

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

Han Wang¹, Hesam Salehipour², Alice Nuz¹, Michael Poon¹, Aurélien L. Ponte⁵, and Nicolas Grisouard¹

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

²Autodesk Research, Ontario, Canada

⁵Ifremer, 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-geostrophic¹. 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 profile². 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.

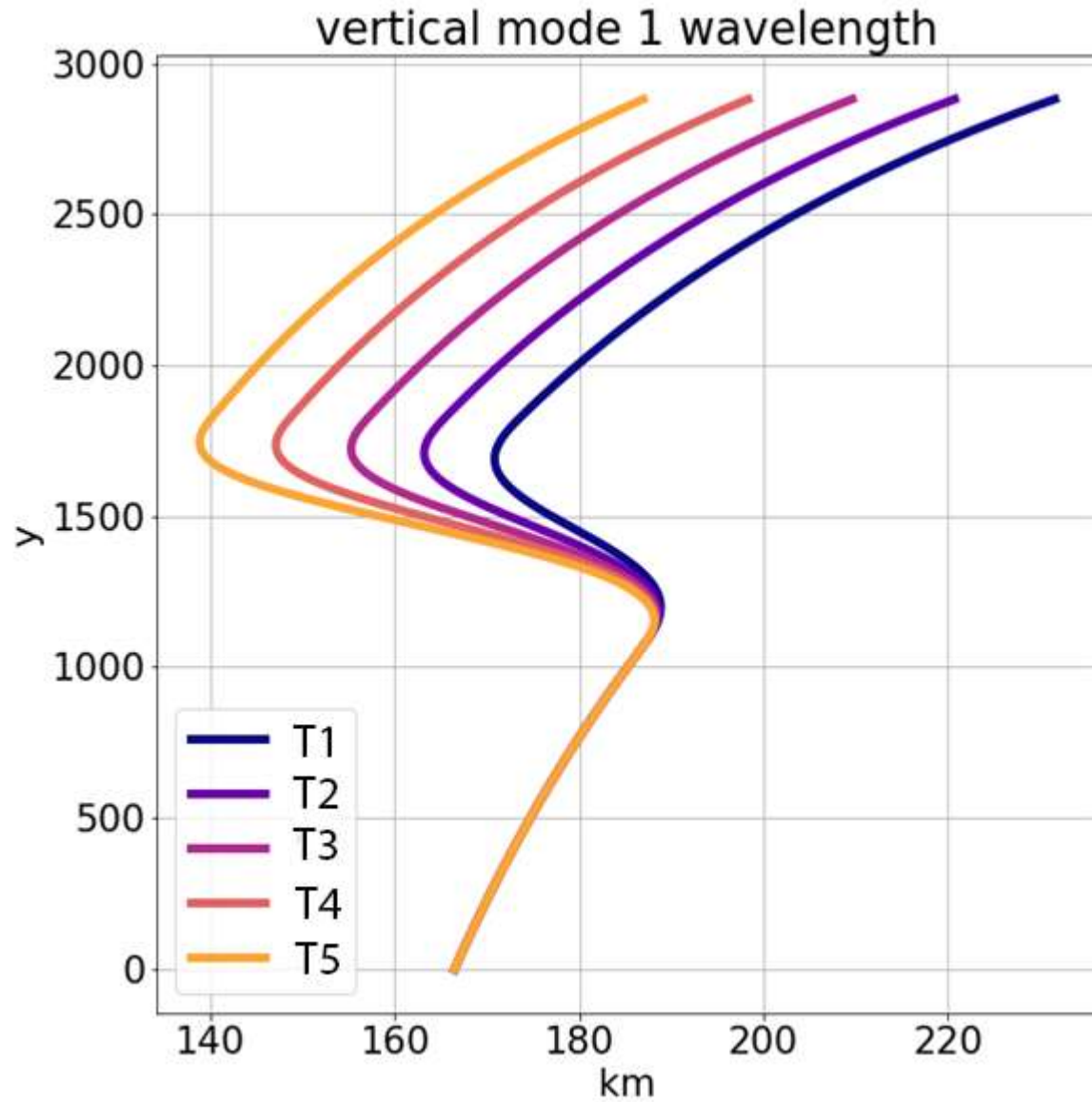
Moreover, the scattering from jets also makes the signals less coherent, as inspected closely by Ponte and Klein³. In SI Movie S3, we present snapshots of normalized vorticities of the TBMs along with the $\eta_{\cos}^{(\text{sim})}$ components in T1-5. There, we can see that as the TBMs become increasingly energetic from T1 to T5, the IT signals are scattered more around or to the north of the jet. As a result, the $\eta_{\cos}^{(\text{sim})}$ patterns are less like plane waves and contain more small-scale features. This factor adds to the complexity of the patterns of η and $\eta_{\cos}^{(\text{sim})}$.

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 publications². 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 $\eta_{\cos}^{(\text{sim})}$ 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 45°N. The Cosine IT Normalized vorticity and KE are averaged over time and the entire simulation domain.

	TBM	TBM KE	Cosine IT	Cosine IT KE
	normalized	(m ² /s ²)	normalized	(m ² /s ²)
	vorticity		vorticity)	
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 $\eta_{\cos}^{(\text{sim})}$ and $\eta_{\cos}^{(\text{gen})}$ are at different scales. It is computed based on the $\eta_{\cos}^{(\text{sim})}$ and $\eta_{\cos}^{(\text{gen})}$ spectra following its definition listed in previous works². Like the spectra, the squared coherences are computed for the up-jet, mid-jet, and down-jet bands separately in this section.

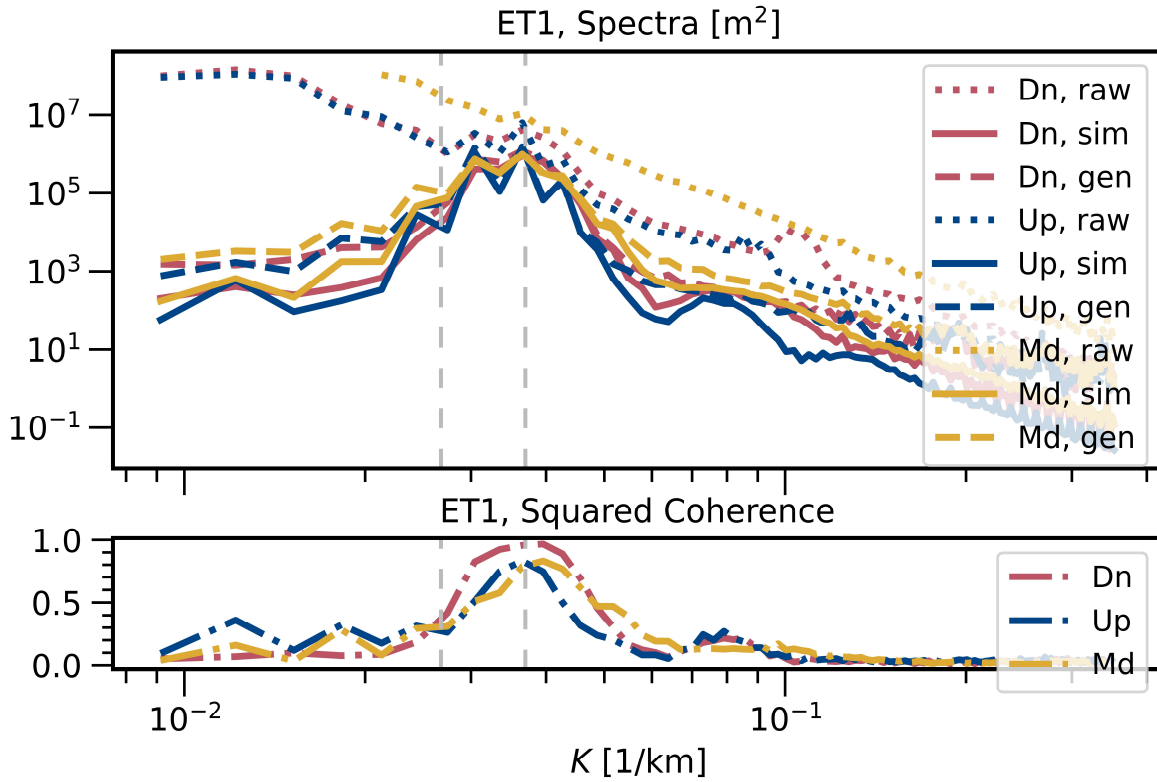
The ET1 run displays excellent spectral behaviors. In SI Fig.S2, the $\eta_{\cos}^{(\text{gen})}$ spectra capture the magnitude and locations of spectral bumps of the $\eta_{\cos}^{(\text{sim})}$ well, and the peaks of the squared coherence are no less than 0.8 in the up-jet, mid-jet, and down-jet bands.

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 $\eta_{\cos}^{(\text{sim})}$ 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 $\eta_{\cos}^{(\text{sim})}$ and $\eta_{\cos}^{(\text{gen})}$ spectra is well less than a decade.

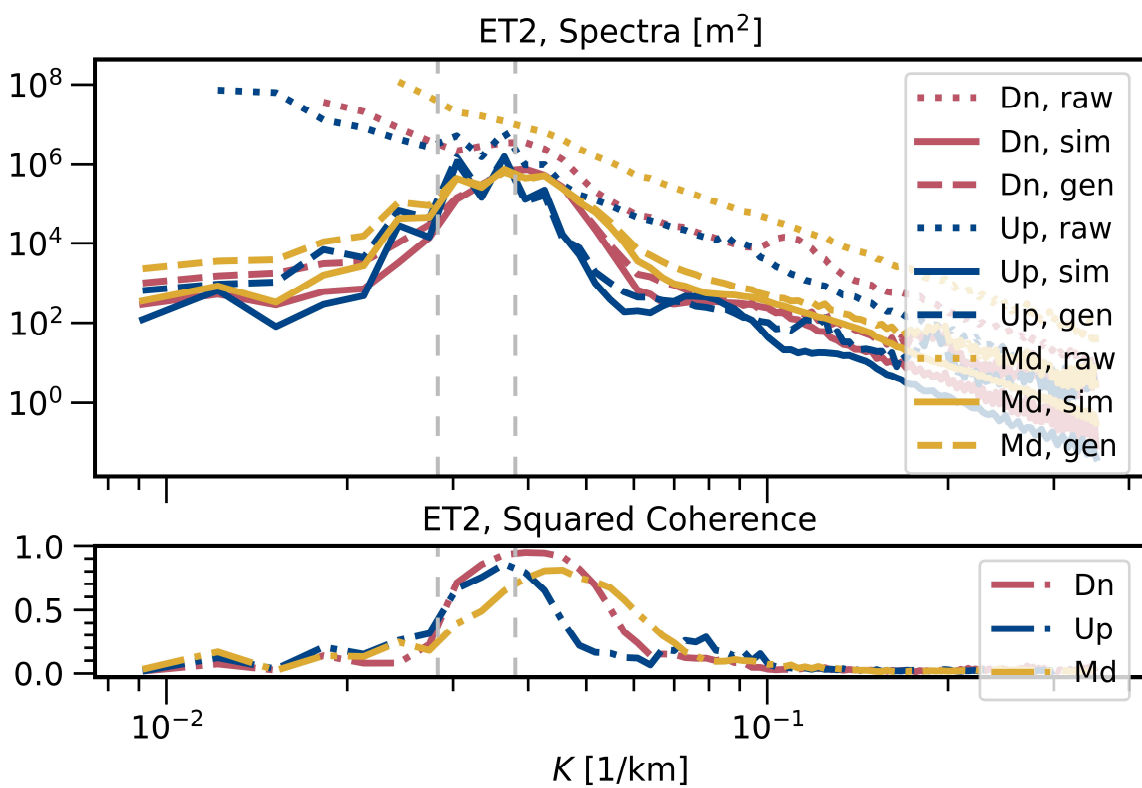
In all the three latitudinal bands, the squared coherences in ET1-4 decay quickly outside the tidal bumps. This is consistent with the fact that outside of the bumps, the $\eta_{\cos}^{(\text{gen})}$ spectra significantly mismatch the $\eta_{\cos}^{(\text{sim})}$ spectra. Spurious signals in $\eta_{\cos}^{(\text{gen})}$ drown out the mode-2 tidal bumps present in $\eta_{\cos}^{(\text{sim})}$ and are especially prominent at large scales, sometimes causing differences between $\eta_{\cos}^{(\text{gen})}$ and $\eta_{\cos}^{(\text{sim})}$ by a factor of 10. As the ground truth spectra $\eta_{\cos}^{(\text{sim})}$ 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.

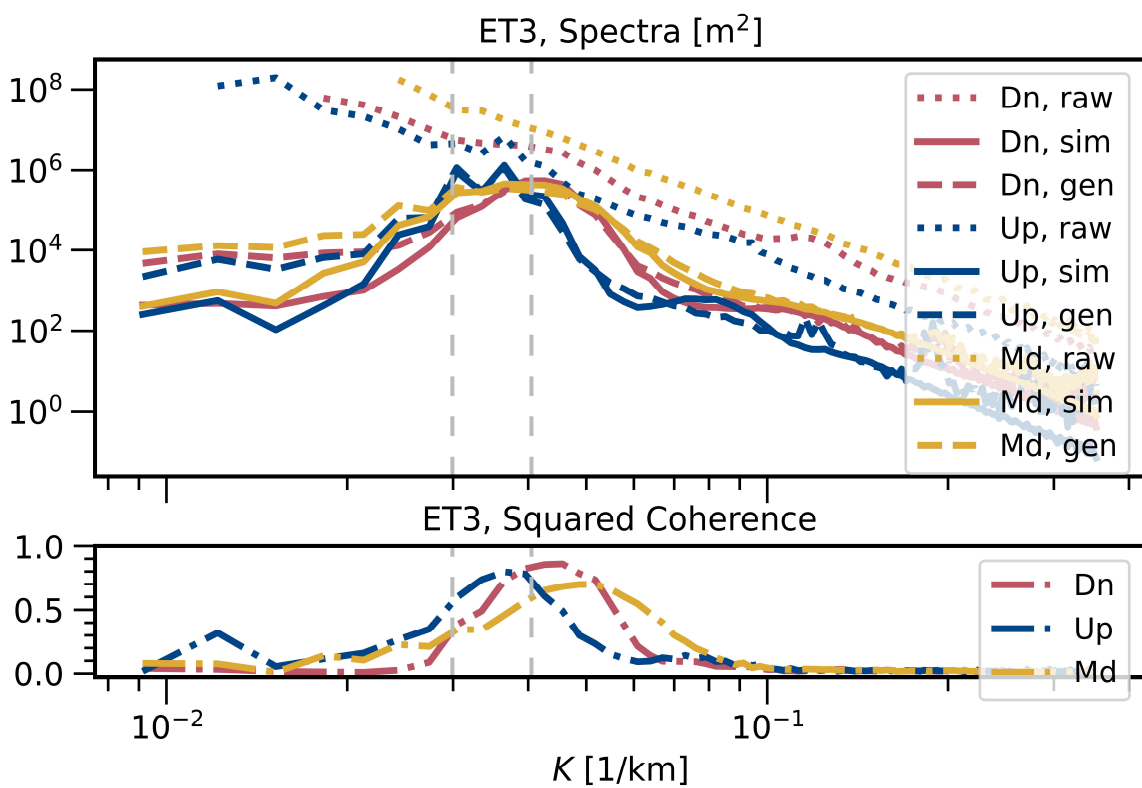
The tidal bumps in the up-jet and down-jet bands are farther apart in higher turbulence runs. In ET1, the tidal bumps in the up-jet and down-jet bands almost completely overlap, while in ET5 the two bumps are shifted apart quite conspicuously. This is consistent with SI Fig.S1, which shows that the tidal 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.



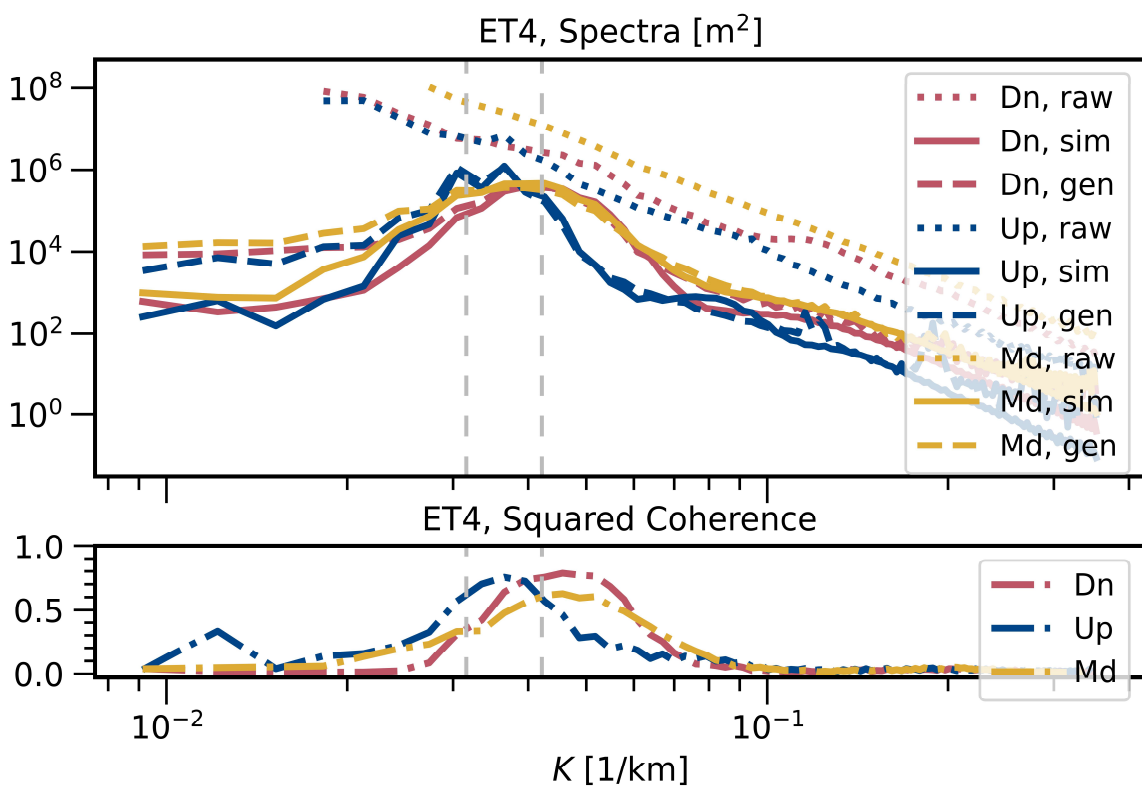
Supporting Information Fig.S2. Spectra and Coherence for the up-jet, mid-jet and down-jet bands in ET1 test set. Compared to Fig.3 in the main article, this figure presents the ET1 test set only, but adds the spectra computed for mid-jet bands (denoted by legend "Md" and coloured yellow) and the squared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



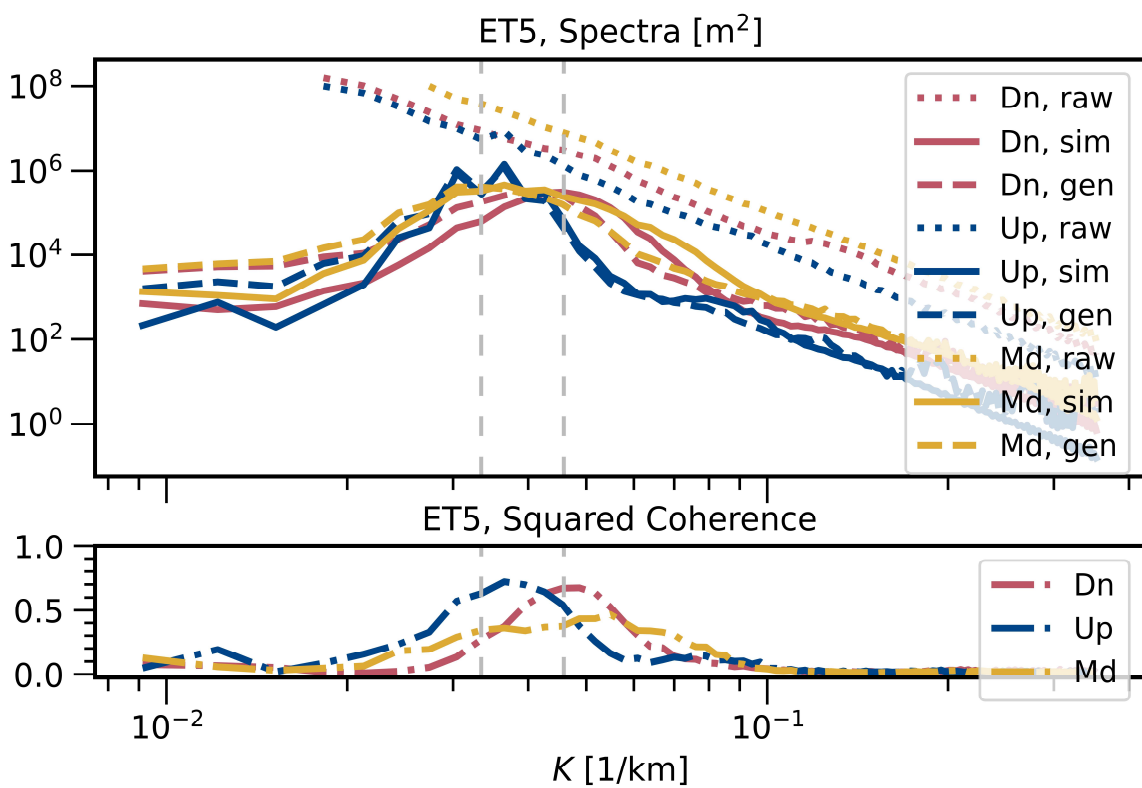
Supporting Information Fig.S3. Similar to SI Fig.S2, but for the ET2 test set.



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.



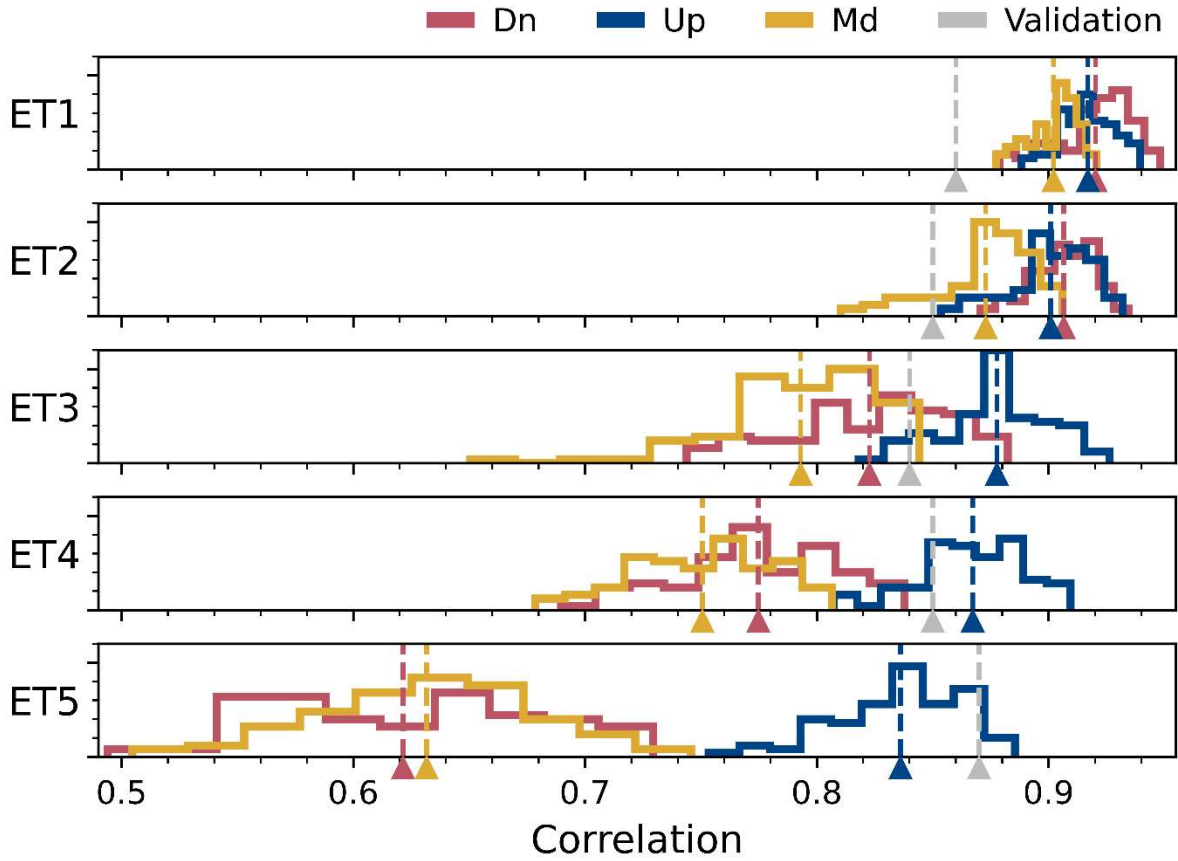
Supporting Information Fig.S6. Similar to SI Fig.S2, but for the ET5 test set.

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 nature⁵. 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.

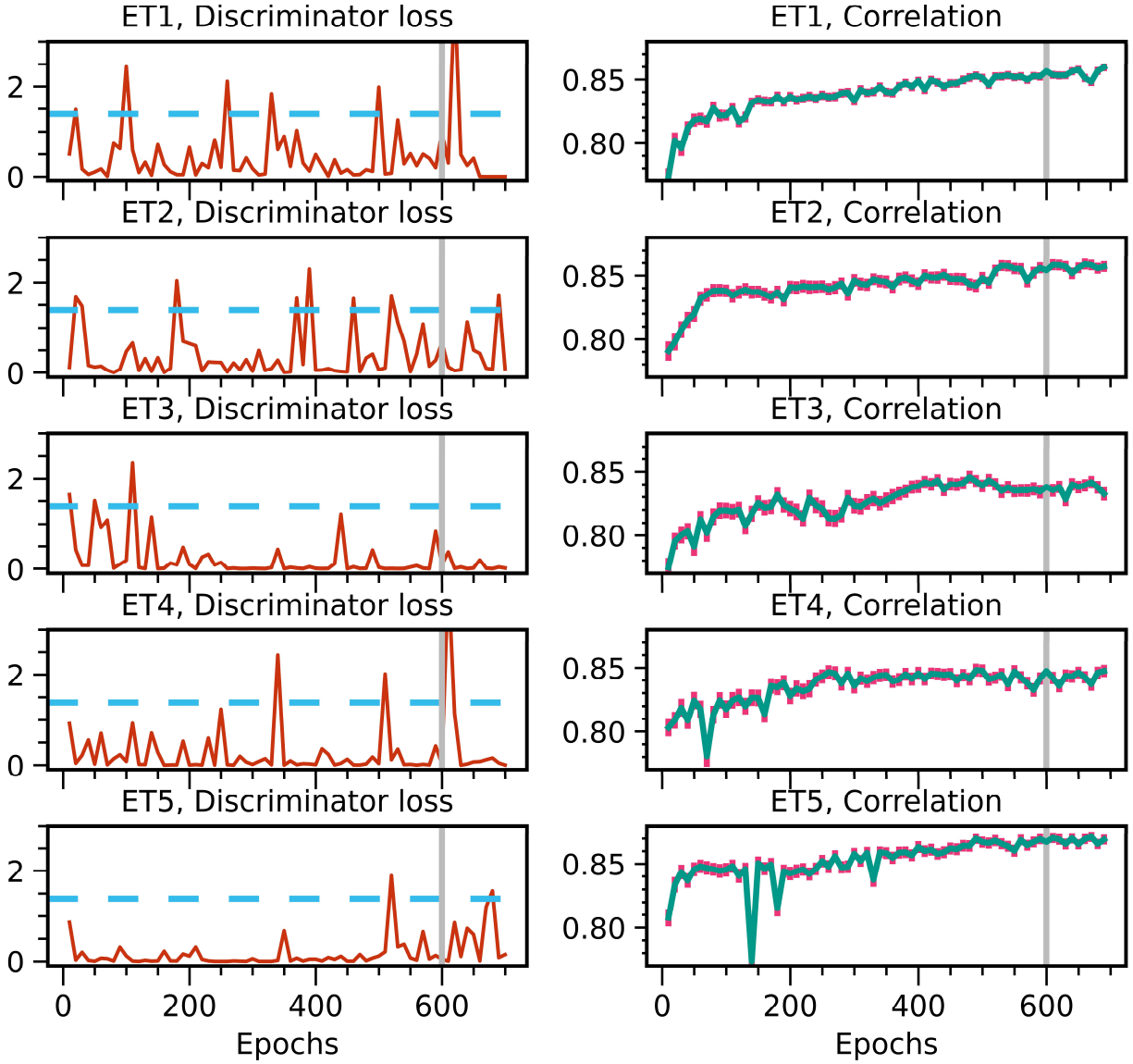
We monitor the training behaviors from two kinds of metrics. First, we monitor the discriminator, the generator, the L1, and the total loss functions respectively as defined in the original publication⁶. Second, we monitor metrics such as correlation factor and relative error

between $\eta_{\cos}^{(\text{sim})}$ and $\eta_{\cos}^{(\text{gen})}$ 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 700th epoch for the ET1-5 runs.

From the definition of discriminator loss in pix2pix⁶, when the discriminator is effectively tossing coins at every judgement, it would have a discriminator loss of $2\log(2)$, 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 noise⁵, or to use a different architecture such as the Wasserstein GAN⁷ 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 $\eta_{\text{cos}}^{(\text{sim})}$ and $\eta_{\text{cos}}^{(\text{gen})}$ 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 $\eta_{\cos}^{(\text{sim})}$ 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 $\Lambda=2$. We refer to Bagrov et al.⁸ for details on the related definitions.

Each square panel of η or $\eta_{\cos}^{(\text{sim})}$ 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 equation 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.

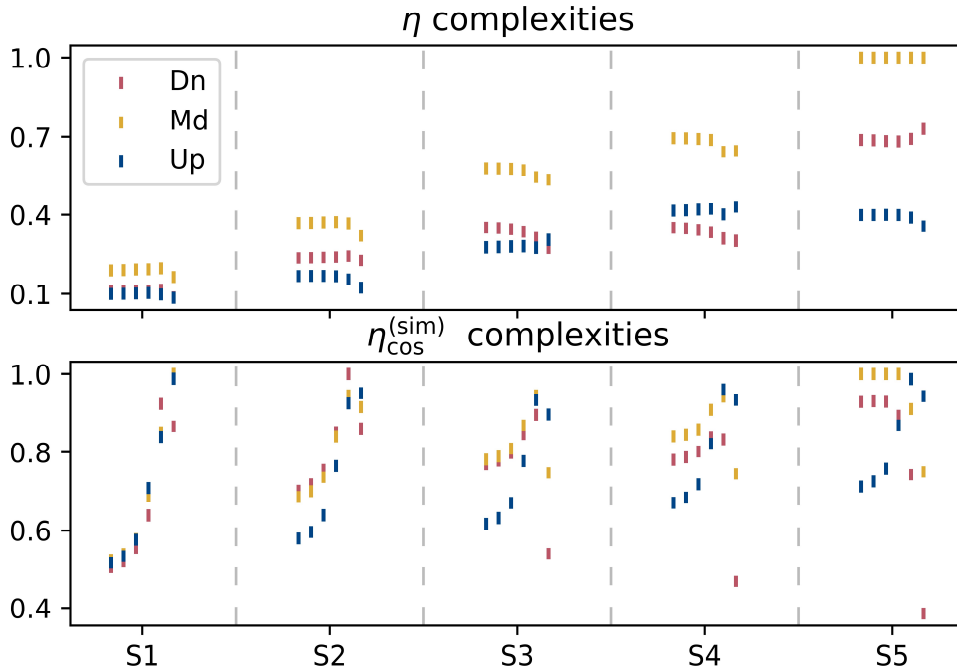
We divide all the panels of η or $\eta_{\cos}^{(\text{sim})}$ into 15 groups (5 turbulence levels and 3 latitudinal bands). Each group contains 100 pairs of $\{\eta, \eta_{\cos}^{(\text{sim})}\}$ panels. The computations of \mathcal{C}_k for η and $\eta_{\cos}^{(\text{sim})}$ are conducted separately. 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 $\eta_{\cos}^{(\text{sim})}$, resulting in 30 different values of averaged \mathcal{C}_k (15 groups each for η and $\eta_{\cos}^{(\text{sim})}$) at each k . Then, we divide \mathcal{C}_k by the maximum value of \mathcal{C}_k among the 15 groups of η or $\eta_{\cos}^{(\text{sim})}$ 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 $\eta_{\cos}^{(\text{sim})}$ panels are 4 km, the metric $\overline{\mathcal{C}}_1$ for $\eta_{\cos}^{(\text{sim})}$ reflects how much the $\eta_{\cos}^{(\text{sim})}$ 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 $\eta_{\text{cos}}^{(\text{sim})}$ 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 $\eta_{\text{cos}}^{(\text{sim})}$ 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 $\eta_{\text{cos}}^{(\text{sim})}$ (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 $\eta_{\cos}^{(\text{gen})}$ and the corresponding $\eta_{\cos}^{(\text{sim})}$, similar to other studies¹⁰, 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 $\eta_{\cos}^{(\text{gen})}$ and $\eta_{\cos}^{(\text{sim})}$ 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 $\eta_{\cos}^{(\text{gen})}$ 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 $\eta_{\cos}^{(\text{gen})}$ against $\eta_{\cos}^{(\text{sim})}$ 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 codes¹³ (hereafter “tutorial codes”). Here, we detail the changes made to the tutorial codes for reproducibility. Some familiarity with the original pix2pix paper⁶ from readers is assumed in the narrations to follow.

First, the $\eta_{\cos}^{(\text{sim})}$ 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 $\eta_{\cos}^{(\text{sim})}$ signals by a uniform factor of 20, after which the max value of $|\eta_{\cos}^{(\text{sim})}|$ 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:

$$\arg \min_G \max_D \mathcal{L}_{\text{CGAN}}(G, D) + \lambda \mathcal{L}_{L1}(G),$$

where $\mathcal{L}_{\text{CGAN}}(G, D)$ 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 images⁶. 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 $\{\eta, 20 \eta_{\cos}^{(\text{sim})}\}$ 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 90° 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 $\eta_{\cos}^{(\text{gen})}$.

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 nomenclature⁶, 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 paper⁶ 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

$\eta_{\cos}^{(\text{sim})}$. 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 $\eta_{\cos}^{(\text{gen})}$ 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 $\eta_{\cos}^{(\text{sim})}$, “Generated” column plots $\eta_{\cos}^{(\text{gen})}$, and “Difference” column plots $(\eta_{\cos}^{(\text{sim})} - \eta_{\cos}^{(\text{gen})})$.

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 $\eta_{\cos}^{(\text{sim})}$. The entire simulation domain is included. Snapshots are ordered by time and separated by $4T$.

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>.