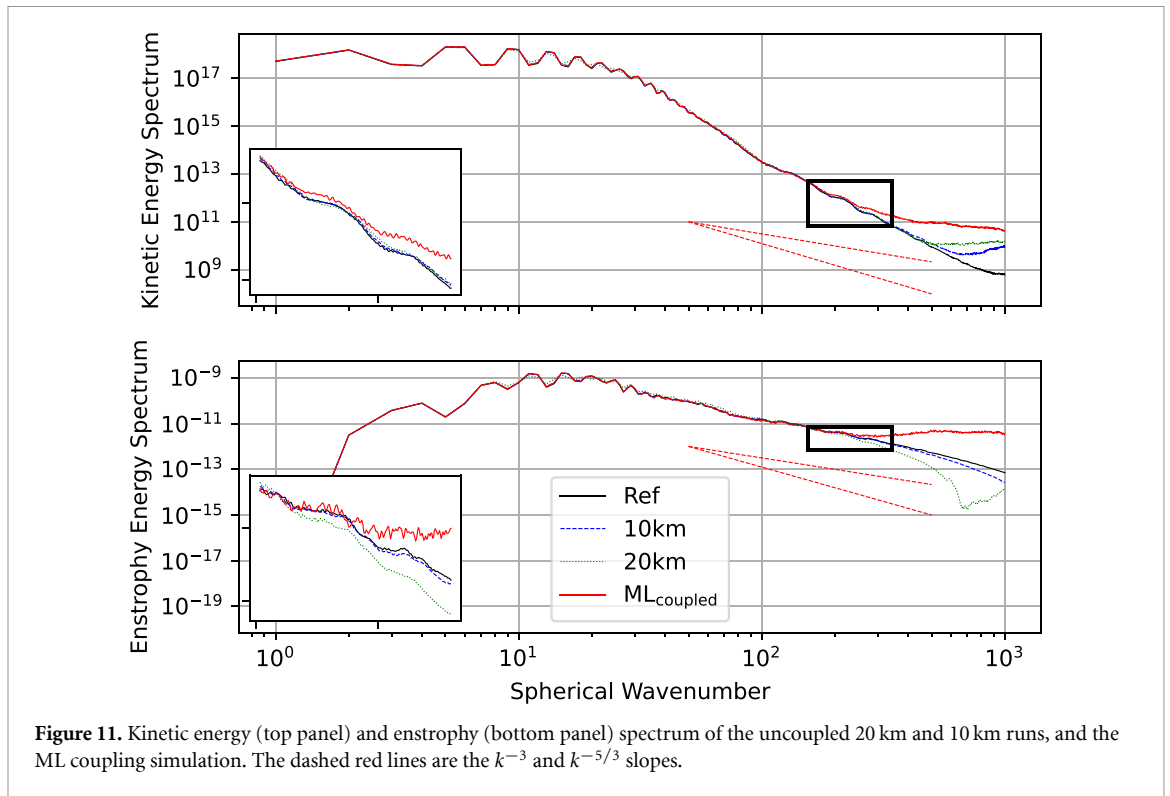


Figure 11. Kinetic energy (top panel) and enstrophy (bottom panel) spectrum of the uncoupled 20 km and 10 km runs, and the ML coupling simulation. The dashed red lines are the k^{-3} and $k^{-5/3}$ slopes.

approximately 80 km wavelength. This reinforces the argument that the ML model is predominantly acting in the smaller scales.

Similarly, the enstrophy spectrum, see figure 11, (lower panel), displays a similar enstrophy cascade up to the 200 km resolution. Around these scales, the uncoupled 20 km simulation is the first to deviate from the theory, followed by the coupled 20 km simulation. In these scales, the uncoupled model loses substantially its enstrophy in higher wavenumbers, while the coupled simulation display an increase for higher resolutions. This indicates that the ML is adding energy into the smaller scales, likely in the form of the previously observed spurious oscillations. Due to the non-dissipative nature of our system, the high wavenumber regime is not dissipated into mixing and, therefore, remains in the domain. Although in a more complex model, these oscillations near grid scale induce spurious mixing of tracers in the ocean [6], these can be solved by an appropriate choice of low pass filter, such as a higher order dissipative operator to dim these higher wavenumber waves.

In order to verify if this increased energy in higher wavenumbers has an overall effect in the energy conservation of our scheme, we evaluate the total energy time evolution (figure 12). In the coupled run, it is observed that the kinetic energy of the system increases up to around day 6 and decreases in the subsequent days. Conversely, both uncoupled simulations show a relatively stable integrated kinetic energy up to day 5, with a slight uptick preceding the instability, followed by a consistent decline in energy, thereafter. The integrated energy variation of the uncoupled run is about as half of the coupled runs, indicating that the ML is inputting energy in the domain. As observed in the energy spectrum, this input of energy is in the form of low resolution (high wavenumbers) waves, likely generated by the spurious oscillations of the ML correction.

4.4. Runtimes

We also measured runtimes of the respective methods for $t = 12$ h of integration, subtracting timing for data input and output, for 32 Nodes on the CPU partition of the Levante HPC system (German Climate Computing Center). While the 10 km run took 63 s and the 20 km simulation 31 s, the ML-coupled version was done in 55 s. Hence, we observe a speedup of 14.5% from the ML-coupled simulation compared to the 10 km run. However, since the focus in this paper was on demonstrating the accuracy of dynamic super-resolution, the ML-corrected model is mostly untuned. In particular, only CPUs were used since we used a shared partition for the ICON-O simulation and the ML model. Moving the ML correction to GPUs and optimizing data transfer would require refactoring of ICON-O but could also greatly improve speedup.

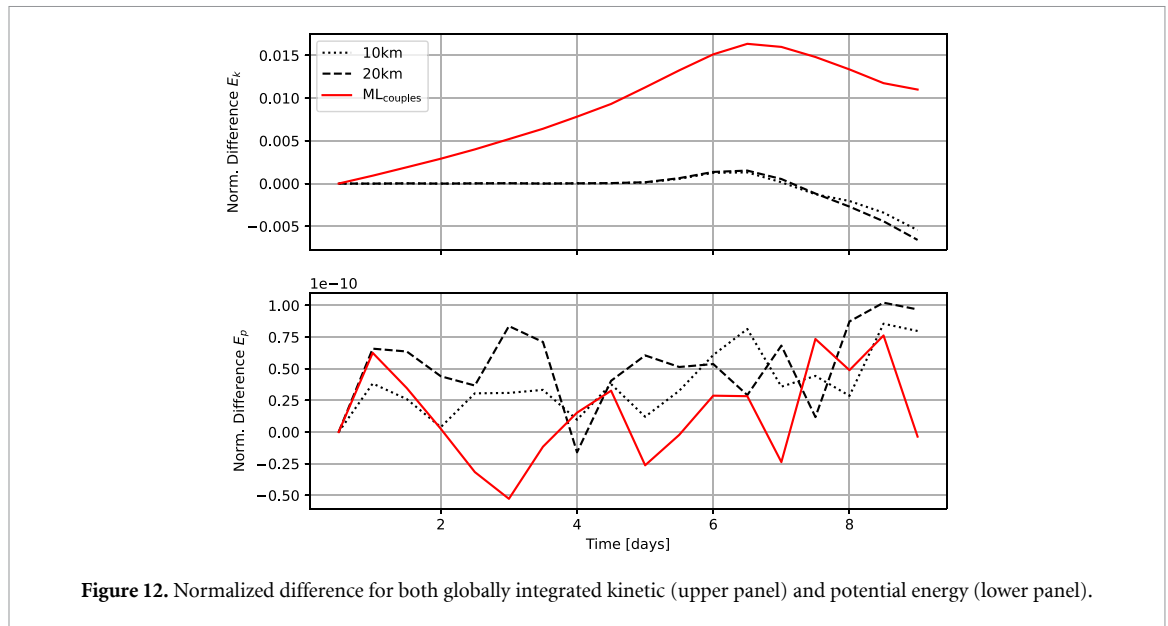


Figure 12. Normalized difference for both globally integrated kinetic (upper panel) and potential energy (lower panel).

5. Conclusions

We present a hybrid modeling approach, combining a mesh based model with an ML correction to incorporate the net impacts from finer scales, unresolved dynamics. Because the simulation is time-dependent and the correction is applied after a set number of time steps have been processed, we call the approach dynamic super-resolution. We demonstrate that the correction allows us to run a global numerical simulation of the SWEs on a twice coarser mesh without loss of accuracy. The continuity between adjacent predictions in the global output confirms the reliability of employing a local ML model approach to correct a global velocity field. As our ML model operates solely through convolutions, adapting it to large-scale fields is straightforward by adjusting the patch size to fit to available computational resources. Employing independent patches enables parallelization of ML predictions for efficient corrections. We emphasize that our approach is based on using a single test case and a single set of parameters in terms of resolution difference and temporal range τ . Both the resolution difference and the temporal range will determine the correlation between the low-resolution and the high-resolution sample, which will directly affect the accuracy of the ML model and our general approach.

The trained ML model effectively reduces the discretization errors of the integration of the SWEs. Additionally, using relative errors in optimizing the parameters of the model ensures consistent model performance across different scales, which is crucial to preserve the stability of the jet of the Galewsky test case. However, the error increases when the output of the ML model is integrated in time in the coupled run. The remaining artifacts produced by the ML model still affect the solution and likely lower the accuracy of the coupling. Additionally, the ML model acts as a source of kinetic energy. Although it does not push the solution to numerical instability, this could become an issue for more complex regimes. Lastly, despite the direct cascade of both energy and enstrophy being respected for large scales, they are affected in the high wavenumbers, likely due to the noise produced by the ML.

In future work, we plan to integrate physical constraints into the ML model to help conserve physical properties in our hybrid approach. In addition, training objectives penalizing artifacts and forcing convergence of the ML correction will help to enhance the stability integrating the ML output in time. Incorporating statistical ML approaches to estimate model uncertainty may offer further insights into stabilizing the ML correction process. We observed a speedup of the coupled simulation compared to a similarly accurate simulation (10 km), even though we only used CPU resources for inference. This shows the potential of dynamic super-resolution to reduce wallclock times of the dynamical cores of earth system models while maintaining high accuracy.








Data availability statement

The data that support the findings of this study are openly available at the following URL/DOI: <https://github.com/FREVA-CLINT/climatereconstructionAI/tree/ICON-CRAI-SR>; <https://gitlab.dkrz.de/icon/icon-model>.

Acknowledgments

This project received funding from the German Federal Ministry of Education and Research (BMBF) under grant 16ME0679K. Supported by the European Union–NextGenerationEU. This work used resources of the Deutsches Klimarechenzentrum (DKRZ) granted by its Scientific Steering Committee (WLA) under project ID bk1372.

ORCID iDs

Maximilian Witte  <https://orcid.org/0000-0002-2923-8907>
Fabrício R Lapolli  <https://orcid.org/0000-0003-2444-3861>
Philip Freese  <https://orcid.org/0000-0002-9838-6321>
Sebastian Götschel  <https://orcid.org/0000-0003-0287-2120>
Daniel Ruprecht  <https://orcid.org/0000-0003-1904-2473>
Peter Korn  <https://orcid.org/0000-0002-7525-5732>
Christopher Kadow  <https://orcid.org/0000-0001-6537-3690>

References

- [1] Aitken A, Ledig C, Theis L, Caballero J, Wang Z and Shi W 2017 Checkerboard artifact free sub-pixel convolution: a note on sub-pixel convolution, resize convolution and convolution resize (arXiv:1707.02937)
- [2] Barthélémy S, Brajard J, Bertino L and Counillon F 2022 Super-resolution data assimilation *Ocean Dyn.* **72** 661–78
- [3] Bretherton C S, Henn B, Kwa A, Brenowitz N D, Watt-Meyer O, McGibbon J, Perkins W A, Clark S K and Harris L 2022 Correcting coarse-grid weather and climate models by machine learning from global storm-resolving simulations *J. Adv. Model. Earth Syst.* **14** e2021MS002794
- [4] Clark S K, Brenowitz N D, Henn B, Kwa A, McGibbon J, Perkins W A, Watt-Meyer O, Bretherton C S and Harris L M 2022 Correcting a 200 km resolution climate model in multiple climates by machine learning from 25 km resolution simulations *J. Adv. Model. Earth Syst.* **14** e2022MS003219
- [5] Dong C, Loy C C, He K and Tang X 2014 Learning a deep convolutional network for image super-resolution *Computer Vision – ECVV 2014* ed D Fleet, T Pajdla, B Schiele and T Tuytelaars (Springer) pp 184–99
- [6] Ducouso N, Le Sommer J, Molines J-M and Bell M 2017 Impact of the “Symmetric Instability of the Computational Kind” at mesoscale- and submesoscale-permitting resolutions *Ocean Modelling* **120** 18–26
- [7] Ekström M, Grose M R and Whetton P H 2015 An appraisal of downscaling methods used in climate change research *Wiley Interdiscip. Rev. Clim. Change* **6** 301–19
- [8] Elfving S, Uchibe E and Doya K 2018 Sigmoid-weighted linear units for neural network function approximation in reinforcement learning *Neural Netw.* **107** 3–11
- [9] Fukami K, Fukagata K and Taira K 2019 Super-resolution reconstruction of turbulent flows with machine learning *J. Fluid Mech.* **870** 106–20
- [10] Fukami K, Fukagata K and Taira K 2023 Super-resolution analysis via machine learning: a survey for fluid flows *Theor. Comput. Fluid Dyn.* **37** 421–44
- [11] Galewsky J, Scott R K, Polvani L M 2004 An initial-value problem for testing numerical models of the global shallow-water equations *Tellus A* **56** 429–40
- [12] Gao S, Liu X, Zeng B, Xu S, Li Y, Luo X, Liu J, Zhen X and Zhang B 2023 Implicit diffusion models for continuous super-resolution *2023 IEEE/CVF Conf. on Computer Vision and Pattern Recognition (CVPR)* pp 10021–30
- [13] He K, Zhang X, Ren S and Sun J 2016 Deep residual learning for image recognition *2016 IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)* pp 770–8
- [14] Hohenegger C et al 2023 ICON-Sapphire: simulating the components of the Earth system and their interactions at kilometer and subkilometer scales *Geosci. Model Dev.* **16** 779–811
- [15] Hortal M and Simmons A J 1991 Use of reduced gaussian grids in spectral models *Mon. Weather Rev.* **119** 1057–74
- [16] Jiang C M, Esmailzadeh S, Azizzadenesheli K, Kashinath K, Mustafa M, Tchelepi H A, Marcus P, Prabhat K and Anandkumar A 2020 Meshfreeflownet: a physics-constrained deep continuous space-time super-resolution framework *Proc. Int. Conf. for High Performance Computing, Networking, Storage and Analysis, SC '20* (IEEE Press)
- [17] Kadow C, Hall D M and Ulbrich U 2020 Artificial intelligence reconstructs missing climate information *Nat. Geosci.* **13** 408–13
- [18] Kashinath K et al 2021 Physics-informed machine learning: case studies for weather and climate modelling *Phil. Trans. R. Soc. A* **379** 20200093
- [19] Khoo J J D, Lim K H and Phang J T S 2020 A review on deep learning super resolution techniques *2020 IEEE 8th Conf. on Systems, Process and Control (ICSPC)* pp 134–9
- [20] Korn P et al 2022 ICON-O: the ocean component of the ICON Earth system model—global simulation characteristics and local telescoping capability *J. Adv. Model. Earth Syst.* **14** 10
- [21] Korn P and Linardakis L 2018 A conservative discretization of the shallow-water equations on triangular grids *J. Comput. Phys.* **375** 871–900
- [22] Kraichnan R H 1967 Inertial ranges in two-dimensional turbulence *Phys. Fluids* **10** 1417–23
- [23] Lam R et al 2023 Learning skillful medium-range global weather forecasting *Science* **382** 1416–21
- [24] Lapolli F R, da Silva Peixoto P and Korn P 2024 Accuracy and stability analysis of horizontal discretizations used in unstructured grid ocean models *Ocean Modelling* **189** 102335
- [25] Ledig C et al 2017 Photo-realistic single image super-resolution using a generative adversarial network *2017 IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)* (IEEE Computer Society) pp 105–14
- [26] Lee D, Martín A F, Bladwell C and Badia S 2024 A comparison of variational upwinding schemes for geophysical fluids and their application to potential enstrophy conserving discretisations *Comput. Math. Appl.* **165** 150–62

- [27] Maraun D *et al* 2010 Precipitation downscaling under climate change: recent developments to bridge the gap between dynamical models and the end user *Rev. Geophys.* **48** 3
- [28] Margenberg N, Jendersie R, Lessig C and Richter T 2024 DNN-MG: A hybrid neural network/finite element method with applications to 3D simulations of the Navier–Stokes equations *Comput. Methods Appl. Mech. Eng.* **420** 116692
- [29] Moghim S and Bras R L 2017 Bias correction of climate modeled temperature and precipitation using artificial neural networks *J. Hydrometeorol.* **18** 1867–84
- [30] Neumann P, Düben P, Adamidis P, Bauer P, Brück M, Kornblueh L, Klocke D, Stevens B, Wedi N and Biercamp J 2019 Assessing the scales in numerical weather and climate predictions: will exascale be the rescue? *Phil. Trans. R. Soc. A* **377** 20180148
- [31] Pant P and Barati Farimani A 2021 Deep learning for efficient reconstruction of high-resolution turbulent DNS data (arXiv:2010.11348)
- [32] Pathak J, Mustafa M, Kashinath K, Motheau E, Kurth T and Day M 2020 Using machine learning to augment coarse-grid computational fluid dynamics simulations (arXiv:2010.00072)
- [33] Pathak J *et al* 2022 Fourcastnet: a global data-driven high-resolution weather model using adaptive fourier neural operators (arXiv:2202.11214)
- [34] Randall D A *et al* 2018 100 years of earth system model development *Meteorol. Monogr.* **59** 12.1–12.66
- [35] Ronneberger O, Fischer P and Brox T 2015 U-net: convolutional networks for biomedical image segmentation
- [36] Schulzweida U, Kornblueh L and Quast R 2019 CDO user guide
- [37] Shi W, Caballero J, Huszar F, Totz J, Aitken A P, Bishop R, Rueckert D and Wang Z 2016 Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network *2016 IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)* (IEEE Computer Society) pp 1874–83
- [38] Staniforth A and Thuburn J 2012 Horizontal grids for global weather and climate prediction models: a review *Q. J. R. Meteorol. Soc.* **138** 1–26
- [39] Tao Y, Gao X, Hsu K, Sorooshian S and Ihler A 2016 A deep neural network modeling framework to reduce bias in satellite precipitation products *J. Hydrometeorol.* **17** 931–45
- [40] Wang Z, Chen J and Hoi S C H 2021 Deep learning for image super-resolution: a survey *IEEE Trans. Pattern Anal. Mach. Intell.* **43** 3365–87
- [41] Xu Y, Wang Y, Hu S and Du Y 2022 Deep convolutional neural networks for bias field correction of brain magnetic resonance images *J. Supercomput.* **78** 17943–68
- [42] Xu Z, Han Y and Yang Z 2019 Dynamical downscaling of regional climate: a review of methods and limitations *Sci. China Earth Sci.* **62** 365–75
- [43] Yuval J and O’Gorman P A 2020 Stable machine-learning parameterization of subgrid processes for climate modeling at a range of resolutions *Nat. Commun.* **11** 3295
- [44] Zhang C, Perezhugin P, Gultekin C, Adcroft A, Fernandez-Granda C and Zanna L 2023 Implementation and evaluation of a machine learned mesoscale eddy parameterization into a numerical ocean circulation model *J. Adv. Model. Earth Syst.* **15** e2023MS003697