A deep learning approach to extract internal tides scattered by geostrophic turbulence

Han Wang[1], Nicolas Grisouard[1], Hesam Salehipour[2], Alice Nuz[3], Michael Poon[1], and Aurelien L.S. Ponte[4]

[1]University of Toronto
[2]Autodesk Research
[3]New York University
[4]Ifremer

November 30, 2022

Abstract

A proper extraction of internal tidal signals is central to the interpretation of Sea Surface Height (SSH) data, yet challenging in upcoming satellite missions, where traditional harmonic analysis may break down at finer observed spatial scales known to contain significant wave-mean interactions. However, the wide swaths featured in such satellite missions render SSH snapshots that are spatially two-dimensional, which allows us to treat the tidal extraction as an image translation problem. We design and train a conditional Generative Adversarial Network, which, given a snapshot of raw SSH from an idealized numerical eddying simulation, generates a snapshot of the embedded tidal component. We test it on synthetic data whose dynamical regimes are different from the data provided during training. Despite the diversity and complexity of data, it accurately extracts tidal components in most individual snapshots considered and reproduces physically meaningful statistical properties.

A deep learning approach to extract internal tides scattered by geostrophic turbulence

Han Wang[1], Nicolas Grisouard[1], Hesam Salehipour[2], Alice Nuz[1*], Michael Poon[1†], and Aurélien L. Ponte[3]

[1]Department of Physics, University of Toronto, Ontario, Canada
[2]Autodesk Research, Ontario, Canada
[3]Ifremer, Plouzané, France

Key Points:

- A deep conditional Generative Adversarial Network is trained to extract tidal components in SSH snapshots.
- Training and testing data are from a set of idealized models where low mode internal tides propagate through a quasi-geostrophic jet.
- The network can extract tidal signals accurately in a snapshot whose underlying dynamics are different from training data.

*Current address, Tandon School of Engineering, New York University, New York, USA
†Current address, Department of Astronomy and Astrophysics, University of Toronto, Ontario, Canada

Corresponding author: Han Wang, hannnwangus@gmail.com

Abstract

A proper extraction of internal tidal signals is central to the interpretation of Sea Surface Height (SSH) data, yet challenging in upcoming satellite missions, where traditional harmonic analysis may break down at finer observed spatial scales known to contain significant wave-mean interactions. However, the wide swaths featured in such satellite missions render SSH snapshots that are spatially two-dimensional, which allows us to treat the tidal extraction as an image translation problem. We design and train a conditional Generative Adversarial Network, which, given a snapshot of raw SSH from an idealized numerical eddying simulation, generates a snapshot of the embedded tidal component. We test it on synthetic data whose dynamical regimes are different from the data provided during training. Despite the diversity and complexity of data, it accurately extracts tidal components in most individual snapshots considered and reproduces physically meaningful statistical properties.

Plain Language Summary

Wide-swath satellite observations of Sea Surface Height (SSH) data at high spatial resolutions will be available in abundance thanks to advances of instrumental technologies. Embedded in the observed SSH are internal tides, a dynamical component that plays a crucial role in ocean circulation. As they are entangled with background currents and eddies, such tidal signals are challenging to extract. Methods that worked with previous-generation altimeters will break down at the resolutions that the new generation promises. On the other hand, the wide satellite swaths provide new opportunities as they allow us to regard the observations as spatially two-dimensional. Here we treat the tidal extraction solely as an image translation problem. We train a deep neural net so that given a snapshot of a raw SSH signal, it produces a "fake" snapshot of the tidal SSH signal that is meant to reproduce the original. The data we use in this article is generated by idealized numerical simulations. Once adapted to realistic data, the network has the potential to become a new tidal extraction tool for satellite observations. More broadly, successes in our experiments can inspire other applications of generative networks to disentangle dynamical components in data where classical analysis may fail.

1 Introduction

Since the launch of TOPEX/Poseidon, oceanographers have used the geostrophic assumption to infer sea surface velocity from SSH. However, while an estimated 90% of the ocean's kinetic energy exists in the form of currents in quasigeostrophic balance (Ferrari & Wunsch, 2009) (hereafter qualified as "balanced"), one still must account for "unbalanced" flows, such as barotropic and baroclinic tides (also called internal tides, hereafter "ITs"), for a refined inference of balanced currents (Fu & Ferrari, 2008). Furthermore, baroclinic tides play a crucial role in ocean mixing (Lien & Gregg, 2001; Whalen et al., 2020), which impacts ocean circulations, and hence the ocean's role in climate change (Jithin & Francis, 2020). Therefore, whether ITs are considered "noise" (e.g., for inferring balanced flows) or "signal" (e.g., for tidally induced mixing), their proper extraction from altimetry data is essential.

For decades, the IT extraction has been conducted via harmonic analysis (Zaron & Rocha, 2018), a method that relies on a close phase relationship (or coherence) between ITs and astronomical forcings (departures from this condition is referred to as "incoherence" (Ponte & Klein, 2015)). Current altimetry has a typical spatial resolution of $O(100)$ km (Ballarotta et al., 2019), which is sufficient to retrieve mode-1 and some of the mode-2 IT wavelengths of semidiurnal tides, along with the dominant turbulent balanced motions (hereafter "TBMs") (Ray & Zaron, 2011). At these scales, the coupling between ITs and TBMs is usually weak and therefore substantial portions of the

ITs are coherent (Egbert & Ray, 2000). Hence, harmonic analysis is in principle sufficient to retrieve the corresponding IT signal.

The next generation of satellite altimetry, in particular the Surface Water Ocean Topography (SWOT) satellite mission, aims to improve the spatial resolutions of the measured data to at least a few tens of km in wavelength (Morrow et al., 2019). A fundamental challenge arises at these smaller scales, namely, the potential inapplicability of traditional harmonic analysis. Indeed, ITs become incoherent (Dunphy et al., 2017; Ponte & Klein, 2015; Dunphy & Lamb, 2014) due to stronger couplings with the TBMs linked to the increased vorticity magnitude (Bühler, 2014). Given the relatively long temporal gap between consecutive measurements of SWOT at the same location, the incoherent signal would be hard to identify using traditional harmonic analysis.

Future altimeters will gather data along wide swaths (two 50 km-wide swaths, 20 km apart in the case of SWOT) as opposed to current linear tracks and as a result they will produce spatially two-dimensional(2D) images. This has motivated the community to regard the extraction of IT signals as an operation on high-resolution 2D snapshots. Current methods rely on exploiting distinct spectral signatures of TBMs and internal waves (H. Torres et al., 2019), or on data assimilation techniques (Metref et al., 2020; Le Guillou et al., 2021).

In this work, we propose instead to regard the IT extraction solely as an image-to-image translation problem, conceiving and tackling the following challenge: *can we discover an algorithm that extracts the SSH signature induced by IT from a raw, instantaneous SSH map*? To answer this challenge, we develop what we call the "Toronto Internal Tide Emulator" (TITE), a deep convolutional neural network that extracts IT signals from individual SSH snapshots. No physical knowledge, statistical properties, or temporal evolution are imparted prior to the training. In general, we find TITE to perform well in most SSH snapshots generated from a set of idealized simulations. We present details about the dataset we use and the development of TITE in section 2, our experiments in section 3, and offer conclusions and discussions in section 4.

2 Methods

2.1 Idealized data supporting TITE's development

Data to support TITE's development are snapshots from a set of idealized numerical simulations, where mode-1 ITs are forced at a fixed tidal period T (12 hours) to propagate through TBMs created by a baroclinically unstable jet (Ponte & Klein, 2015; Ponte et al., 2020). The SSH signatures of TBMs in these simulations are generally larger than those induced by ITs, and exhibit a significant overlap in spatial scales at $O(100)$ km with ITs. Spatial filtering is thus difficult, an issue that is also faced by satellite altimetry in oceanic regions such as the Gulf Stream or Drake Passage, where powerful TBMs exist (Rocha et al., 2016; Richman et al., 2012).

We run the model under five different initial meridional density contrasts. With increasing contrast, the baroclinic jet becomes more unstable and creates a more vigorous baroclinic eddy field. The spectra induced by these eddies follow a geostrophic turbulence law (Ponte & Klein, 2015; Charney, 1971), and are thus identified as TBMs. In ascending order of stationary surface kinetic energy levels of TBM (hereafter referred to as "turbulence levels"), we label the five simulations as T1 to T5. See Text S1 in Supporting Information for more details on the numerical setup. IT snapshots are computed online via harmonic fits over time series that are $2T$ long and sampled every 300 seconds, or $T/144$. For simplicity, we only study $\eta_{\cos}^{(\mathrm{sim})}$, the cosine component of ITs from the simulations, defined as

$$\eta_{\cos}^{(\mathrm{sim})}(x,y,t) = \frac{1}{T}\int_{t-2T}^{t} \eta(x,y,t')\cos\left(\frac{2\pi}{T}t'\right)\mathrm{d}t', \tag{1}$$

where x, y are the zonal and meridional coordinates, respectively, and η denotes raw SSH. For each snapshot, we cut out three square panels covering three fixed latitudinal bands, labeled as "down-jet", "mid-jet" and "up-jet" bands, as illustrated in Fig. 1. One hundred snapshots are captured every $4T$ for each simulation in T1-5, resulting in 1500 pairs of $\left\{\eta,\ \eta_{\cos}^{(\mathrm{sim})}\right\}$ panels (5 runs, 3 latitudinal bands, and 100 snapshots) altogether.

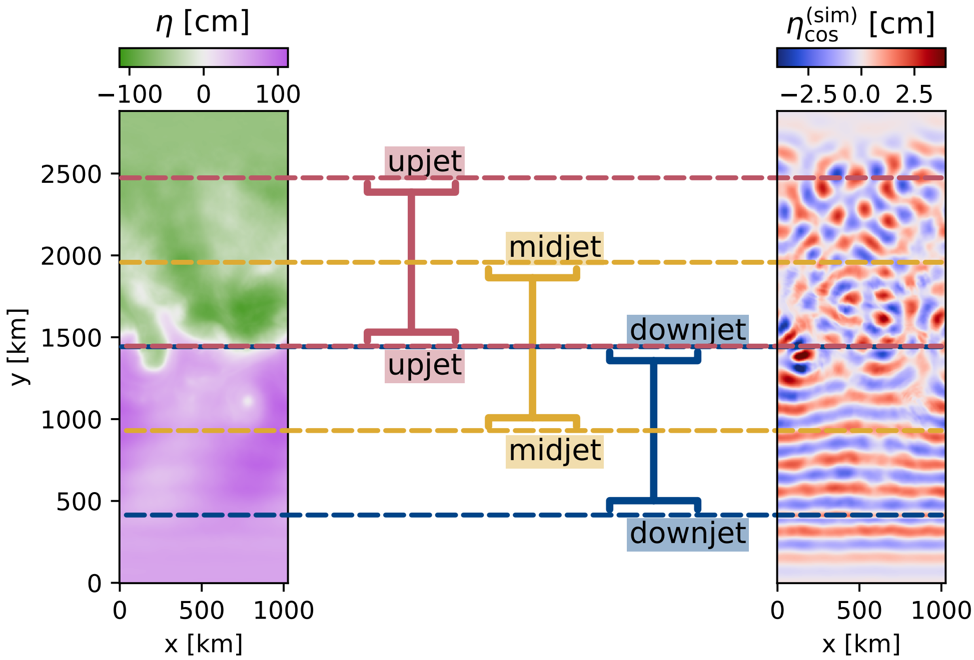


Figure 1. The "down-jet", "mid-jet" and "up-jet" bands plotted over a snapshot of η (left) and $\eta_{\cos}^{(\mathrm{sim})}$ (right), sampled from T3 at day 2120. The "mid-jet" band is centred around the baroclinic jet. ITs are forced to the south of "up-jet" bands, and as the ITs propagates northward and loses coherence due to interactions with the TBM, the $\eta_{\cos}^{(\mathrm{sim})}$ patterns are less reminiscent of plane waves in the "down-jet" band than in the "up-jet" band.

2.2 Deep-learning algorithm designed to extract tidal signals

During the design of the TITE runs, we implicitly apply four assumptions: (1) there is abundant spatial information, (2) all snapshots are statistically independent from each other, (3) a raw SSH functionally determines its IT component, but properties of the functional dependence are unknown, and (4) there exists abundant data where ITs are already extracted from the raw SSH. Discussions about these assumptions are included at the end of article.

TITE is based on a popular conditional Generative Adversarial Network (hereafter referred to as "cGAN") (Isola et al., 2017). As the name implies, a cGAN consists of two parts, namely, a conditional generator (hereafter "generator") that learns how to manufacture a "fake" image that's conditioned on an "input image", and a discriminator that tries to determine if an image is "genuine" (i.e., paired to the input image in the training data), or fake (i.e., created by the generator). Either part is on its own a convolutional neural network, and during training, the two parts compete against each other to co-evolve (Mirza & Osindero, 2014; Goodfellow et al., 2014). We denote the cosine IT panels generated from TITE as $\eta_{\cos}^{(\mathrm{gen})}$; following our notations, the *input* image would be η, the *genuine* image would be $\eta_{\cos}^{(\mathrm{sim})}$, and the *fake* image would be $\eta_{\cos}^{(\mathrm{gen})}$. As reflected in this general workflow, during training, other than the paired panels, no further information is given to TITE.

The particular cGAN we adapt to TITE is called "pix2pix" (Isola et al., 2017), applications of which range from artistic creations (ml4a, 2017) to scientific problems such as remote sensing image classifications (Lebedev et al., 2018). Our codes are adapted from the code downloaded from TensorFlow Tutorials (Tensorflow, n.d.). We refer to the original publication for details of pix2pix (Isola et al., 2017), and to Text S7 in Supporting Information for details on the changes we made to the original codes. Here, we mention a few relevant traits.

The generator and the discriminator have around 10^4 and 2000 convolutional layers respectively, each layer containing a 2-by-2 kernel to be learned during training. The considerable number of model parameters makes TITE a black box, as in the case of many deep learning algorithms.

Prior to each epoch, training images are randomly reshuffled in time, cropped, flipped, and rotated. Here, an epoch means the duration it takes for the cGAN to iterate over all data in the training set once. The random cropping, rotation and flipping are intended to roughly mimic realistic situations where we don't have a priori knowledge of the observer's orientation/location about IT generation sites and direction of propagation. By randomly reshuffling in time, we enforce that every panel pair at every snapshot in the simulation be sequentially independent from the others. This means that any temporal information in the simulations is unknown to the pix2pix kernel, in line with our assumption (2) made previously in this section.

As the fully convolutional U-Net structure inherited from pix2pix (Isola et al., 2017) in the generator can be applied to images of arbitrary sizes in principle, when producing Movies S1 and S2 in Supporting Information , we directly apply the trained TITE onto rectangular input images, even though TITE is trained on square images illustrated in Fig. 1. This versatility on the shapes of input images would be useful for along-swath satellite products.

We systematically run our code with TensorFlow 2.3.0 under Python 3.7. One hundred training epochs with 960 pairs of $\left\{\eta, \eta_{\cos}^{(\mathrm{sim})}\right\}$ in the training set take about 1.5 hours with a NVIDIA GP100 GPU. For all the TITE runs in the article, we choose to present the results after 600 training epochs. Details on how we decide on the cut-off epoch are provided in Text S4 in Supporting Information.

2.3 Division of data to training, testing and validation sets

As a first check on whether TITE could achieve any success at all, we randomly select 20% of all 1500 pairs of $\left\{\eta, \eta_{\cos}^{(\mathrm{sim})}\right\}$ panels from T1-5 to form a so-called validation set, and use the rest as the training set. During training, TITE has access to all pairs of $\left\{\eta, \eta_{\cos}^{(\mathrm{sim})}\right\}$ in the training set, but none from the validation set. After 600 epochs, the training phase is over, and we apply the trained TITE into snapshots in the validation set. The mean correlation between $\eta_{\cos}^{(\mathrm{sim})}$ and $\eta_{\cos}^{(\mathrm{gen})}$ in the validation set turns out to be 0.85, which suggests that the generated $\eta_{\cos}^{(\mathrm{gen})}$ reasonably resemble the ground truths $\eta_{\cos}^{(\mathrm{sim})}$. However, under this division, the training set contains turbulence levels that are statistically similar to the validation set on which the trained TITE is applied, and the good

Table 1. Mean correlation factors of validation and test sets in the ET1-5 runs†.

TITE run	Validation set, all	Test set, all	Test set, down-jet	Test set, mid-jet	Test set, up-jet
ET1	0.86	0.91	0.92	0.90	0.92
ET2	0.85	0.89	0.90	0.87	0.90
ET3	0.84	0.83	0.82	0.79	0.88
ET4	0.85	0.80	0.77	0.75	0.87
ET5	0.87	0.70	0.62	0.63	0.84

†The second and third columns present mean correlation factors averaged over all panels in the validation sets and test sets respectively. The last three columns present mean correlation factors averaged over down-jet, mid-jet, and up-jet bands in the test sets respectively.

correlation factors could be caused by overfitting. To address this possibility, we challenge TITE to extract $\eta_{\cos}^{(\mathrm{sim})}$ signals linked to a different turbulence level as those employed for its training.

Specifically, in what we refer to as the "ET1 run", we reserve a *test* set, which contains all 300 pairs of panels from the simulation T1 and *none* from T2, T3, T4 or T5. Among the remaining panels from T2-5, we randomly select 80% pairs for the training set, and reserve the other 20% for the validation set. The validation and test sets are both inaccessible to TITE during training, but crucially, in terms of average turbulence levels, the training set is similar to the validation set, yet *different* from the test set. Similarly, we carry out ET2-5 runs, following the same logic, where the test sets are panels from the simulations T2-5 respectively.

3 Performance of TITE

In this section, we evaluate the performance of TITE from several statistical metrics and we discuss the causes of relatively decreased performance when they arise. All metrics are computed using standard methods and detailed in Text S6 in Supporting Information .

We first investigate how close $\eta_{\cos}^{(\mathrm{gen})}$ is to the ground truth $\eta_{\cos}^{(\mathrm{sim})}$ by measuring the correlation between the two. The mean correlation factors in the test and validation sets of the ET1-5 runs are listed in Table 1 (first three columns). The highly correlated predictions of TITE in the test set in ET1-4 are especially interesting, as turbulence levels of the test set are different from that of the training set. There is however a relatively sharper drop in the mean correlation from ET4 to ET5.

The test instances associated with the highest and lowest correlations among ET1-5 are presented in Fig. 2. In the test instance with the highest (lowest) correlation that belongs to ET1 (ET5), the ratio between the root mean square of $\left(\eta_{\cos}^{(\mathrm{sim})} - \eta_{\cos}^{(\mathrm{gen})}\right)$ and the root mean square of $\eta_{\cos}^{(\mathrm{sim})}$ is 0.12 (4.77). In Movie S1 in Supporting Information, we re-order all the shuffled test instances of ET1 in time. Considering that the snapshots are randomly shuffled and hence the temporal evolution of these images is unknown to TITE, this reconstructed temporal continuity is remarkable. Nevertheless, for the strongly turbulent flows of T5 that ET5 tests, the evolution of $\eta_{\cos}^{(\mathrm{gen})}$ bears little semblance to $\eta_{\cos}^{(\mathrm{sim})}$ (Movie S2 in Supporting Information). This observation, together with the lower correlation factors of ET5 (Table 1), suggest a categoric difference between ET5 and ET1-ET4.

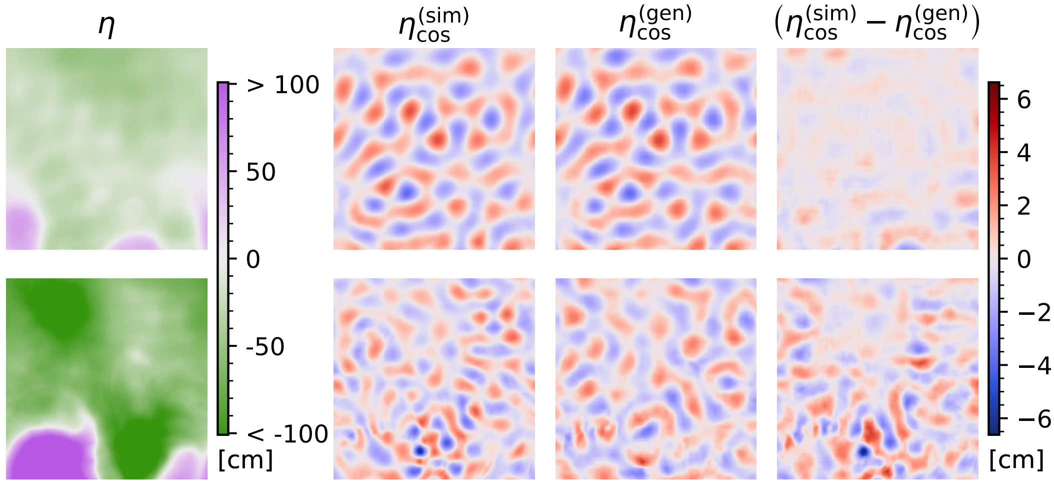


Figure 2. Individual tests with the highest and lowest correlations. For legibility reasons, we omit spatial axis labels, see fig. 1 for their definitions. The upper row corresponds to the test instance that has the highest correlation among the ET1–ET5 runs. It belongs to the ET1 run and has a correlation factor of 0.95. The lower row corresponds to the test instance with the lowest correlation. It belongs to the ET5 run and has a correlation factor of 0.4.

To gain more insight about the relative failures in ET5, we conduct a spectral analysis that focuses on comparing ET4 and ET5. The wavenumber spectra for the down-jet and up-jet bands are computed separately for $\eta_{\cos}^{(\mathrm{sim})}$ and $\eta_{\cos}^{(\mathrm{gen})}$ in the test set of ET4 and ET5, and presented in Fig. 3. The spectra for mid-jet bands are omitted for readability here and attached in Text S2 in Supporting Information .

Prominent bumps appear near the wavenumbers corresponding to mode-1 tidal wavelengths (See Text S1 in Supporting Information) in all the spectra of $\eta_{\cos}^{(\mathrm{sim})}$ (Solid lines in Fig. 3). These bumps are somewhat broad, and their locations are noticeably different between the down-jet and up-jet bands. This is expected, as the density profiles and the Coriolis parameter both vary with latitude, which modulates the mode-1 tidal wavelength (See Text S1 and Fig. S1 in Supporting Information). Such variations can be found in satellite observations too (Ray & Zaron, 2011). Interestingly, in ET4, the locations of spectral bumps in the $\eta_{\cos}^{(\mathrm{gen})}$ spectra also vary between the down-jet and up-jet bands, in a manner such that they closely overlap with bumps of the $\eta_{\cos}^{(\mathrm{sim})}$ spectra at both bands. This implies that in the ET4 run, the trained TITE identifies the dominant wavelength even as it varies. In other words, TITE can identify patterns at varying spatial scales.

In the ET5 run, the $\eta_{\cos}^{(\mathrm{gen})}$ spectra fail to trace the location of the bumps in the down-jet bands, which is qualitatively different from ET4. The performance in up-jet bands appears as good as ET4, which may be attributed to the fact that the mode-1 tidal wavelengths to the south of the jets are the same in all five simulations.

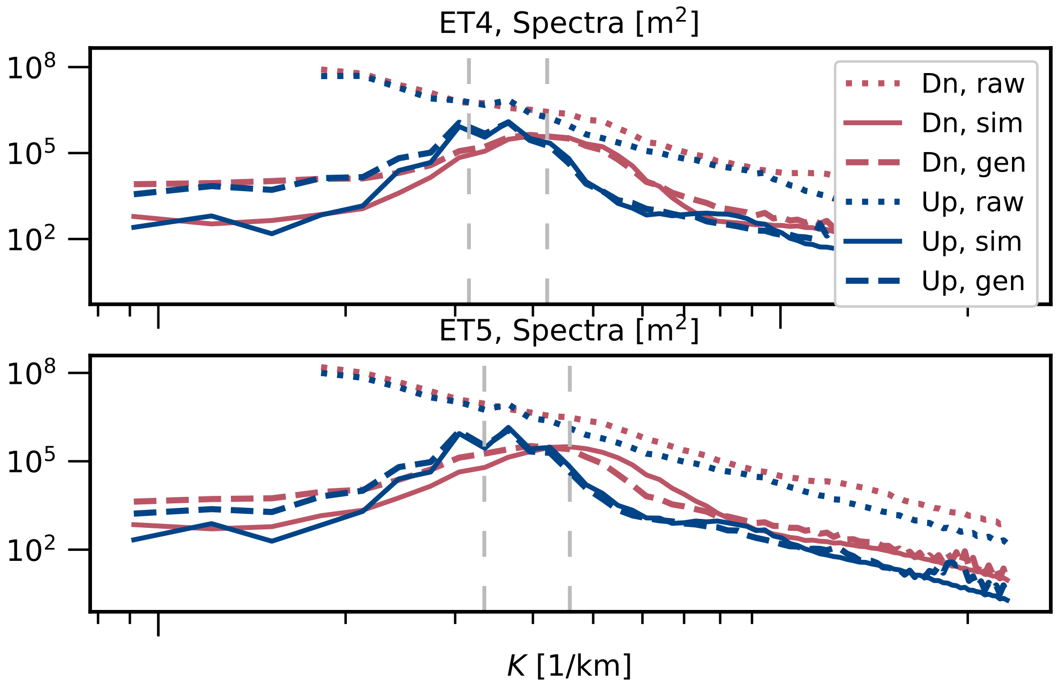


Figure 3. Spectra for the down-jet and up-jet bands in ET4 and ET5 test set. In the legends "Up", "Dn" denote the down-jet and up-jet bands respectively. "raw", "sim", and "gen" denote spectra computed from panels of $\eta, \eta_{\cos}^{(\text{sim})}$, and $\eta_{\cos}^{(\text{gen})}$, respectively. "$K$" denotes the horizontal wavenumber magnitude. The vertical dashed lines mark the largest and smallest mode-1 tidal wavenumbers over the simulation domain at initial time, following Figure S1 in Supporting Information. Raw spectra higher than $2 \times 10^8 \text{m}^2$ at large scales are omitted. Higher wavenumbers are omitted.

One might be tempted to think that overfitting is the cause of the good performance in ET1-4, and vice-versa when the performance decreases in ET5. Indeed, as listed in Table S1 in Supporting Information, the kinetic energy and normalized vorticity (absolute values of surface vorticities normalized by the local Coriolis frequency) for the TBM and IT all increase from T1 to T5, and in terms of these dynamical metrics, the training set of ET5 is less diverse compared to, say, the training set of ET4 that spans a wider range of these metrics. This explanation based on overfitting is also consistent with the fact that the ET5 run has the highest mean correlation in the validation set (second column in Table 1).

However, if overfitting was the only factor, then TITE should perform poorly in the ET1 test set too, which is not the case. In fact, the ET1 run produces the best mean correlation in the test set among ET1-5; in Text S2 in Supporting Information, we show that the ET1 test set also demonstrates excellent spectral behaviours. Moreover, the mean correlations in the *test* sets are *higher* than in the *validation* sets in ET1 and ET2 (Table 1). Therefore, we postulate that a more crucial factor at play is the turbulence levels of the data themselves: higher turbulence levels appear to decrease TITE's prediction accuracies. In the ET1 test set, the turbulence levels are lower, and TITE performs well despite the possible impacts from overfitting. In the ET1 and ET2 runs, the test data are at a lower turbulence level than the validation data, and TITE generates better predictions in the test sets than in the validation sets, even though the training set includes the turbulence levels in the validation set and excludes the turbulence levels in the test set.

It is not too surprising that higher turbulence levels make the IT extraction more challenging. As explained in Text S1 in Supporting Information , stronger scatterings of ITs from TBMs induce more longitudinal variations as well as small-scale features in the IT components. In addition, the tidal wavelengths vary more latitudinally due to increased density gradients, which increases the diversity of dominant spatial scales of IT signals across the domain and time. Both factors add complexities to the η and $\eta_{\cos}^{(\mathrm{sim})}$ patterns. In Text S5 in Supporting Information, we show that a generically defined metric of pattern complexities introduced by Bagrov et al. (2020) generally increases under stronger TBMs as we expected.

The difficulty associated with vigorous turbulence levels is also reflected in the relatively worse performance of TITE in the mid-jet bands centered around the turbulence. In the last three columns of Table 1, the correlations for the down-jet, mid-jet and up-jet bands are presented separately for the test sets in ET1-5. Within each of ET1-5, the up-jet bands have a higher mean correlation than the mid-jet bands. As the turbulence level increases, this difference gets more pronounced. The degraded performance at mid-jet bands is also reflected from the "square coherences" in Text S2 in Supporting Information.

We note that despite the relative lack of prediction accuracy under higher turbulence levels, in our data, TITE would still outperform simple spatial filtering methods that would break down due to the strong TBMs superimposing the ITs around tidal wavelengths (Text S2 in Supporting Information), or harmonic analysis that would not work due to the strong incoherence and the temporal interval of $4T$.

4 Conclusions and Discussions

We designed a novel technique based on a deep neural network algorithm to extract internal tides that are entangled with geostrophic turbulence. We trained and validated TITE using randomly shuffled simulation snapshots that were categorically different from the dynamic regime of the testing data. The testing data sets are designed in a way that classical methods such as harmonic fits or spectral filtering could not extract tidal signals accurately, and yet in most test cases, TITE can still 1) extract IT signals that agree well with ground truths in a deterministic sense, and 2) capture the dominant tidal energy in the wavenumber spectra, even when it varies temporally and latitudinally. When TITE does not perform as well, the main cause seems to be the high complexities of the patterns linked to stronger turbulent motions. Overall, we believe that this work provides a fresh angle on how to disentangle dynamical components from two-dimensional data via a deep learning approach. Some discussions are offered below.

Although we make no claim about TITE or cGANs in general as being the best possible algorithms to specifically achieve our goal, we found it superior to other deep learning methods we investigated, which include several types of decision trees regressors, long short-term memory networks, and U-Net structures without a discriminator. We did not attempt to optimize model parameters such as numbers of layers or learning rates, among others. More recent variations of pix2pix such as pix2pixHD (Park et al., 2019) could also outperform our current implementation. Moreover, as mentioned in section 3, the generated images always contain spurious signals outside the dominant tidal bump, which remains to be resolved. We leave these as thoughts for future work.

In this work, TITE only extracts the cosine IT signals. The generalization to the sinusoidal IT signals, which are defined by replacing $\cos(2\pi t'/T)$ in equation (1) with $\sin(2\pi t'/T)$, should be straightforward. With both cosine and sinusoidal IT signals, phase information can be retrieved. One may also study the performance of TITE for extractions of signals at higher tidal frequencies that correspond to smaller spatial scales. Pix2pix has been observed to be capable of capturing fine features in images (Isola et al., 2017), and smaller scales don't necessarily make the problem more challenging to TITE.

So far, TITE has only been developed by the idealized simulations T1-T5 with a single baroclinic jet and single tidal frequency, simplistic boundary conditions, flat topography, an absence of air. As an ongoing work, we are investigating the effects of including snapshots from a global ocean GCM.

With SWOT in mind, we may reassess the four assumptions stated in section 2.2. All images used in this work have a 4 km horizontal resolution that resolves the tides adequately, addressing assumption (1). In preparation for satellite data that suffer from measurement noises and more limited resolutions, we may coarse-grain and augment the training data with the type of noises expected in SWOT (Gaultier et al., 2016) and investigate the impacts. Assumption (2), motivated by the incoherence of ITs and the relatively long sampling intervals of SWOT, is satisfied by the design of the TITE architecture, and by the frequent random shuffling of snapshots during training. However, complete statistical independence between ITs and TBMs can be overly strict for several reasons, ranging from a higher temporal sampling at high latitude, to the possibility of "filling in the time gaps" with other sources of data such as those from assimilated models or in-situ instruments (d'Ovidio et al., 2019). From the overall satisfactory performance of TITE, the assumption (3) appears to be satisfied in our simulation outputs, perhaps due to simplistic simulation settings, such as a perfectly harmonic incoming IT signal, or simple boundary conditions. Under more realistic configurations, a functional dependence might not be guaranteed. On the other hand, the assumption (3) can also be overly strict, considering recent progress in the theory of IT/TBM interactions (H. S. Torres et al., 2018; Savva & Vanneste, 2018; Savva et al., 2021). The assumption (4) relies on the premise that there will be pre-processed training data (presumably from highly skilled model outputs) that mimic the dynamics to be sampled by SWOT. Productions of such data are receiving significant attention within the modelling communities (Zaron & Rocha, 2018; Rocha et al., 2016; Arbic et al., 2010; Shchepetkin & McWilliams, 2005; Savage et al., 2017). Overall, to make TITE eventually applicable to SWOT and other satellite missions in the future, more work is required, especially in coordination with different communities.

Acknowledgments

H.W., A.N., M.P and N.G. acknowledge the financial support of the Canadian Space Agency [14SUSWOTTO] and of the Natural Sciences and Engineering Research Council of Canada (NSERC) [RGPIN-2015-03684]. A.L.P was supported by ANR project number 17-CE01-0006-01 and by project DIEGO of the CNES/TOSCA SWOT program. This work was made possible by the facilities of the Shared Hierarchical Academic Research Computing Network (SHARCNET: www.sharcnet.ca) and Compute/Calcul Canada. The work has benefited from scientific discussions with Laure Zanna, Brian Arbic, Michael W. Lever, and Ilia A. Iakovlev.

Simulation results used in this study to train and test TITE are published on Scholars Portal Dataverse (Ponte et al., 2020). Codes defining the architecture of TITE are available on Github via link `https://github.com/hannnwang/Pix2Pix_TITE_examples`.

References

Arbic, B. K., Wallcraft, A. J., & Metzger, E. J. (2010). Concurrent simulation of the eddying general circulation and tides in a global ocean model. *Ocean Modelling*, *32*(3-4), 175–187.

Bagrov, A. A., Iakovlev, I. A., Iliasov, A. A., Katsnelson, M. I., & Mazurenko, V. V.

(2020). Multiscale structural complexity of natural patterns. *Proceedings of the National Academy of Sciences*, *117*(48), 30241–30251.

Ballarotta, M., Ubelmann, C., Pujol, M.-I., Taburet, G., Fournier, F., Legeais, J.-F., ... others (2019). On the resolutions of ocean altimetry maps. *Ocean Science*, *15*(4), 1091–1109.

Bühler, O. (2014). *Waves and Mean Flows* (2nd ed.). Cambridge: Cambridge University Press. Retrieved from `http://ebooks.cambridge.org/ref/id/CBO9781107478701` doi: 10.1017/CBO9781107478701

Charney, J. G. (1971). Geostrophic turbulence. *Journal of the Atmospheric Sciences*, *28*(6), 1087–1095.

Dunphy, M., & Lamb, K. G. (2014). Focusing and vertical mode scattering of the first mode internal tide by mesoscale eddy interaction. *Journal of Geophysical Research: Oceans*, *119*(1), 523–536.

Dunphy, M., Ponte, A. L., Klein, P., & Le Gentil, S. (2017). Low-mode internal tide propagation in a turbulent eddy field. *Journal of Physical Oceanography*, *47*(3), 649–665.

d'Ovidio, F., Pascual, A., Wang, J., Doglioli, A. M., Jing, Z., Moreau, S., ... others (2019). Frontiers in fine-scale in situ studies: Opportunities during the swot fast sampling phase. *Frontiers in Marine Science*, *6*, 168.

Egbert, G., & Ray, R. (2000). Significant dissipation of tidal energy in the deep ocean inferred from satellite altimeter data. *Nature*, *405*(6788), 775–778.

Ferrari, R., & Wunsch, C. (2009). Ocean circulation kinetic energy: Reservoirs, sources, and sinks. *Annual Review of Fluid Mechanics*, *41*, 253–282.

Fu, L.-L., & Ferrari, R. (2008). Observing oceanic submesoscale processes from space. *Eos, Transactions American Geophysical Union*, *89*(48), 488–488.

Gaultier, L., Ubelmann, C., & Fu, L.-L. (2016). The challenge of using future swot data for oceanic field reconstruction. *Journal of Atmospheric and Oceanic Technology*, *33*(1), 119–126.

Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., ... Bengio, Y. (2014). Generative adversarial nets. *Advances in neural information processing systems*, *27*.

Isola, P., Zhu, J.-Y., Zhou, T., & Efros, A. A. (2017). Image-to-image translation with conditional adversarial networks. In *Proceedings of the ieee conference on computer vision and pattern recognition* (pp. 1125–1134).

Jithin, A., & Francis, P. (2020). Role of internal tide mixing in keeping the deep andaman sea warmer than the bay of bengal. *Scientific reports*, *10*(1), 1–10.

Lebedev, M., Vizilter, Y. V., Vygolov, O., Knyaz, V., & Rubis, A. Y. (2018). Change detection in remote sensing images using conditional adversarial networks. *International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences*, *42*(2).

Le Guillou, F., Metref, S., Cosme, E., Ubelmann, C., Ballarotta, M., Le Sommer, J., & Verron, J. (2021). Mapping altimetry in the forthcoming swot era by back-and-forth nudging a one-layer quasigeostrophic model. *Journal of Atmospheric and Oceanic Technology*, *38*(4), 697–710.

Lien, R.-C., & Gregg, M. (2001). Observations of turbulence in a tidal beam and across a coastal ridge. *Journal of Geophysical Research: Oceans*, *106*(C3), 4575–4591.

Metref, S., Cosme, E., Le Guillou, F., Le Sommer, J., Brankart, J.-M., & Verron, J. (2020). Wide-swath altimetric satellite data assimilation with correlated-error reduction. *Frontiers in Marine Science*, *6*, 822.

Mirza, M., & Osindero, S. (2014). Conditional generative adversarial nets. *arXiv preprint arXiv:1411.1784*.

ml4a. (2017). *Machine Learning for Artists.* Retrieved from `https://ml4a.github.io/guides/Pix2Pix/`

Morrow, R., Fu, L.-L., Ardhuin, F., Benkiran, M., Chapron, B., Cosme, E., ... others (2019). Global observations of fine-scale ocean surface topography with the surface water and ocean topography (swot) mission. *Frontiers in Marine Science*, *6*, 232.

Park, T., Liu, M.-Y., Wang, T.-C., & Zhu, J.-Y. (2019). Semantic image synthesis with spatially-adaptive normalization. In *Proceedings of the ieee/cvf conference on computer vision and pattern recognition* (pp. 2337–2346).

Ponte, A. L., & Klein, P. (2015). Incoherent signature of internal tides on sea level in idealized numerical simulations. *Geophysical Research Letters*, *42*(5), 1520–1526.

Ponte, A. L., Le Gentil, S., & Grisouard, N. (2020). *Internal tides and geostrophic turbulence in a Boussinesq re-entrant channel.* Scholars Portal Dataverse. Retrieved from `https://doi.org/10.5683/SP2/HU58SG` doi: 10.5683/SP2/HU58SG

Ray, R. D., & Zaron, E. D. (2011). Non-stationary internal tides observed with satellite altimetry. *Geophysical Research Letters*, *38*(17).

Richman, J. G., Arbic, B. K., Shriver, J. F., Metzger, E. J., & Wallcraft, A. J. (2012). Inferring dynamics from the wavenumber spectra of an eddying global ocean model with embedded tides. *Journal of Geophysical Research: Oceans*, *117*(C12).

Rocha, C. B., Chereskin, T. K., Gille, S. T., & Menemenlis, D. (2016). Mesoscale to submesoscale wavenumber spectra in drake passage. *Journal of Physical Oceanography*, *46*(2), 601–620.

Savage, A. C., Arbic, B. K., Alford, M. H., Ansong, J. K., Farrar, J. T., Menemenlis, D., ... others (2017). Spectral decomposition of internal gravity wave sea surface height in global models. *Journal of Geophysical Research: Oceans*, *122*(10), 7803–7821.

Savva, M. A., Kafiabad, H. A., & Vanneste, J. (2021, jun). Inertia-gravity-wave scattering by three-dimensional geostrophic turbulence. *J. Fluid Mech.*, *916*, A6. Retrieved from `https://www.cambridge.org/core/product/identifier/S0022112021002056/type/journal{\_}article` doi: 10.1017/jfm.2021.205

Savva, M. A., & Vanneste, J. (2018). Scattering of internal tides by barotropic quasigeostrophic flows. *Journal of Fluid Mechanics*, *856*, 504–530.

Shchepetkin, A. F., & McWilliams, J. C. (2005). The regional oceanic modeling system (roms): a split-explicit, free-surface, topography-following-coordinate oceanic model. *Ocean modelling*, *9*(4), 347–404.

Tensorflow. (n.d.). *pix2pix: Image-to-image translation with a conditional GAN.* Retrieved from `https://www.tensorflow.org/tutorials/generative/pix2pix` (Accessed on Sep.20, 2020.)

Torres, H., Klein, P., Siegelman, L., Qiu, B., Chen, S., Ubelmann, C., ... Fu, L.-L. (2019). Diagnosing ocean-wave-turbulence interactions from space. *Geophysical Research Letters*, *46*(15), 8933–8942.

Torres, H. S., Klein, P., Menemenlis, D., Qiu, B., Su, Z., Wang, J., ... Fu, L.-L. (2018). Partitioning ocean motions into balanced motions and internal gravity waves: A modeling study in anticipation of future space missions. *Journal of Geophysical Research: Oceans*, *123*(11), 8084–8105.

Whalen, C. B., de Lavergne, C., Garabato, A. C. N., Klymak, J. M., Mackinnon, J. A., & Sheen, K. L. (2020). Internal wave-driven mixing: governing processes and consequences for climate. *Nature Reviews Earth & Environment*, *1*(11), 606–621.

Zaron, E. D., & Rocha, C. B. (2018). Internal gravity waves and meso/submesoscale currents in the ocean: anticipating high-resolution observations from the swot swath altimeter mission. *Bulletin of the American Meteorological Society*, *99*(9), ES155–ES157.

AGU PUBLICATIONS

Geophysical Research Letters

Supporting Information for

A deep learning approach to extract internal tides scattered by geostrophic turbulence

Han Wang[1], Hesam Salehipour[2], Alice Nuz[1], Michael Poon[1], Aurélien L. Ponte[5], and Nicolas Grisouard[1]

[1]Department of physics, University of Toronto, Ontario, Canada

[2]Autodesk Research, Ontario, Canada

[5]Ifremer, Plouzané, France

Contents of this file

Text S1 to S6

Figures S1 to S8

Tables S1 to S2

Additional Supporting Information (Files uploaded separately)

Captions for Movies S1 to S6

Text S1. Numerical simulations to produce snapshots in T1-5

The T1-5 simulations are based on a beta-plane centred around 45°N. The mode-1 IT is forced to the south of a baroclinically unstable jet centred in the middle section of the computational domain, and propagates northwards[1,2]. All snapshots used in the development of TITE are publicly available, per the Data Availability Statement. Here, we summarize the relevant features of the simulations.

The baroclinically unstable jet is simulated in a zonal beta-plane channel centred at 45°N based on the primitive equation code CROCO (https://www.croco-ocean.org, v1628). Initial density profiles are different at the northern and southern ends of the domain. During a spin-up phase, the associated initial meridional density gradient undergoes geostrophic adjustment, eventually creating a zonal jet in thermal wind balance. This jet is baroclinically unstable and a zonal perturbation triggers this destabilization, resulting in low-frequency TBMs that we can reasonably describe as quasi-geostrophic[1]. Subsequently, relaxation towards unperturbed initial conditions maintains the TBMs. Statistical equilibrium is reached after O(100 days).

Starting at day 2000, a zonally uniform mode-1 internal tide of 12-hour period is forced within a narrow area south of the jet. Outgoing internal tides are damped in regions extending by 300 km from the southern and northern boundaries to prevent reflections back into the domain. Zonally, periodic conditions are enforced. All snapshots included in T1-5 are captured starting at day 2100. The latitudes covered by the three panels shown in Figure 1 in the main text are sufficiently away from the IT-radiating and damping regions at the southern and northern ends of the domain.

To create different levels of turbulent energy in T1 to T5, the meridional initial density gradient is modulated by changing the northern profile[2]. The TBM components are extracted online via a sliding average, replacing $\cos\left(\frac{2\pi}{T}t'\right)$ in equation (1) in the main text with a constant factor of $\frac{1}{2}$.

We compute the normalized vorticity and horizontal surface kinetic energy for the TBM components, as presented in the first three columns in Supporting Information(SI) Table S1. The TBM normalized vorticity and kinetic energy increase significantly from T1 to T5, with the normalized vorticity well bounded by 0.2.

Even though the wave amplitudes forced to the south of the turbulent jet are the same in T1-5, the IT energetics are different between simulations, due to different strengths of scattering from interactions with the TBMs. We extract the cosine IT components of surface velocities by replacing η with surface velocity components in equation (1) and compute the corresponding normalized vorticity and horizontal kinetic energy, as listed in the last two columns in SI Table S1. From T1 to T4, the kinetic energy increases. However, the kinetic energy stays about the same from T4 to T5 while getting more concentrated at smaller scales, as suggested by the increase in their respective normalized vorticities.

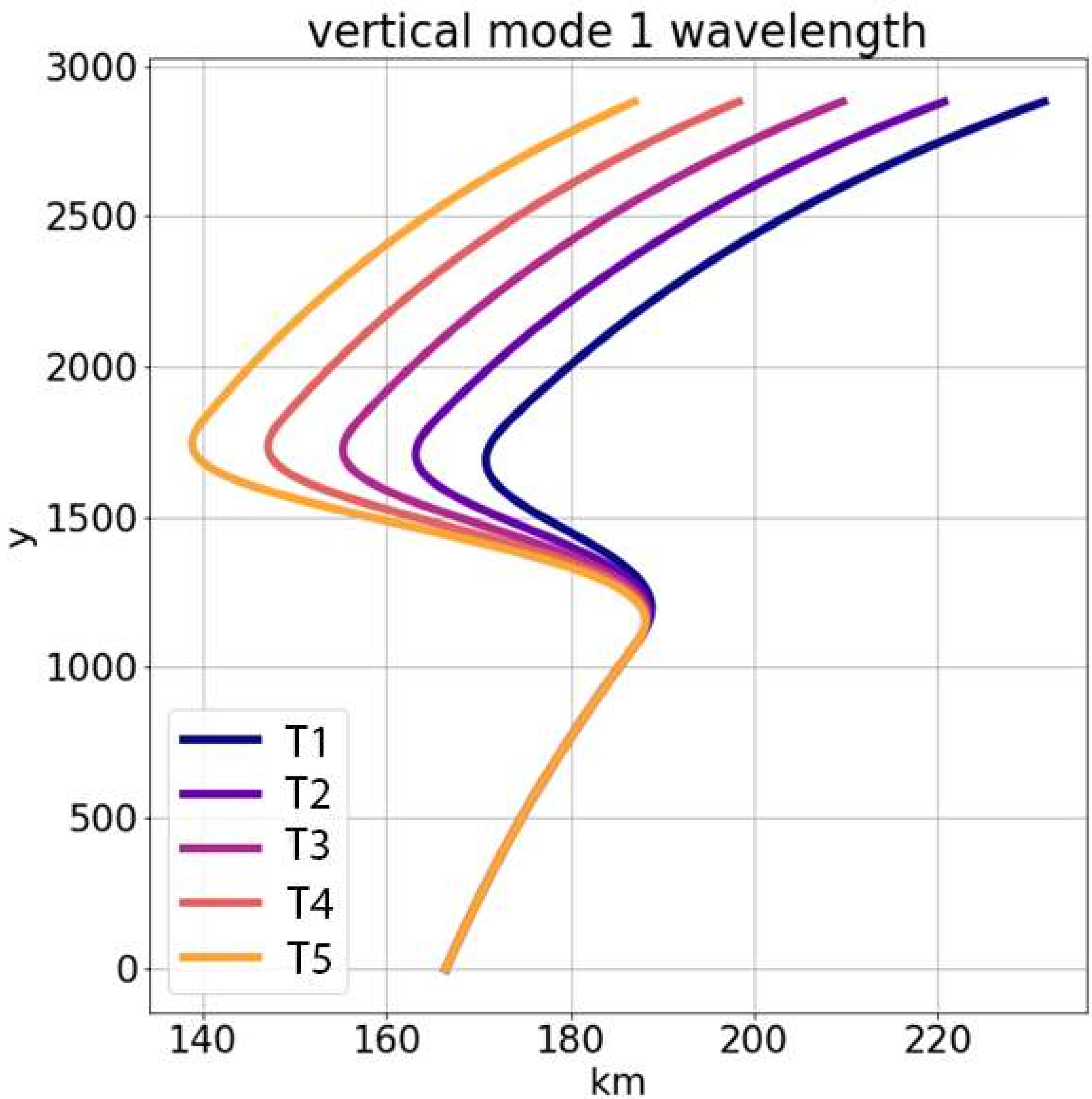
Moreover, the scattering from jets also makes the signals less coherent, as inspected closely by Ponte and Klein[3]. In SI Movie S3, we present snapshots of normalized vorticities of the TBMs along with the $\eta_{\cos}^{(\mathrm{sim})}$components in T1-5. There, we can see that as the TBMs become increasingly energetic from T1 to T5, the IT signals are scattered more around or to the north of the jet. As a result, the $\eta_{\cos}^{(\mathrm{sim})}$ patterns are less like plane waves and contain more small-scale features. This factor adds to the complexity of the patterns of η and $\eta_{\cos}^{(\mathrm{sim})}$.

In T1-5, the ITs are much weaker than the TBM in kinetic energy or normalized vorticity, which enables the linearized analysis conducted in past publications[2]. As a result, the internal tides are dominated by tidal wavelengths consistent with the dispersion relationship of the modal equations and the eigenvalue corresponding to the first vertical mode in the Sturm-

Liouville problem for surface fields[1,4]. The variations of density profiles in T1-5 result in variations of the tidal wavelength profiles. In the northern half of the domain, the wavelengths at higher-turbulence simulations are generally smaller than lower-turbulence simulations, as reflected in SI Fig.S1. The meridional and temporal variations of density profiles also lead to variations of the tidal wavelengths in latitude (SI Fig.S1) and time within each simulation. As density gradients are stronger in simulations at higher turbulence levels, the variation of tidal wavelengths, and hence the dominant length scales of tidal patterns, are also larger, which is another cause of the higher complexity of η and $\eta_{\cos}^{(\mathrm{sim})}$ patterns.

As mid-jet panels are centered around the baroclinic jet, the density gradients and TBMs there are on average stronger than those in the up-jet and down-jet panels. Hence, within each simulation, in the mid-jet panels, the two effects described above (scatterings of ITs and variations of tidal wavelengths) are stronger.

To sum up, the simulations correspond to a regime where TBMs, whose relative vorticities are well bounded by 1, are stronger than the ITs. The TBMs and ITs overlap significantly in spatial scales. The density profile is varied between different simulations. As a result, the TBMs and ITs become more energetic as reflected by the dynamical metrics listed in SI Table S1, and the IT wavelength profiles shift towards smaller scales, as demonstrated in Extended Data Figure 1. Enhanced scattering of ITs from TBMs causes IT incoherence, and IT patterns lose resemblance to plane waves. The density profile is varied within each simulation temporally and latitudinally, which result in corresponding variations of the IT wavelength profile. Stronger TBMs are accompanied by shaper density gradients, leading to more variations of the IT wavelengths.



Supporting Information Fig.S1. Mode-1 wavelengths at day 1 as a function of meridional profiles. As the TBM develops, the wavelength jumps observed in the central part of the domain become smoother.

Supporting Information Table S1. Dynamical metrics of T1-T5. KE denotes "kinetic energy". The TBM normalized vorticity and KE are averaged over time, longitude, and jet width, which we define as 800 km around 45ºN. The Cosine IT Normalized vorticity and KE are averaged over time and the entire simulation domain.

	TBM normalized vorticity	TBM KE (m^2/s^2)	Cosine IT normalized vorticity)	Cosine IT KE (m^2/s^2)
Simulation				
T1	0.06	0.04	2.3×10^{-3}	2.0×10^{-3}
T2	0.09	0.08	2.8×10^{-3}	2.3×10^{-3}
T3	0.12	0.12	1.6×10^{-2}	2.5×10^{-3}
T4	0.13	0.15	2.0×10^{-2}	2.6×10^{-3}
T5	0.14	0.20	2.2×10^{-2}	2.6×10^{-3}

Text S2. Detailed spectral behaviors in ET1-5 test sets

We attach in SI Fig.S2-6 the spectra and squared coherence of up-jet, mid-jet and down-jet bands in ET1-5 test sets. The squared coherence (i.e., normalized cross spectra) reflects how linearly related $\eta_{\cos}^{(\mathrm{sim})}$ and $\eta_{\cos}^{(\mathrm{gen})}$ are at different scales. It is computed based on the $\eta_{\cos}^{(\mathrm{sim})}$ and $\eta_{\cos}^{(\mathrm{gen})}$ spectra following its definition listed in previous works[2]. Like the spectra, the squared coherences are computed for the up-jet, mid-jet, and down-jet bands separately in this section.

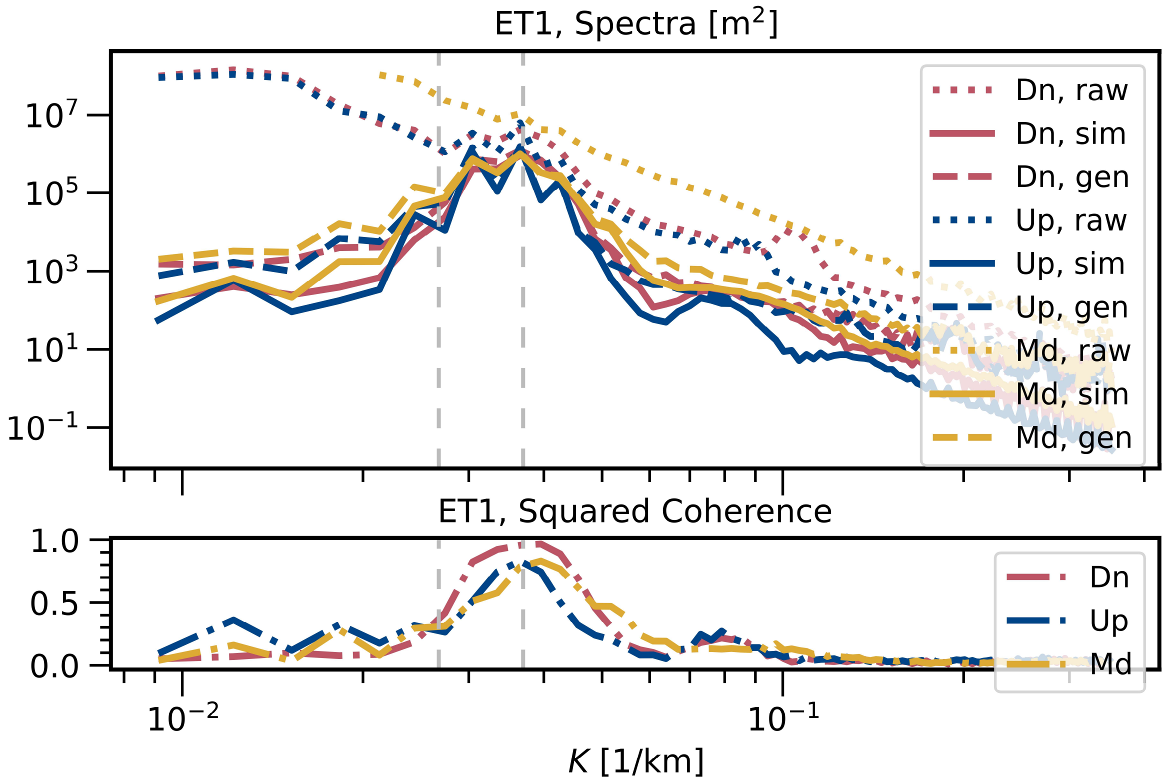
The ET1 run displays excellent spectral behaviors. In SI Fig.S2, the $\eta_{\cos}^{(\mathrm{gen})}$ spectra capture the magnitude and locations of spectral bumps of the $\eta_{\cos}^{(\mathrm{sim})}$ well, and the peaks of the squared coherence are no less than 0.8 in the up-jet, mid-jet, and down-jet bands.

In a relatively worse run (ET5) at a worse band (mid-jet band), the TITE would still outperform a simple spatial filter. This can be seen by comparing the dotted and solid yellow lines in SI Fig.S6, which correspond to the spectra of η and $\eta_{\cos}^{(\mathrm{sim})}$ signals respectively. In the η spectra, due to the strong TBMs that co-exist with ITs around the tidal wavenumbers, there is no noticeable tidal bump around the tidal wavenumber. Thus, unless one makes strong assumptions or utilizes a-priori physical knowledge, no information about the tidal wavenumber or the magnitude of tidal motions could be gained from any spatial filters applied onto η. However, TITE is still able to capture the magnitude of the spectral bumps of ITs, and the shifting between the tidal bumps of $\eta_{\cos}^{(\mathrm{sim})}$ and $\eta_{\cos}^{(\mathrm{gen})}$ spectra is well less than a decade.

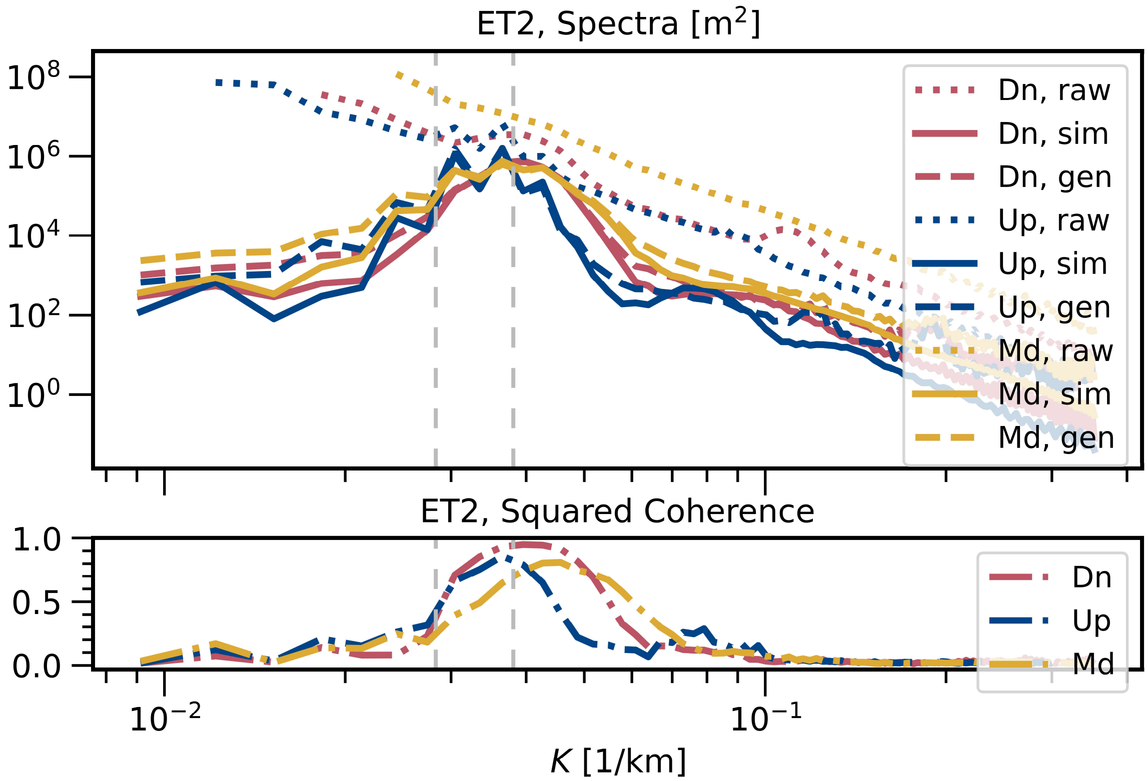
In all the three latitudinal bands, the squared coherences in ET1-4 decay quickly outside the tidal bumps. This is consistent with the fact that outside of the bumps, the $\eta_{\cos}^{(\mathrm{gen})}$spectra significantly mismatch the $\eta_{\cos}^{(\mathrm{sim})}$ spectra. Spurious signals in $\eta_{\cos}^{(\mathrm{gen})}$ drown out the mode-2 tidal bumps present in $\eta_{\cos}^{(\mathrm{sim})}$ and are especially prominent at large scales, sometimes causing differences between $\eta_{\cos}^{(\mathrm{gen})}$ and $\eta_{\cos}^{(\mathrm{sim})}$ by a factor of 10. As the ground truth spectra $\eta_{\cos}^{(\mathrm{sim})}$ are orders of magnitudes lower outside the tidal bumps, this decreased performance outside the bumps is not our major concern in this project. To alleviate it, we could incorporate spectral forcing in the architecture design of TITE, which is left for future work.

The squared coherences for the mid-jet band at ET5 are lower than 0.5 at all wavenumbers, suggesting the relatively poor performance of TITE in ET5 amidst the jet. Similarly, in ET2-4, the squared coherences for the mid-jet bands also peak lower than the down-jet or up-jets. An explanation of such decreased behavior around mid-jet bands is described in the main text.

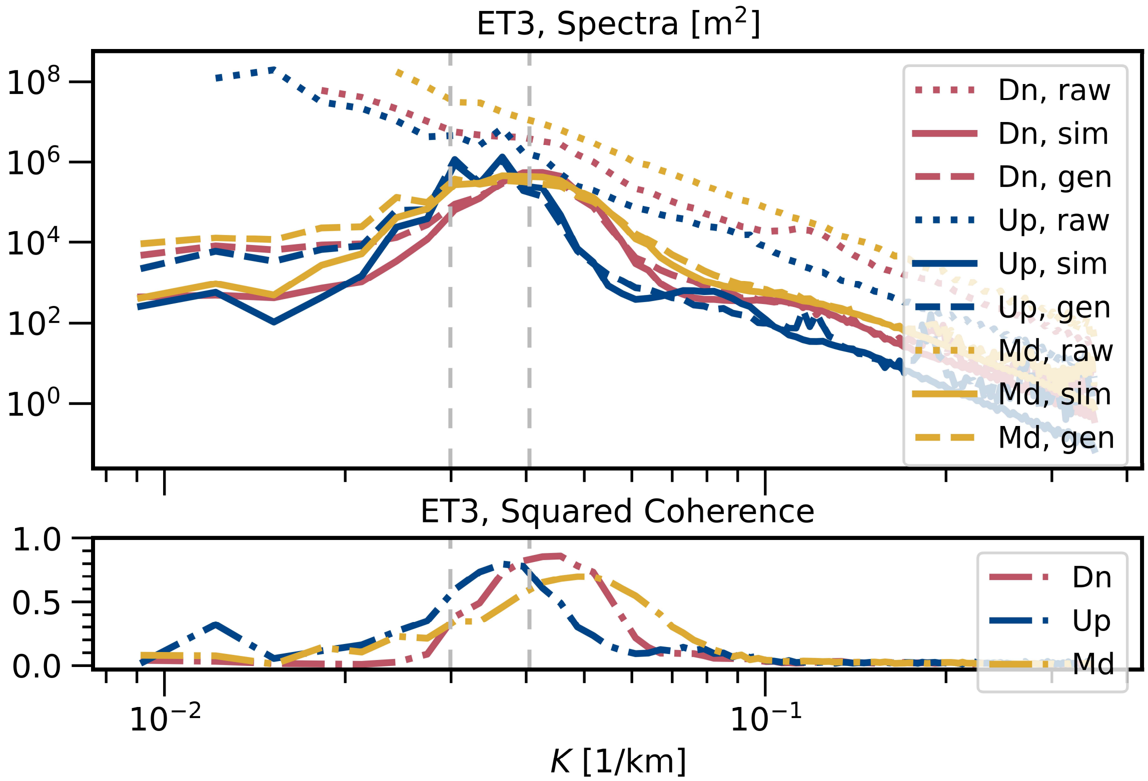
The tidal bumps in the up-jet and down-jet bands are farther apart in higher turbulence runs. In ET1, the tidal bumps in the up-jet and down-jet bands almost completely overlap, while in ET5 the two bumps are shifted apart quite conspicuously. This is consistent with SI Fig.S1, which shows that the tidal wavenumbers vary more in higher turbulence runs. As discussed in the main text, the stronger variations of tidal wavelengths may be part of why the higher turbulence runs are intrinsically more challenging to TITE.



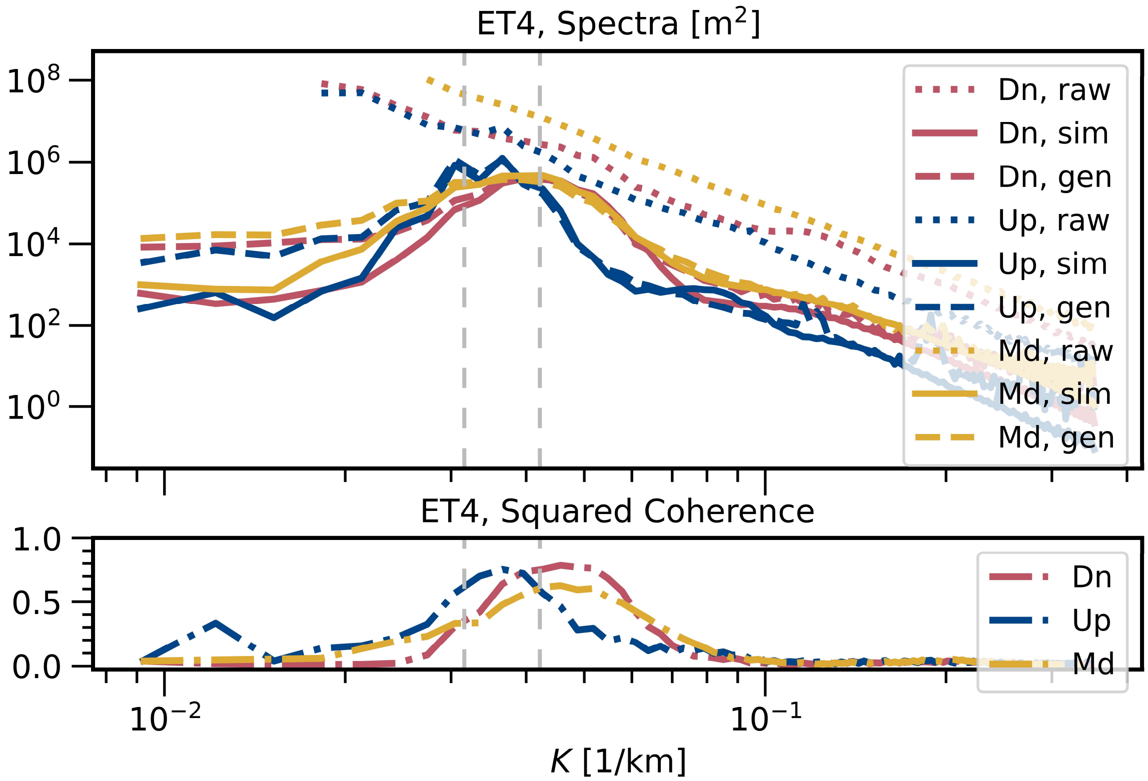
Supporting Information Fig.S2. Spectra and Coherence for the up-jet, mid-jet and down-jet bands in ET1 test set. Compared to Fig.3 in the main article, this figure presents the ET1 test set only, but adds the spectra computed for mid-jet bands (denoted by legend "Md" and coloured yellow) and the squared coherence for the three bands (lower row).



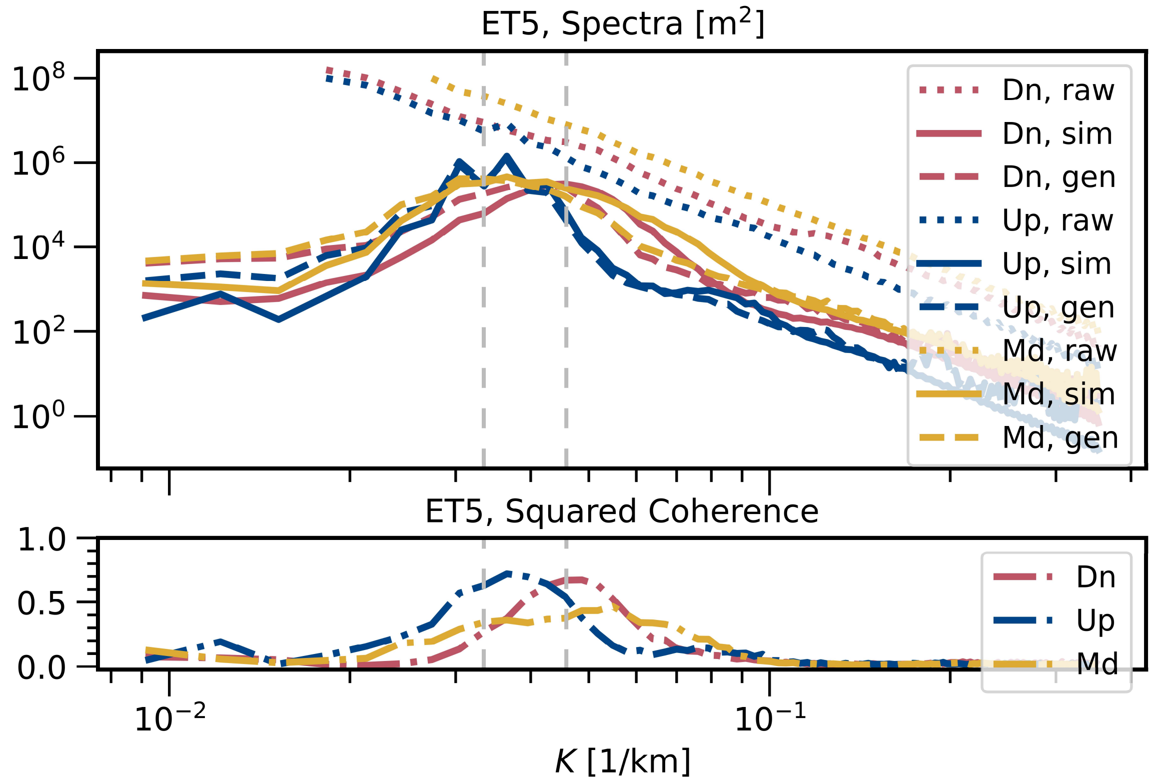
Supporting Information Fig.S3. Similar to SI Fig.S2, but for the ET2 test set.



Supporting Information Fig.S4. Similar to SI Fig.S2, but for the ET3 test set.



Supporting Information Fig.S5. Similar to SI Fig.S2, but for the ET4 test set.



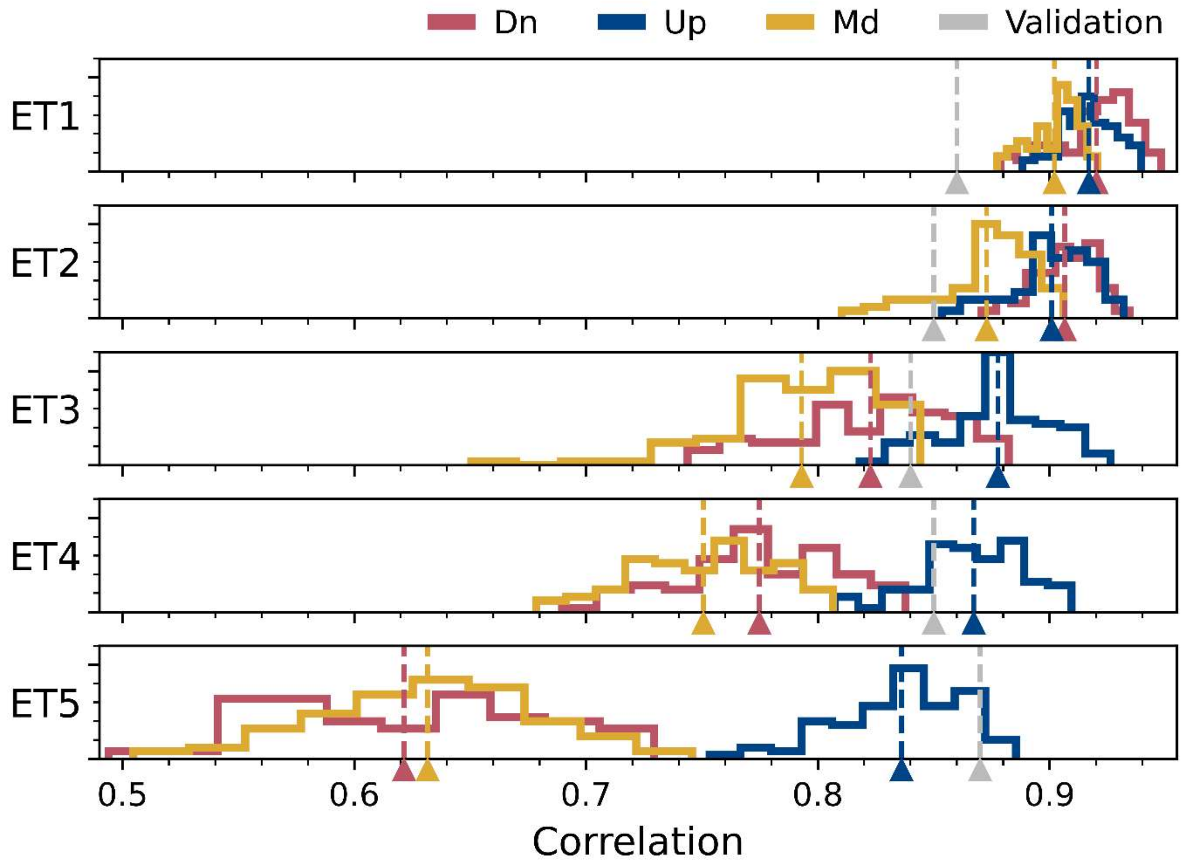
Supporting Information Fig.S6. Similar to SI Fig.S2, but for the ET5 test set.

Text S3. Statistics of correlation factors

The mean correlation values averaged over different subsets of test/validation instances are listed in SI Table S2. The histograms of the correlations in the test sets are presented in SI Fig.S7. From either the table or the histogram, the general trend of correlation to deteriorate as turbulence level gets higher, and the sharper drop from ET4 to ET5 can be observed, which are mentioned in the main text. Computations of the correlation factors are detailed in SI Text S6.

Supporting Information Table S2.   Mean correlation factors of validation and test sets in the ES1-5 runs. The second and third columns present mean correlation factors averaged over all panels in the validation sets and test sets respectively. The last three columns present mean correlation factors averaged over down-jet, mid-jet, and up-jet bands in the test sets respectively.

	Validation set, all	**Test set, all**	**Test set, down-jet**	**Test set, mid-jet**	**Test set, up-jet**
TITE run					
ET1	0.86	0.91	0.92	0.90	0.92
ET2	0.85	0.89	0.90	0.87	0.90
ET3	0.84	0.83	0.82	0.79	0.88
ET4	0.85	0.80	0.77	0.75	0.87
ET5	0.87	0.70	0.62	0.63	0.84



Supporting Information Fig.S7. Histogram and mean (denoted by the vertical dashed lines) of correlation factors in the test cases of ET1-5, presented for down-jet, up-jet, and mid-jet panels separately. The three panels are denoted by the colors marked in the legends, where "Dn","Up" and "Md" denote the down-jet, up-jet and mid-jet panels respectively. The mean correlations of the validation sets (averaged over all available panels) are presented in the dashed gray vertical lines for reference. When histogram is plotted, each group is divided into 10 bins. Vertical axis group denote number counts in each bin, with axis limits fixed at 0 and 30.

Text S4. Monitoring the training and deciding the stopping epoch

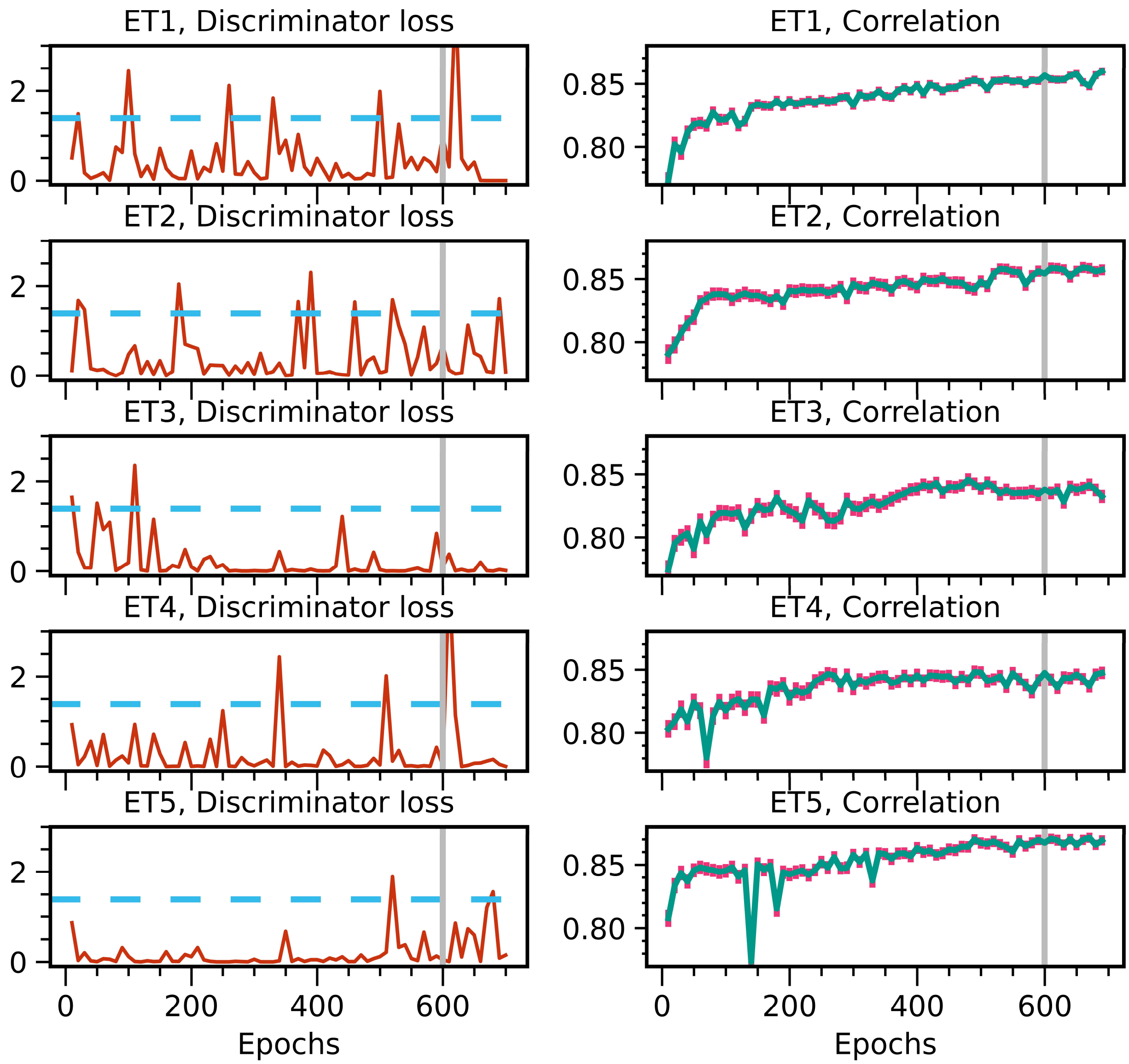
The stopping criteria during a GAN training is a delicate issue, as the convergence of GAN is hard to identify due to its fleeting nature[5]. In this work, the analysis of ET1-5 in the main article are all conducted right after 600 epochs. We do not claim that it is the optimal stopping epoch for these runs but observe that there is no definite sign of model collapse around the 600th epoch.

We monitor the training behaviors from two kinds of metrics. First, we monitor the discriminator, the generator, the L1, and the total loss functions respectively as defined in the original publication[6]. Second, we monitor metrics such as correlation factor and relative error between $\eta_{\cos}^{(\mathrm{sim})}$ and $\eta_{\cos}^{(\mathrm{gen})}$ in the validation set. In SI Fig.S8, we present the evolution of discriminator loss and the correlation factor in the validation set up to the 700^{th} epoch for the ET1-5 runs.

From the definition of discriminator loss in pix2pix[6], when the discriminator is effectively tossing coins at every judgement, it would have a discriminator loss of 2log(2), which is marked by the horizontal dashed line in SI Fig.S8. Observing the discriminator loss, we find that in the ET1-4 runs, the race between the discriminator and the generator appears healthily close, as the discriminator loss frequently surges above the coin-tossing line, which are then recovered back below the line in a few dozen epochs. This suggests that the discriminator and the generator are likely indeed co-evolving. Observing the correlation in the validation set of ET1-4, we see that the correlation generally stabilizes after 300th epoch with a slight tendency to increase afterwards.

The evolution of discriminator loss of the ET5 run appears less ideal. Up to the 500th epoch, the discriminator loss is always well below the coin-tossing line. In principle, this indicates a potential GAN collapse: the generator can almost never cheat the discriminator and may not be able to learn due to vanishing gradients. Between the 500th and the 700th epoch, the discriminator loss starts to surge occasionally above the coin-tossing line, which indicates that the generator may have somehow still evolved well enough to cheat the discriminator. Hence, we decide to stop at the 600th epoch, by which time the generator starts to sometimes prevail, to stay safely away from the potential collapse in earlier epochs. We note that even though the discriminator evolution is less ideal in ET5, the evolution of the correlation factor in the validation set appears to show similar behaviors as ET1-4 (bottom row, right column of SI Fig.S8), in that it stabilizes after around 300 epochs. As there are no signs of model collapsing from the evolution of the correlation, it is likely that the GAN did not collapse after all; perhaps the small bumps of the discriminator loss in the first 500 epochs in the ET5 run are sufficient to prevent vanishing gradients for the generator.

Left as future work, we can try to pace the improvement of the discriminator's performance by adding noise[5], or to use a different architecture such as the Wasserstein GAN[7] to address potentially vanishing gradients.



Supporting Information Fig.S8. Evolution of discriminator loss and correlation factor in validation set during the training of ET1-5. The loss and the correlation are recorded every 10 epochs, starting at the 10th epoch, and ending at the 700th epoch. Gray vertical lines mark the 600th epoch, which is the stopping epoch for the analysis in the main article. The dashed horizontal line in the left columns denote the level at which the discriminator is tossing coins. Correlation factors shown in the right columns are computed between $\eta_{\cos}^{(\mathrm{sim})}$ and $\eta_{\cos}^{(\mathrm{gen})}$ in the validation set. The line plots present the mean correlation factor of all validation instances, with error bars marking one standard deviation.

Text S5. Multi-scale structural complexity of simulation snapshots

Recently, a generic metric for complexity of image patterns called “multi-scale structural complexity” is proposed in Bagrov et al[8]. Briefly speaking, this metric measures on how much variation is induced every time one coarse-grains the image at interest. Here, we present that this metric computed over the η or $\eta_{\cos}^{(\mathrm{sim})}$ panels captured from simulations T1-5 agrees with our physical understandings about the impacts of stronger TBMs.

We follow the notations in Bagrov et al.[8] throughout this section. After an image is coarse-grained by k times, the quantity $\mathcal{C}_k$ is intended to measure how much variation is induced if one further coarse grain the image by one step. We coarse-grain the images under the same discrete decimation scheme as in Bagrov et al.[8] At each coarse-graining step, we set the filter parameter Λ=2. We refer to Bagrov et al.[8] for details on the related definitions.

Each square panel of η or $\eta_{\cos}^{(\mathrm{sim})}$ used by TITE is originally 258-by-258 pixels. For simplicity, we delete the first and last rows and columns of each panel, resulting in panels at 256-by-256 pixels. The new width(length) of the panels (i.e., 256) is a power of 2, which makes the computations quicker without significantly sacrificing original information.

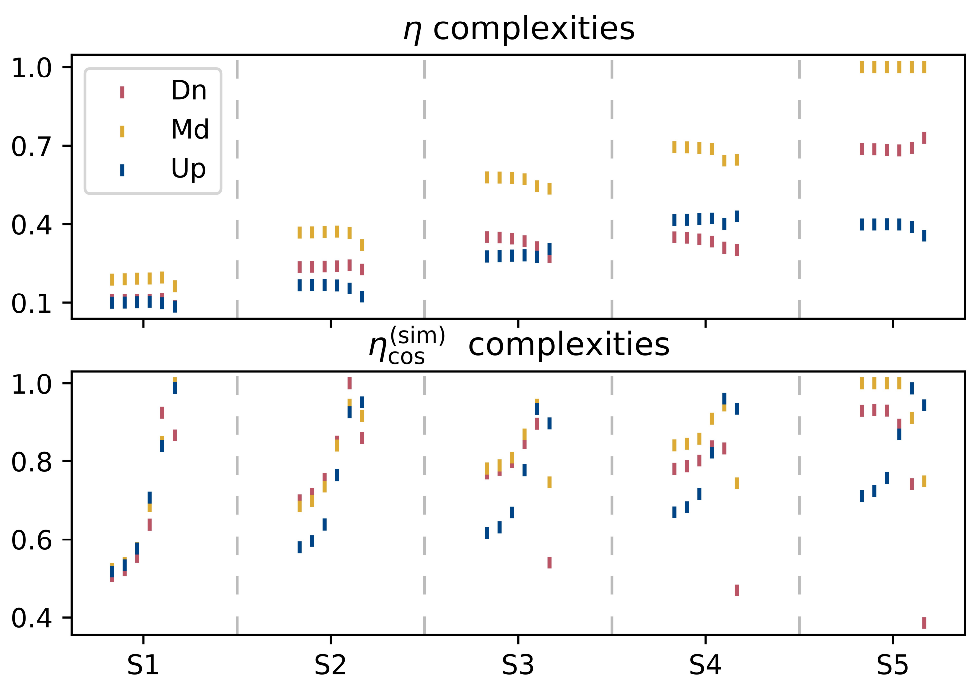
The images are then coarse-grained by 6 times, and $\mathcal{C}_k$ at each step are recorded. For our purpose, the complexity $\mathcal{C}_k$ at each k individually is more informative than the summation of $\mathcal{C}_k$ over k, denoted in Bagrov et al. as $\mathcal{C}$. One can prove that the equ ation 4 in Bagrov et al. can be simplified as $\mathcal{C} = 0.5(O_{0,0} - O_{k,k})$[9], which smudges out contributions from the intermediate coarse-graining steps, while here we are more interested in the complexities related to different individual spatial scales.

We divide all the panels of η or $\eta_{\cos}^{(\mathrm{sim})}$ into 15 groups (5 turbulence levels and 3 latitudinal bands). Each group contains 100 pairs of $\left\{\eta, \eta_{\cos}^{(\mathrm{sim})}\right\}$ panels. The computations of $\mathcal{C}_k$ for η and $\eta_{\cos}^{(\mathrm{sim})}$ are conducted separately. For presentation purpose, we average and normalize $\mathcal{C}_k$. Specifically, in each group and at each k, we compute the mean of $\mathcal{C}_k$ over all the 100 panels of η or $\eta_{\cos}^{(\mathrm{sim})}$, resulting in 30 different values of averaged $\mathcal{C}_k$ (15 groups each for η and $\eta_{\cos}^{(\mathrm{sim})}$) at each k. Then, we divide $\mathcal{C}_k$ by the maximum value of $\mathcal{C}_k$ among the 15 groups of η or $\eta_{\cos}^{(\mathrm{sim})}$ separately. The averaged and normalized $\mathcal{C}_k$ is denoted as $\overline{\mathcal{C}_k}$, which ranges from 0 to 1.

In the upper row of SI Fig.S9, $\overline{\mathcal{C}_k}$ increases consistently from T1 to T5, and the mid-jet bands always contain higher $\overline{\mathcal{C}_k}$ within each simulation. This agrees with our expectation: vigorous TBMs would make the TBM components in the raw η patterns more complicated.

Moreover, stronger TBMs are linked to increased complexities of IT patterns due to two mechanisms explained in SI Text S1: 1. increased scatterings, which lead to more longitudinal variations as well as small-scale features, and 2. increased density gradients, which lead to more variations of dominant tidal wavelengths. In the lower row of SI Fig.S9, we see that at $k = 1{,}2{,}3{,}4$ (first four vertical lines plotted in each group), $\overline{\mathcal{C}_k}$ increase from T1 to T5 and are highest at mid-jet panels within each simulation, in agreement with the stronger TBMs. At $k = 5{,}6$ (last two vertical lines in each group), the tendency of $\overline{\mathcal{C}_k}$ appears quite random. Noting that the resolution of the $\eta_{\cos}^{(\mathrm{sim})}$ panels are 4 km, the metric $\overline{\mathcal{C}_1}$ for $\eta_{\cos}^{(\mathrm{sim})}$ reflects how much the $\eta_{\cos}^{(\mathrm{sim})}$ panels change when coarse-grained from a 4 km resolution to an 8 km resolution. Similarly, $\overline{\mathcal{C}_6}$ reflects how much change occurs when coarse graining is done from a 128 km resolution to a 256 km resolution. As the dominant tidal wavelengths are between 135 km and 230 km (see SI Fig.S1), at $k = 6$, the images are coarse-grained across the dominant spatial scale of the patterns, and large-scale (>256 km) features left afterwards are not affected by the two mechanisms mentioned before. Therefore, $\overline{\mathcal{C}_6}$ is expected to be insensitive to the strength of TBMs. As for the erratic tendency of $\overline{\mathcal{C}_5}$, we cannot find a physical explanation, but note that this does not contradict with our conjecture that the $\eta_{\cos}^{(\mathrm{sim})}$ patterns are more complex under higher TBMs in general, given the consistent tendencies at $k = 1,2,3,4$.

We don't claim that $\mathcal{C}_k$ from Bagrov et al. is the most reflective metric on image complexity in our case. A metric defined on a more refined coarse-graining process could be more meaningful for the $\eta_{\cos}^{(\mathrm{sim})}$ panels. For example, one could modify how the coarse-graining is conducted, so that a new metric captures how much the image changes from resolutions at 128 km to 256 km more incrementally (say, coarse-graining by 4km at each step). Such a more refined metric may be able to detect the impacts of the variations of tidal wavelengths. This is left for future explorations.



Supporting Information Fig.S9. Multi-scale structural complexity of panels used in T1-5. The up-jet, mid-jet, and down-jet bands (denoted as "Up", "Md" and "Dn" in legends) for η (upper row) and $\eta_{\cos}^{(\mathrm{sim})}$(lower row) are presented for simulations T1-5 (horizontal axis) separately. For each latitudinal band at each simulation, six vertical line markers are presented, which sequentially correspond to $\overline{\mathcal{C}_k}$ at k=1,2,3,4,5,6. For example, in the group of six yellow vertical markers at the upper right corner in the upper row, the first vertical marker denotes the mean of $\overline{\mathcal{C}_1}$ computed from mid-jet bands of η in T5.

Text S6. Statistical metrics

The correlation factors and 1D spectra are computed from standard approaches. Specifically, for one panel of $\eta_{\cos}^{(\mathrm{gen})}$ and the corresponding $\eta_{\cos}^{(\mathrm{sim})}$, similar to other studies[10], we compute the correlation factor between the two arrays flattened from the two images. The mean correlation factors are averaged over all correlation factors in the data sets at interest. Take the fourth column (titled as "test set, down-jet") in Table 2 as an example. In each of the ET1-5 runs, we single out the 100 test instances belonging to down-jet panels, compute the Pearson correlation between $\eta_{\cos}^{(\mathrm{gen})}$ and $\eta_{\cos}^{(\mathrm{sim})}$ in each instance, and then average the 100 correlation factors to get the mean correlation. The maximum, minimum and standard deviation of correlation factors are computed similarly and recorded in SI 2.

Our 1D spectra are computed from 2D spectra via a numerical azimuthal averaging used in other studies [11,12]. The 2D spectra are computed over collective statistics of the down-jet, mid-jet, or up-jet panels in the test set separately. For example, in the ET5 run, the 2D spectra for the generated down-jet panels are computed from the 100 $\eta_{\cos}^{(\mathrm{gen})}$ instances from the down-jet panels in the test set. A Hanning window in the latitudinal direction is applied at each panel prior to conducting the 2D fast Fourier transforms.

In addition, we have also computed relative error of $\eta_{\cos}^{(\mathrm{gen})}$ against $\eta_{\cos}^{(\mathrm{sim})}$ for each test instance. The relative error turns out to be larger than 0.3 for each test instance in the five runs. This non-negligible relative error is consistent with the spurious large-scale signals discussed in the main article.

Text S7. Changes to the Tensorflow Tutorial code

TITE is modified from the Tensorflow Tutorial codes[13] (hereafter "tutorial codes"). Here, we detail the changes made to the tutorial codes for reproducibility. Some familiarity with the original pix2pix paper[6] from readers is assumed in the narrations to follow.

First, the $\eta_{\cos}^{(\mathrm{sim})}$ fields (ground truth) are weaker in amplitude than η (inputs) due to our simulation configurations. By trial and error, we find that this imbalance of magnitudes between inputs and outputs often destabilizes the training. To alleviate this issue, we multiply the $\eta_{\cos}^{(\mathrm{sim})}$ signals by a uniform factor of 20, after which the max value of $|\eta_{\cos}^{(\mathrm{sim})}|$ is around 78% of the max value of $|\eta|$ among all simulation snapshots we use. The other modifications we make are not essential for the training to succeed, and are rather finer improvements of training behaviours, to simplify the algorithm, or are inspired by challenges to be faced in future satellite altimetric data.

As explained in Isola et al.[6], the objective function during the training can be expressed as:

$$\arg \min_{G} \max_{D} \mathcal{L}_{cGAN}(G, D) + \lambda\, \mathcal{L}_{L1}(G),$$

where $\mathcal{L}_{cGAN}(G, D)$ is the classic minmax cGAN loss, and $\mathcal{L}_{L1}(G)$ is the L1 loss, which controls the impact of overall L1 error of generated images[6]. We change the parameter λ from 10^2 to 10^3, which improves the mean correlation in the validation set by around 0.09 in all the ES1-5 runs and appears to stabilize the training. Increasing λ to 10^4 or 10^5 does not significantly change the outcomes.

As the inputs and outputs in our application are both scalar fields, we store all the panels as single-precision 2D numerical arrays rather than image-formatted files. We modified the input pipeline in the tutorial code accordingly, and the number of input and output channels is reduced from 3 (for RGB) to 1. Hence, we save some computational costs. This scalar approach is equivalent to using int32 grayscale images, and for convenience we still refer to the scalar arrays as "images" in the article. All image panels plotted in this paper are contours of the scalar fields, and the colormaps in plots are picked only for readability or aesthetic purposes. Occasionally, colours saturate in plots as an artifact from the way we define the colormaps (e.g., input fields in Fig.2), though not in our data. For normalization, we find the maximum value among all pixels in the η snapshots and divide $\{\eta, 20\,\eta_{\cos}^{(\mathrm{sim})}\}$ by this maximum value, so that all data is bounded by 1.

Prior to each epoch, training images are randomly reshuffled in time, cropped, flipped, and rotated. The random reshuffles, crops and horizontal flips are inherited from the tutorial code, whereas the random rotations and vertical flips are added by us. For random rotation, we randomly rotate each panel by 90º in either clockwise or counterclockwise directions. "Random cropping" means that we interpolate the images from a 258-by-258 to a 286-by-286 pixels grid, and within it, randomly crop a square panel of 256-by-256 pixels. All these manipulations are synchronized between the inputs η and outputs $\eta_{\cos}^{(\mathrm{gen})}$.

During random cropping, the pixel number choices of 286-by-286 and 256-by-256 are inherited from Isola et al.[6] We keep these choices for the following reasons. First, having the pixel number to be powers of 2 after cropping simplifies the downsampling steps in the generators' architecture as it helps avoid zero-paddings. Second, cropping from a 286-by-286 image to a 256-by-256 image deletes about 20% of all pixels, which is an appropriate cropping rate. The cropped images would still span over a few tidal wavelengths and thus retain the IT patterns, and yet, as the cropping causes the images to lose about 10% of the pixels in the longitudinal direction, the exact zonal periodic condition would be excluded during TITE's training, which corresponds to challenges in realistic situations

Other data augmentations (random rotations and flipping) of the training images also introduce to TITE challenges motivated by realistic situations. For example, in the simulations, ITs are forced at the southern boundary of the domain, and propagate northward. If all snapshots are upright, then during training, TITE might learn that the ITs always propagate northward, and use that knowledge during testing. But after random rotations and flipping are introduced, such information would be unavailable to TITE, which corresponds to realistic situations where one doesn't necessarily know the IT generation sites a-priori when extracting IT signals. We also experimented on TITE runs where random rotations and flipping are suppressed, and did not see any qualitative changes in TITE's performance.

Following the original nomenclature[6], our discriminator architecture can be expressed as C64-C128-C256-C512-C512-C512. The main difference between this and the architecture recommended in the original paper[6] is that at one step, our discriminator treats a whole image at once, while the original code applies a "patchGAN", which divides the image into different patches regarded independent from each other and treats each patch separately. While the patchGAN contains less convolutional layers and are less costly, one must decide on the size of the individual patches prior to the training. We haven't investigated how to pick the patch size in our problem yet. Thus, for design simplicity, we make the patch size equal to the image size of

$\eta_{\cos}^{(\mathrm{sim})}$. To investigate the impacts of this change, we have also tried using the patchGAN with the 70-by-70 patch size adopted in the tutorial code, and the mean correlation and spectral properties of $\eta_{\cos}^{(\mathrm{gen})}$ stay similar.

Additional Supporting Information

Caption for Movie S1: Performance of TITE on T1 data after trained on data from T2, T3, T4 and T5. All snapshots are re-arranged in order of time. "Input" column plots η, "Truth" column plots $\eta_{\cos}^{(\mathrm{sim})}$, "Generated" column plots $\eta_{\cos}^{(\mathrm{gen})}$, and "Difference" column plots ($\eta_{\cos}^{(\mathrm{sim})} - \eta_{\cos}^{(\mathrm{gen})}$).

Caption for Movie S2: Similar to Movie S1, but for the performance of TITE on T2 data after trained on data from T2,T3, T4 and T5.

Caption for Movie S3: Similar to Movie S1, but for the performance of TITE on T3 data after trained on data from T1,T2, T4 and T5.

Caption for Movie S4: Similar to Movie S1, but for the performance of TITE on T4 data after trained on data from T1,T2, T3 and T5.

Caption for Movie S5: Similar to Movie S1, but for the performance of TITE on T5 data after trained on data from T1,T2, T3 and T4.

Caption for Movie S6: Illustration of simulations T1-5. Five columns correspond to five simulations respectively. The upper row plots local Rossby number, defined as relative vorticities divided by Coriolis parameter. Lower row plots $\eta_{\cos}^{(\mathrm{sim})}$. The entire simulation domain is included. Snapshots are ordered by time and separated by 4*T.*

References

1. Dunphy, M., Ponte, A. L., Klein, P. & Le Gentil, S. Low-Mode Internal Tide Propagation in a Turbulent Eddy Field. *J. Phys. Oceanogr.* **47**, 649–665 (2017).
2. Ponte, A. L., Klein, P., Dunphy, M. & Le Gentil, S. Low-mode internal tides and balanced dynamics disentanglement in altimetric observations: Synergy with surface density observations: SLOW VERSUS FAST SIGNATURES ON SEA LEVEL. *J. Geophys. Res. Oceans* **122**, 2143–2155 (2017).
3. Ponte, A. L. & Klein, P. Incoherent signature of internal tides on sea level in idealized numerical simulations. *Geophysical Research Letters* **42**, 1520–1526 (2015).

4. GILL, A. E. Atmosphere-Ocean Dynamics. *Int. Geophys. Ser.* **30**, 662p (1982).

5. Arjovsky, M. & Bottou, L. Towards Principled Methods for Training Generative Adversarial Networks. *arXiv:1701.04862 [cs, stat]* (2017).

6. Isola, P., Zhu, J.-Y., Zhou, T. & Efros, A. A. Image-to-Image Translation with Conditional Adversarial Networks. *arXiv:1611.07004 [cs]* (2018).

7. Arjovsky, M., Chintala, S. & Bottou, L. Wasserstein GAN. *arXiv:1701.07875 [cs, stat]* (2017).

8. Bagrov, A. A., Iakovlev, I. A., Iliasov, A. A., Katsnelson, M. I. & Mazurenko, V. V. Multiscale structural complexity of natural patterns. *Proc Natl Acad Sci USA* **117**, 30241–30251 (2020).

9. Github, IakovlevIA. *structural-complexity,* Issue 3. https://github.com/IakovlevIA/structural-complexity/issues/3. Accessed on August 4, 2021.

10. Torres, H. S. *et al.* Diagnosing Ocean-Wave-Turbulence Interactions From Space. *Geophys. Res. Lett.* **46**, 8933–8942 (2019).

11. Torres, H. S. *et al.* Partitioning Ocean Motions Into Balanced Motions and Internal Gravity Waves: A Modeling Study in Anticipation of Future Space Missions. *Journal of Geophysical Research: Oceans* **123**, 8084–8105 (2018).

12. Savage, A. C. *et al.* Spectral decomposition of internal gravity wave sea surface height in global models. *Journal of Geophysical Research: Oceans* **122**, 7803–7821 (2017).

13. Pix2Pix | TensorFlow Core. *TensorFlow* https://www.tensorflow.org/tutorials/generative/pix2pix.