4-13-2023

Accelerating Atmospheric Gravity Wave Simulations using Machine Learning: Kelvin-Helmholtz Instability and Mountain Wave Sources Driving Gravity Wave Breaking and Secondary Gravity Wave Generation

Wenjun Dong
Embry-Riddle Aeronautical University, dongw1@erau.edu

David Fritts
GATS Boulder

Alan Z. Liu
Embry Riddle Aeronautical University - Daytona Beach, liuz2@erau.edu

Hanli Liu
National Center for Atmospheric Research

Jonathan Snively
Embry-Riddle Aeronautical University, snivelyj@erau.edu

Scholarly Commons Citation
Dong, W., Fritts, D., Liu, A. Z., Liu, H., & Snively, J. (2023). Accelerating Atmospheric Gravity Wave Simulations using Machine Learning: Kelvin-Helmholtz Instability and Mountain Wave Sources Driving Gravity Wave Breaking and Secondary Gravity Wave Generation. *Research Square,* (). https://doi.org/10.21203/rs.3.rs-2790920/v2

Accelerating Atmospheric Gravity Wave Simulations using Machine Learning: Kelvin-Helmholtz Instability and Mountain Wave Sources Driving Gravity Wave Breaking and Secondary Gravity Wave Generation

Wenjun DONG
Embry-Riddle Aeronautical University https://orcid.org/0000-0003-0773-5862
David Fritts
GATS Boulder
Thomas Lund
GATS Boulder
Alan Liu (liuz2@erau.edu)
Embry-Riddle Aeronautical University https://orcid.org/0000-0002-1834-7120
Hanli Liu
National Center for Atmospheric Research https://orcid.org/0000-0002-6370-0704
Jonathan Snively
Embry-Riddle Aeronautical University https://orcid.org/0000-0002-7616-439X

Article

Keywords: Machine Learning, Kelvin-Helmholtz instability, Gravity Wave

Posted Date: April 13th, 2023

DOI: https://doi.org/10.21203/rs.3.rs-2790920/v2

Additional Declarations: There is **NO** Competing Interest.

Accelerating Atmospheric Gravity Wave Simulations using Machine Learning: Kelvin-Helmholtz Instability and Mountain Wave Sources Driving Gravity Wave Breaking and Secondary Gravity Wave Generation

Wenjun Dong[1,2,3], David C. Fritts[1,2], Thomas S. Lund[2], Alan Z. Liu[1], Han-Li Liu[3], and Jonathan Snively[1]

[1]Center for Space and Atmospheric Research (CSAR) and Department of Physical Sciences, Embry-Riddle Aeronautical University, Daytona Beach, FL, USA

[2]Global Atmospheric Technologies and Sciences (GATS), Boulder, CO, USA

[3]High Altitude Observatory, National Center for Atmospheric Research, Boulder, CO, USA

Corresponding author: Alan Z. Liu

Email: liuz2@erau.edu

Key Words: Machine Learning, Kelvin-Helmholtz instability, Gravity Wave

Key Points:

1. A machine learning model (CAMNet) tailored for nonlinear gravity wave simulations is developed.
2. CAMNet can achieve a several order-of-magnitude acceleration relative to physics-based model without sacrificing accuracy.
3. CAMNet opens a new window to improve the parameterization of primary and secondary GWs in the global atmospheric models.

Abstract

Gravity waves (GWs) and their associated multi-scale dynamics are known to play fundamental roles in energy and momentum transport and deposition processes throughout the atmosphere. We describe an initial, two-dimensional (2-D), machine learning model – the Compressible Atmosphere Model Network (CAMNet) - intended as a first step toward a more general, three-dimensional, highly-efficient, model for applications to nonlinear GW dynamics description. CAMNet employs a physics-informed neural operator to dramatically accelerate GW and secondary GW (SGW) simulations applied to two GW sources to date. CAMNet is trained on high-resolution simulations by the state-of-the-art model Complex Geometry Compressible Atmosphere Model (CGCAM). Two initial applications to a Kelvin-Helmholtz instability source and mountain wave generation, propagation, breaking, and SGW generation in two wind environments are described here. Results show that CAMNet can capture the key 2-D dynamics modeled by CGCAM with high precision. Spectral characteristics of primary and SGWs estimated by CAMNet agree well with those from CGCAM. Our results show that CAMNet can achieve a several order-of-magnitude acceleration relative to CGCAM without sacrificing accuracy and suggests a potential for machine learning to enable efficient and accurate descriptions of primary and secondary GWs in global atmospheric models.

Plain Language Summary

Atmospheric gravity waves (GWs) are well described by the Navier-Stokes equations, but solving these equations including small scale remains daunting, limited by the very high computational cost of resolving the smallest spatial-temporal features in a global context. To address this challenge, we developed a machine learning model called CAMNet. Our model demonstrates that neural networks can be trained on high-resolution compressible atmospheric model data and then used to simulate gravity wave evolution. Importantly, initial results show that using such trained model can achieve computational savings of >1000 times compared to a physics-based simulation while still achieve highly accurate results. These findings are exciting, as they suggest that CAMNet can overcome the limitations of current GW parameterizations and provide a promising avenue for studying the effects of sub-grid-scale processes in atmospheric science and properly incorporating them in global models. The development of CAMNet opens up major new opportunities for improving effective model resolution, accuracy, and efficiency.

1. Introduction

Gravity waves (GWs) play prominent roles throughout Earth's atmosphere. They are generated at lower altitudes by various primary sources including airflow over topography (i.e., mountain waves, MWs), convection, and jet streams. Additional GWs are generated due to strong GW/mean-flow interactions described as "self-acceleration" (SA) dynamics and by resonant and off-resonant wave-wave interactions that can take many forms. These diverse GWs play central roles in Earth's atmospheric dynamics and climate by transporting energy, pseudo-momentum, and constituents over depths extending into the thermosphere (Fritts & Alexander, 2003).

Accounting for larger- and smaller-scale GW transports and influences remains a challenging problem due to the complex physics involved and the need for high-resolution simulations to describe detailed responses where these are important. Traditional simulation methods, such as finite difference or finite volume methods, can be computationally expensive, and quantifying sub-grid scale processes has been, and remains, a major challenge. Multiple parameterization schemes spanning 40 years have aimed to account for GW pseudo-momentum deposition for various GW sources, discretely or spectrally (i.e., linearly or nonlinearly), from the surface into the thermosphere (e.g., Lindzen, 1981; Holton, 1982; Palmer et al., 1986; Fritts and Lu, 1993). More recent schemes have built on these earlier efforts and insights (Warner and McIntyre, 1996; Hines, 1997; Alexander & Dunkerton, 1999; Yiğit et al., 2008; Eckermann et al., 2015; Amemiya & Sato, 2016; Gettleman et al., 2019; Miyoshi & Yiğit., 2019; Ribstein et al., 2022).

Importantly, all these various schemes are based on simplified, often linear or weakly nonlinear, mathematical models and/or empirical relations that are significant approximations having limited quantitative predictive abilities. As such, they introduce significant model uncertainties and biases in predictions of middle and upper atmosphere responses (Pedatella et al., 2014). Additionally, parameter settings in these schemes may require adjustments for different models, model configurations, and/or model resolutions.

Parameterizations addressing GWs that are partially resolved (the "gray zone") that maintain physical consistency between the resolved and parameterized dynamics are promising (Vosper, 2015; Vosper et al., 2016), but can also be challenging (Liu, 2019). Such efforts attempt to represent the complex and highly nonlinear physics of GWs. However, there remain many aspects of GW dynamics, e.g., SA dynamics, local instabilities and breaking, multi-scale interactions, and secondary GW (SGW) generation that become increasingly important at increasing altitudes, but that cannot be addressed by linear theory or existing GW parameterization schemes.

The recent boom in hardware and software developments relevant to machine learning (ML) has motivated some efforts to examine the possible benefits that ML can bring to GW parameterization (e.g., Chantry et al., 2021; Espinosa et al., 2022). ML methods offer several potential advantages over traditional parameterization schemes. These include the following:

1) ML applications can learn complex, nonlinear relationships directly from data, without the need for pre-determined equations or assumptions. This makes ML methods well-suited for problems exhibiting highly nonlinear dynamics;

2) ML can identify complex patterns and relationships in the data that may be difficult to discern using traditional methods. This can lead to improved accuracy and transferability of the resulting flow descriptions; and

3) ML has the potential to dramatically reduce the computational cost associated with descriptions of nonlinear GW dynamics by replacing traditional parameterizations with highly-efficient, data-driven models.

Such methods have been shown to yield significant benefits in several applications to date. Chantry et al. (2021) trained a neural network on an upgraded version of an existing parameterization scheme that yielded improved results describing GW drag in a numerical weather prediction (NWP) system. In another study, Espinosa et al. (2022) developed an artificial neural network to emulate the pseudo-momentum forcing described in a traditional GW parameterization in an idealized climate model. By coupling the climate model with their ML-based GW parameterization, they were able to accurately reproduce the quasi-biennial oscillation, a well-known atmospheric phenomenon. However, these ML-based GW parameterizations rely on traditional GW parameterizations, hence inherit their assumptions and simplifications. They also remain unable to represent the true, highly nonlinear GW dynamics.

As mentioned earlier, the dynamics of GWs are governed by the Navier-Stokes equations. In recent years, several ML-based solvers for partial differential equations (PDEs) have been proposed to approximate or improve various numerical methods. The most explored of these can be divided into two categories: physics-informed neural networks (PINN, e.g., Maziar et al., 2019; Wandel et al., 2022) and neural operators (NOs, e.g., Lu et al., 2019; Li et al., 2020; Xiong et al., 2023). PINN uses a neural network as the solution function and optimizes a loss function to minimize violation of the given equation. However, it experiences difficulties in propagating information from initial or boundary conditions to unseen parts of the interior and to future times. NOs are better suited for solving PDEs and have been successfully used in flow prediction (e.g., Lu et al., 2019; Li et al., 2020; Xiong et al., 2023). However, they require large volumes of simulation data. Recently, physics-informed neural operators, e.g., both the physics-informed Deep Operator Network proposed by Goswami et al(2022) and physics-informed Fourier neural operator proposed by Li et al (2022) employ both data and physics losses on operator learning to overcome the shortcomings of purely PINN or data-driven learning. Our ML approach is also based on this physics-informed neural operator.

While MLs offer potential advantages for GW simulation, there are also challenges that must be addressed. The primary challenge is the limited availability of high-quality data for training and validation. To address this need, we utilize the Complex Geometry Compressible Atmospheric Model (CGCAM), a finite volume model that has been used extensively to study GW dynamics and their instabilities in the Mesosphere and Lower Thermosphere (MLT) at very high resolution, see, e.g., Dong et al., 2020, 2021,2022,2023; Fritts et al., 2020, 2021,2022a, 2022b; Lund et al., 2021). CGCAM is capable of capturing highly nonlinear GW and GW-related dynamics yielding high-fidelity GW training datasets for ML-based approaches.

Inspired by previous research in this field, this study investigates the application of ML algorithms for simulating GW dynamics. Our focus is on the potential of ML to improve the efficiency of GW simulations while capturing their highly nonlinear dynamics with high fidelity, specifically including instabilities, breaking, and SGW generation. Our approach employs the Compressible Atmosphere Model Network (CAMNet) model based on CGCAM that solves the compressible Navier-Stokes equations in the Complex Geometry Compressible Atmosphere Model (CGCAM). To explore the performance of CAMNet, we train the model using single-channel inputs instead of multi-channel inputs as explored by Pathak et al (2022), with a focus on improving training efficiency. We expect this study to provide valuable insights into the potential benefits of ML in modeling GW and instability dynamics and to inform future research in this area. When optimized, this method will be extended to three-dimensions in order to achieve much more efficient descriptions of GW effects in global models while achieving sufficient accuracy.

2. Method

2.1 Complex Geometry Compressible Atmospheric Model (CGCAM)

The Complex Geometry Compressible Atmosphere Model (CGCAM) solves the two/three-dimensional (2/3-D) nonlinear and compressible Navier-Stokes equations written in strong conservation law (flux) form as follows:

$$\frac{\partial \rho}{\partial t}+\frac{\partial\left(\rho u_j\right)}{\partial x_j}=0 \tag{1}$$

$$\frac{\partial(\rho u_i)}{\partial t}+\frac{\partial\left(\rho u_i u_j\right)}{\partial x_j}=-\frac{\partial p}{\partial x_i}-\rho g\delta_{i3}+\frac{\partial \sigma_{ij}}{\partial x_j} \tag{2}$$

$$\frac{\partial(\rho E)}{\partial t}+\frac{\partial\left[(\rho E+p)u_j\right]}{\partial x_j}=-\rho g u_3+\frac{\partial\left(u_i\sigma_{ij}\right)}{\partial x_j}-\frac{\partial q_j}{\partial x_j} \tag{3}$$

where σ_{ij} and q_j are the viscous stress and thermal conduction, respectively, defined as

$$\sigma_{ij} = \mu\left[\left(\frac{\partial u_i}{\partial x_j}+\frac{\partial u_j}{\partial x_i}\right)-\frac{2}{3}\left(\frac{\partial u_k}{\partial x_k}\right)\delta_{ij}\right] \tag{4}$$

$$q_j = -\kappa\frac{\partial T}{\partial x_j} \tag{5}$$

and where μ is the dynamic viscosity, κ is the thermal conductivity, δ_{ij} is the Kronecker delta, ρ is density, and g is the gravitational acceleration. μ and κ depend on the temperature through Sutherland's Law (White, 1974).

The solution variables are ρ, the momentum per unit volume, ρu_i *or* $(\rho u, \rho v, \rho w)$, and the total energy $E=e+u_k u_k/2=c_v T+u_k u_k/2$, with velocity components (u_i, u_j, u_k) along (x, y, z). Also $c_v=\frac{R}{\gamma-1}$ is the specific heat at constant volume and T is the temperature. The compressible equation set is discretized using a second-order finite-volume scheme identical to the method discussed by Felten & Lund (2006). Time advancement is achieved via a third-order accurate Runge-Kutta scheme. Additional details for CGCAM are provided by Dong et al. (2020) and Lund et al. (2020). CGCAM has been successfully used in various studies on GW generation, breaking, and SGW generation (e.g., Dong et al., 2020,2021,2022; Fritts et al., 2020;2021; 2022a; 2022b; Lund et al., 2021).

2.2 Compressible Atmospheric Model Network (CAMNet)

CAMNet is a hybrid machine learning model that combines data-driven and physics-informed approaches. It is based on the Adaptive Fourier Neural Operator (AFNO) proposed by Guibas et al., 2021. AFNO is a Fourier transform-based token-mixing scheme with a vision transformer backbone (Dosoviskiy et al., 2020). AFNO is based on the Fourier neural operator (FNO) that learns in a resolution-invariant manner and has shown success in modeling challenging partial differential equations such as fluid dynamics (Li et al., 2020). The Fourier architecture of AFNO applies a fast Fourier Transform (FFT) to the data and applies its fully connected layers in Fourier space before performing an inverse FFT back to real space. Moreover, the Fourier architecture has been demonstrated the ability to perform zero-shot super-resolution, predicting on higher-resolution data having only seen low resolution data. The introduction of vision transformer enables it to model long-range dependencies well and yields a state-of-the-art high-resolution model that resolves fine-grained features and scales well with resolution and size of dataset. AFNO enables training high-fidelity data-driven models as truly unprecedented resolution (Pathak et al., 2022).

The power of AFNO stems from its ability to combine linear integral operators, implemented through the Fourier transform, with non-linear activation functions, enabling it to learn highly non-linear operators. This is similar to standard Multi-Layer Perceptron (MLP) and Convolutional Neural Network (CNN), where linear multiplications are combined with non-linear activations to learn highly non-linear functions. Although AFNO truncates higher frequency modes in the Fourier layer, Li et al. (2020) argue that the entire operator can still approximate functions with the full frequency range, due to the function being represented in a high-dimensional channel space. The non-linear decoder network then recovers the higher frequency modes when projecting back to the desired dimension. In our case of predicting multi-scale GW dynamics, the Fourier layer truncation of high-frequency information resulted in poor small-scale structure prediction. To address this, we added a convolutional layer, which is able to amplify high-frequency components and complement the information truncated by the Fourier layer.

Additionally, CAMNet further extends the AFNO architecture by incorporating physics information from the Navier-Stokes equations, which govern the GW dynamics. These equations are used to create a loss function that captures the violation of these laws, and Fourier derivatives (Li et al., 2021) are used to compute the derivatives for the physics constraints, as the automatic differentiation in PyTorch is very memory-intensive for this type of architecture. The physical constraints reduce the demand for training datasets and improve the generalization and physical validity of CAMNet learning compared to purely data-driven methods.

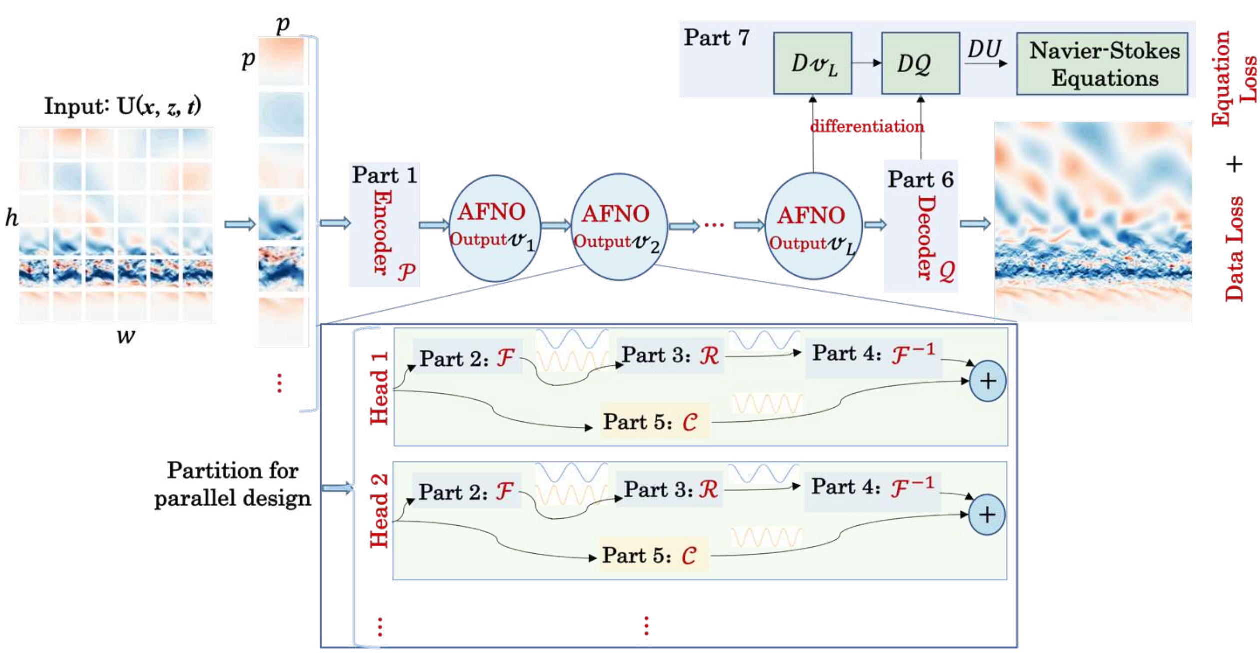

Figure 1. CAMNet architecture that utilizes the modified Adaptive Fourier Neural Operator (AFNO) and follows a patch-based approach. The input frame is divided into a $h \times w$ grid of patches, each of size $p \times p$, and encoded in a higher dimensional space with position embedding is added to form a sequence of tokens. These tokens are then mixed spatially using AFNO, which is repeated for L layers, and then a decoder reconstructs the patches for the next frame. A single AFNO layer is composed of multiple heads for parallel processing. Parts 2-4 of the layer applies the Fourier transform $\boldsymbol{\mathcal{F}}$ to the input, followed by a linear transform $\boldsymbol{\mathcal{R}}$ that acts on the lower-frequency Fourier modes and filters out the higher-frequency modes, this is then followed by an inverse Fourier transform $\boldsymbol{\mathcal{F}^{-1}}$. Part 5 stands for a convolutional layer that is used to capture the higher-frequency modes that are missed by Part 2-4. Both AFNO outputs $\mathcal{v}_L$ and reconstructed U are functions, and their derivatives $D\mathcal{v}_L$ and DU can be computed at any query points x and z. The solutions are constrained by the Navier-Stokes equations that govern GW dynamics. For more information, refer to the corresponding texts.

The CAMNet architecture is shown in **Figure 1**. The CAMNet is currently trained on horizontal wind U, so CAMNet's input is $U(x, z, t)$. The CAMNet architecture consists of 7 parts. Below, we present a detailed computational implementation of each part. Similar to the iterative update strategy of each Fourier layer in FNO (Li et al., 2020), the improved iterative update strategy of each Fourier layer in CAMNet can be expressed as follows:

$$\mathcal{v}_1 = \overbrace{\mathcal{P}}^{Part\ 1} U(x,z,t) \tag{6}$$

$$\mathcal{v}_{l+1} = \sigma\left(\overbrace{\mathcal{C}}^{part\ 5} \mathcal{v}_l + \overbrace{\mathcal{F}^{-1}}^{part\ 2} \overbrace{\mathcal{R}}^{part\ 3} \overbrace{\mathcal{F}}^{part\ 4} \mathcal{v}_l \right), for\ l = 1,2,\cdots,L \tag{7}$$

$$U(x,z,t+1) = \overbrace{\mathcal{Q}}^{part\ 6} \mathcal{v}_L \tag{8}$$

where $\mathcal{P}$ and $\mathcal{Q}$ are encoder and decoder that are realized by two neural networks that projects $U(x, z, t)$ to hidden representation v_l and projects the representation back to the solution $U(x, z, t+1)$. σ is a nonlinear activation function. The additional term $\mathcal{C}$ is a convolutional layer that acts on v_l. v_l denotes the output of the l-th Fourier layer of AFNO. $\mathcal{F}$ denotes the Fourier transform that acts on v_l, $\mathcal{R}$ is a linear transform layer that acts on $\mathcal{F}(v_l)$ to handle its low-frequency modes. $\mathcal{F}^{-1}$ is an inverse Fourier transform that acts on $\mathcal{R}(\mathcal{F}(v_l))$. The details of each part are provided below:

Part 1: Encoder $\mathcal{P}$ The encoder is implemented using the token embedding layer in the Vision Transformer architecture proposed by Dosovitskiy et al. (2020). This layer applies a linear projection to each patch to obtain a fixed-sized vector, which is then concatenated with positional embeddings representing the spatial location of the patch. The resulting sequence of vectors serves as the input to the sequent neural network layers.

Part 2-4: Fourier Transform $\mathcal{F}$, Linear Transform $\mathcal{R}$, and Inverse Fourier Transform $\mathcal{F}^{-1}$

As stated in Li et al., 2020, since the inputs and outputs of partial differential equations (PDEs, such as Navier-Stokes equations) are continuous functions, it is more efficient to represent them in Fourier space and perform global convolution. This is due to the quasi-linear computational complexity and global properties of Fourier transform, making it a more efficient approach. The convolution in the spatial domain is equivalent to the pointwise multiplication in the Fourier domain. To capture global features in input data, a Fourier transform is first applied to the inputs, followed by a Linear Transform $\mathcal{R}$ that acts on the lower-frequency Fourier modes by assigning weights to them. These weights will be updated during the training. Finally, an inverse Fourier transform is performed to obtain the output.

Part 5: High-frequency Information Compensation $\mathcal{C}$ In each Fourier layer, we utilize a convolutional layer to extract high-frequency information because it can amplify high-frequency components. Therefore, we train a convolutional layer $\boldsymbol{C}$ on the outputs of Part 1 to extract their high-frequency information. As a complement to Parts 2-4, the convolutional layer enables the forward prediction of high-frequency information.

Part 6: Decoder $\boldsymbol{Q}$ Given two future states independently predicted by Parts 2-4 and Part 5, we combine them and train a non-linear decoder using a multi-Layer perceptron layer with a tanh activation function to transform the AFNO outputs back into U.

Part 7: Physics Informed Loss $\boldsymbol{\mathcal{L}_p}$

As CAMNet is currently trained on horizontal wind U, we utilize only the momentum flux equation (1) for the physics-informed part, assuming a constant density ρ over time. Thus, equation (1) can be simplified to

$$\frac{\partial(\rho U)}{\partial x} + \frac{\partial(\rho W)}{\partial z} = 0 \tag{9}$$

In equation (9), the vertical wind W is obtained from CGCAM simulations at each prediction time step, and ρ is set to its initial values. The only physics-informed variable to be calculated in CAMNet is the derivative of U with respect to x. To enable multi-variable predictions in the future, the physics-informed part will need to involve the complete Navier-Stokes equations.

The loss function of CAMNet for optimizing equations (6), (7), and (8) is defined as

$$\mathcal{L} = \alpha\mathcal{L}_{data} + \beta\mathcal{L}_p = \alpha\left|U - \widehat{U}\right| + \beta\left|\frac{\partial(\rho U)}{\partial x} + \frac{\partial(\rho W)}{\partial z}\right| \tag{10}$$

where α and β control the weights of data-driven and physics-informed part in loss functions, respectively.

Table 1: Key model parameters used in CAMNet model and training.

Hyperparameter	**Value**
Batch Size	32
Learning rate	5×10^{-4}
Learning rate schedule	Cosine
Patch size (Case 1/2)	$10 \times 10/8 \times 8$
Number of AFNO layers	8
Heads number	8
Heads depth	6
AFNO embedding dimension	768
Activation function	GeLU
Dropout	0

CAMNet is highly optimized so that it can be trained efficiently on massively parallel GPU resources. The initial application of CAMNet has shown that CAMNet can achieve order-of-magnitude speedup over numerical model CGCAM. We performed two cases to evaluate the performance of CAMNet (see "Result" section), and the relevant model parameters are presented in **Table 1**.

3. Results

This section describes our initial efforts using CAMNet to accelerate simulations of GWs arising from two very different sources constrained to a two-dimensional (2-D) domain. Case 1 describes the generation of initial GWs by large-scale, shear-induced Kelvin-Helmholtz Instability (KHI) and the successive generation of SGWs at much larger scales that readily propagate to much higher altitudes. Case 2 describes mountain waves arising from flow over idealized terrain, their attainment of large amplitudes, breaking, and generation of SGWs that likewise attain very high altitudes. CAMNet wind fields and spectra are compared with high-resolution 2-D CGCAM simulations in both cases.

3.1 Case 1: Gravity Waves emitted from Kelvin–Helmholtz Instability

We explore here the potential of CAMNet for modeling GWs emitted from Kelvin-Helmholtz Instability (KHI) described by Dong et al.(2023). We use CGCAM to generate training and testing data for CAMNet. The initial background winds (see Figure 2) are specified as

$$U(z) = U_0 cos\left[\frac{\pi(z - z_0)}{15\ km}\right] tanh\left(\frac{z - z_0}{h}\right) \tag{11}$$

CGCAM simulations are performed for a computational domain having dimensions 180 km $\times$ 180 km $(x,\ z)$ with resolutions of 50 m at the shear center, with exponential mesh stretching approaching the upper and lower boundaries to reduce computational demands. Periodic boundary conditions are used at the lateral boundaries. An isothermal no-stress wall condition is used at the lower boundary, and a characteristic radiation condition is used at the upper boundary. The vertical boundary conditions are supplemented with sponge layers having 20-km depths to further ensure no reflected GWs. After excluding irrelevant data in the sponge layers, the variable U are stored on a grid of dimension of 2000×1000.

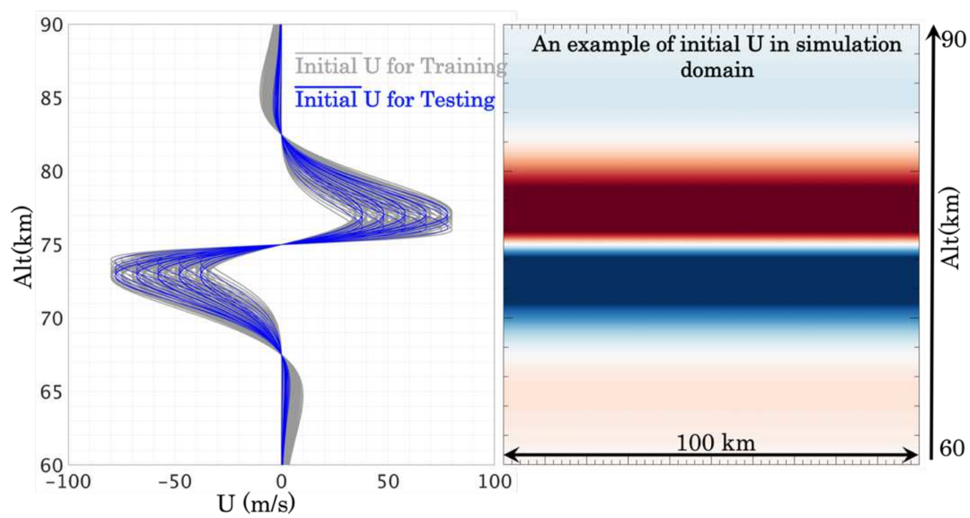

Figure 2: Initial conditions for generating training and testing data (left) and an example of initial U in the simulation domain (right) for the KHI cases.

Given the initial conditions, CAMNet is required to simulate the future states of variable U at $t \in \{1,2,3,\cdots,50\}$ min for a suite of initial conditions. A total of 200 cases are generated by varying U_0 and h in Equation 11, and the corresponding outputs of CGCAM serves as the true reference solutions for each case. The CGCAM simulations for each case were run for 50 minutes at an interval of 1 minute. The 200 CGCAM cases were then split into a training set of 180 cases and a testing set of 20 cases. The training set is used to train CAMNet, and the testing set is used to evaluate the model's performance. Keeping the testing set separate from the training set is crucial to obtain an unbiased estimate of the model performance. All samples have a grid of 2000×1000. The CAMNet training is implemented in a multi-GPU environment with 4 V100 GPUs. Convergence is observed after approximately 250 epochs during the training process.

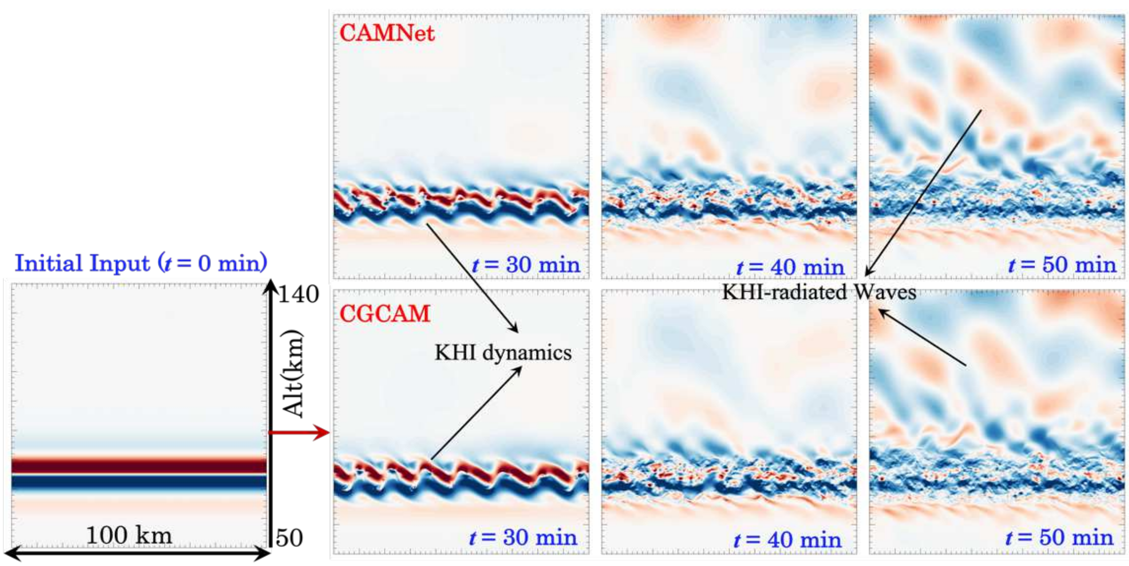

Figure 3. KHI and GWs predicted by CGCAM and CAMNet.

Figure 3 displays the variable U employed for both the CAMNet and CGCAM simulations during the model test, and we observe a high level of consistency between two models. The ML model CAMNet can capture small- and large-scale structures qualitatively, with clear evidence of KHI and KHI-radiated GWs seen in both CGCAM and CAMNet. The initial strong shear produces deep and broad KH billows that break down after ~30 min, leading to the emergence of small-amplitude GWs above the KHI altitude. By 50 min, increasing GWs and apparent SGWs are seen to propagate with group velocities along the iso-phase lines that extend away from the KHI shear layer, causing GWs to propagate to higher altitudes and achieve larger amplitudes. Both models suggest that GWs are continuously emitted from the KHI and turbulence dynamics, with well-defined spatial structures having orientations and spatial scales that agree closely between the CGCAM and CAMNet fields.

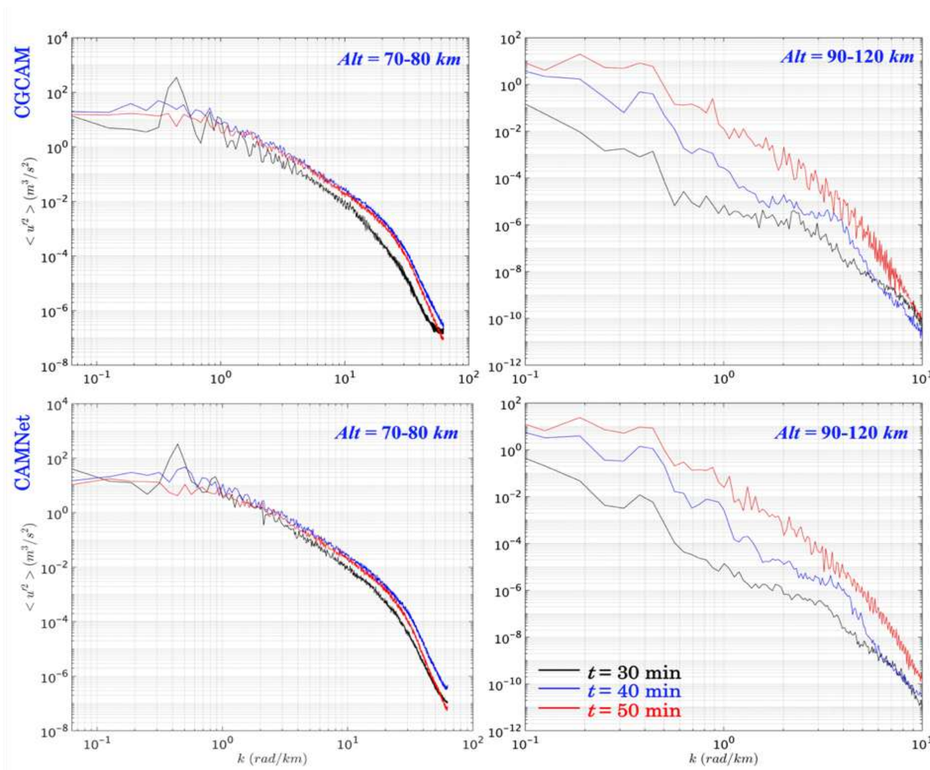

Figure 4. Case 1 spectral characteristics predicted by CGCAM (top) and CAMNet (bottom) at 70-80 km (left) and 90-120 km (right) altitude ranges at 30, 40, and 50 min model simulation times (see time labels at lower right).

The spectral properties of GWs can reveal important details about their sources, such as the altitude and vertical extent of the source region, and the dominant wavelengths and frequencies of the GWs generated. This information can be used to improve parameterizations of GW sources in global climate models, which are critical for predicting the response of the atmosphere to changing climate conditions. The spectral structures of horizontal wind disturbance u' are calculated and found to exhibit high consistency between CGCAM and CAMNet. The spectra of u'^2 computed at the KHI region and higher altitudes at t = 30, 40, and 50 min are displayed in Figure 4. The spectra reflect the characteristics of KHI dynamics. In the KHI region, the onset of strong KHI and 2-D turbulence yields spectral slopes approaching -5/3 corresponding to wavenumbers of ~1-6 rad/km. Spectral amplitudes fall sharply beyond wavenumbers of ~20 rad/km and exhibit steeper slopes approaching -7 in the viscous range. At higher altitudes, the -5/3 slope corresponds to wavenumbers of ~0.8-3 rad/km and -7 slopes correspond to wavenumbers of 3-10 rad/km. These spectra suggest that small-scale structures discussed above are well resolved at these times.

These results suggest that CAMNet has promising potential as an alternative to CGCAM for simulating KHI and KHI-radiated GWs. Notably, the time cost of a single KHI case simulation using CAMNet was approximately 0.8 seconds on a single A100 (80GB). This represents a significant acceleration (by > 2000) compared with the 30 minutes needed by CGCAM when using 36 CPU cores.

3.2 Case 2: Mountain Waves Generation, Propagation, and Breaking

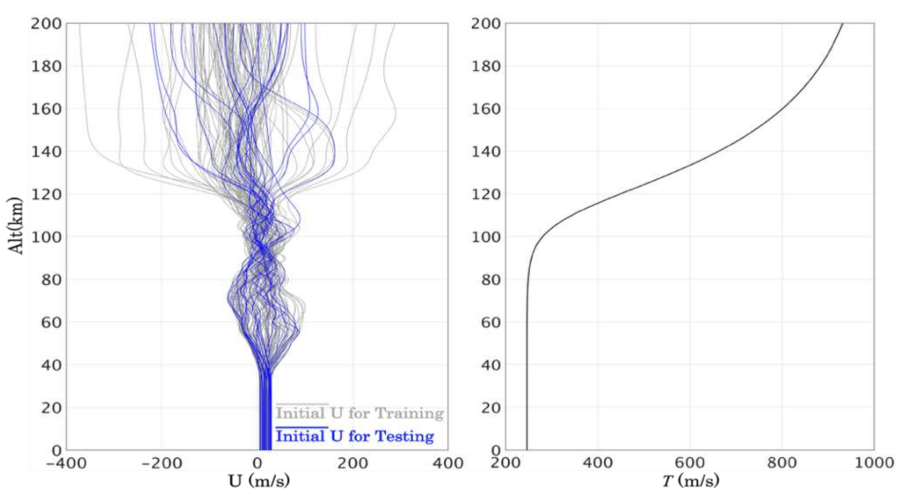

Figure 5. Initial mean winds U (left) and temperature T (right) as the inputs of CGCAM for generating training and testing data for MW cases.

Our intent in Case 2 was to explore CAMNet capabilities for modeling MW generation, propagation, breaking, and their radiation of SGWs. MW breaking is one of the strongest sources of SGWs (Lund et al., 2020). As in Case 1, CAMNet training employs CGCAM simulation data. The CGCAM simulations cover a computational domain extending 700×220 km (x and z) at a resolution of 1 and 0.5 km in x and z. As in Case 1, the lateral boundary is periodic. At the lower boundary, a Gaussian terrain of peak height 4 km and half-width of 30 km is used, and a characteristic radiation condition is used at the upper boundary. Sponge layers of 20 km and 50 km are added to the vertical and lateral boundaries, respectively, to ensure absorption of outgoing GWs. The variable U are stored on a grid having dimensions of 600×400 after irrelevant values in the sponge layers are excluded.

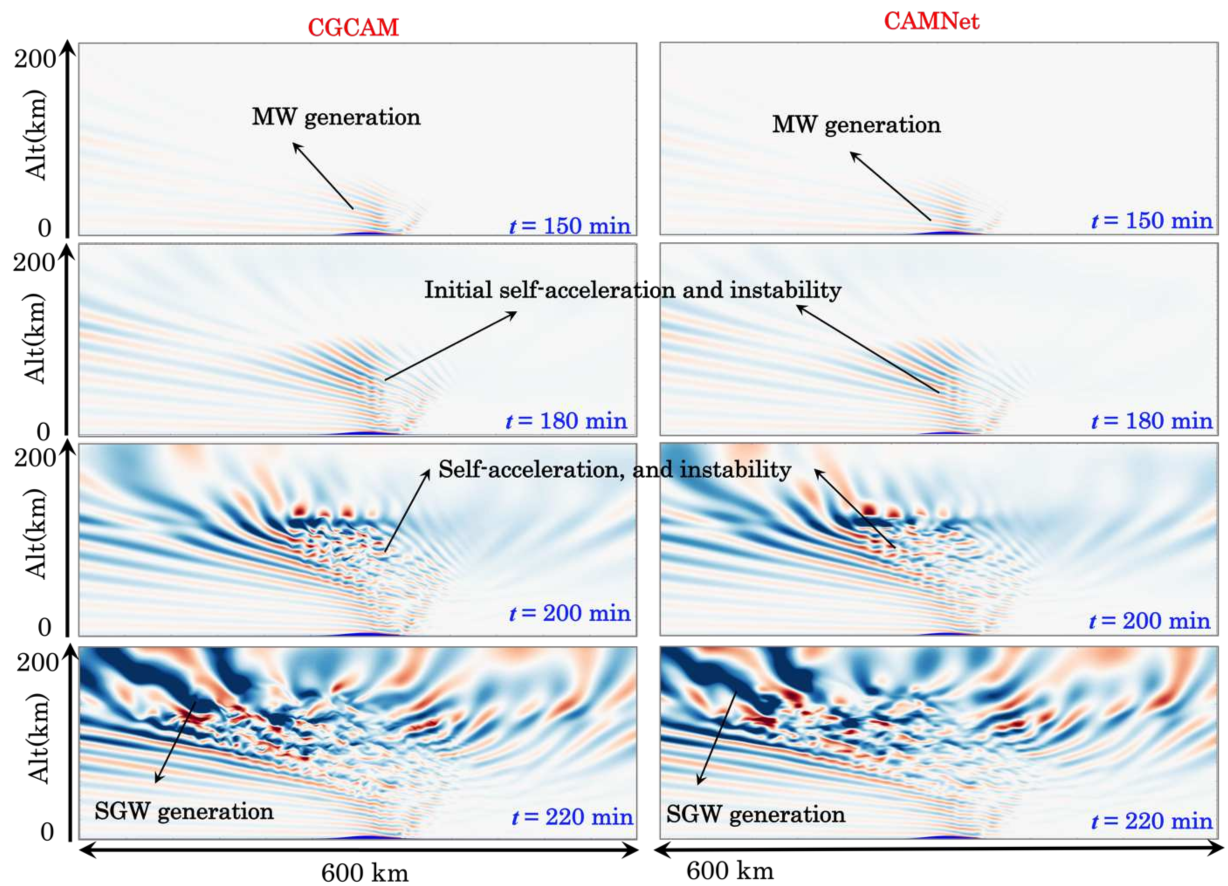

Figure 6. MW evolutions predicted by CGCAM (left) and CAMNet (right). The initial condition is a horizontal wind of 25 m/s at the surface flowing over a Gaussian Mountain with a height of 4 km and a half width of 30 km (see Figure 4).

Given the samples of initial conditions, CAMNet is trained to reproduce the future states of variable U at $t \in \{140,145,150,\cdots,220\}$ min. Note that we start from t = 140 mins to avoid CAMNet being trained with non-physical data produced by CGCAM at early simulation times. A total of 200 cases are generated by varying the initial wind field, with a random wind field randomly extracted from the HWM14 at $30^{\circ}S, 70^{\circ}W$ (Andes Lidar Observatory) at 00:00 on 200 days among 365 days. The initial temperature field is simplified as used in Dong et al. (2020). Winds at lower altitudes from HWM14 are consistently lower than actual observations, and thus a correction is needed to enable simulation of MW generation. To account for this discrepancy, we randomly assign wind values ranging from 0-30 m/s at these lower altitudes to facilitate the occurrence of MWs. The initial fields were assumed to be uniform over the domain. The initial wind and temperature fields are shown in Figure 5. The corresponding output of CGCAM serves as the true reference solution for each case. CGCAM simulations for each case were run for 220 minutes at an interval of 5 minutes. The CGCAM simulations were then split into a training set of 180 cases and a testing set of 20 cases. All samples have a grid of 600×400. The training of CAMNet is implemented in a multi-GPU environment with 4 V100. Convergence is observed after approximately 320 epochs during training process.

The CAMNet model achieves excellent skill in modeling MW generation, propagation, breaking, and SGW generation. As an illustrative example, we choose a case with a uniform initial background wind of 25 m/s. We begin with an overview of the major features of the MW and SGW evolution from t = 150-220 min. Figure 6 shows U at $t = 150, 180, 200, 220$ min generated by CAMNet and the corresponding CGCAM results at these times. Considering CAMNet results first, the earliest responses at t =150 min reveal MW generation at lower altitudes and their extension into the MLT. At t =180 min, initial SA dynamics and instabilities are seen at lower altitudes.

At $t = 200$ min, there is evidence for strong SGW excitation in the MW breaking regions. The MW field and its associated instabilities and SGWs continue to intensify to $t = 200$ min. The CAMNet results approximate the CGCAM ground truth remarkably well over 220 min. Additionally, high consistency is found between CGCAM and CAMNet in the u' spectra, which are shown in the first and second rows of Figure 8, respectively.

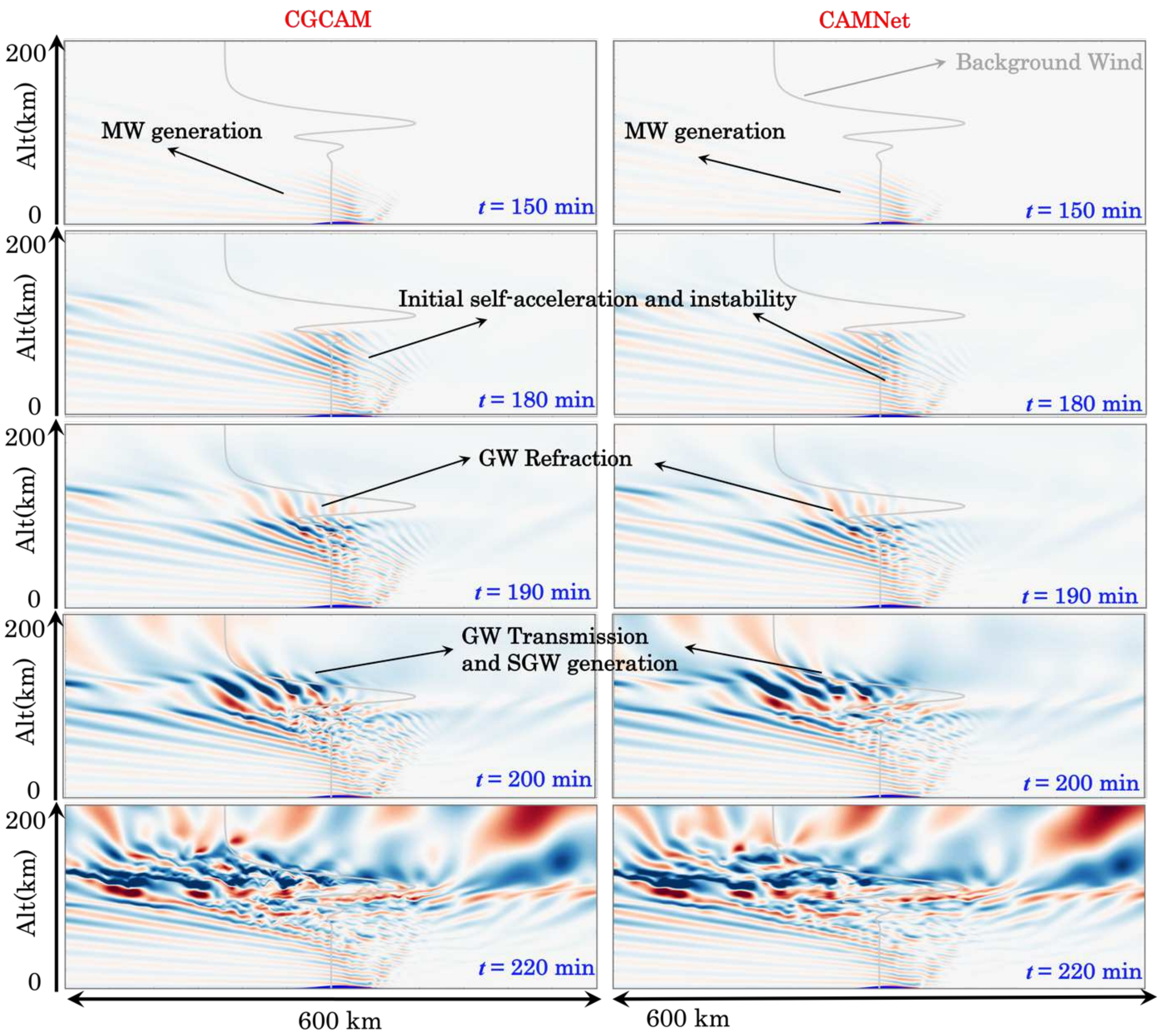

Figure 7. Same as Figure 6, but for a tidal wind background.

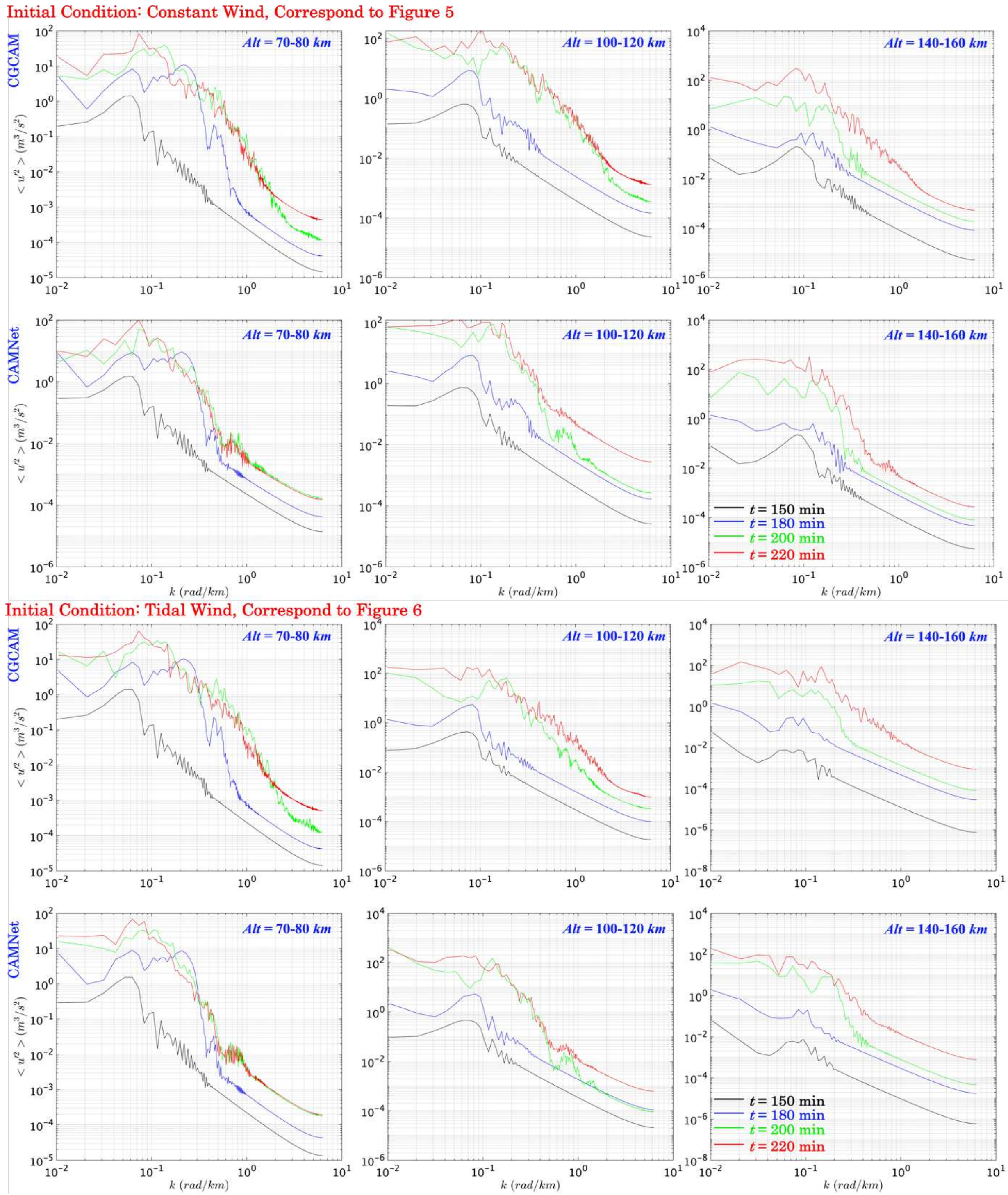

Figure 8. Case 2 spectral characteristics predicted by CGCAM (first and third rows) and CAMNet (second and fourth rows) at 70-80 km (left), 100-120 km (middle) and 90-120 km (right) at 150,180,200, and 220 min (see time labels at lower right).

As a further test of the generalization ability of the trained CAMNet model, we chose an initial wind field that includes a tidal wind field (represented by grey lines in Figure 7) that was not part of the 200 cases used for training and testing. CAMNet results were compared to those of CGCAM, and also showed high consistency in capturing the MW dynamics (see Figures 7 and 8), including the following:

1. Major MW responses, including MW generation, propagation, breaking, local instabilities and dissipation, and SGW generation;
2. strong MW breaking, instabilities, and 2-D turbulence dynamics;
3. SGWs modulated by tidal winds and having large scales and large influences extending into the thermosphere. Responses include refraction by tidal winds, and reduced SGW responses at higher altitudes relative to the case of no tidal wind (as shown in Figures 5 and 6); and
4. CAMNet exhibits highly consistent MW spectral characteristics with CGCAM (see the third and fourth rows in Figure 7).

This case demonstrates that CAMNet exhibits good generalization abilities.

CAMNet demonstrates promising potential as a competitive alternative to CGCAM for simulating MW generation, propagation, and breaking. A single MW case simulation using CAMNet takes approximately 0.5 seconds on a single A100 (80GB), in this case ~4000 times faster than the corresponding CGCAM simulation using 36 CPU cores, which took around 40 minutes.

4. Discussion

GWs play significant roles in the transport of energy and pseudo-momentum through Earth's atmosphere. However, current numerical weather prediction and climate models lack the resolution needed to describe the smaller-scale GWs and SGWs accounting for the large majority of GW energy and pseudo-momentum fluxes to higher altitudes. Thus, parameterizations are employed to represent unresolved GW influences in most global atmospheric models. These typically have three primary components: (1) specification of GWs at source levels, (2) GWs propagation with altitude, and (3) GW dissipation and further parameterized body forcing and mixing.

The primary function of GW parameterizations as currently applied in global models is to compute the wave-driven force on the mean flow, where the mean flow refers to the grid-box mean, and the waves are intended to represent sub-grid unresolved GW anomalies. These parameterizations treat GWs as linear, hydrostatic, vertically propagating waves in a steady ambient environment with Boussinesq governing equations. However, these assumptions and simplifications limit the representation of many observed GW characteristics, which are physically well understood (e.g., Erying et al., 2006, Hertzog et al., 2012; Eckermann et al., 2015; Stephan et al., 2016). Some ML-based GW parameterization schemes have been proposed to enhance the accuracy of GW parameterization (e.g., Chantry et al., 2021; Espinosa et al., 2022). These ML-based GW parameterization schemes rely on traditional GW parameterizations and hence are limited by the assumptions and simplifications inherent in them. CAMNet offers several advantages over traditional and previous ML-based GW parameterization schemes. The training process of CAMNet does not rely on any existing GW parameterization schemes, thus it is not limited by their assumptions and simplifications. CAMNet is trained with high-resolution simulation data from the state-of-the-art atmospheric model CGCAM, which are accurate numerical solutions of the Navier-Stokes equations. Well-trained CAMNet is capable of resolving multi-scale and highly nonlinear GW dynamics, such as instability and turbulence dynamics, self-acceleration, GW breaking, and SGW generations, at much faster speed. To the best of our knowledge, CAMNet is the first ML-based approach that can directly simulate highly nonlinear GW dynamics.

To further enhance the application of CAMNet's exceptional GW simulation capabilities, CAMNet will be optimized and extended 1) from single-variable simulation to multi-variable simulation to better capture the complex nonlinear interactions among various GW variables; and 2) from 2D to 3D to more accurately model the horizontal evolution characteristics of GWs, which are only fully displayed in 3D cases.

5. Summary

In this paper, we developed a machine learning model solving the compressible Navier-Stokes equations in our Complex Geometry Compressible Atmosphere Model (CGCAM) named CAMNet. CAMNet is a hybrid machine learning model that combines data-driven and physics-informed approaches. It is based on the Adaptive Fourier Neural Operator (AFNO) proposed by Guibas et al., 2021, with modifications tailored to our simulations. The main improvements include: 1) the addition of convolutional layer branch to compensate for high-frequency components truncated by the Fourier layers, making the model more robust in resolving multi-scale dynamics, and 2) the incorporation of physical information from the Navier-Stokes equation in CGCAM. The CAMNet feedback neural network utilizes a loss function that combines both the physical information in Navier-Stokes equations and the data loss from CGCAM simulations. This approach reduces the need for extensive training data and improves the model generalization ability.

We evaluated the performance of CAMNet with two test cases: the first one described the KHI and the associated GW radiation as explored in Dong et al., 2023, and the second one addressed the generation, propagation, and breaking of MWs, which are one of the most important GW components. Both cases involve small-scale instability and turbulence dynamics, as well as larger-scale GWs. Our main findings from the simulations are as follows:

1) CAMNet shows excellent skill on simulating the formation and intensification of KHI, and KHI-radiated GWs. Those results qualitatively match the ground truth remarkably well over a period of 50 mins.
2) CAMNet has excellent skill on the simulations of MW generation, propagation, and breaking. CAMNet captures the major MW responses, including MW

generation, propagation into MLT, strong MW breaking via vortex ring formation, strong instability and turbulence dynamics, intense SGW generation, and strong modulation by tidal wind.

3) CAMNet can be trained on high-resolution CGCAM simulations under various conditions and has the potential to significantly accelerate GW simulations while maintaining high accuracy.

Accounting for the GW forcing in global atmospheric models is challenging due to their limited model resolutions. A well-trained CAMNet can produce the simulations orders of magnitude faster than CGCAM without any noticeable accuracy loss. This has two important implications. Firstly, high-resolution GW simulations can be generated within seconds, thus enabling estimation of well-calibrated and constrained uncertainties regarding unresolved GW scales with higher confidence compared to current global models that have severely simplified GW parameterization schemes due to computational cost. Secondly, CAMNet is suitable for rapidly testing hypotheses about mechanisms of GW forcing and their predictability. Moreover, there is potential to develop a software library of well-trained CAMNet models to be applied to a broad range of conditions. The well-trained models have the potential to become a viable alternative to current GW parameterizations in global models.

Data availability

Outputs from model simulations used in this study are archived on the ERAU High Performance Computing System Vega and can be made available upon publication.

Code availability

The CGCAM code utilized in this study is proprietary and can only be accessed through collaboration. The original AFNO code is available at https://github.com/lonestar686/AdaptiveFourierNeuralOperator.

Acknowledgement

Research described here was supported by the Air Force Office of Scientific Research (AFOSR) grant FA9550-18-1-0009 and NSF grants AGS-1759471, AGS-2032678, AGS-2131350, and AGS-2128443. We also acknowledge Embry-Riddle Aeronautical University and the National Center for Atmospheric Research for access to supercomputer platforms that allowed the CGCAM and CAMNet simulations reported here. The National Center for Atmospheric Research is a major facility sponsored by the National Science Foundation under Cooperative Agreement No. 1852977.

References

Alexander, M. J., & Dunkerton, T. J. (1999). A spectral parameterization of mean-flow forcing due to breaking gravity waves. Journal of the Atmospheric Sciences, 56(24), 4167-4182. https://doi.org/10.1175/1520-0469(1999)056%3C4167:ASPOMF%3E2.0.CO;2.

Amemiya, A., & Sato, K. (2016). A New Gravity Wave Parameterization Including Three-Dimensional Propagation. Journal of the Meteorological Society of Japan, https://doi.org/10.2151/jmsj.2016-013.

Chantry, M., Hatfield, S., Dueben, P., Polichtchouk, I., & Palmer, T. (2021). Machine learning emulation of gravity wave drag in numerical weather forecasting. Journal of Advances in Modeling Earth Systems, 13, e2021MS002477. https://doi.org/10.1029/2021MS002477.

Dong, W., Fritts, D. C., Lund, T. S., Wieland, S. A., & Zhang, S. (2020). Self-acceleration and instability of gravity wave packets: 2. two-dimensional packet propagation, instability dynamics, and transient flow responses. Journal of Geophysical Research: Atmospheres, 125, e2019JD030691. https://doi.org/10.1029/2019JD030691.

Dong, W., Fritts, D. C., Thomas, G. E., & Lund, T. S. (2021). Modeling Responses of Polar Mesospheric Clouds to Gravity Wave and Instability Dynamics and

Induced Large-Scale Motions. Journal of Geophysical Research: Atmospheres, 126, https://doi.org/10.1029/2021JD034643.

Dong, W., Fritts, D. C., Hickey, M. P., Liu, A. Z., Lund, T. S., Zhang, S., et al. (2022). Modeling studies of gravity wave dynamics in highly structured environments: Reflection, trapping, instability, momentum transport, secondary gravity waves, and induced flow responses. Journal of Geophysical Research: Atmospheres, 127, e2021JD035894. https://doi.org/10.1029/2021JD035894.

Dong, W., Fritts, D. C., Liu A. Z., Lund, T.S., & Hiu, H. (2023a). Gravity Waves Emitted from Kelvin-Helmholtz Instabilities. Geophysical Research Letters, https://doi.org/10.1029/2022GL102674.

Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., ... & Houlsby, N. (2020). An image is worth 16x16 words: Transformers for image recognition at scale. https://doi.org/10.48550/arXiv.2010.11929.

Eyring, V., et al. (2006), Assessment of temperature, trace species, and ozone in chemistry-climate model simulations of the recent past, J. Geophys. Res., 111, D22308, https://doi.org/10.1029/2006JD007327.

Eckermann, S. D., Broutman, D., & Knight, H. (2015). Effects of Horizontal Geometrical Spreading on the Parameterization of Orographic Gravity Wave Drag. Part II: Analytical Solutions. Journal of the Atmospheric Sciences, 72(6), 2348-2365. https://journals.ametsoc.org/view/journals/atsc/72/6/jas-d-14-0148.1.xml

Espinosa, Z. I., Sheshadri, A., Cain, G. R., Gerber, E. P., & DallaSanta, K. J. (2022). Machine learning gravity wave parameterization generalizes to capture the QBO and response to increased CO2. Geophysical Research Letters, 49, e2022GL098174. https://doi.org/10.1029/2022GL098174.

Fritts, D. C., & Lu, W. (1993). Spectral estimates of gravity wave energy and momentum fluxes. Part II: Parameterization of wave forcing and variability. Journal of Atmospheric Sciences, 50(22), 3695-3713. https://doi.org/10.1175/1520-0469(1993)050%3C3695:SEOGWE%3E2.0.CO;2.

Fritts, D. C., and Alexander, M. J. (2003), Gravity wave dynamics and effects in the middle atmosphere, Rev. Geophys., 41, 1003, thhps://doi.org/10.1029/2001RG000106.

Fritts, D. C., Dong, W., Lund, T. S., Wieland, S., & Laughman, B. (2020). Self-Acceleration and instability of gravity wave packets: 3. Three-Dimensional packet propagation, secondary gravity waves, momentum transport, and transient mean forcing in tidal winds. Journal of Geophysical Research: Atmospheres, 125(3). https://doi.org/10.1029/2019JD030692.

Fritts, D. C., Lund, A. C., Lund, T. S., & Yudin, V. (2022a). Impacts of limited model resolution on the representation of mountain wave and secondary gravity wave dynamics in local and global models. 1: Mountain waves in the stratosphere and mesosphere. Journal of Geophysical Research: Atmospheres, 127, e2021JD035990. https://doi.org/10.1029/2021JD035990.

Fritts, D. C., Lund, A. C., Lund, T. S., & Yudin, V. (2022b). Impacts of limited model resolution on the representation of mountain wave and secondary wave dynamics in local and global models: 2. Mountain wave and secondary wave evolutions in the thermosphere. Journal of Geophysical Research: Atmospheres, 127, e2021JD036035. https://doi.org/10.1029/2021JD036035.

Gettelman, A., Mills, M. J., Kinnison, D. E., Garcia, R. R., Smith, A. K., Marsh, D. R., et al. (2019). The whole atmosphere community climate model version 6 (WACCM6). Journal of Geophysical Research: Atmospheres, 124, 12380– 12403. https://doi.org/10.1029/2019JD030943.

Goswami, S., Bora, A., Yu, Y., & Karniadakis, G. E. (2022). Physics-informed neural operators. https://doi.org/10.48550/arXiv.2207.05748.

Hertzog, A.,, G. Boccara, R. A. Vincent, F. Vial, and P. Cocquerez (2008). Estimation of Gravity Wave Momentum Flux and Phase Speeds from Quasi-Lagrangian Stratospheric Balloon Flights. Part II: Results from the Vorcore Campaign in Antarctica. J. Atmos. Sci., 65, 3056–3070. https://doi.org/10.1175/2008JAS2710.1.

Hines, C. O. (1997). Doppler-spread parameterization of gravity-wave momentum deposition in the middle atmosphere. Part 1: Basic formulation. Journal of Atmospheric and Solar-Terrestrial Physics, 59(4), 371-386. https://doi.org/10.1016/S1364-6826(96)00079-X.

Holton, J. R.(1982), The role of gravity wave induced drag and diffusion in the momentum budget of the mesosphere, J. Atmos. Sci., 39, 791– 799, https://doi.org/10.1175/1520-0469(1982)039%3C0791:TROGWI%3E2.0.CO;2.

Li Z, Kovachki N, Azizzadenesheli K, Liu B, Bhattacharya K, Stuart A, Anandkumar A (2020). Fourier neural operator for parametric partial differential equations. https://doi.org/10.48550/arXiv.2010.08895.

Li, Z., Zheng, H., Kovachki, N., Jin, D., Chen, H., Liu, B., ... & Anandkumar, A. (2021). Physics-informed neural operator for learning partial differential equations. https://doi.org/10.48550/arXiv.2111.03794.

Liu H (2019), Quantifying gravity wave forcing using scale invariance, Nature Communications. https://doi.org/10.1038/s41467-019-10527-z.

Lindzen, R. S. (1981). Turbulence and stress owing to gravity wave and tidal breakdown. Journal of Geophysical Research: Oceans, 86(C10), 9707-9714. https://doi.org/10.1029/JC086iC10p09707.

Lund, T.S., Fritts, D.C., Wan, K., Laughman, B., Liu, H. (2021), Numerical simulation of mountain waves over the Southern Andes. Part 1: Mountain wave and secondary wave character, evolutions, and breaking. Journal of Atmospheric Sciences, 4337-4356, https://doi.org/10.1175/JAS-D-19-0356.1.

Maziar Raissi, Paris Perdikaris, and George E Karniadakis (2019). Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. Journal of Computational Physics, 378:686–70, https://doi.org/10.1016/j.jcp.2018.10.045.

Miyoshi, Y. and Yiğit, E. (2019), Impact of gravity wave drag on the thermospheric circulation: implementation of a nonlinear gravity wave parameterization in a whole-atmosphere model, Ann. Geophys., 37, 955–969, https://doi.org/10.5194/angeo-37-955-2019.

Kruse, C. G., J. H. Richter, M. J. Alexander, J. T. Bacmeister, C. Heale, and J. Wei: Gravity wave drag parameterizations for Earth's atmosphere. Chapter 11 of AGU Book Titled "Fast physics in large-scale models: parameterization, evaluation, and observations." Accepted 15 Dec 2021 in AGU Books. https://commons.erau.edu/cgi/viewcontent.cgi?article=3130&context=publication.

Garcia, R. R., Smith, A. K., Kinnison, D. E., Cámara, Á. d. l., & Murphy, D. J. (2017). Modification of the Gravity Wave Parameterization in the Whole Atmosphere Community Climate Model: Motivation and Results, Journal of the Atmospheric Sciences, 74(1), 275-291. Retrieved Feb 1, 2023, from https://journals.ametsoc.org/view/journals/atsc/74/1/jas-d-16-0104.1.xml.

Guibas, J., Mardani, M., Li, Z., Tao, A., Anandkumar, A., & Catanzaro, B. (2021). Adaptive fourier neural operators: Efficient token mixers for transformers. https://doi.org/10.48550/arXiv.2111.13587.

Pedatella, N. M. et al (2014). The neutral dynamics during the 2009 sudden stratosphere warming simulated by different whole atmosphere models. J. Geophys. Res. 119, 1306–1324.

Pathak J, Subramanian S, Harrington P, Raja S, Chattopadhyay A, Mardani M, Kurth T, Hall D, Li Z, Azizzadenesheli K, Hassanzadeh P (2022). FourCastNet: A global data-driven high-resolution weather model using adaptive fourier neural operators. https://doi.org/10.48550/arXiv.2202.11214.

Palmer, T. N., Shutts, G. J., & Swinbank, R. (1986). Alleviation of a systematic westerly bias in general circulation and numerical weather prediction models through an orographic gravity wave drag parametrization. Quarterly Journal of the Royal Meteorological Society, 112(474), 1001-1039. https://doi.org/10.1002/qj.49711247406.

Ribstein, B., Millet, C., Lott, F., & de la Cámara, A. (2022). Can we improve the realism of gravity wave parameterizations by imposing sources at all altitudes in the atmosphere? Journal of Advances in Modeling Earth Systems, 14, e2021MS002563. https://doi.org/10.1029/2021MS002563.

Stephan, C., Alexander, M. J., & Richter, J. H., (2016). Characteristics of Gravity Waves from Convection and Implications for Their Parameterization in Global Circulation Models. J. Atmos. Sci., 73, 2729-2742, https://doi.org/10.1175/JAS-D-15-0303.1.

Vosper, S. B. (2015). Mountain waves and wakes generated by South Georgia: Implications for drag parameterization. Quarterly Journal of the Royal Meteorological Society, 141(692), 2813– 2827. https://doi.org/10.1002/qj.2566.

Vosper, S. B., Brown, A. R., & Webster, S. (2016). Orographic drag on islands in the NWP mountain grey zone. Quarterly Journal of the Royal Meteorological Society, 142, 3128– 3137. https://doi.org/10.1002/qj.2894.

Wandel, N., Weinmann, M., Neidlin, M., & Klein, R. (2022). Spline-pinn: Approaching pdes without data using fast, physics-informed hermite-spline cnns. In Proceedings of the AAAI Conference on Artificial Intelligence (Vol. 36, No. 8, pp. 8529-8538).

Warner, C. D., & McIntyre, M. E. (1996). On the propagation and dissipation of gravity wave spectra through a realistic middle atmosphere. Journal of Atmospheric Sciences, 53(22), 3213-3235. https://doi.org/10.1175/1520-0469(1996)053%3C3213:OTPADO%3E2.0.CO;2.

Xiong, W., Huang, X., Zhang, Z., Deng, R., Sun, P., & Tian, Y. (2023). Koopman neural operator as a mesh-free solver of non-linear partial differential equations. https://doi.org/10.48550/arXiv.2301.10022.

Yiğit, E., Aylward, A. D., and Medvedev, A. S. (2008), Parameterization of the effects of vertically propagating gravity waves for thermosphere general circulation models: Sensitivity study, J. Geophys. Res., 113, D19106, https://doi.org/10.1029/2008JD010135.