Supplement of Ocean Sci., 22, 2249–2265, 2026
https://doi.org/10.5194/os-22-2249-2026-supplement

Ocean Science
Open Access

Supplement of

Bottom topography effects on the internal wave climate in the Ionian Sea

Florian Kokoszka et al.

Correspondence to: Federico Falcini (federico.falcini@cnr.it)

S1. Spectral significance

We assess the robustness of the differences between binned mean spectra (by Δ-slope or roughness) by simulating small variations in the sampling. This is done using a bootstrap approach, where each group of spectra is repeatedly reconstructed by randomly sampling with replacement spectra, while keeping the same number of samples. This does not alter the data themselves, but slightly changes the contribution of individual profiles to the mean spectrum. In this context, "small variations" refer to these changes in weighting within the existing dataset, without introducing new data or modifying the spectra. By repeating this process many times (3000 bootstrap realizations were generated by resampling spectra with replacement within each group while preserving the original sample size), we evaluate whether the observed differences between bins remain consistent or depend strongly on the specific composition of the original sample. The uncertainty associated with the logarithmic spectral differences is reported as the 95% bootstrap confidence interval.

In **Figures S1a,b** the right panel shows the difference between two binned (Δ-slope bins) mean spectra, expressed as the logarithmic ratio between them. The grey shaded area represents the uncertainty estimated from the bootstrap procedure: it reflects how much the difference varies when the groups are slightly resampled. The horizontal line at zero corresponds to no difference between the two spectra. When the grey interval does not include zero, the difference is considered robust, meaning that it remains consistent despite sampling variability. In contrast, if the grey area touches or crosses zero, the difference is not statistically distinguishable from zero and may arise from sampling noise. Green points highlight the wavenumbers where the difference remains significant after accounting for multiple testing. For this, at each wavenumber, we test whether the difference between the two spectra is significant. Since this is done simultaneously over many wavenumbers, some points may appear significant purely by chance. To account for this chance, we apply a False Discovery Rate (FDR) correction: among all the wavenumbers flagged as significant, only a small and controlled fraction are expected to be false positives. The green points therefore indicate differences that remain statistically robust even after controlling for this multiple testing effect. Green points may occur on both sides of zero when the two spectra cross, or when the sign of the spectral difference changes with vertical scale; they should therefore be interpreted as local significant differences, while the physical interpretation is based on the band-integrated statistics reported in Tables S1 and S2. This indicates spectral changes that are unlikely to be due to random variability alone.

In **Table S1** we summarize the differences between three physical bands (large scales 640–160 m; intermediate scales: 160–80 m; small scales 80–10 m). This summary quantifies the spectral differences between adjacent bins over selected vertical wavenumber bands. For each band, the reported effect corresponds to the logarithmic ratio between the mean spectra of the two bins, indicating whether spectral energy increases or decreases from one bin to the next. The 95% confidence interval, estimated via bootstrap resampling, reflects the uncertainty associated with this difference: if the interval does not include zero, the change is considered robust. The p-value, obtained from a permutation test, evaluates whether such a difference could arise by chance when no real distinction exists between the groups. Together, these metrics provide a simple and quantitative assessment of whether the observed spectral variations over specific scale ranges are statistically significant and physically meaningful.

Table S1. Statistical comparison of adjacent spectral groups (by Δ-slope bins) integrated over three vertical wavenumber bands representative of large scales (640–160 m), intermediate scales (160–80 m), and small scales (80–10 m). For each pairwise comparison, the table reports the logarithmic spectral difference between groups (Effect_log10), the associated 95% bootstrap confidence interval (CI_low, CI_high), and the permutation-test p-value. Positive values indicate enhanced spectral energy in the second group relative to the first, while negative values indicate reduced energy.

Type	**Pair**	**Band**	**Effect_log10**	**CI_low**	**CI_high**	**p_value**
SH	SH 1 vs 2	large_scales	-0.029	-0.057	0.003	0.0567
SH	SH 1 vs 2	inter_scales	0.004	-0.034	0.041	0.8620
SH	SH 1 vs 2	small_scales	-0.042	-0.079	-0.003	0.0220
SH	SH 2 vs 3	large_scales	-0.067	-0.092	-0.042	0.0000
SH	SH 2 vs 3	inter_scales	-0.061	-0.096	-0.026	0.0003
SH	SH 2 vs 3	small_scales	-0.089	-0.117	-0.063	0.0000
SH	SH 3 vs 4	large_scales	-0.087	-0.112	-0.065	0.0000
SH	SH 3 vs 4	inter_scales	-0.060	-0.089	-0.032	0.0000
SH	SH 3 vs 4	small_scales	0.004	-0.022	0.030	0.7107
SH	SH 4 vs 5	large_scales	-0.112	-0.135	-0.089	0.0000
SH	SH 4 vs 5	inter_scales	-0.117	-0.149	-0.085	0.0000
SH	SH 4 vs 5	small_scales	-0.044	-0.066	-0.022	0.0003
SH	SH 5 vs 6	large_scales	0.001	-0.027	0.030	0.9427
SH	SH 5 vs 6	inter_scales	0.137	0.097	0.177	0.0000
SH	SH 5 vs 6	small_scales	0.069	0.042	0.095	0.0000
ST	ST 1 vs 2	large_scales	-0.267	-0.364	-0.163	0.0000
ST	ST 1 vs 2	inter_scales	-0.232	-0.323	-0.131	0.0000
ST	ST 1 vs 2	small_scales	-0.138	-0.195	-0.079	0.0000
ST	ST 2 vs 3	large_scales	0.043	-0.022	0.108	0.1883
ST	ST 2 vs 3	inter_scales	0.038	-0.026	0.104	0.2580
ST	ST 2 vs 3	small_scales	0.063	0.036	0.089	0.0000
ST	ST 3 vs 4	large_scales	0.076	0.011	0.138	0.0250
ST	ST 3 vs 4	inter_scales	0.080	0.014	0.142	0.0177
ST	ST 3 vs 4	small_scales	0.072	0.042	0.102	0.0000
ST	ST 4 vs 5	large_scales	0.113	0.051	0.178	0.0003
ST	ST 4 vs 5	inter_scales	0.126	0.065	0.192	0.0000
ST	ST 4 vs 5	small_scales	0.051	0.026	0.080	0.0003
ST	ST 5 vs 6	large_scales	0.403	0.321	0.479	0.0000
ST	ST 5 vs 6	inter_scales	0.411	0.329	0.488	0.0000
ST	ST 5 vs 6	small_scales	0.221	0.190	0.252	0.0000

For shear (SH), the differences between bins are not always significant and do not follow a smooth trend. However, clear and robust changes appear at intermediate and small scales. In particular, there is a strong increase from bins 3 to 5, followed by a marked decrease between bins 5 and 6. This suggests a non-linear evolution, with distinct transitions rather than a gradual change. For strain (ST), the signal is clearer and more consistent. Most differences are significant across all scales, with larger amplitudes than for shear. The spectrum shows alternating increases and decreases between bins, indicating a strong and structured response to the binning parameter. The results show that the spectral variations are real and not due to noise, with strain displaying a more coherent behavior than shear.

Note that the same analysis is also provided for spectra organized by roughness bins, leading to qualitatively similar and coherent results (**Figures S2a,b; Table S2**).

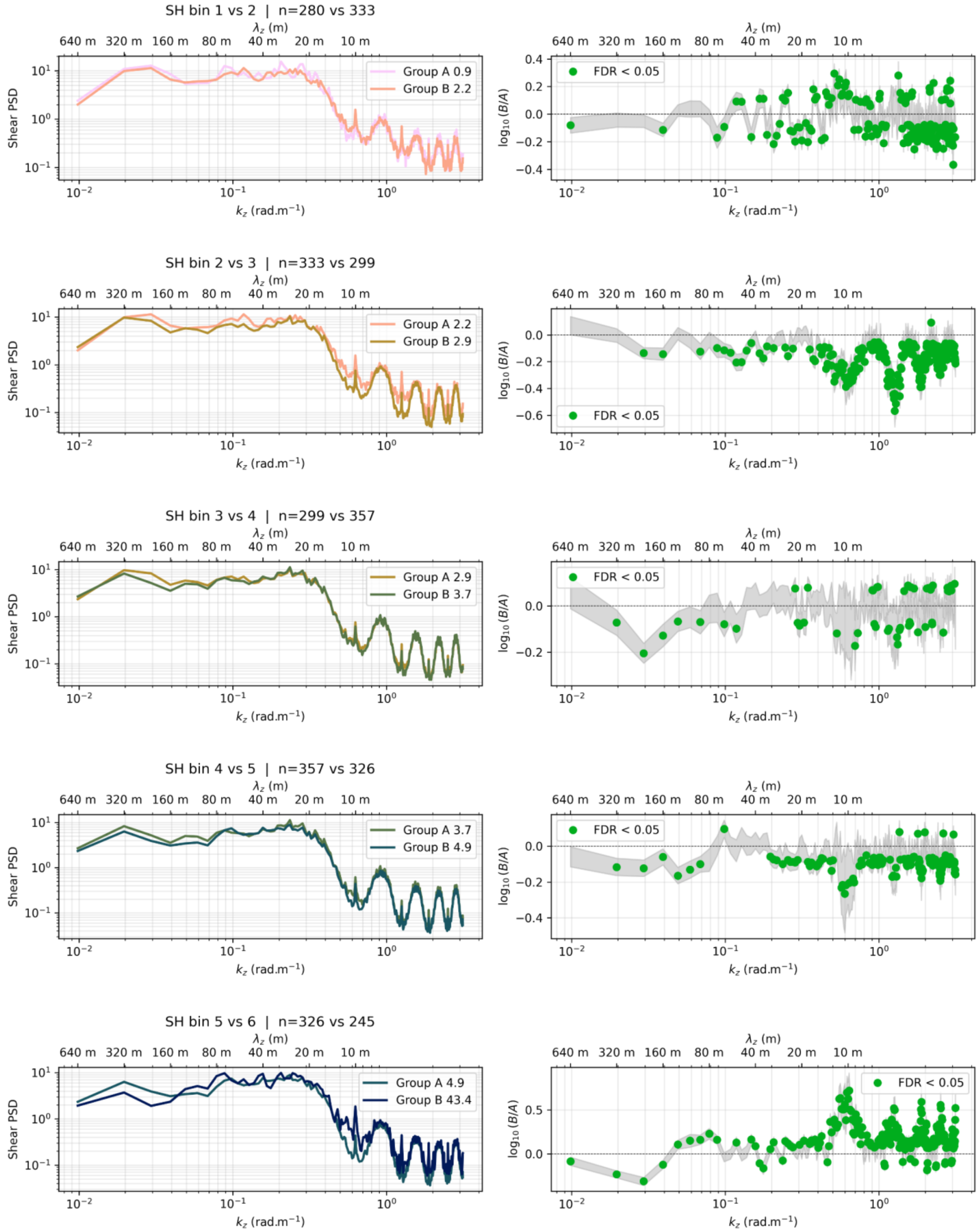

Figure S1a. Representative statistical comparisons between adjacent Δ-slope bins for shear spectra. Left panels show the mean spectra for the two groups, while right panels show the logarithmic spectral difference together with bootstrap confidence intervals (grey shading) and FDR-corrected significant wavenumbers (green markers). The grey shaded area represents the 95% bootstrap confidence interval of the logarithmic spectral difference.

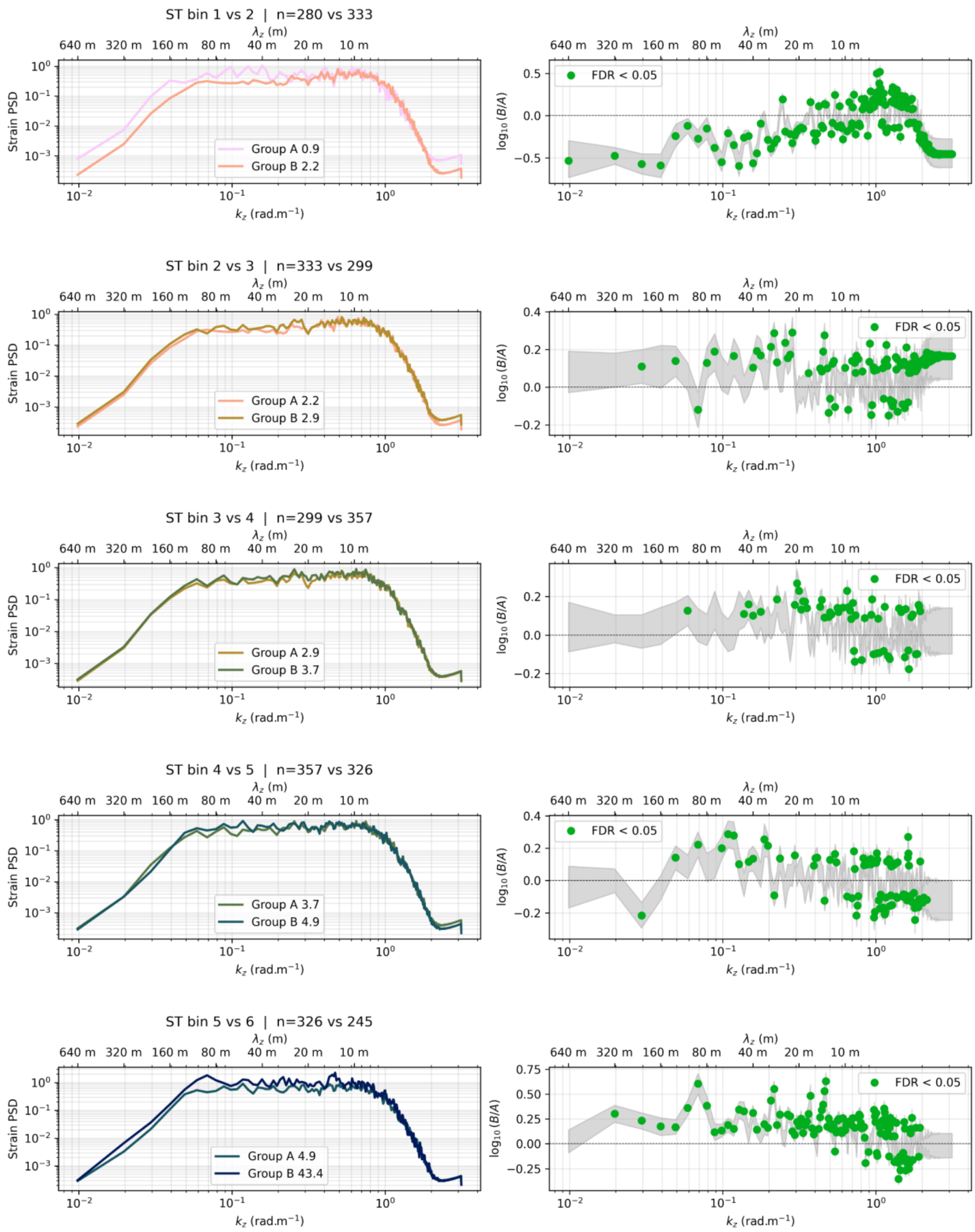

Figure S1b. Representative statistical comparisons between adjacent Δ-slope bins for strain spectra. Left panels show the mean spectra for the two groups, while right panels show the logarithmic spectral difference together with bootstrap confidence intervals (grey shading) and FDR-corrected significant wavenumbers (green markers). The grey shaded area represents the 95% bootstrap confidence interval of the logarithmic spectral difference.

Table S2. Statistical comparison of adjacent spectral groups (by roughness bins) integrated over three vertical wavenumber bands representative of large scales (640–160 m), intermediate scales (160–80 m), and small scales (80–10 m). For each pairwise comparison, the table reports the logarithmic spectral difference between groups (Effect_log10), the associated 95% bootstrap confidence interval (CI_low, CI_high), and the permutation-test p-value. Positive values indicate enhanced spectral energy in the second group relative to the first, while negative values indicate reduced energy.

Type	**Pair**	**Band**	**Effect_log10**	**CI_low**	**CI_high**	**p_value**
SH	SH 1 vs 2	large_scales	-0.011	-0.044	0.023	0.5433
SH	SH 1 vs 2	inter_scales	0.080	0.037	0.123	0.0010
SH	SH 1 vs 2	small_scales	-0.017	-0.038	0.005	0.1297
SH	SH 2 vs 3	large_scales	0.063	0.021	0.108	0.0037
SH	SH 2 vs 3	inter_scales	-0.006	-0.048	0.035	0.7973
SH	SH 2 vs 3	small_scales	0.105	0.079	0.129	0.0000
SH	SH 3 vs 4	large_scales	-0.010	-0.047	0.028	0.5943
SH	SH 3 vs 4	inter_scales	-0.028	-0.060	0.003	0.0867
SH	SH 3 vs 4	small_scales	0.070	0.042	0.098	0.0000
SH	SH 4 vs 5	large_scales	0.082	0.055	0.110	0.0000
SH	SH 4 vs 5	inter_scales	0.175	0.143	0.205	0.0000
SH	SH 4 vs 5	small_scales	0.180	0.160	0.202	0.0000
SH	SH 5 vs 6	large_scales	-0.190	-0.230	-0.151	0.0000
SH	SH 5 vs 6	inter_scales	-0.168	-0.208	-0.127	0.0000
SH	SH 5 vs 6	small_scales	-0.457	-0.493	-0.424	0.0000
ST	ST 1 vs 2	large_scales	-0.120	-0.218	-0.028	0.0077
ST	ST 1 vs 2	inter_scales	-0.104	-0.205	-0.008	0.0267
ST	ST 1 vs 2	small_scales	0.044	0.011	0.077	0.0123
ST	ST 2 vs 3	large_scales	0.339	0.243	0.437	0.0000
ST	ST 2 vs 3	inter_scales	0.331	0.237	0.430	0.0000
ST	ST 2 vs 3	small_scales	0.065	0.025	0.105	0.0013
ST	ST 3 vs 4	large_scales	-0.153	-0.216	-0.090	0.0000
ST	ST 3 vs 4	inter_scales	-0.164	-0.228	-0.098	0.0000
ST	ST 3 vs 4	small_scales	-0.097	-0.134	-0.061	0.0000
ST	ST 4 vs 5	large_scales	-0.230	-0.297	-0.163	0.0000
ST	ST 4 vs 5	inter_scales	-0.235	-0.301	-0.170	0.0000
ST	ST 4 vs 5	small_scales	-0.072	-0.102	-0.041	0.0000
ST	ST 5 vs 6	large_scales	0.664	0.571	0.750	0.0000
ST	ST 5 vs 6	inter_scales	0.680	0.593	0.769	0.0000
ST	ST 5 vs 6	small_scales	0.329	0.286	0.368	0.0000

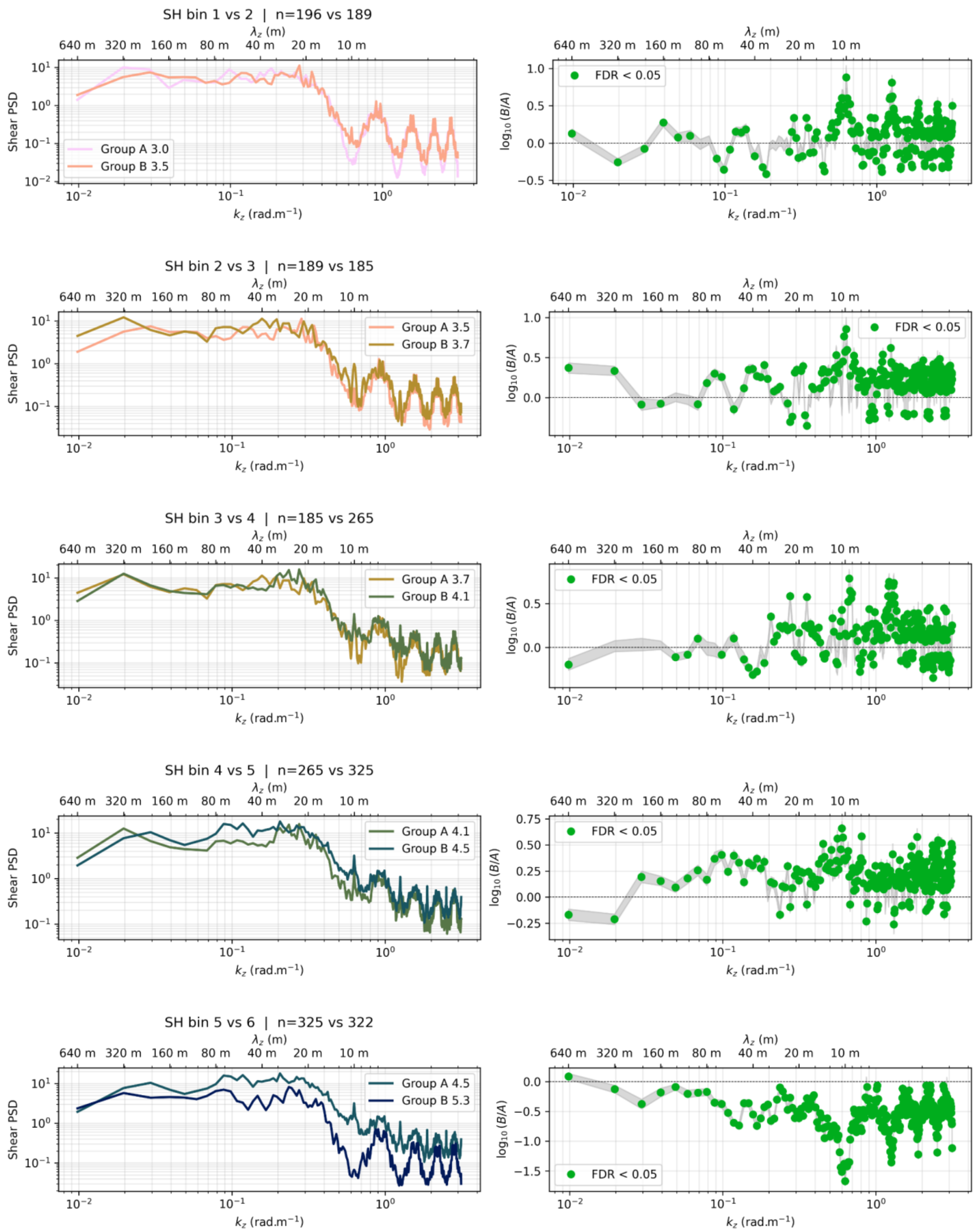

Figure S2a. Representative statistical comparisons between adjacent roughness bins for shear spectra. Left panels show the mean spectra for the two groups, while right panels show the logarithmic spectral difference together with bootstrap confidence intervals (grey shading) and FDR-corrected significant wavenumbers (green markers). The grey shaded area represents the 95% bootstrap confidence interval of the logarithmic spectral difference.

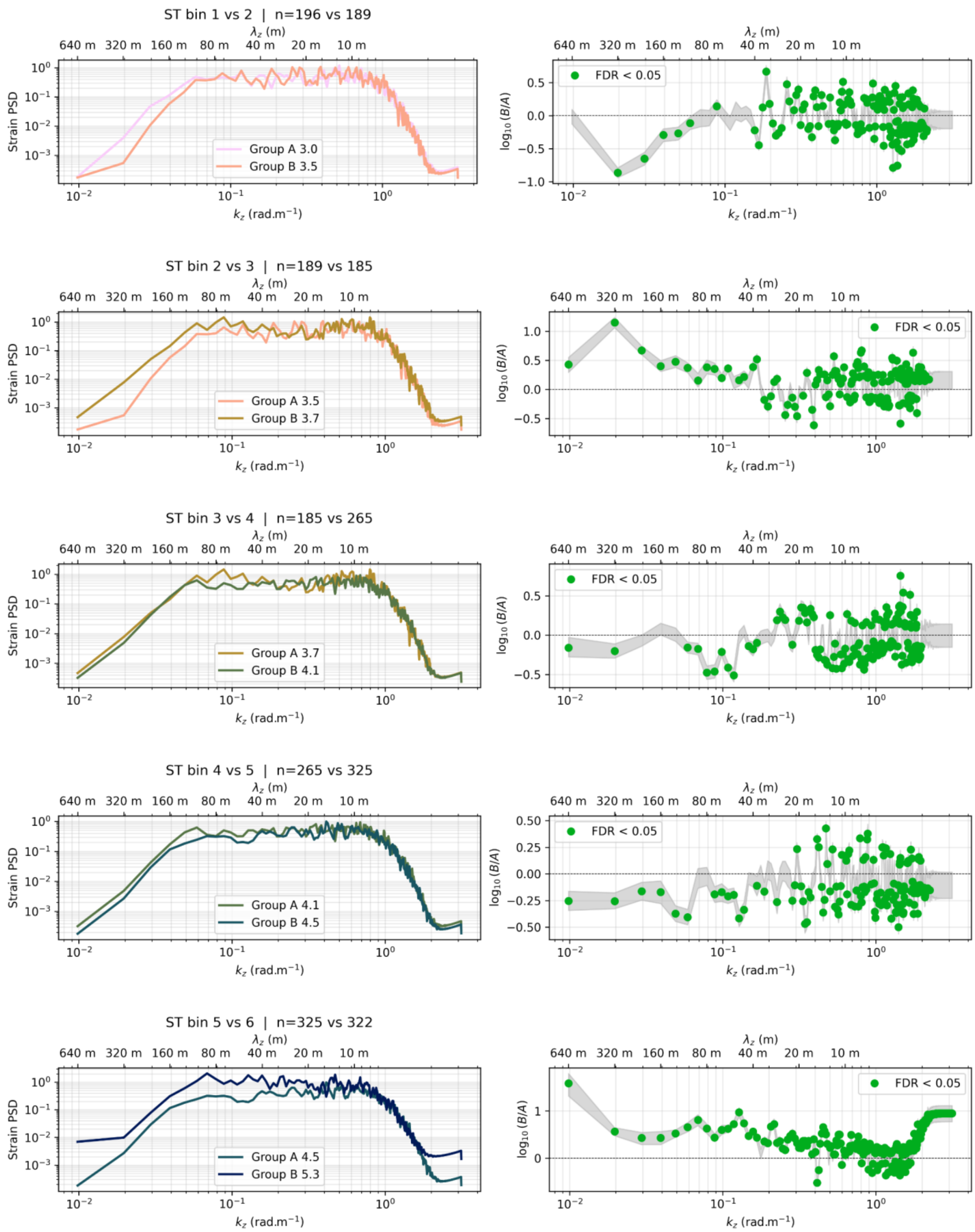

Figure S2b. Representative statistical comparisons between adjacent roughness bins for strain spectra. Left panels show the mean spectra for the two groups, while right panels show the logarithmic spectral difference together with bootstrap confidence intervals (grey shading) and FDR-corrected significant wavenumbers (green markers). The grey shaded area represents the 95% bootstrap confidence interval of the logarithmic spectral difference.

S2. Sensitivity of the shear-to-strain spectral ratios to bandwidth separation

For each individual spectral window i whose centre is located within 960 m above the bottom we calculate shear and strain spectral ratios relative to their respective Garrett–Munk (GM) reference spectra, hereafter $R_{sh}(i, k_z) = S_{sh}/S_{sh,GM}$ and $R_{st}(i, k_z) = S_{st}/S_{st,GM}$, where S_{sh},S_{st} are shear and strain spectra respectively, and GM denotes their reference model. These ratios quantify the local enhancement or depletion of shear and strain variance at each vertical wavenumber. These ratios are then combined together as $R = R_{sh}/R_{st}$, documenting a ratio representative of the shear-to-strain in its two-dimensional matrix form. For each spectrum i, we obtain a matrix $R(i, k_{z,sh}, k_{z,st})$ containing all the ratios possible between spectral components, allowing to test the homogeneity of the ratios among choices of vertical wavenumbers in $k_{z,sh}$ and $k_{z,st}$. For every single ratio R, variants of the median are diagnosed for each bandwidth configuration listed in Table S3, each bandwidth representing a subset. Spearman correlations will be then tested between each median series of R associated to each subset.

For those, several spectral subsets are subsequently defined (**Table S3**) by selecting specific vertical wavelength ranges independently for shear and strain. The corresponding submatrices are extracted, and their median value is computed to provide an estimate representing R associated with different choices of spectral bands. Different families of spectral configurations are considered to assess the robustness of R estimates to band selection (**Table S3**).

- The reference case (***REF***) uses the nominal asymmetric configuration of our analysis, with a broad shear range (640–107 m) chosen to maximize signal while avoiding the noisiest small scales, and a strain range (128–10 m) consistent with the high-pass filtering (160m) applied to remove large-scale contributions.
- *St_redux* configurations isolate the effect of progressively restricting the strain range while keeping the shear band fixed: ***St_redux1*** (128–30 m) and ***St_redux2*** (128–64 m) test the sensitivity of the estimates to the inclusion of shorter vertical wavelengths in strain, which may carry more variability or residual noise.
- *Sh_redux* configurations perform the complementary test by progressively restricting the shear range while keeping the last restrained strain band (128–64 m): ***Sh_redux1*** (320–107 m) and ***Sh_redux2*** (160–107 m) evaluate how the removal of larger shear scales impacts the resulting estimates.
- A strictly consistent configuration (***SAME_redux***) is also considered, using an identical and conservative band for both shear and strain (160–107 m). This provides a reference case free of asymmetry between variables, albeit over a reduced spectral range.
- Three additional "same-band" configurations (***same1***: 160–80 m, ***same2***: 107–64 m, ***same3***: 80–30 m) explore progressively smaller common bands. These cases are used as sensitivity tests to assess how the estimates behave when both variables are restricted to narrower and shorter-wavelength ranges, approaching the limits where noise and resolution effects may become significant.

Table S3. Summary of the spectral bandwidth configurations used to assess the sensitivity of the *R* diagnostic to the choice of vertical wavenumber ranges. The reference configuration (**REF**) corresponds to the nominal asymmetric setup used in the main analysis, while the **St_redux** and **Sh_redux** cases progressively restrict the strain and shear wavelength ranges, respectively. **SAME_redux** uses identical conservative bands for both variables, and **same1–same3** explore progressively smaller common wavelength ranges. SH and ST indicate the number of spectral points retained within each wavelength interval. These configurations are used to evaluate the robustness and consistency of the diagnostic across different spectral band selections.

Configuration	SH	ST
REF	5 [640–107 m]	60 [128–10 m]
St_redux1	5 [640–107 m]	17 [128–30 m]
St_redux2	5 [640–107 m]	6 [128–64 m]
Sh_redux1	4 [320–107 m]	6 [128–64 m]
Sh_redux2	2 [160–107 m]	6 [128–64 m]
SAME_redux	2 [160–107 m]	2 [160–107 m]
same1	5 [160–80 m]	5 [160–80 m]
same2	5 [107–64 m]	5 [107–64 m]
same3	14 [80–30 m]	14 [80–30 m]

To assess the consistency between the different band selections, we compare medians across all individual spectra. First, we compute pairwise Spearman correlations between subsets to evaluate whether they capture the same variability, independently of amplitude. Then, we examine the relative ordering between subsets for each profile, by checking whether one subset systematically yields higher or lower values than the others. This combination of covariance (correlation) and ordering analyses allows us to evaluate whether using different, including asymmetric, spectral bands alters the structure of the shear-to-strain diagnostic or simply affects its local amplitude without introducing a systematic bias.

The correlations between subsets are consistently high (**Table S4**; mean Spearman $\rho \approx 0.75$), showing that *R* estimates evolve coherently across a wide range of bandwidth choices. In particular, the strong agreement between nearby configurations (e.g. successive reductions in shear or strain ranges) indicates that the distributions we obtained are not strongly sensitive to the exact bandwidth definition. At the same time, there is no stable ordering across subsets (fraction ordered = 0; REF > St_redux1 > St_redux2 > Sh_redux1 > Sh_redux2 > SAME_redux > same1 > same2 > same3 = 0.0). This means that no configuration systematically gives higher or lower values than the others. The relative ranking between subsets changes from one spectrum to another, indicating that differences between configurations remain secondary compared to the shared variability. Together, these results support that the distribution is robust to the choice of spectral bands. The spectrum is not strictly single-frequency, but it retains a coherent structure across scales.

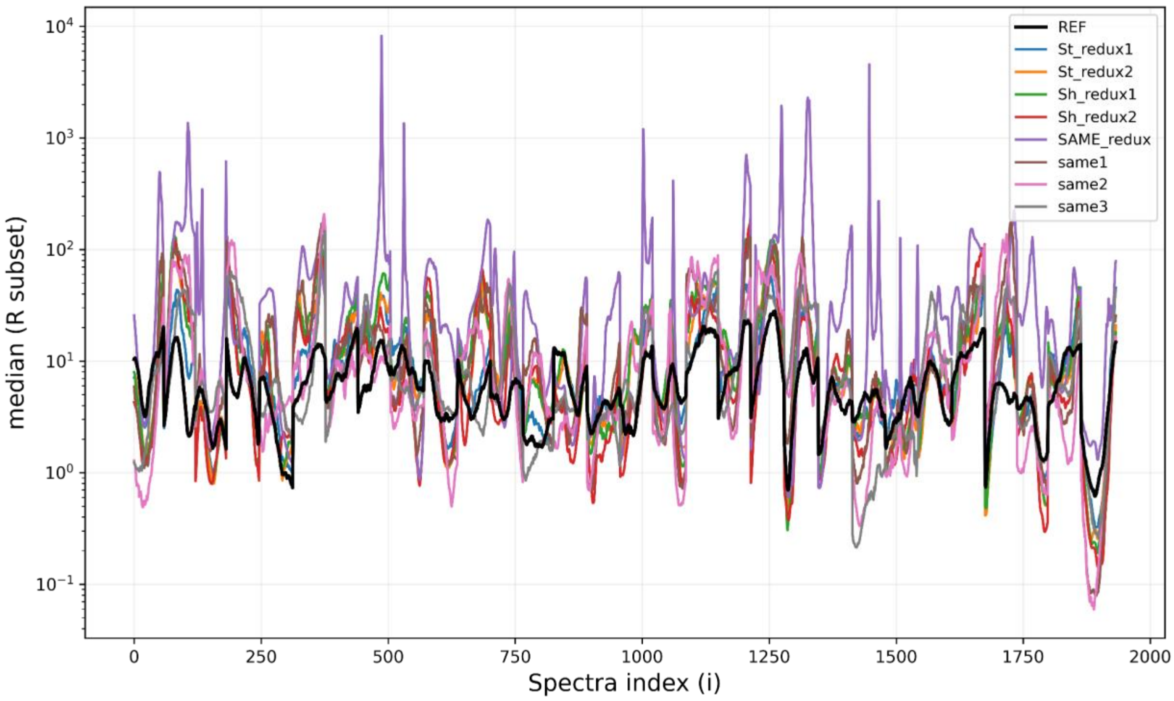

Figure S3. **a)** Median of *R* ratio across spectral subsets. Each curve represents the median value computed independently for a given spectral subset configuration (e.g., reduced or same-band integrations; **Table S3**), obtained from each spectral window i. The comparison highlights the robustness and variability of the shear-to-strain ratio across different choices of shear and strain integration ranges. Here, i is the index of the pooled list of all spectra used in the test, obtained by stacking spectra from all stations and all sliding-window positions whose centre is located within 960 m above the bottom; **b)** Comparison between Rω* obtained with the reference configuration (REF) and the conservative same-band configuration (SAME_redux). SAME_redux uses identical shear and strain wavelength bands and therefore changes the absolute values of Rω*, as expected from the reduced spectral bandwidth. However, the broad distribution of Rω* remains consistent with REF, indicating that the main interpretation based on the spatial organization of the diagnostic is not changed.

Table S4. Pairwise Spearman rank correlations between *R* ratio estimates obtained from the different spectral bandwidth configurations described in **Table S3**. High correlation values indicate that the relative organization and variability of the diagnostic remain largely consistent across bandwidth choices, supporting the robustness of the observed patterns against changes in spectral integration ranges.

	REF	**St_redux1**	**St_redux2**	**Sh_redux1**	**Sh_redux2**	**SAME_redux**	**same1**	**same2**
St_redux1	0.87							
St_redux2	0.76	0.89						
Sh_redux1	0.77	0.89	0.99					
Sh_redux2	0.69	0.83	0.92	0.93				
SAME_redux	0.54	0.63	0.69	0.69	0.76			
same1	0.61	0.78	0.87	0.87	0.91	0.77		
same2	0.47	0.67	0.8	0.8	0.79	0.52	0.86	
same3	0.62	0.8	0.7	0.7	0.66	0.49	0.72	0.78

S3. Lee wave

We estimated a characteristic topographic forcing frequency as $\omega = |U|\, k_{eff}$, where the effective horizontal wavenumber $k_{eff} = 2\pi\, s \,/\, H_{rms}$ is estimated with s, i.e., the bathymetrical slope, and H_{rms} is the root-mean-square height of the detrended bathymetry. The velocity $|U|$ and stratification were taken from the observed profiles and the resulting ω was normalized by N and f, and referenced to distance above the bottom and averaged in bins of 5m.

In **Figure S4** the normalized values of ω/N remain low, typically between 0 and ~0.2–0.25, indicating that the diagnosed frequencies are well below the buoyancy frequency. In contrast, ω/f shows a broader distribution, but most values remain below 1, i.e. in the sub-inertial regime, with only a few localized cases approaching or slightly exceeding unity.
Overall, these results indicate that the mean-flow/topography interaction does not fall within the dynamical regime required for lee-wave generation, and therefore lee waves are unlikely to be the dominant mechanism controlling the observed variability.

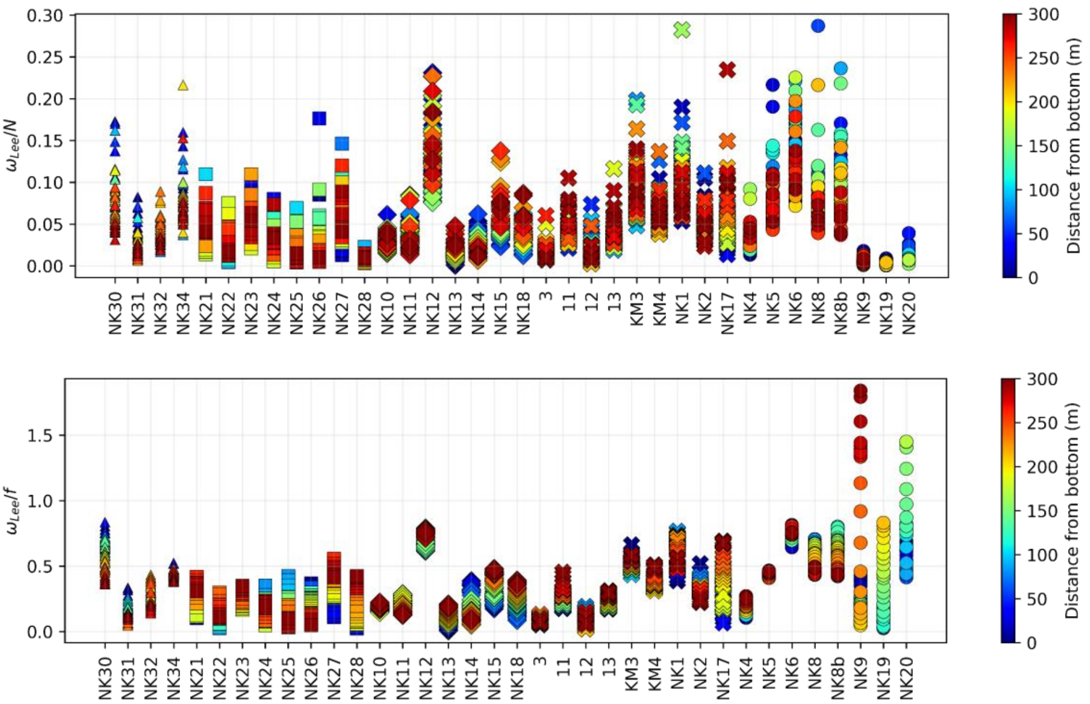

Figure S4. Normalized characteristic topographic forcing frequency estimated from the interaction between the observed flow and bottom topography. The upper panel shows the forcing frequency normalized by stratification, while the lower panel shows the same quantity normalized by the Coriolis frequency. Values are organized by station and colored according to distance from the bottom. Most values remain well below the buoyancy frequency and largely within the subinertial regime relative to the Coriolis frequency, suggesting that the observed flow–topography interaction does not predominantly fall within a classical lee-wave generation regime.