

# Appendix

---

## Even more results

*Torture the data, and it will confess to anything."*

- Ronald Coase

### 1.1 INTRODUCTION

In this appendix, you will find the offline results of phases 5 and 6, as well as all the online results involving parameterizations trained on jet-driven flows or evaluated on online simulations of jet flows. Due to the substantial number of results generated, it should be noted that all the missing offline results and online results that involve eddy-trained parameterizations or testing with eddy datasets are available in my GitHub repository. Indeed, less than a third of the results are shown in this master thesis.

Throughout this work, ensuring the reproducibility of my results was my top priority. Therefore, all the codes and notebooks created for this work or based on the benchmarking framework of Ross et al., 2023 have been thoroughly documented. This documentation not only makes it easier for those interested to understand how to use the benchmarking framework, but also allows for a comprehensive overview of all the work done. One of my proudest achievements in terms of code is the master thesis notebook, which is concise, well documented, and enables easy dataset generation, training of new parameterizations, and both offline and online testing.

<https://github.com/VikVador/Ocean-subgrid-parameterizations-in-an-idealized-model-using-machine-learning>

# PHASE I



Figure 18: |Offline - Phase 1| This table summarizes offline results, including correlations (columns 1 and 2) and mean-squared errors (columns 3 and 4), for parameterizations trained on **eddies unique (5000)** and evaluated on dataset **eddies offline**. On the right, input and output details for each parameterization are provided. The last row showcases results from the 3 FCNN parameterizations presented in Ross et al., 2023.

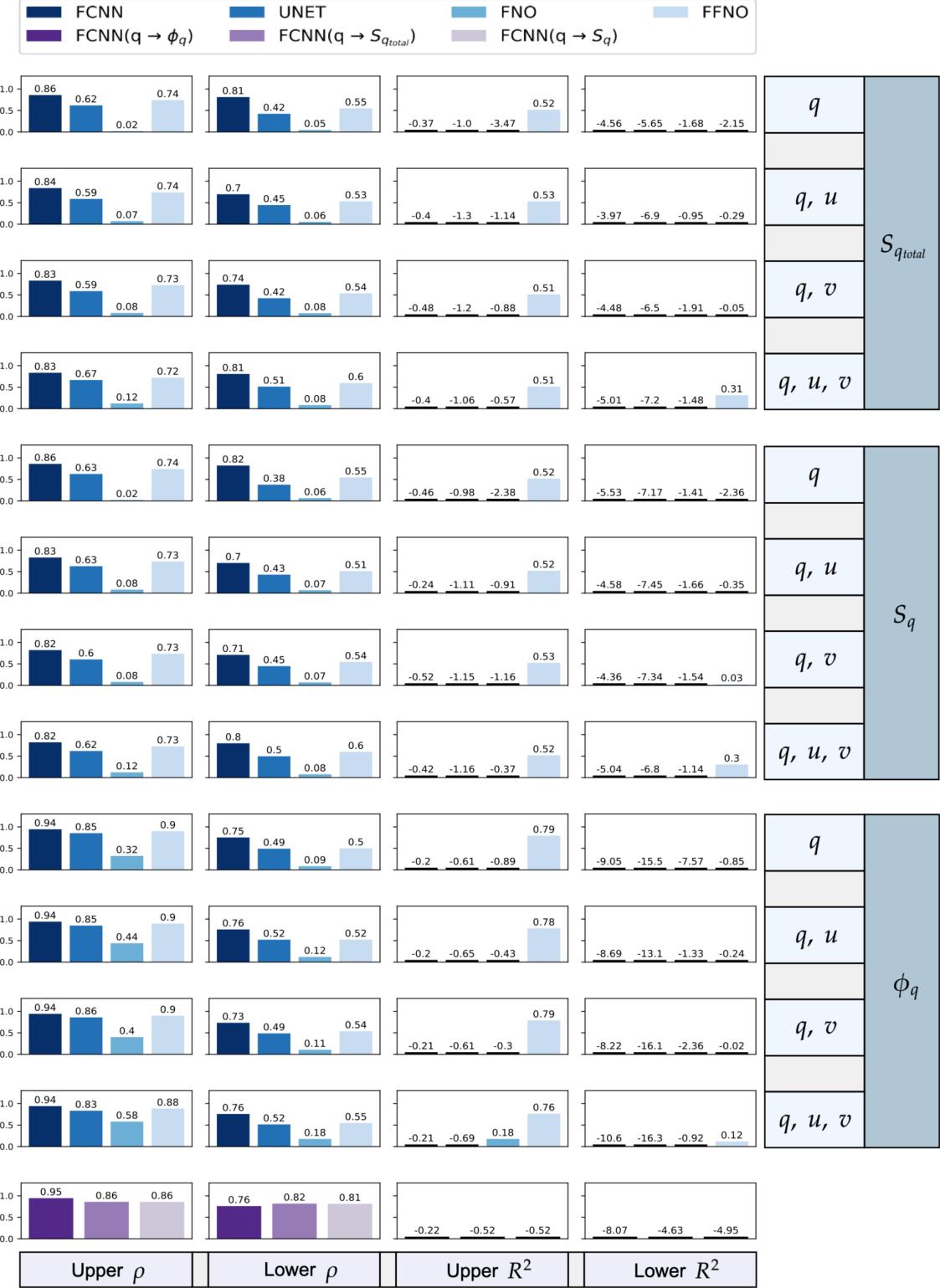


Figure 19: |Offline - Phase 1| This table summarizes offline results, including correlations (columns 1 and 2) and mean-squared errors (columns 3 and 4), for parameterizations trained on **eddies unique (5000)** and evaluated on dataset **jets offline**. On the right, input and output details for each parameterization are provided. The last row showcases results from the 3 FCNN parameterizations presented in Ross et al., 2023.

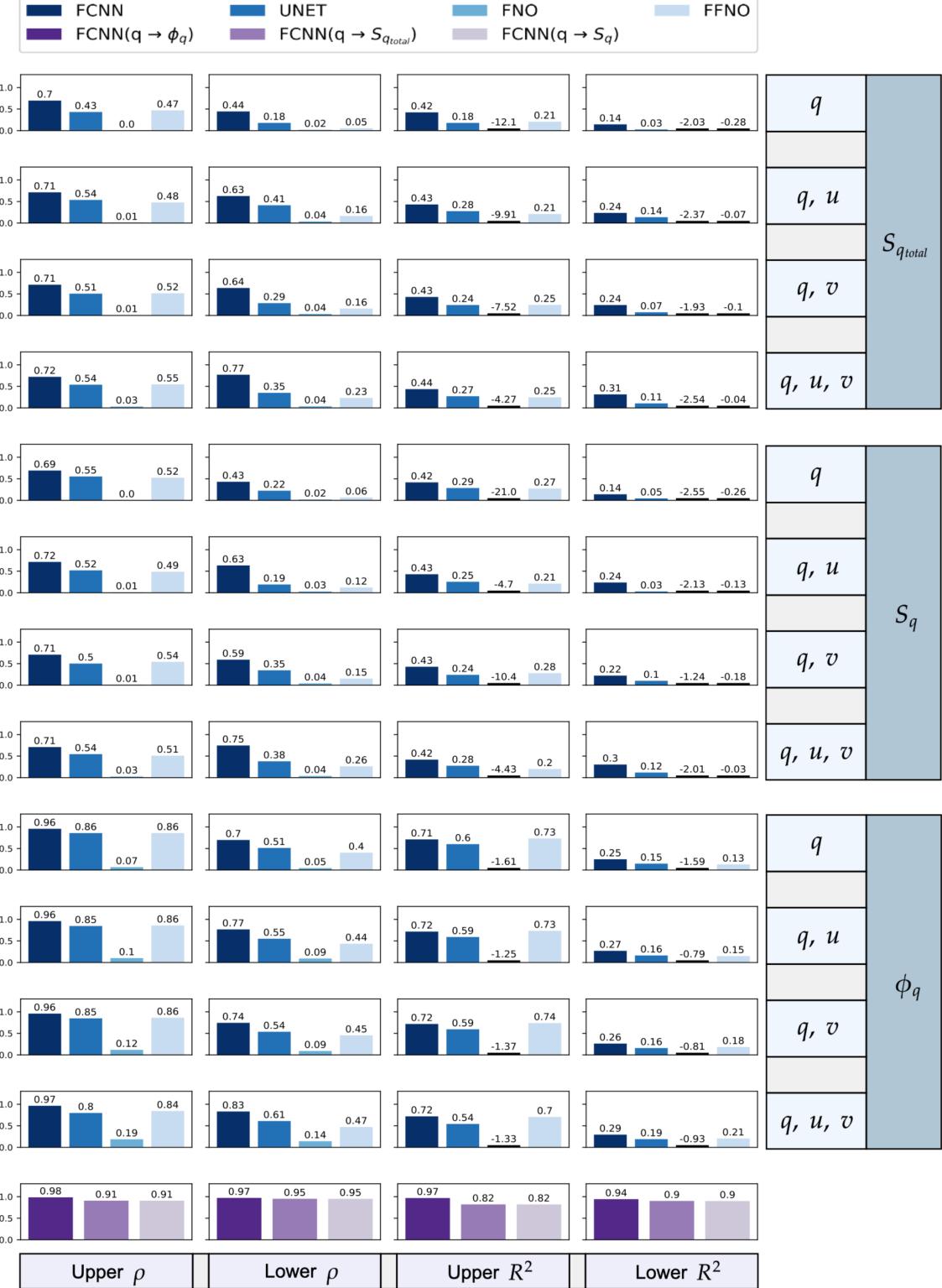


Figure 20: |Offline - Phase 1| This table summarizes offline results, including correlations (columns 1 and 2) and mean-squared errors (columns 3 and 4), for parameterizations trained on jets unique (5000) and evaluated on dataset eddies offline. On the right, input and output details for each parameterization are provided. The last row showcases results from the 3 FCNN parameterizations presented in Ross et al., 2023.

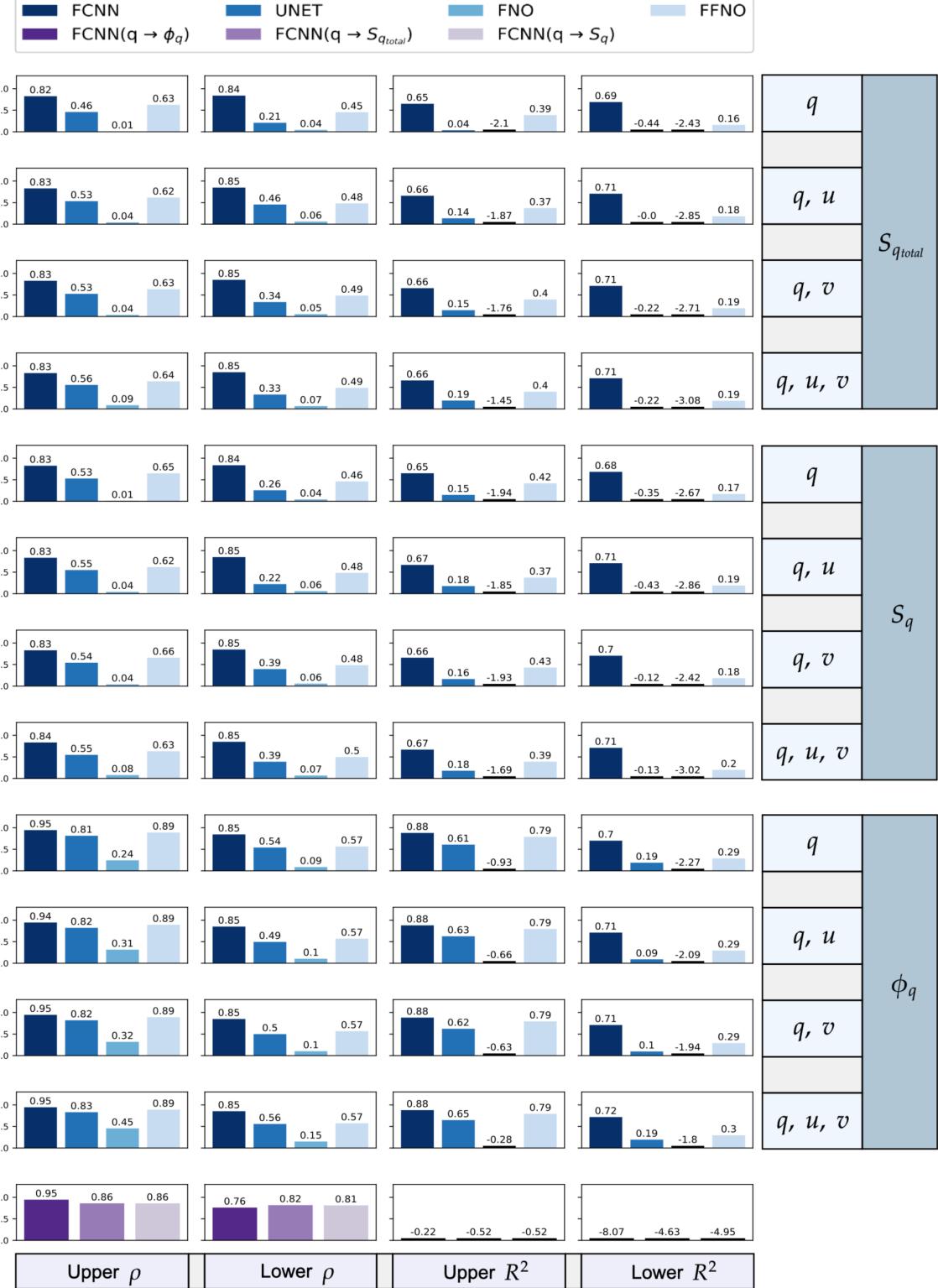


Figure 21: |Offline - Phase 1| This table summarizes offline results, including correlations (columns 1 and 2) and mean-squared errors (columns 3 and 4), for parameterizations trained on jets **unique (5000)** and evaluated on dataset jets **offline**. On the right, input and output details for each parameterization are provided. The last row showcases results from the 3 FCNN parameterizations presented in Ross et al., 2023.

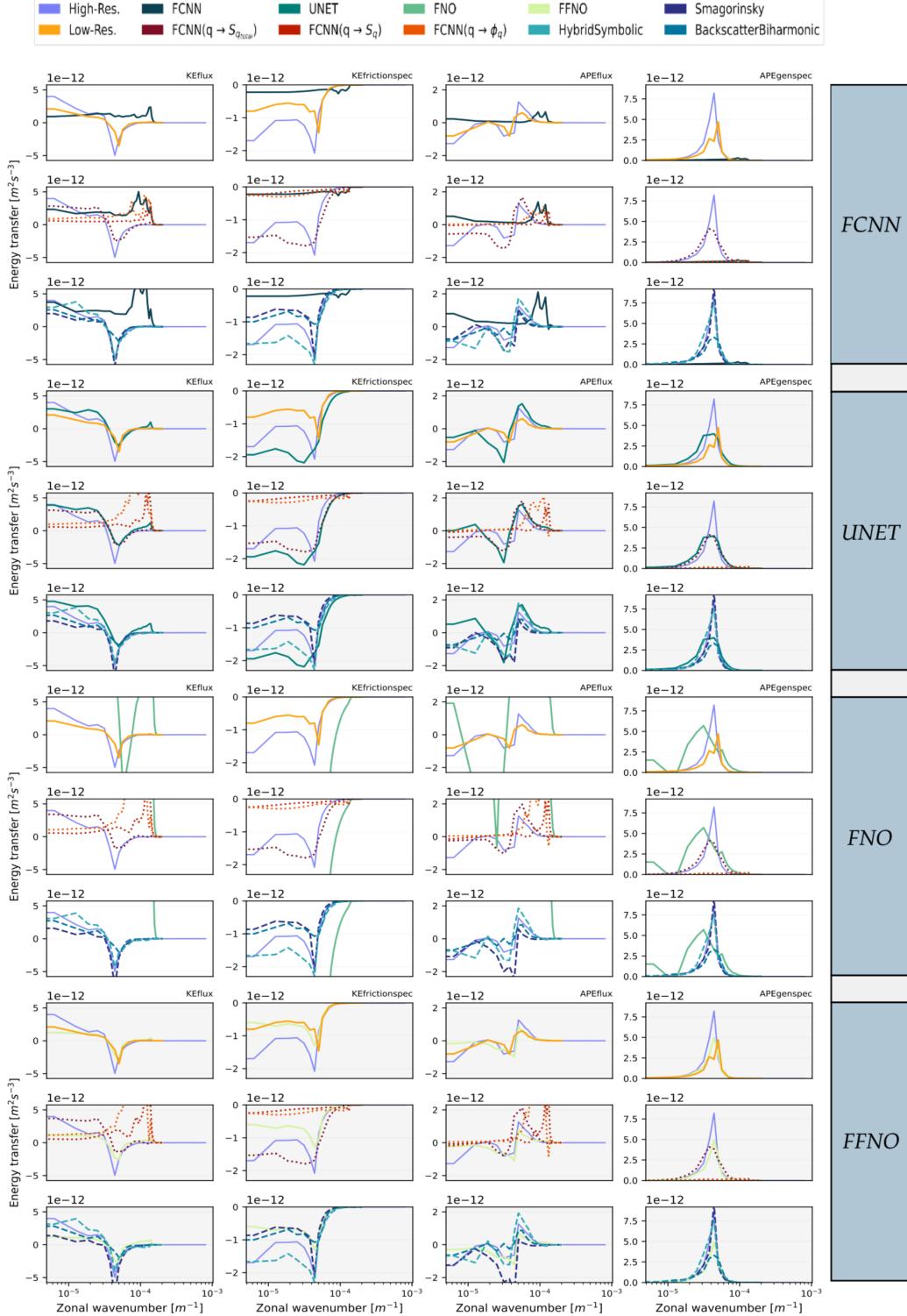


Figure 22: |Online - Phase 1 - Energy budget| This table displays energy spectra for KEflux, KEfrictionspec, APEflux, and APEgenspec using parameterizations grouped in Tab.1, these were trained on **eddies unique (5000)** and tested on **eddies online**. Each parameterization spectrum is compared against high-resolution and various low-resolution simulations, including neural networks from Ross et al., 2023 and analytical parameterizations from Smagorinsky, 1963; Jansen and Held, 2014; Zanna and Bolton, 2020.

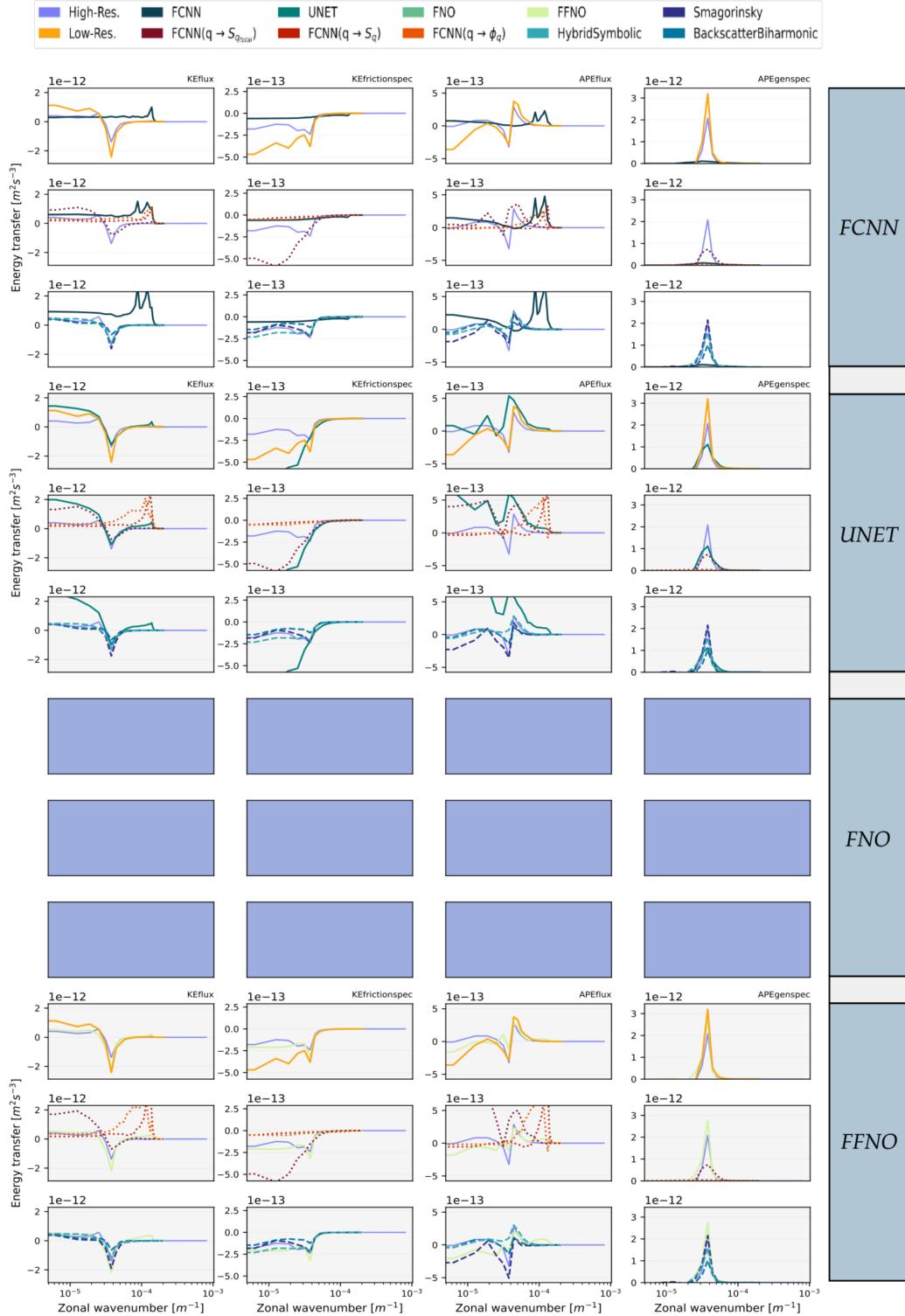


Figure 23: |Online - Phase 1 - Energy budget| This table displays energy spectra for KEflux, KEfrictionspec, APEflux, and APegaspec using parameterizations grouped in Tab.1, these were trained on **eddies unique (5000)** and tested on **jets online**. Each parameterization spectrum is compared against high-resolution and various low-resolution simulations, including neural networks from Ross et al., 2023 and analytical parameterizations from Smagorinsky, 1963; Jansen and Held, 2014; Zanna and Bolton, 2020.

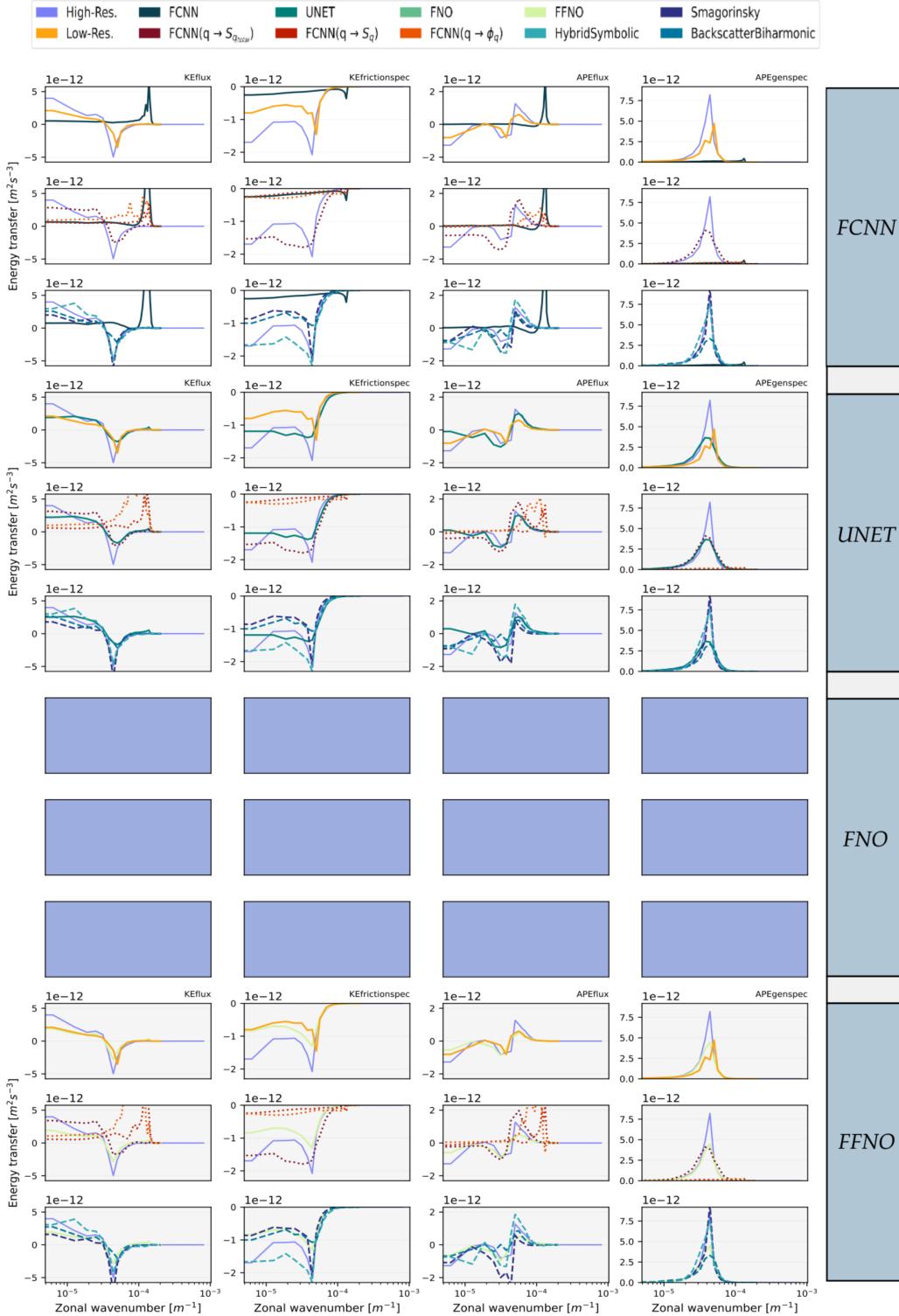


Figure 24: |Online - Phase 1 - Energy budget| This table displays energy spectra for KEflux, KEfrictionspec, APEflux, and APEgenspec using parameterizations grouped in Tab.1, these were trained on jets unique (5000) and tested on eddies online. Each parameterization spectrum is compared against high-resolution and various low-resolution simulations, including neural networks from Ross et al., 2023 and analytical parameterizations from Smagorinsky, 1963; Jansen and Held, 2014; Zanna and Bolton, 2020.

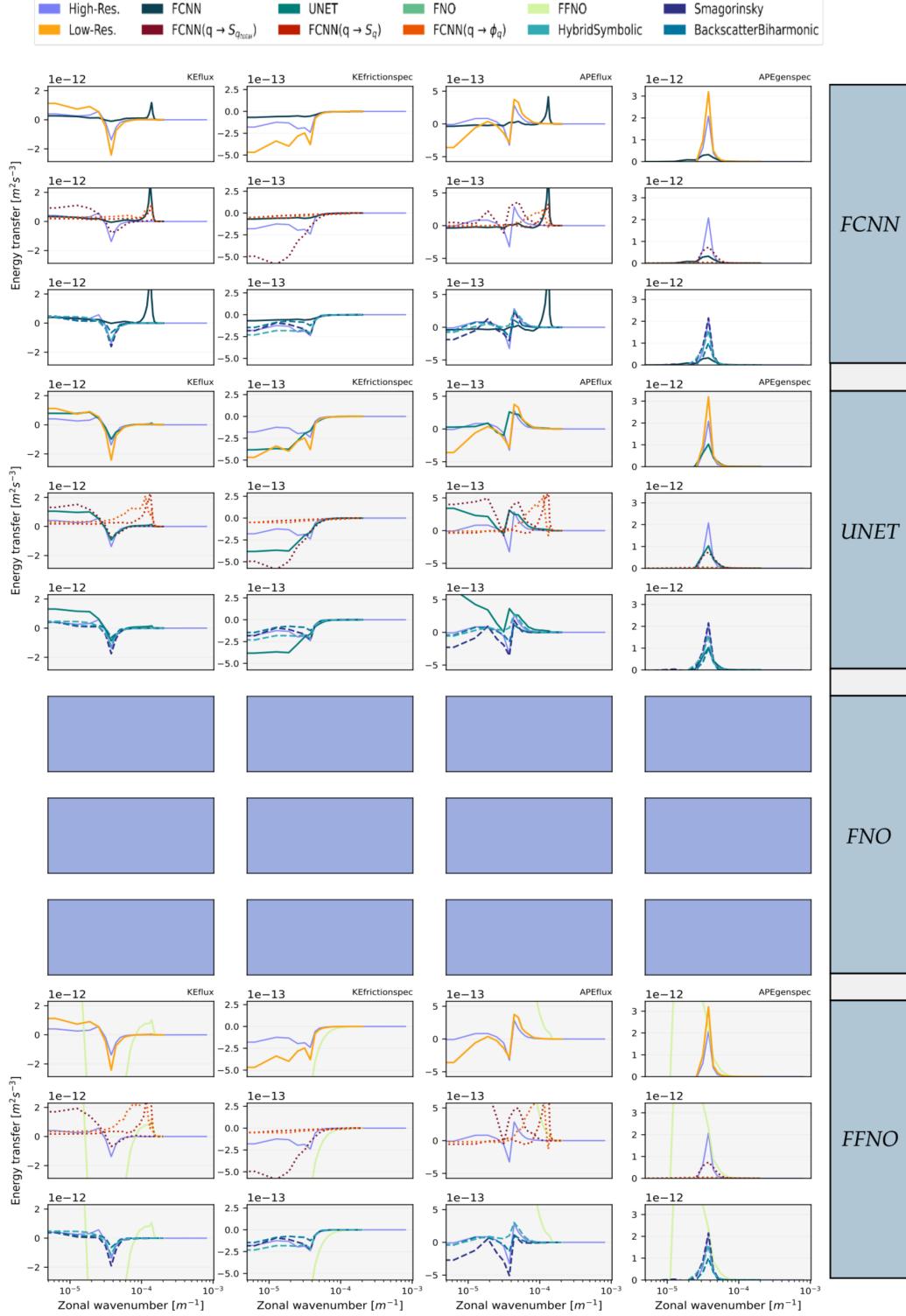


Figure 25: |Online - Phase 1 - Energy budget| This table displays energy spectra for KEflux, KEfrictionspec, APEflux, and APEnspec using parameterizations grouped in Tab.1, these were trained on **jets unique (5000)** and tested on **jets online**. Each parameterization spectrum is compared against high-resolution and various low-resolution simulations, including neural networks from Ross et al., 2023 and analytical parameterizations from Smagorinsky, 1963; Jansen and Held, 2014; Zanna and Bolton, 2020.



Figure 26: |Online - Phase 1 - Similarities| This table provides a summary of the Earth mover’s distance, reformulated as a similarity metric for various flow quantities represented in either spectral or spatiotemporal domains. A value approaching 1 indicates strong agreement between the distribution obtained from high-resolution simulations and the current observations. Negative values are considered unfavorable, and values lower than -0.5 are disregarded. The tested parameterizations are grouped in Tab.1, they are trained on **eddies unique (5000)** and tested on **eddies online**. For comparison, the results of neural networks (Ross et al., 2023) and analytical parameterizations are also presented (Smagorinsky, 1963; Jansen and Held, 2014; Zanna and Bolton, 2020).

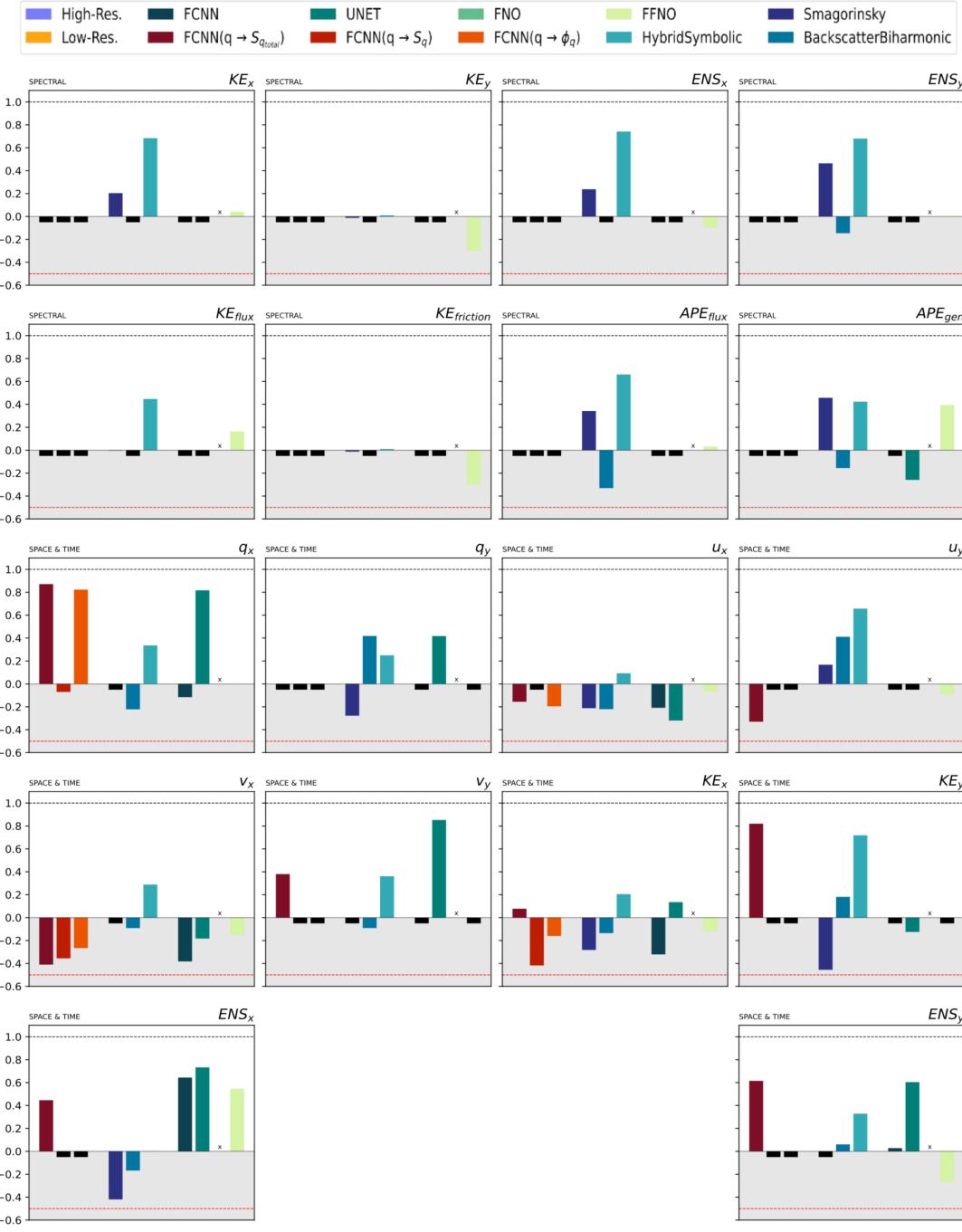


Figure 27: |Online - Phase 1 - Similarities| This table provides a summary of the Earth mover’s distance, reformulated as a similarity metric for various flow quantities represented in either spectral or spatiotemporal domains. A value approaching 1 indicates strong agreement between the distribution obtained from high-resolution simulations and the current observations. Negative values are considered unfavorable, and values lower than -0.5 are disregarded. The tested parameterizations are grouped in Tab.1, they are trained on **eddies unique (5000)** and tested on **jets online**. For comparison, the results of neural networks (Ross et al., 2023) and analytical parameterizations are also presented (Smagorinsky, 1963; Jansen and Held, 2014; Zanna and Bolton, 2020).

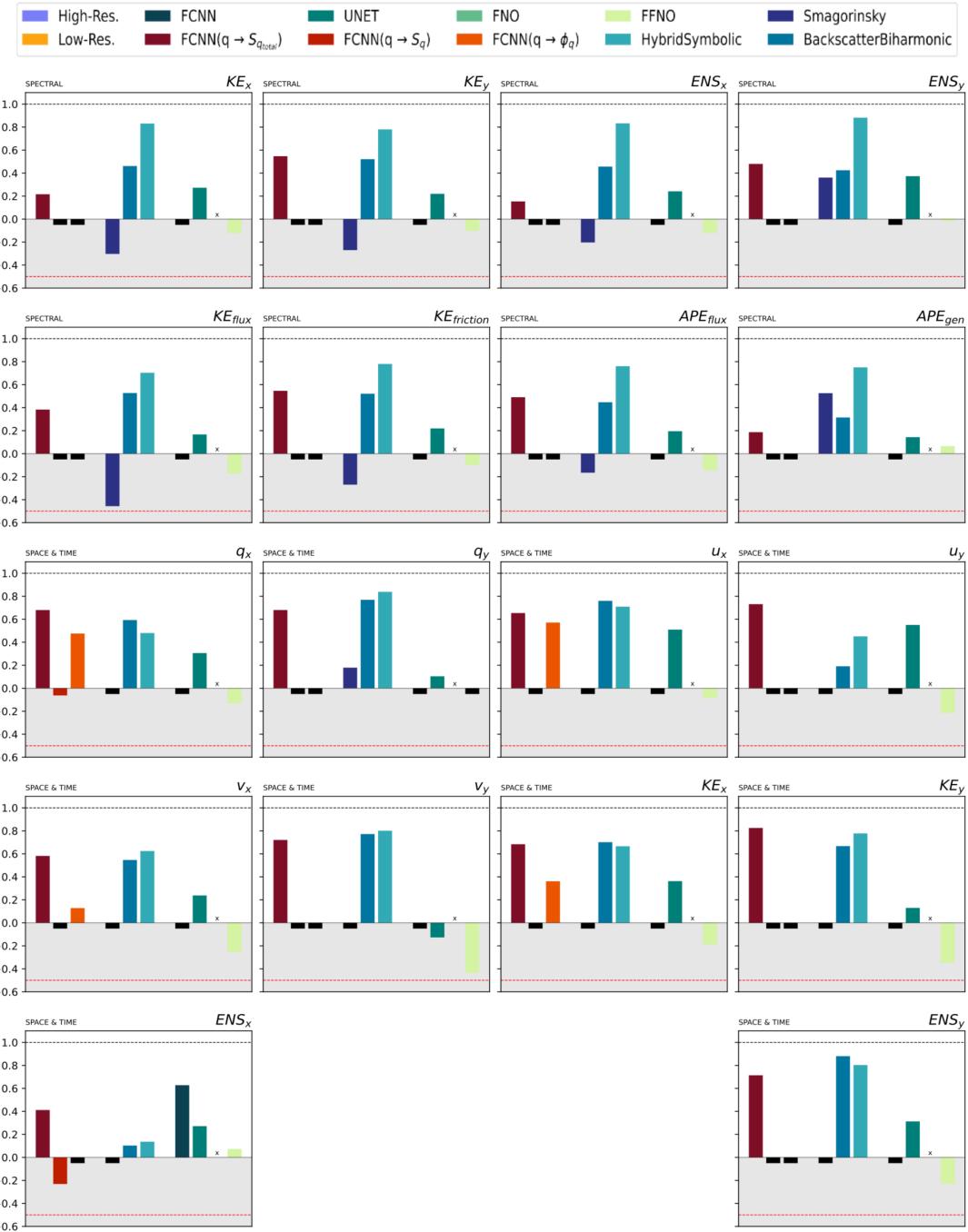


Figure 28: |Online - Phase 1 - Similarities| This table provides a summary of the Earth mover’s distance, reformulated as a similarity metric for various flow quantities represented in either spectral or spatiotemporal domains. A value approaching 1 indicates strong agreement between the distribution obtained from high-resolution simulations and the current observations. Negative values are considered unfavorable, and values lower than -0.5 are disregarded. The tested parameterizations are grouped in Tab.1, they are trained on **jets unique (5000)** and tested on **eddies online**. For comparison, the results of neural networks (Ross et al., 2023) and analytical parameterizations are also presented (Smagorinsky, 1963; Jansen and Held, 2014; Zanna and Bolton, 2020).

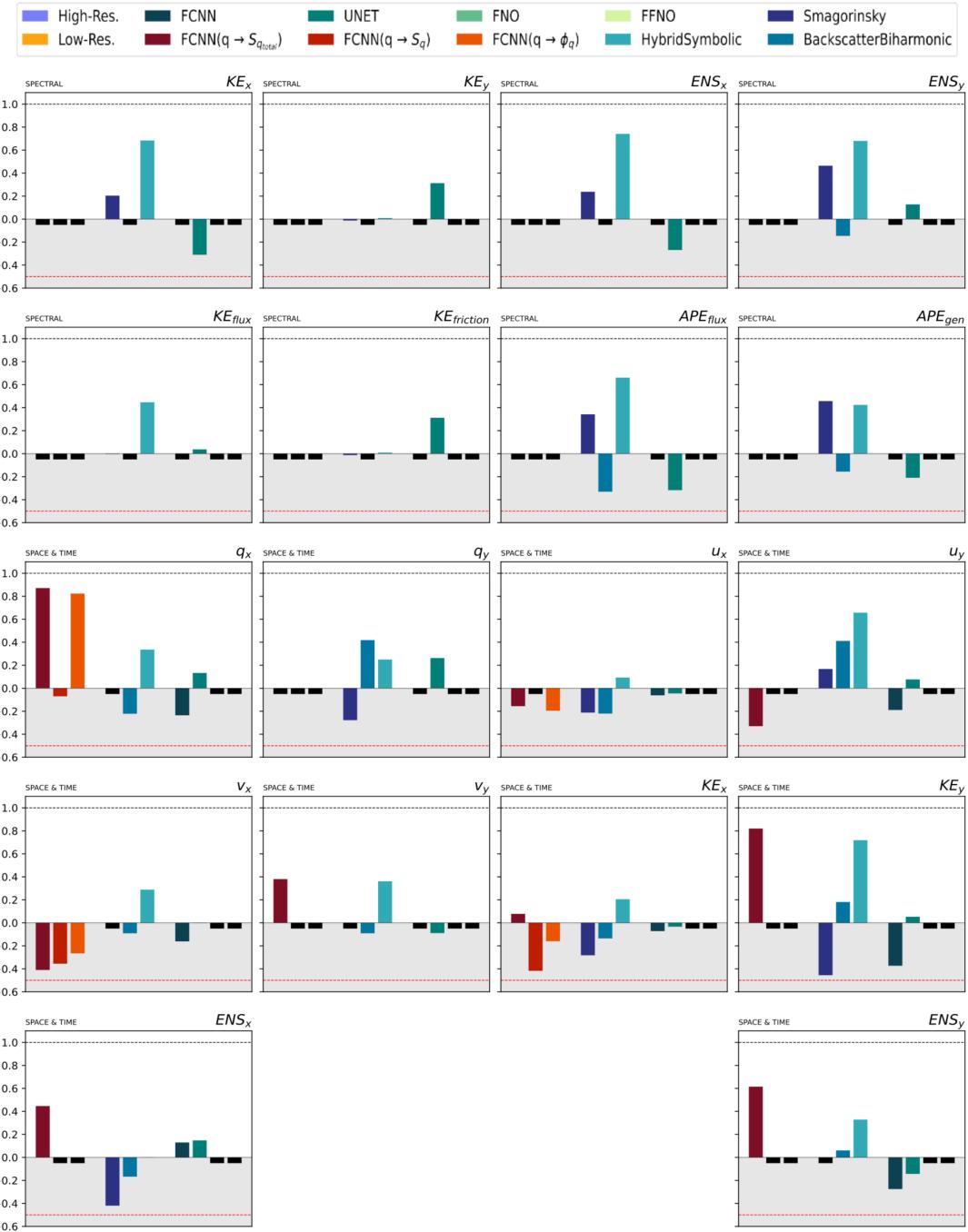


Figure 29: |Online - Phase 1 - Similarities| This table provides a summary of the Earth mover's distance, reformulated as a similarity metric for various flow quantities represented in either spectral or spatiotemporal domains. A value approaching 1 indicates strong agreement between the distribution obtained from high-resolution simulations and the current observations. Negative values are considered unfavorable, and values lower than -0.5 are disregarded. The tested parameterizations are grouped in Tab.1, they are trained on **jets unique (5000)** and tested on **jets online**. For comparison, the results of neural networks (Ross et al., 2023) and analytical parameterizations are also presented (Smagorinsky, 1963; Jansen and Held, 2014; Zanna and Bolton, 2020).

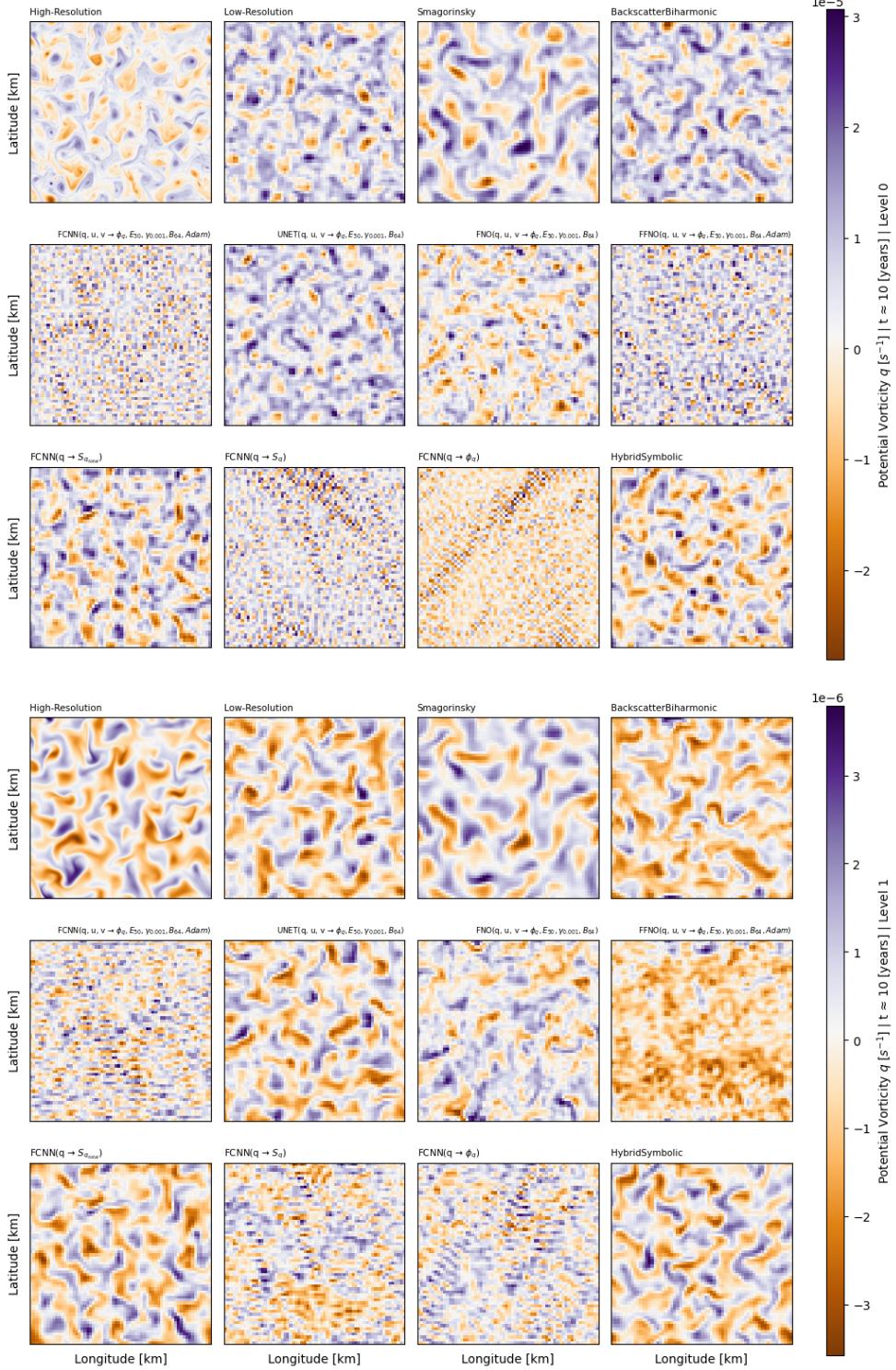


Figure 30: |Online - Phase 1 - Potential vorticity| Visualization of potential vorticity  $q$  for both upper (first three rows) and lower (last three rows) layers across different simulation types, indicated at the top of each image. Each image represents the  $q$  value spanning the entire computational domain after 10 years of simulations. The objective is to emphasize and visualize simulations that lose their physical relevance, becoming mere pixel grids, and to illustrate the divergence from the high-resolution simulation. Furthermore, the evaluated parameterizations are detailed in Tab.1, they were trained using **eddies unique (5000)** and tested on **eddies online**.

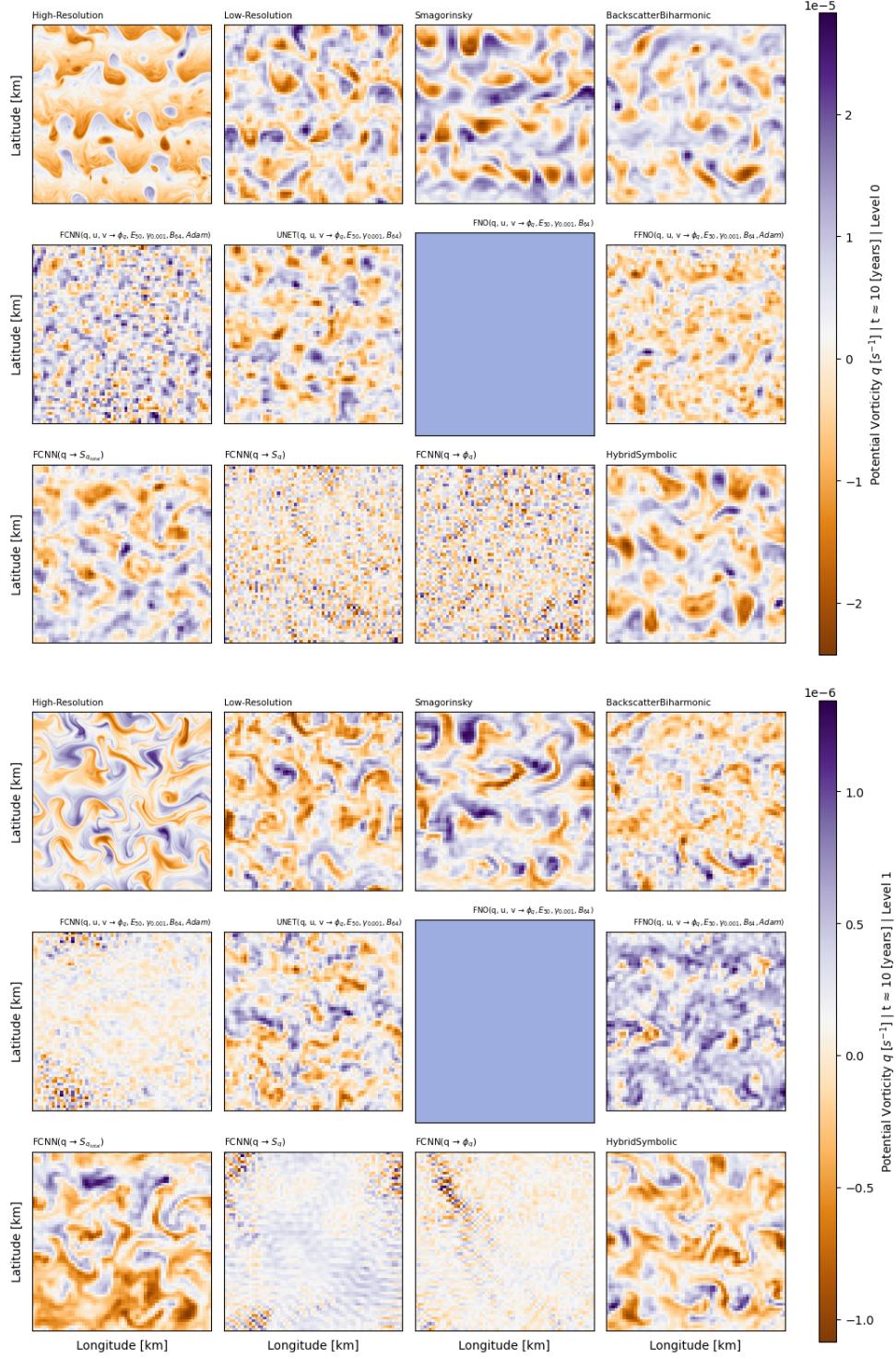


Figure 31: |Online - Phase 1 - Potential vorticity| Visualization of potential vorticity  $q$  for both upper (first three rows) and lower (last three rows) layers across different simulation types, indicated at the top of each image. Each image represents the  $q$  value spanning the entire computational domain after 10 years of simulations. The objective is to emphasize and visualize simulations that lose their physical relevance, becoming mere pixel grids, and to illustrate the divergence from the high-resolution simulation. Furthermore, the evaluated parameterizations are detailed in Tab.1, they were trained using **eddies unique (5000)** and tested on **jets online**.

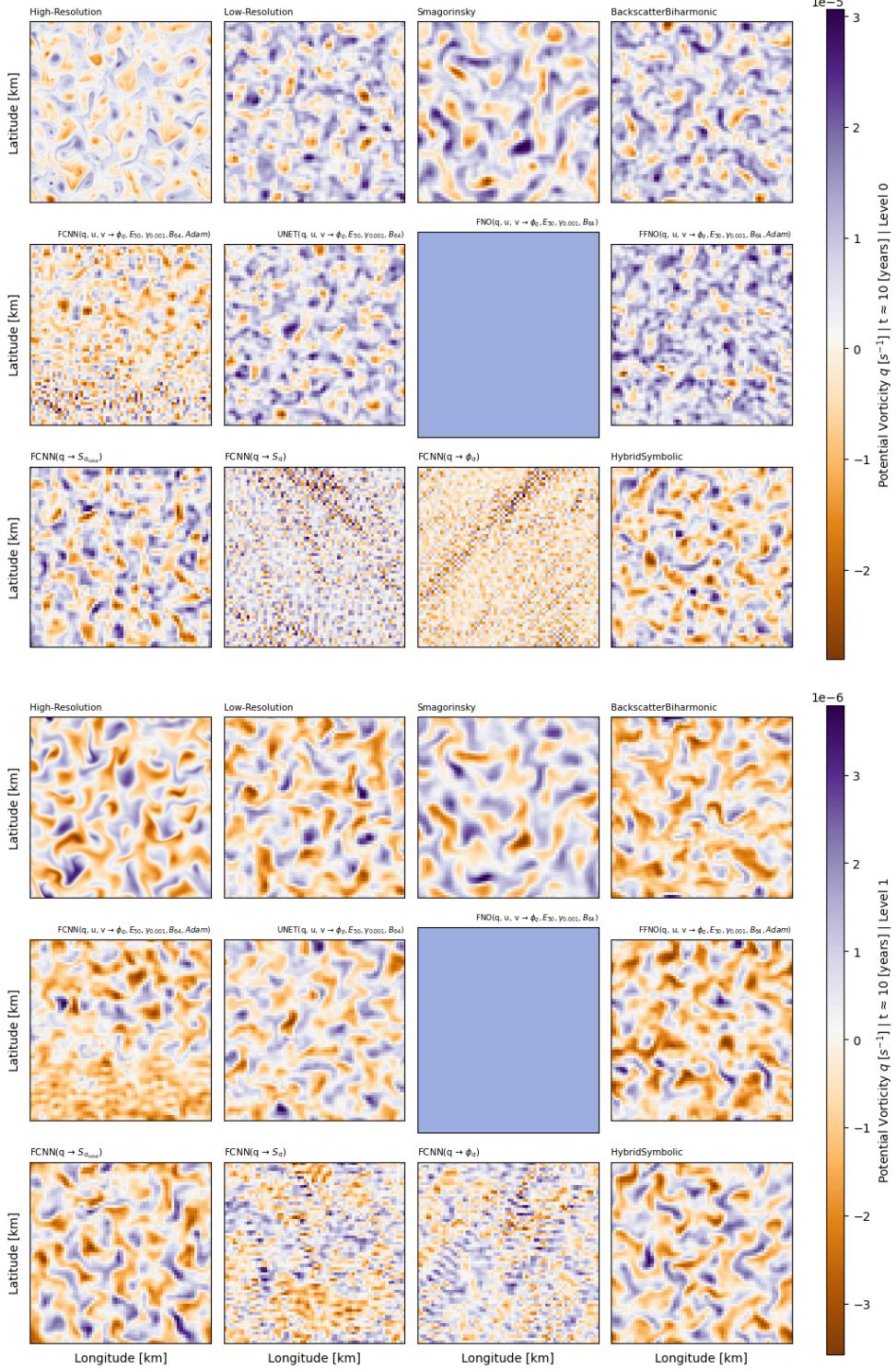


Figure 32: |Online - Phase 1 - Potential vorticity| Visualization of potential vorticity  $q$  for both upper (first three rows) and lower (last three rows) layers across different simulation types, indicated at the top of each image. Each image represents the  $q$  value spanning the entire computational domain after 10 years of simulations. The objective is to emphasize and visualize simulations that lose their physical relevance, becoming mere pixel grids, and to illustrate the divergence from the high-resolution simulation. Furthermore, the evaluated parameterizations are detailed in Tab.1, they were trained using **jets unique (5000)** and tested on **eddies online**.

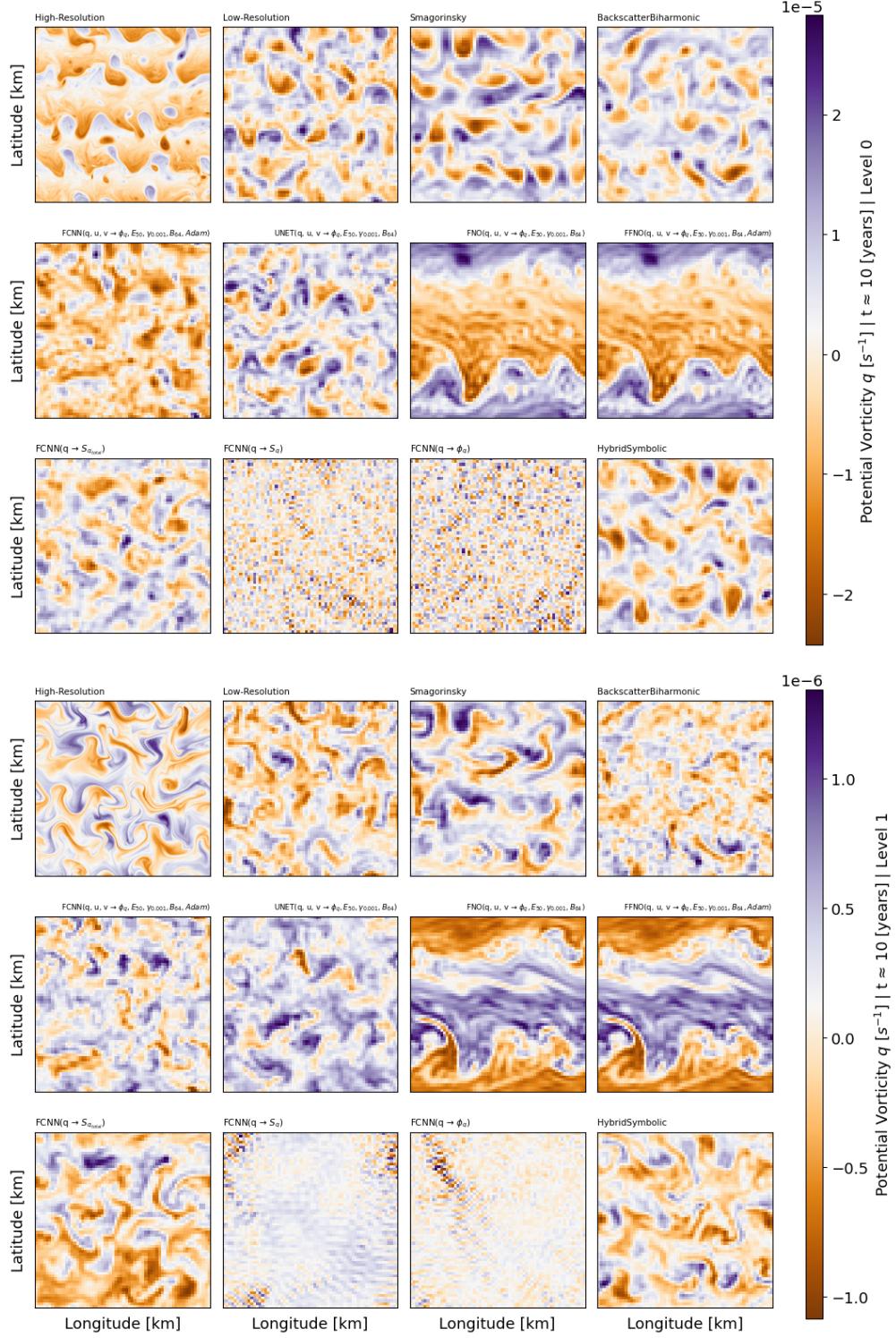


Figure 33: |Online - Phase 1 - Potential vorticity| Visualization of potential vorticity  $q$  for both upper (first three rows) and lower (last three rows) layers across different simulation types, indicated at the top of each image. Each image represents the  $q$  value spanning the entire computational domain after 10 years of simulations. The objective is to emphasize and visualize simulations that lose their physical relevance, becoming mere pixel grids, and to illustrate the divergence from the high-resolution simulation. Furthermore, the evaluated parameterizations are detailed in Tab.1, they were trained using **jets unique (5000)** and tested on **jets online**.

# PHASE II

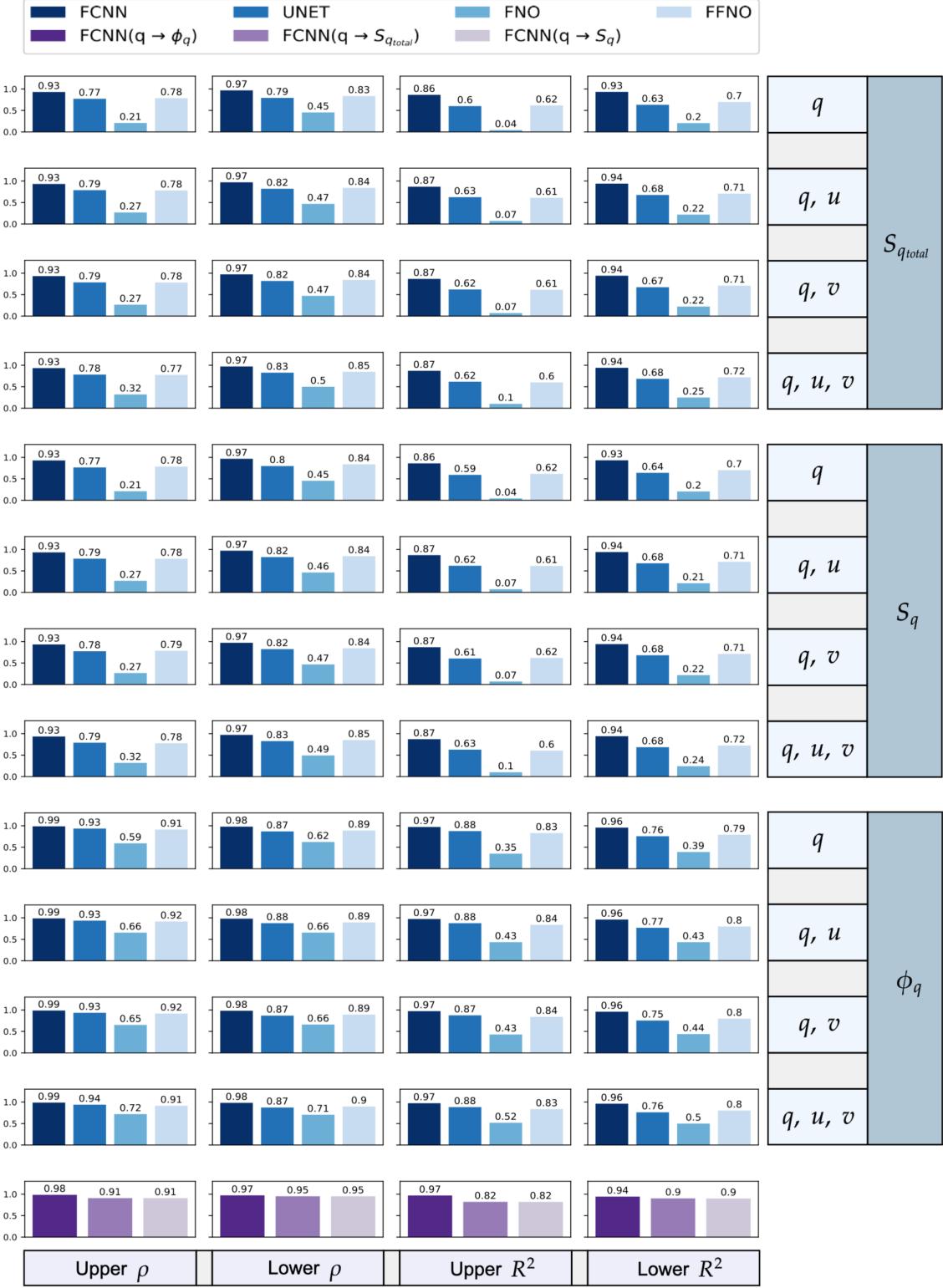


Figure 34: |Offline - Phase 2| This table summarizes offline results, including correlations (columns 1 and 2) and mean-squared errors (columns 3 and 4), for parameterizations trained on **eddies mixed (5000)** and evaluated on dataset **eddies offline**. On the right, input and output details for each parameterization are provided. The last row showcases results from the 3 FCNN parameterizations presented in Ross et al., 2023.

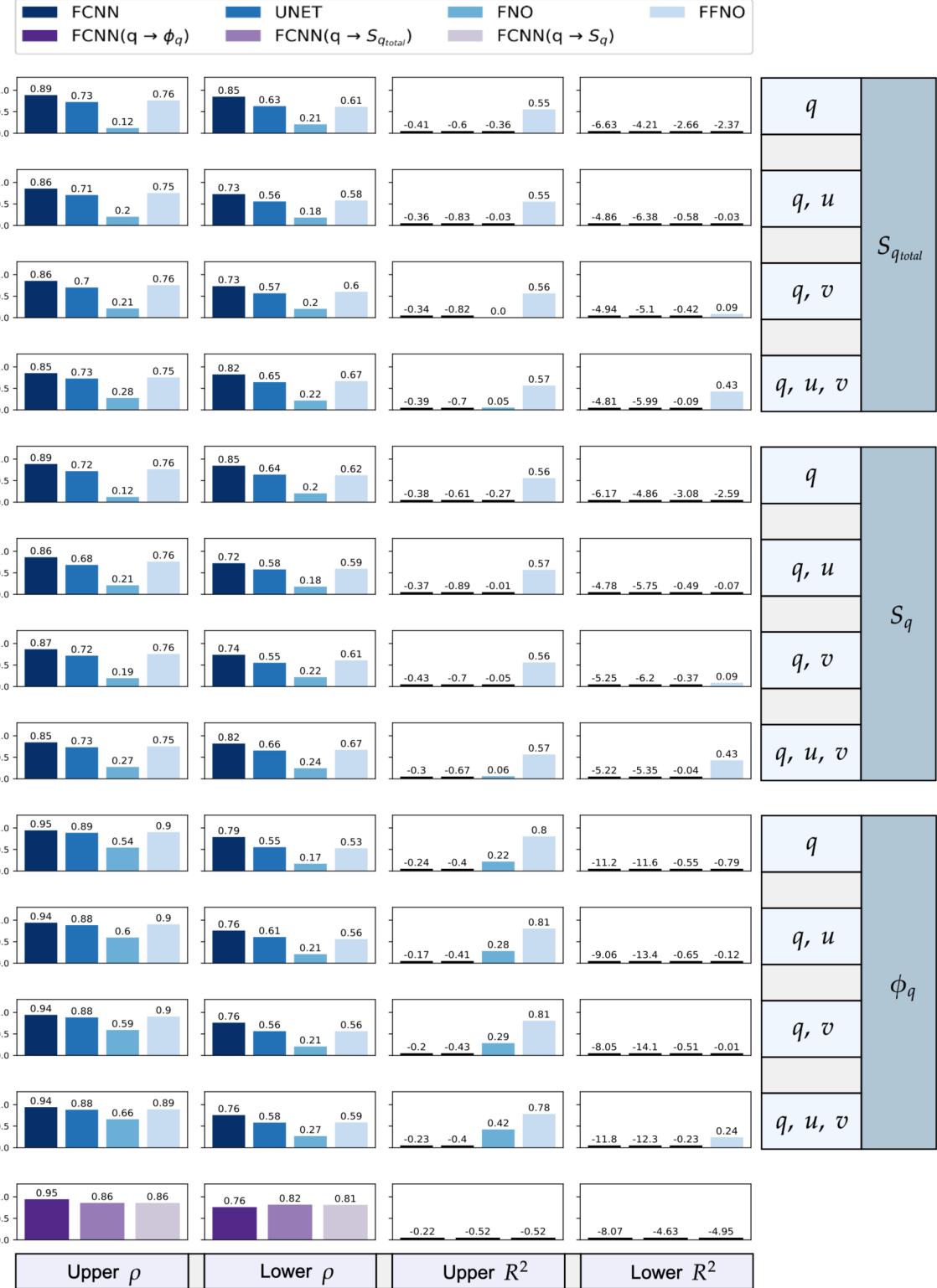


Figure 35: |Offline - Phase 2| This table summarizes offline results, including correlations (columns 1 and 2) and mean-squared errors (columns 3 and 4), for parameterizations trained on **eddies mixed (5000)** and evaluated on dataset **jets offline**. On the right, input and output details for each parameterization are provided. The last row showcases results from the 3 FCNN parameterizations presented in Ross et al., 2023.



Figure 36: |Offline - Phase 2| This table summarizes offline results, including correlations (columns 1 and 2) and mean-squared errors (columns 3 and 4), for parameterizations trained on jets mixed (5000) and evaluated on dataset eddies offline. On the right, input and output details for each parameterization are provided. The last row showcases results from the 3 FCNN parameterizations presented in Ross et al., 2023.



Figure 37: |Offline - Phase 2| This table summarizes offline results, including correlations (columns 1 and 2) and mean-squared errors (columns 3 and 4), for parameterizations trained on jets mixed (5000) and evaluated on dataset jets offline. On the right, input and output details for each parameterization are provided. The last row showcases results from the 3 FCNN parameterizations presented in Ross et al., 2023.

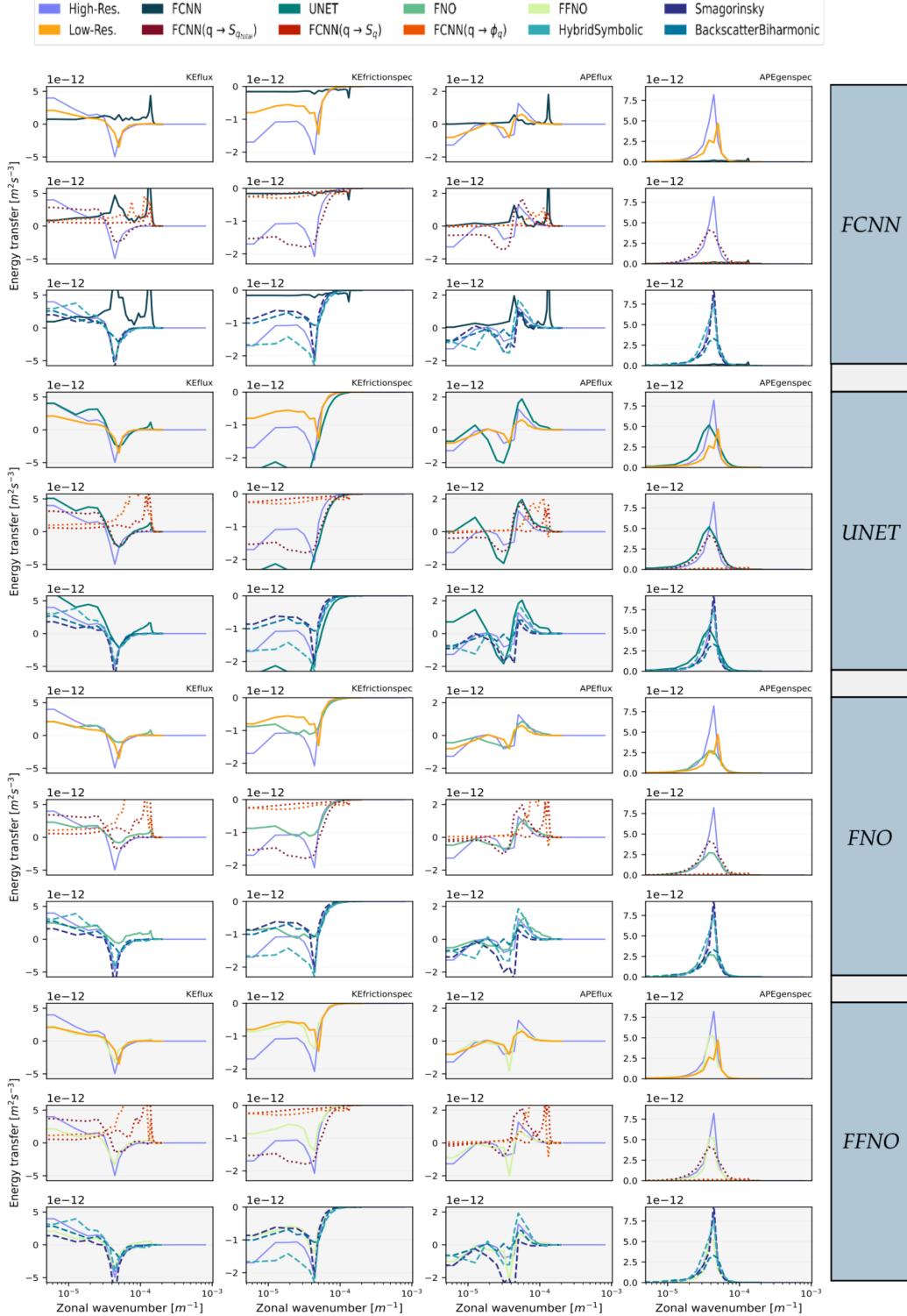


Figure 38: |Online - Phase 2 - Energy budget| This table displays energy spectra for KEflux, KEfrictionspec, APEflux, and APEgenspec using parameterizations grouped in Tab.2, these were trained on **eddies mixed (5000)** and tested on **eddies online**. Each parameterization spectrum is compared against high-resolution and various low-resolution simulations, including neural networks from Ross et al., 2023 and analytical parameterizations from Smagorinsky, 1963; Jansen and Held, 2014; Zanna and Bolton, 2020.

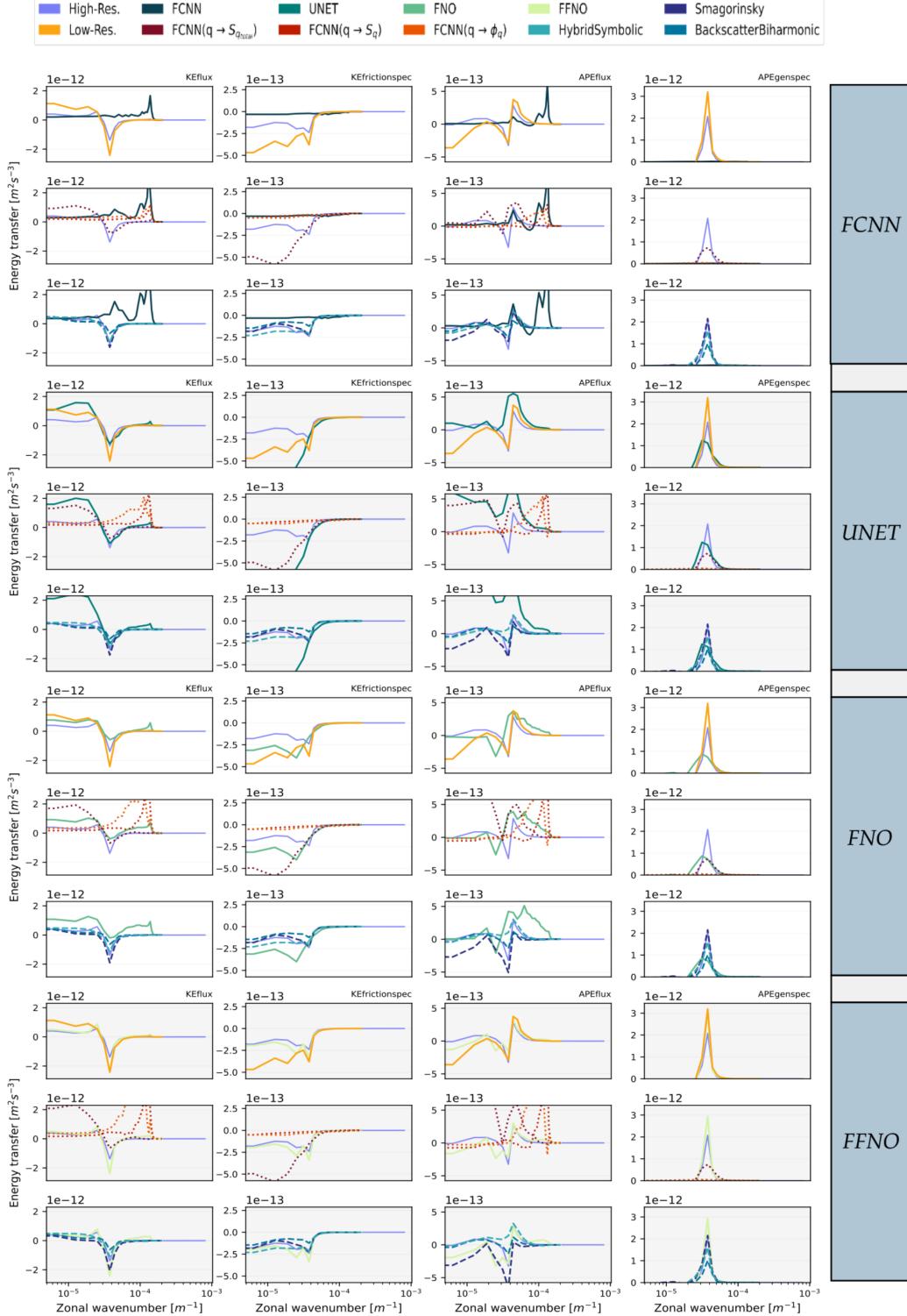


Figure 39: |Online - Phase 2 - Energy budget| This table displays energy spectra for KEflux, KEfrictionspec, APEflux, and APegaspec using parameterizations grouped in Tab.2, these were trained on **eddies mixed (5000)** and tested on **jets online**. Each parameterization spectrum is compared against high-resolution and various low-resolution simulations, including neural networks from Ross et al., 2023 and analytical parameterizations from Smagorinsky, 1963; Jansen and Held, 2014; Zanna and Bolton, 2020.

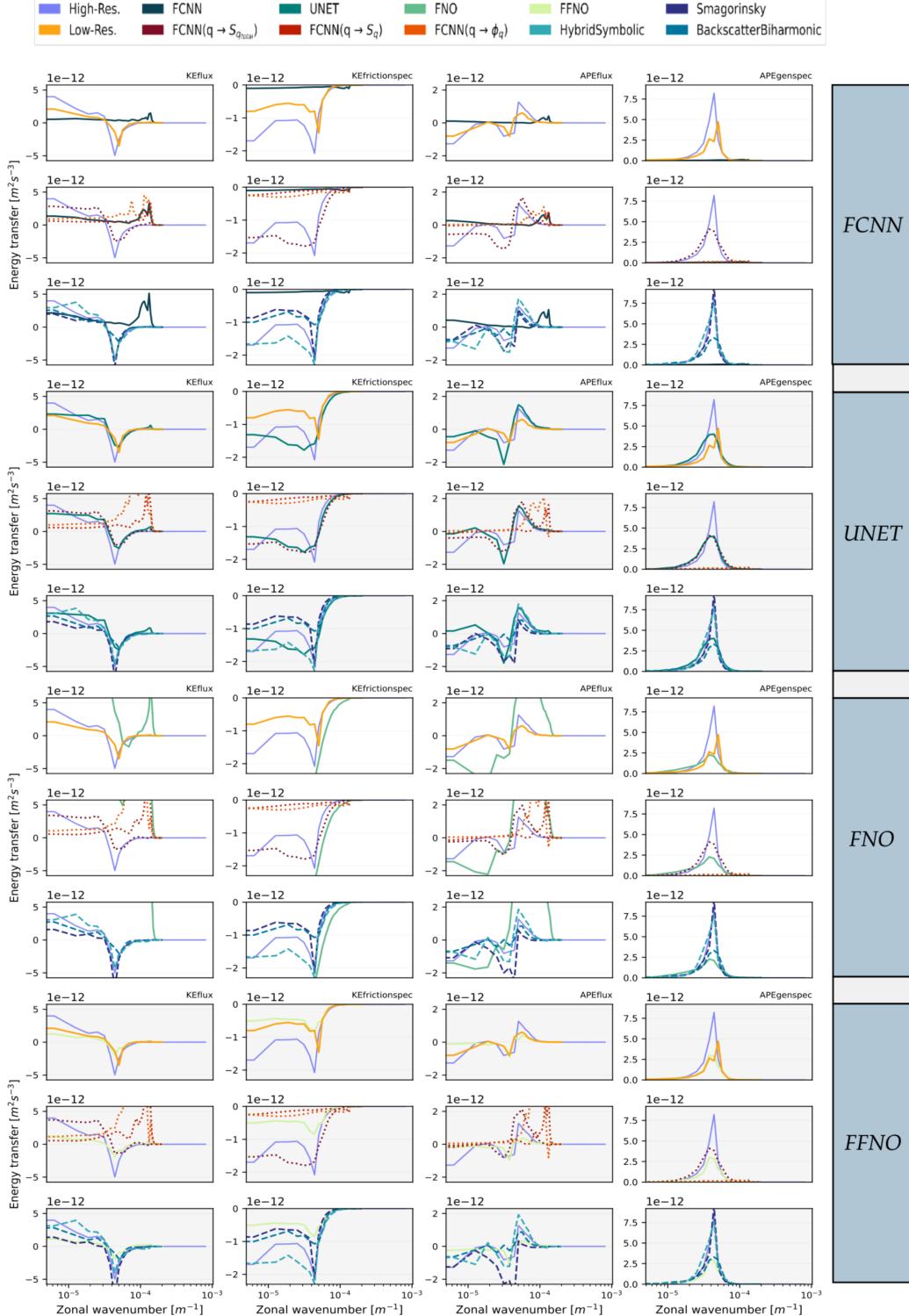


Figure 40: |Online - Phase 2 - Energy budget| This table displays energy spectra for KEflux, KEfrictionspec, APEflux, and APEnegspec using parameterizations grouped in Tab.2, these were trained on jets mixed (5000) and tested on eddies online. Each parameterization spectrum is compared against high-resolution and various low-resolution simulations, including neural networks from Ross et al., 2023 and analytical parameterizations from Smagorinsky, 1963; Jansen and Held, 2014; Zanna and Bolton, 2020.

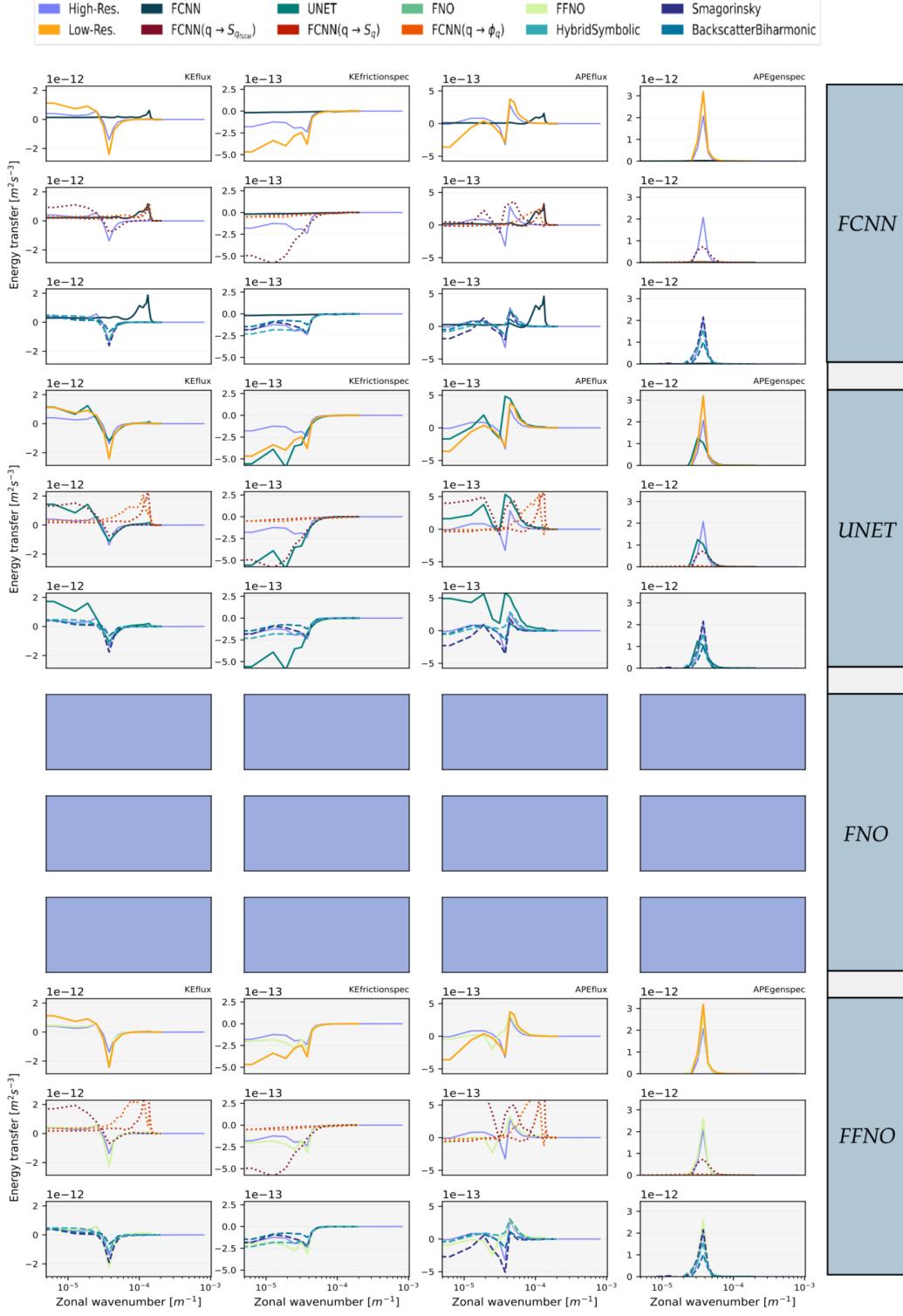


Figure 41: |Online - Phase 2 - Energy budget| This table displays energy spectra for **KEflux**, **KEfrictionspec**, **APEflux**, and **APEngenspec** using parameterizations grouped in Tab.2, these were trained on **jets mixed (5000)** and tested on **jets online**. Each parameterization spectrum is compared against high-resolution and various low-resolution simulations, including neural networks from Ross et al., 2023 and analytical parameterizations from Smagorinsky, 1963; Jansen and Held, 2014; Zanna and Bolton, 2020.

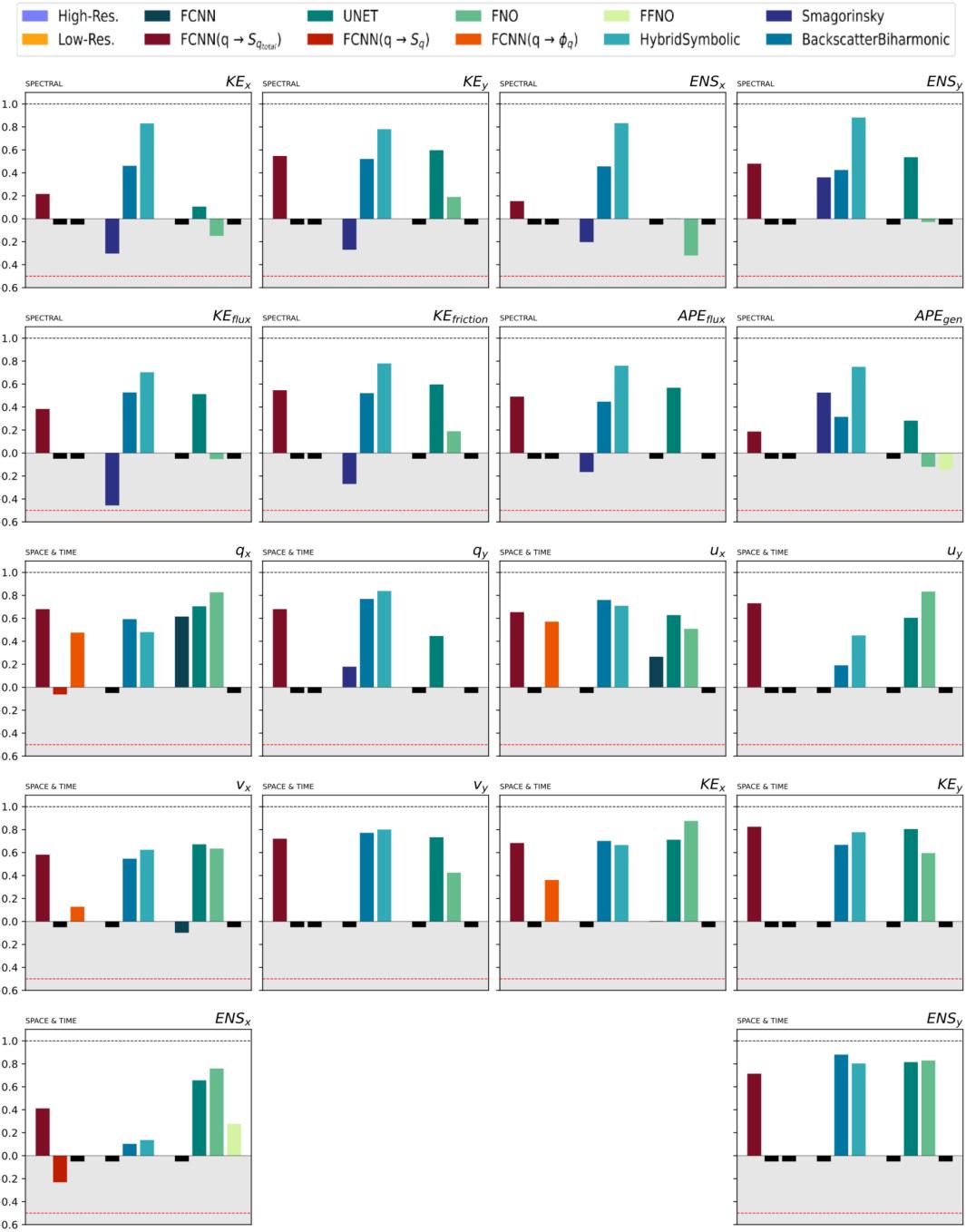


Figure 42: |Online - Phase 2 - Similarities| This table provides a summary of the Earth mover’s distance, reformulated as a similarity metric for various flow quantities represented in either spectral or spatiotemporal domains. A value approaching 1 indicates strong agreement between the distribution obtained from high-resolution simulations and the current observations. Negative values are considered unfavorable, and values lower than -0.5 are disregarded. The tested parameterizations are grouped in Tab.2, they are trained on **eddies mixed (5000)** and tested on **eddies online**. For comparison, the results of neural networks (Ross et al., 2023) and analytical parameterizations are also presented (Smagorinsky, 1963; Jansen and Held, 2014; Zanna and Bolton, 2020).

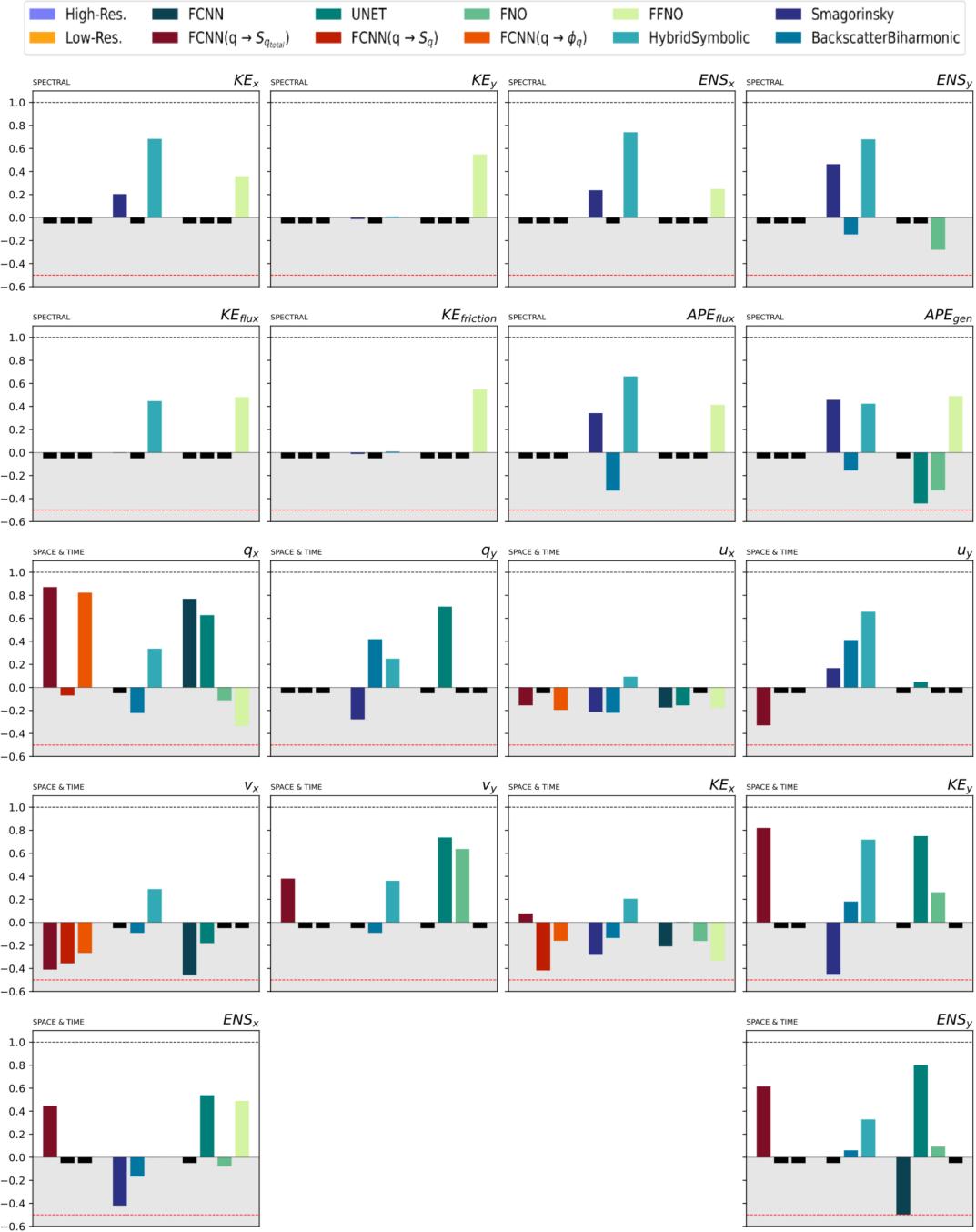


Figure 43: |Online - Phase 2 - Similarities| This table provides a summary of the Earth mover's distance, reformulated as a similarity metric for various flow quantities represented in either spectral or spatiotemporal domains. A value approaching 1 indicates strong agreement between the distribution obtained from high-resolution simulations and the current observations. Negative values are considered unfavorable, and values lower than -0.5 are disregarded. The tested parameterizations are grouped in Tab.2, they are trained on **eddies mixed (5000)** and tested on **jets online**. For comparison, the results of neural networks (Ross et al., 2023) and analytical parameterizations are also presented (Smagorinsky, 1963; Jansen and Held, 2014; Zanna and Bolton, 2020).

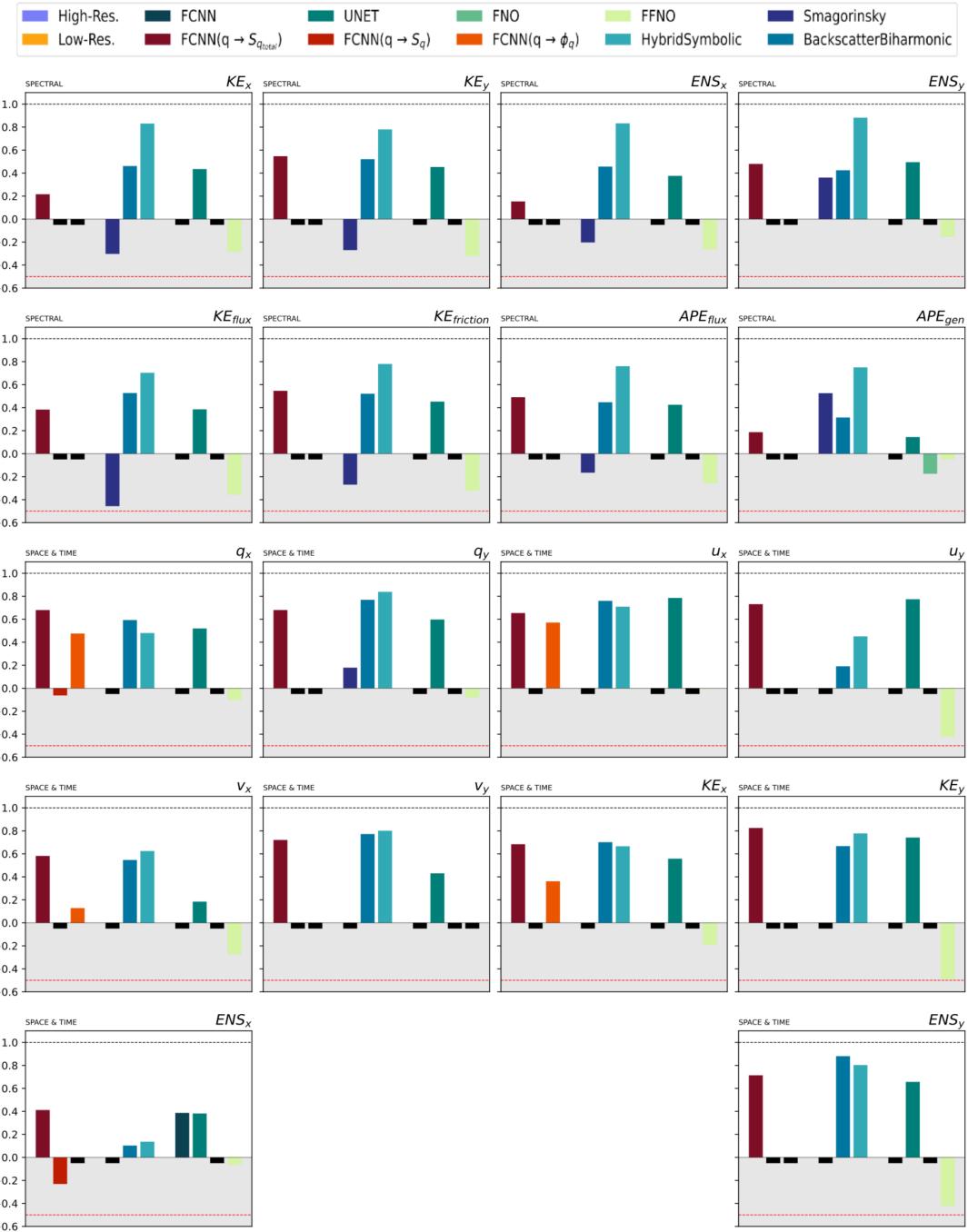


Figure 44: |Online - Phase 2 - Similarities| This table provides a summary of the Earth mover's distance, reformulated as a similarity metric for various flow quantities represented in either spectral or spatiotemporal domains. A value approaching 1 indicates strong agreement between the distribution obtained from high-resolution simulations and the current observations. Negative values are considered unfavorable, and values lower than -0.5 are disregarded. The tested parameterizations are grouped in Tab.2, they are trained on **jets mixed (5000)** and tested on **eddies online**. For comparison, the results of neural networks (Ross et al., 2023) and analytical parameterizations are also presented (Smagorinsky, 1963; Jansen and Held, 2014; Zanna and Bolton, 2020).

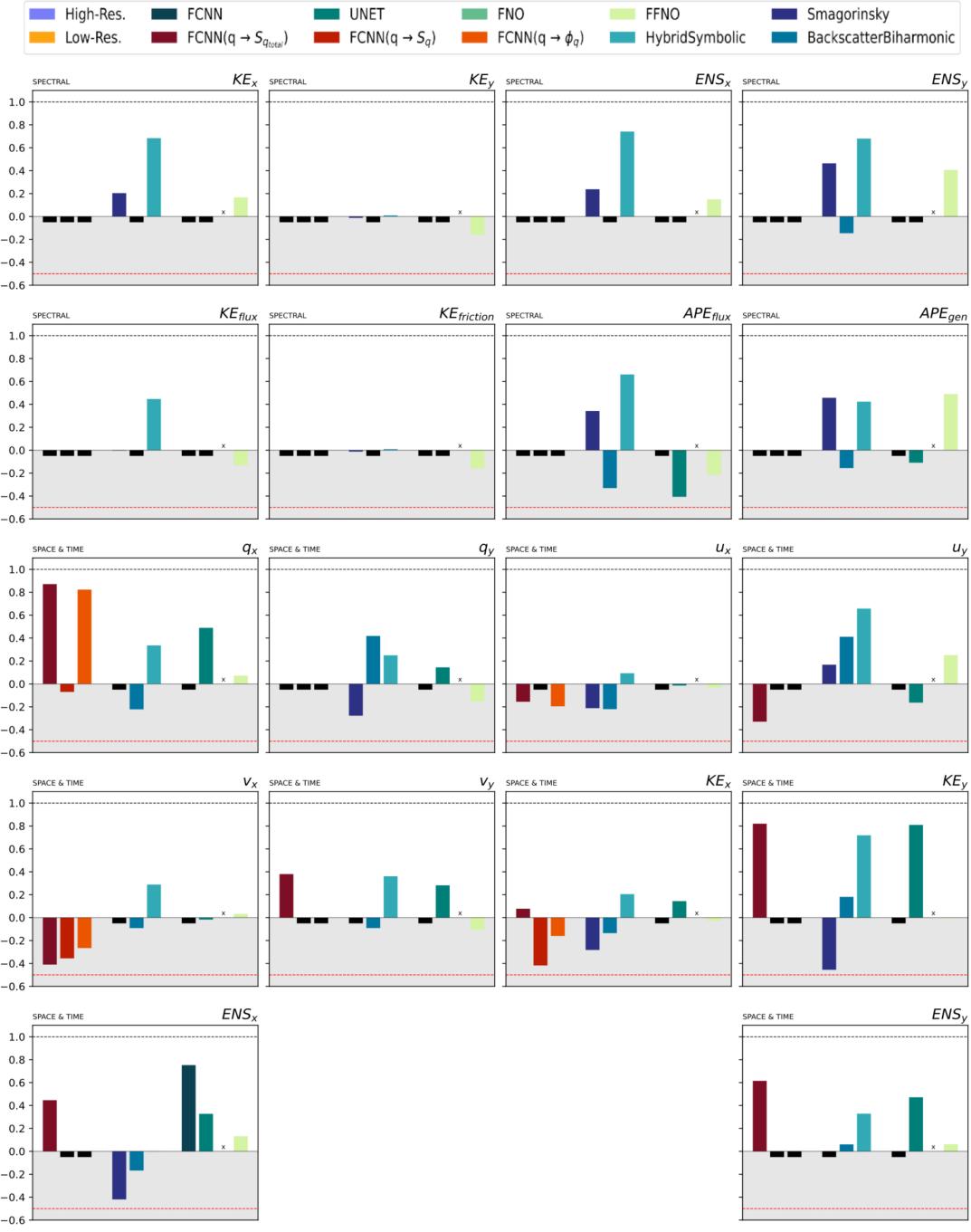


Figure 45: |Online - Phase 2 - Similarities| This table provides a summary of the Earth mover's distance, reformulated as a similarity metric for various flow quantities represented in either spectral or spatiotemporal domains. A value approaching 1 indicates strong agreement between the distribution obtained from high-resolution simulations and the current observations. Negative values are considered unfavorable, and values lower than -0.5 are disregarded. The tested parameterizations are grouped in Tab.2, they are trained on **jets mixed (5000)** and tested on **jets online**. For comparison, the results of neural networks (Ross et al., 2023) and analytical parameterizations are also presented (Smagorinsky, 1963; Jansen and Held, 2014; Zanna and Bolton, 2020).

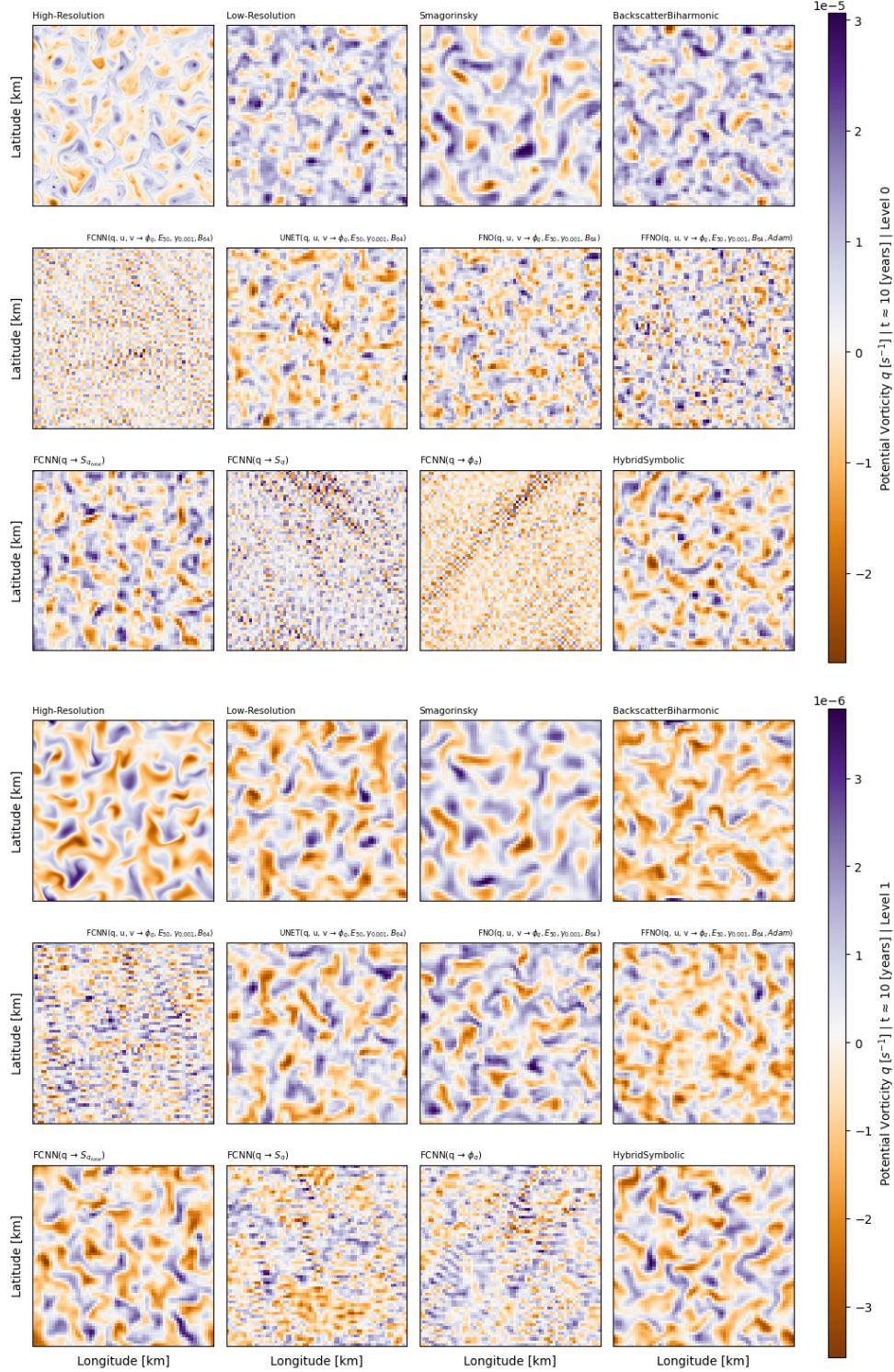
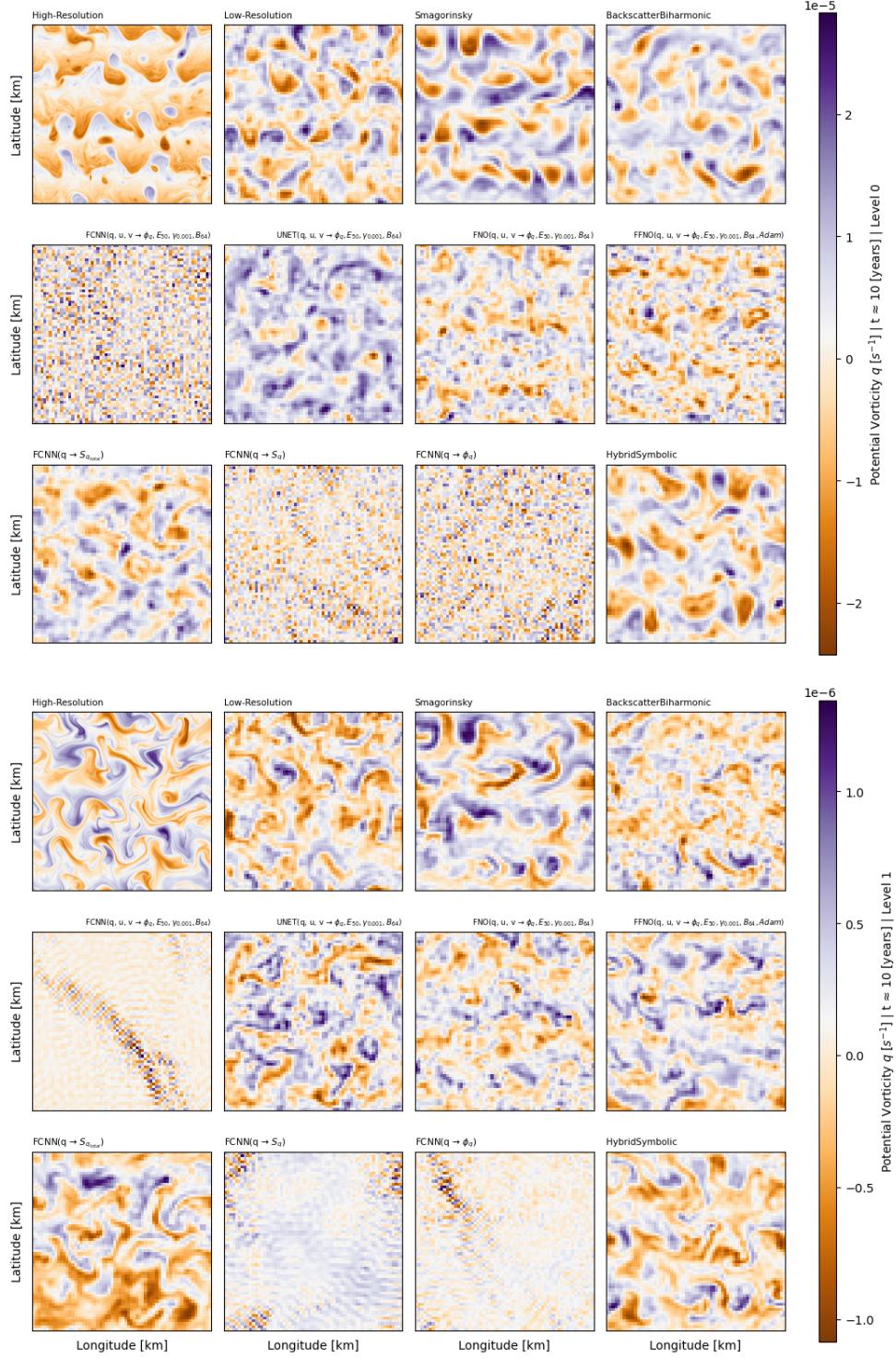


Figure 46: |Online - Phase 2 - Potential vorticity| Visualization of potential vorticity  $q$  for both upper (first three rows) and lower (last three rows) layers across different simulation types, indicated at the top of each image. Each image represents the  $q$  value spanning the entire computational domain after 10 years of simulations. The objective is to emphasize and visualize simulations that lose their physical relevance, becoming mere pixel grids, and to illustrate the divergence from the high-resolution simulation. Furthermore, the evaluated parameterizations are detailed in Tab.2, they were trained using **eddies mixed (5000)** and tested on **eddies online**.



**Figure 47: |Online - Phase 2 - Potential vorticity|** Visualization of potential vorticity  $q$  for both upper (first three rows) and lower (last three rows) layers across different simulation types, indicated at the top of each image. Each image represents the  $q$  value spanning the entire computational domain after 10 years of simulations. The objective is to emphasize and visualize simulations that lose their physical relevance, becoming mere pixel grids, and to illustrate the divergence from the high-resolution simulation. Furthermore, the evaluated parameterizations are detailed in Tab.2, they were trained using **eddies mixed (5000)** and tested on **jets online**.

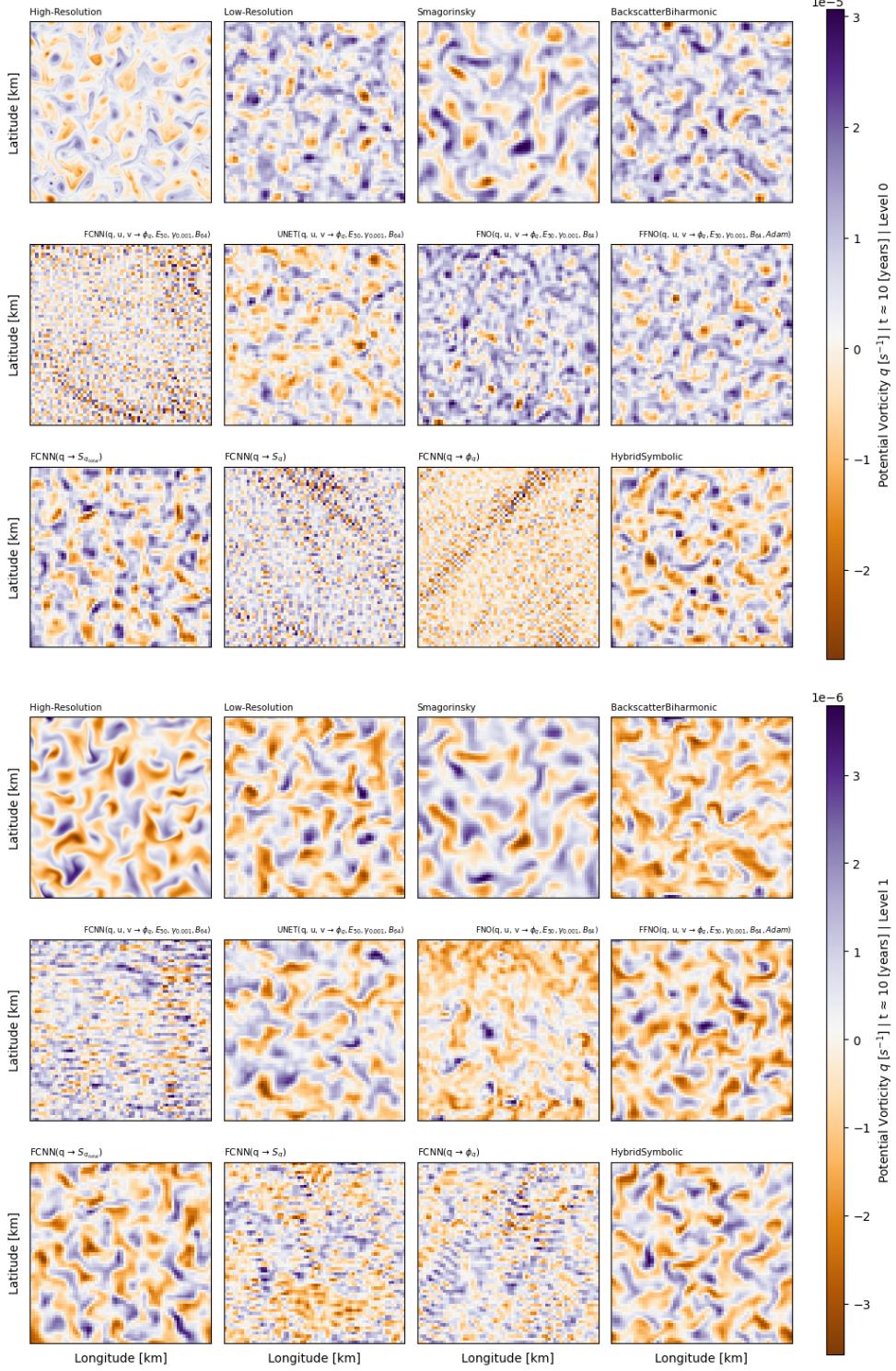
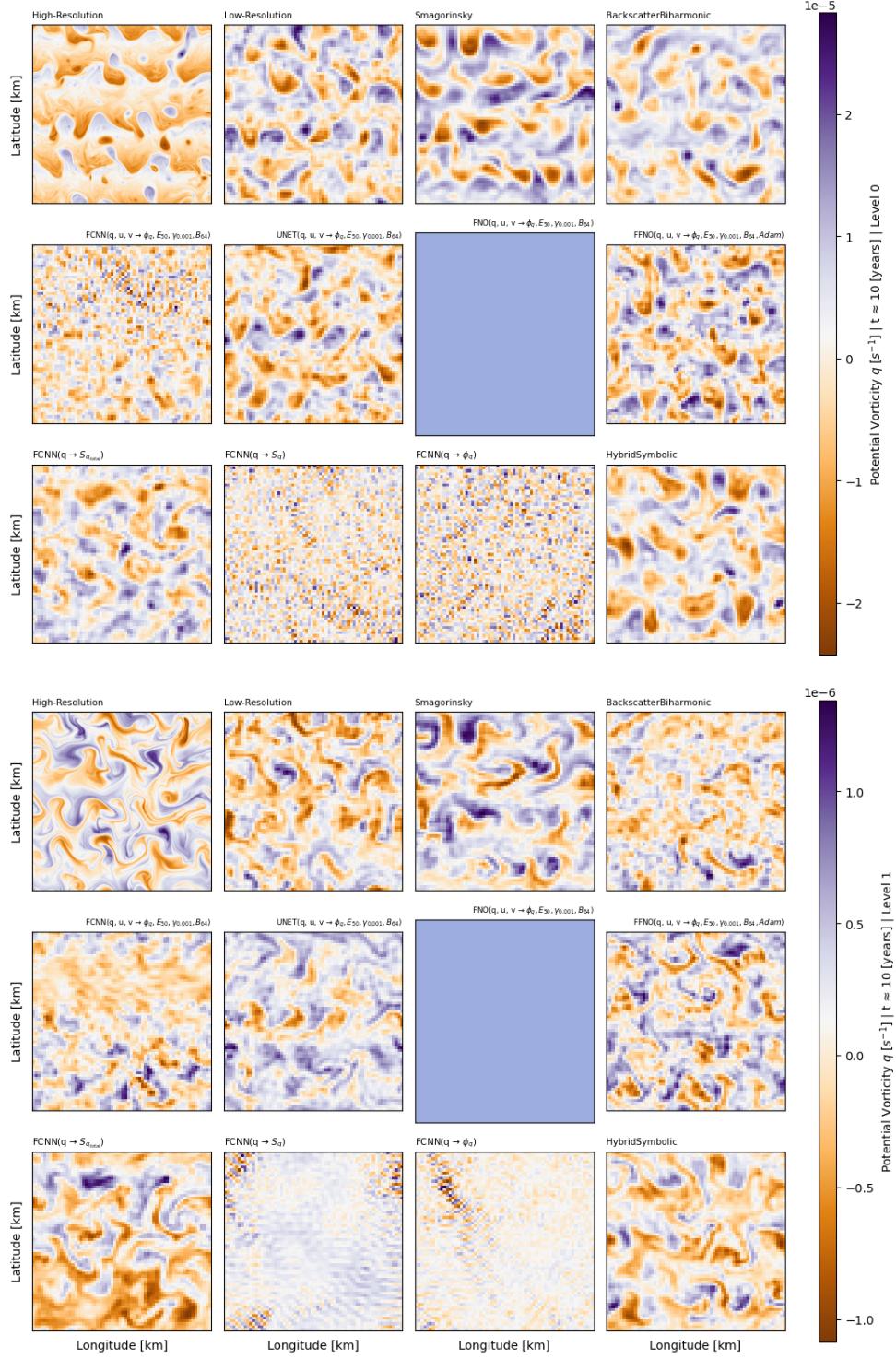


Figure 48: |Online - Phase 2 - Potential vorticity| Visualization of potential vorticity  $q$  for both upper (first three rows) and lower (last three rows) layers across different simulation types, indicated at the top of each image. Each image represents the  $q$  value spanning the entire computational domain after 10 years of simulations. The objective is to emphasize and visualize simulations that lose their physical relevance, becoming mere pixel grids, and to illustrate the divergence from the high-resolution simulation. Furthermore, the evaluated parameterizations are detailed in Tab.2, they were trained using **jets mixed (5000)** and tested on **eddies online**.



**Figure 49: |Online - Phase 2 - Potential vorticity|** Visualization of potential vorticity  $q$  for both upper (first three rows) and lower (last three rows) layers across different simulation types, indicated at the top of each image. Each image represents the  $q$  value spanning the entire computational domain after 10 years of simulations. The objective is to emphasize and visualize simulations that lose their physical relevance, becoming mere pixel grids, and to illustrate the divergence from the high-resolution simulation. Furthermore, the evaluated parameterizations are detailed in Tab.2, they were trained using **jets mixed (5000)** and tested on **jets online**.

# PHASE III



Figure 50: |Offline - Phase 3| This table summarizes offline results, including correlations (columns 1 and 2) and mean-squared errors (columns 3 and 4), for parameterizations trained on **eddies mixed**, evaluated on dataset **eddies offline** and predicting subgrid flux  $\Phi_q$  (see Eq. 23). On the right, input details for each parameterization are supplied, along with the number of samples used for training. The final row showcases outcomes from the three FCNN parameterizations introduced in Ross et al., 2023.

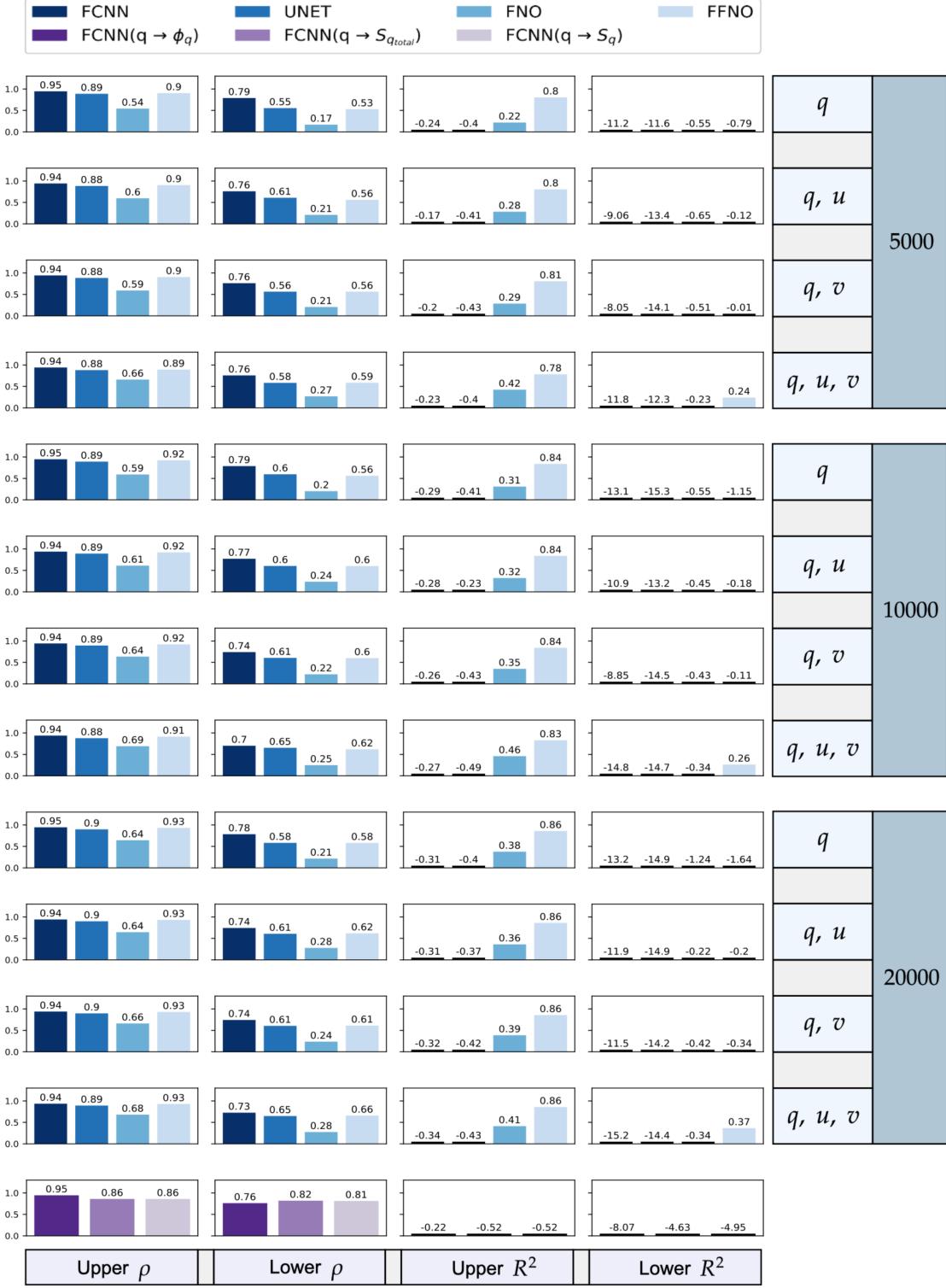


Figure 51: |Offline - Phase 3| This table summarizes offline results, including correlations (columns 1 and 2) and mean-squared errors (columns 3 and 4), for parameterizations trained on **eddies mixed**, evaluated on dataset **jets offline** and predicting subgrid flux  $\Phi_q$  (see Eq. 23). On the right, input details for each parameterization are supplied, along with the number of samples used for training. The final row showcases outcomes from the three FCNN parameterizations introduced in Ross et al., 2023.

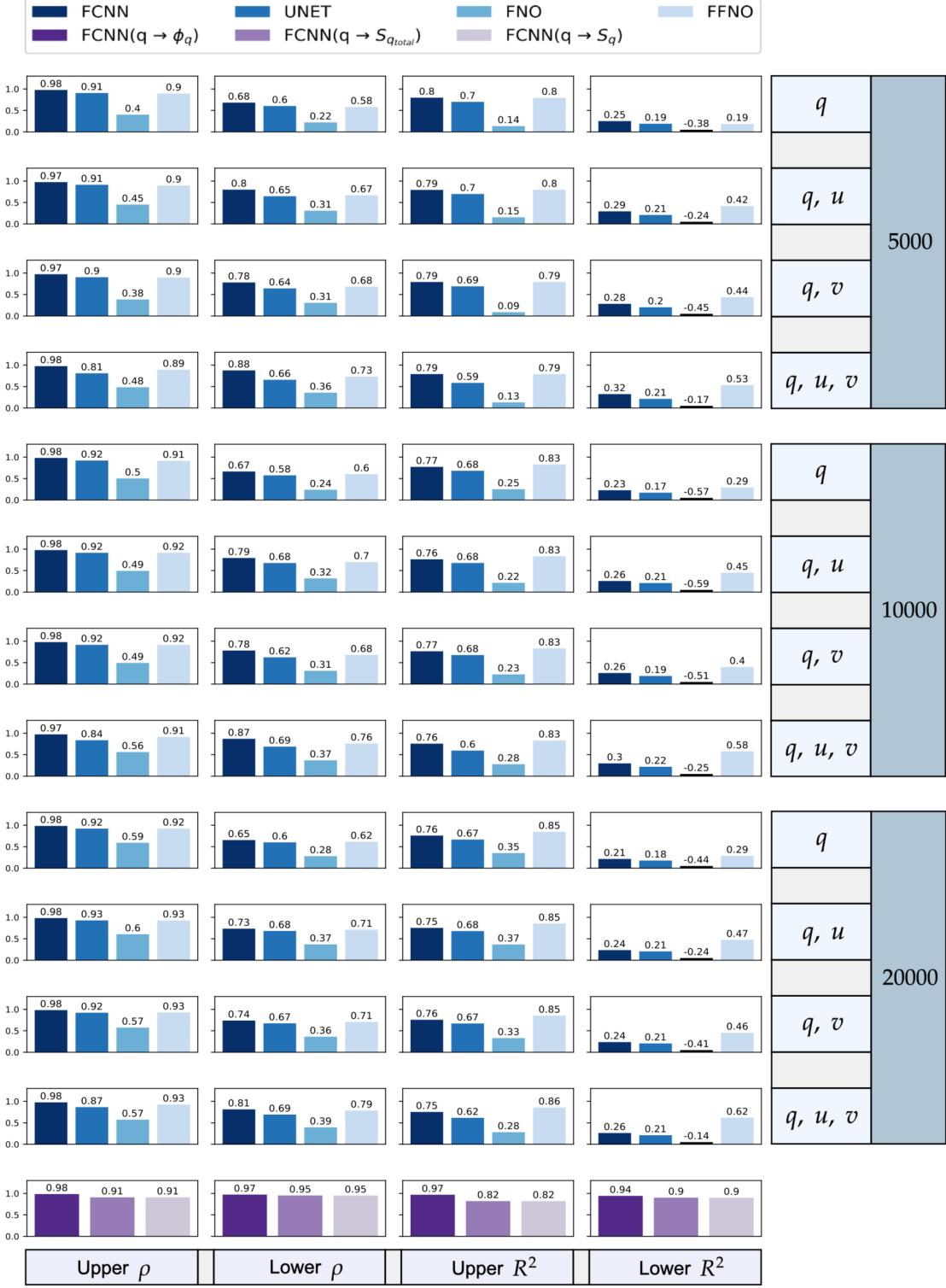


Figure 52: |Offline - Phase 3| This table summarizes offline results, including correlations ( $\rho$ ) and mean-squared errors ( $R^2$ ), for parameterizations trained on jets mixed, evaluated on dataset eddies offline and predicting subgrid flux  $\Phi_q$  (see Eq. 23). On the right, input details for each parameterization are supplied, along with the number of samples used for training. The final row showcases outcomes from the three FCNN parameterizations introduced in Ross et al., 2023.

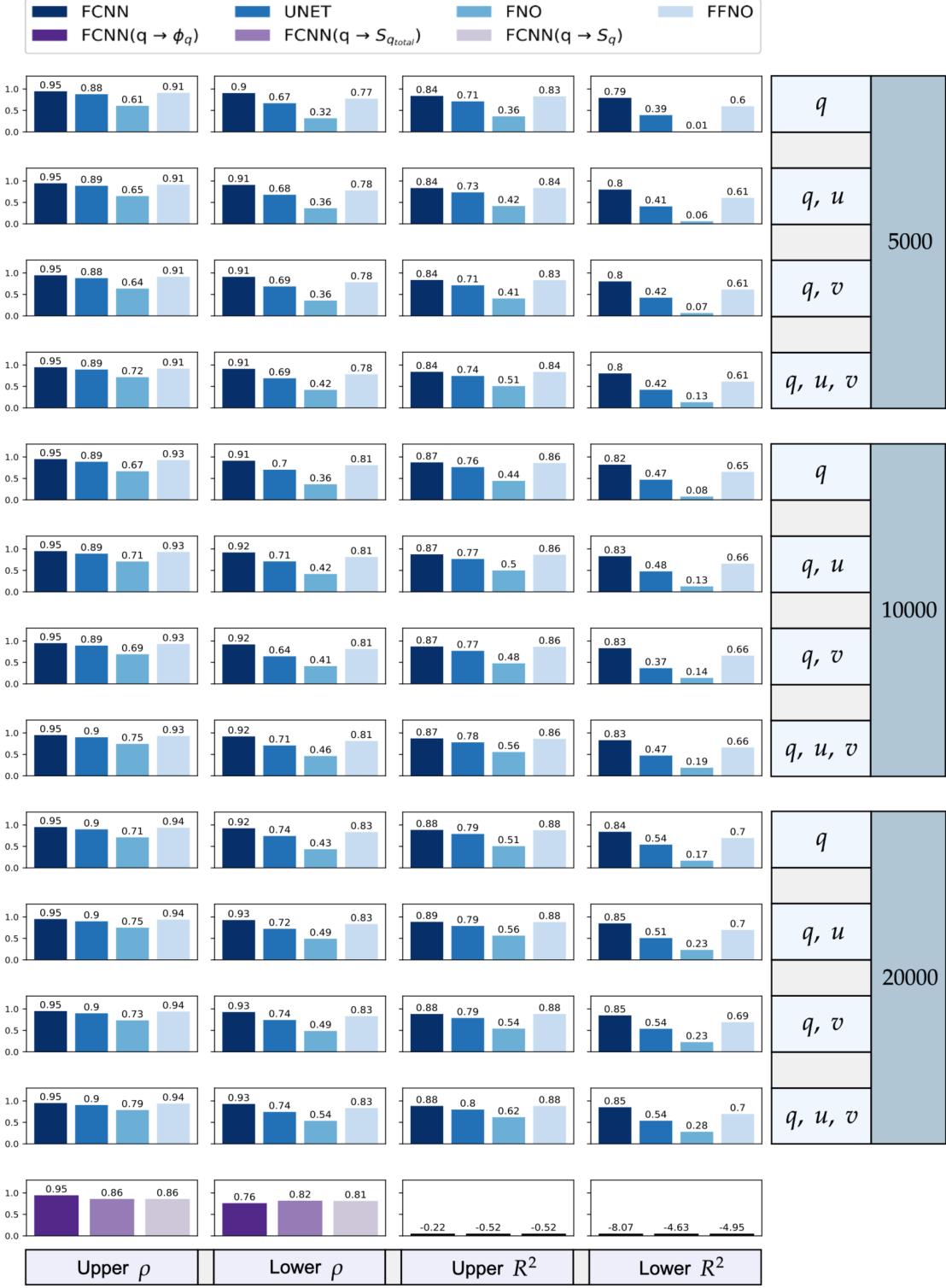


Figure 53: |Offline - Phase 3| This table summarizes offline results, including correlations (columns 1 and 2) and mean-squared errors (columns 3 and 4), for parameterizations trained on **jets mixed**, evaluated on dataset **jets offline** and predicting subgrid flux  $\Phi_q$  (see Eq. 23). On the right, input details for each parameterization are supplied, along with the number of samples used for training. The final row showcases outcomes from the three FCNN parameterizations introduced in Ross et al., 2023.

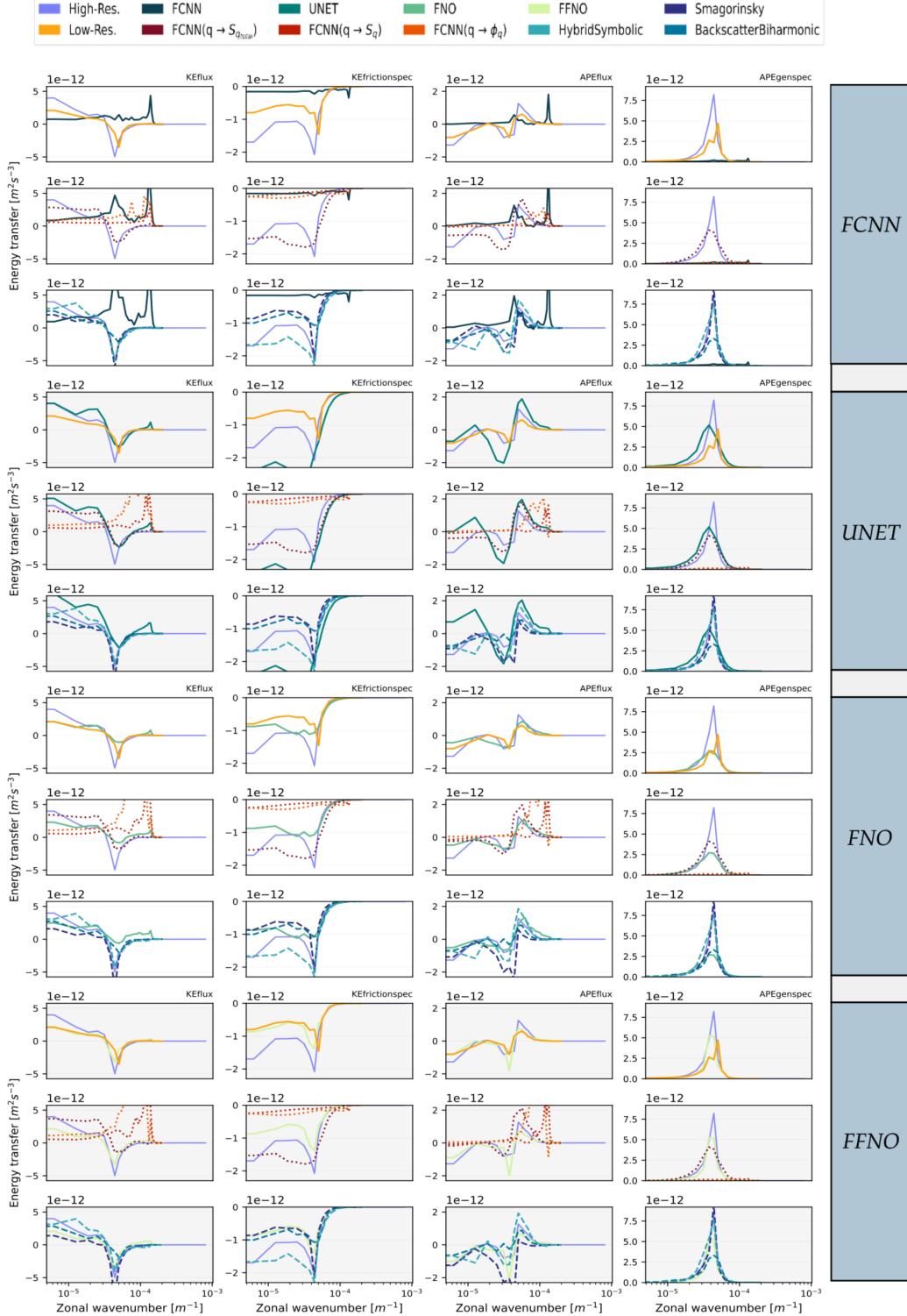


Figure 54: |Online - Phase 3 - Energy budget| This table displays energy spectra for KEflux, KEfrictionspec, APEflux, and APEgenspec using parameterizations grouped in Tab.1, these were trained on **eddies mixed (20000)** and tested on **eddies online**. Each parameterization spectrum is compared against high-resolution and various low-resolution simulations, including neural networks from Ross et al., 2023 and analytical parameterizations from Smagorinsky, 1963; Jansen and Held, 2014; Zanna and Bolton, 2020.

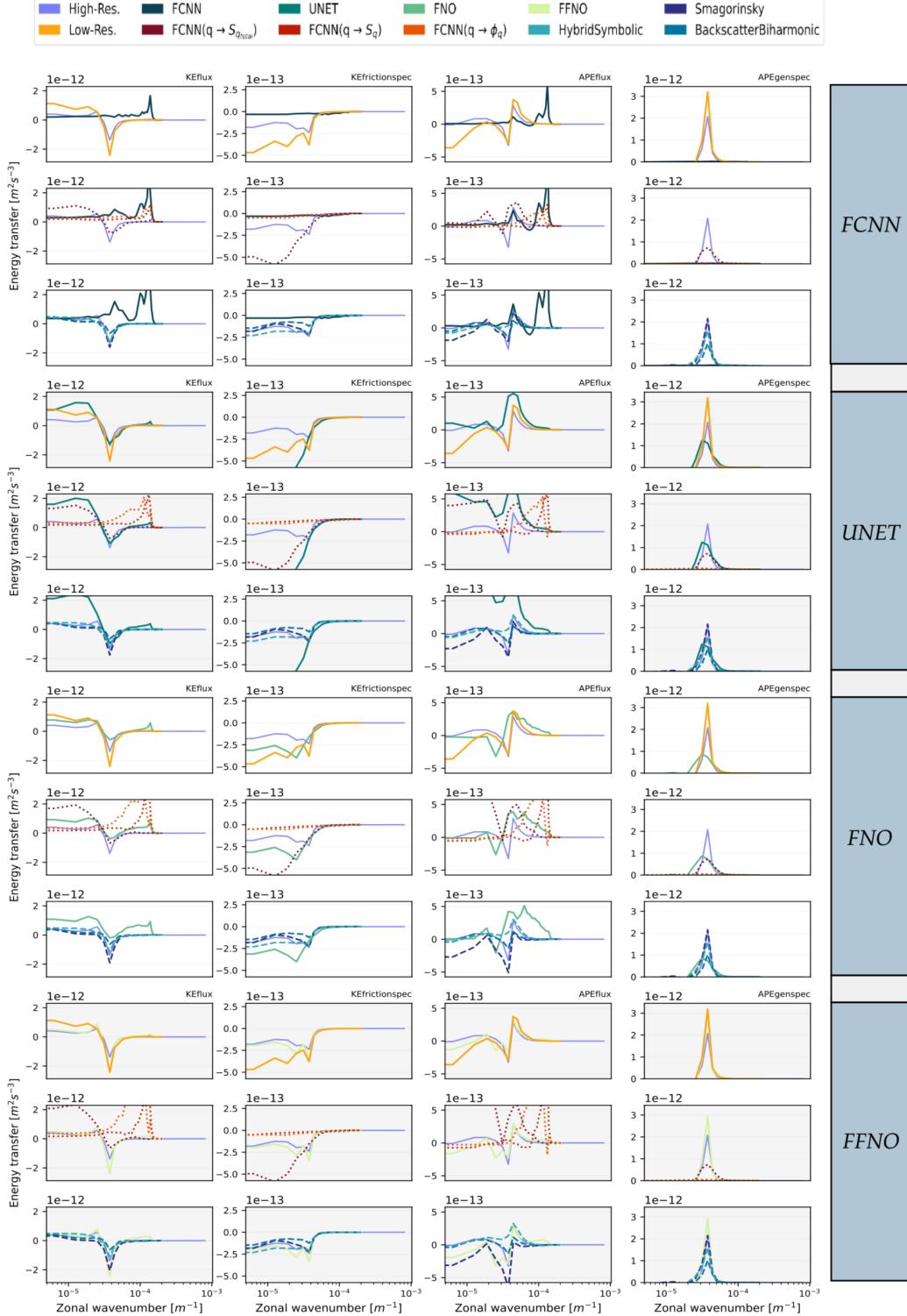


Figure 55: |Online - Phase 3 - Energy budget| This table displays energy spectra for KEflux, KEfrictionspec, APEflux, and APegaspec using parameterizations grouped in Tab.1, these were trained on **eddies mixed (20000)** and tested on **jets online**. Each parameterization spectrum is compared against high-resolution and various low-resolution simulations, including neural networks from Ross et al., 2023 and analytical parameterizations from Smagorinsky, 1963; Jansen and Held, 2014; Zanna and Bolton, 2020.

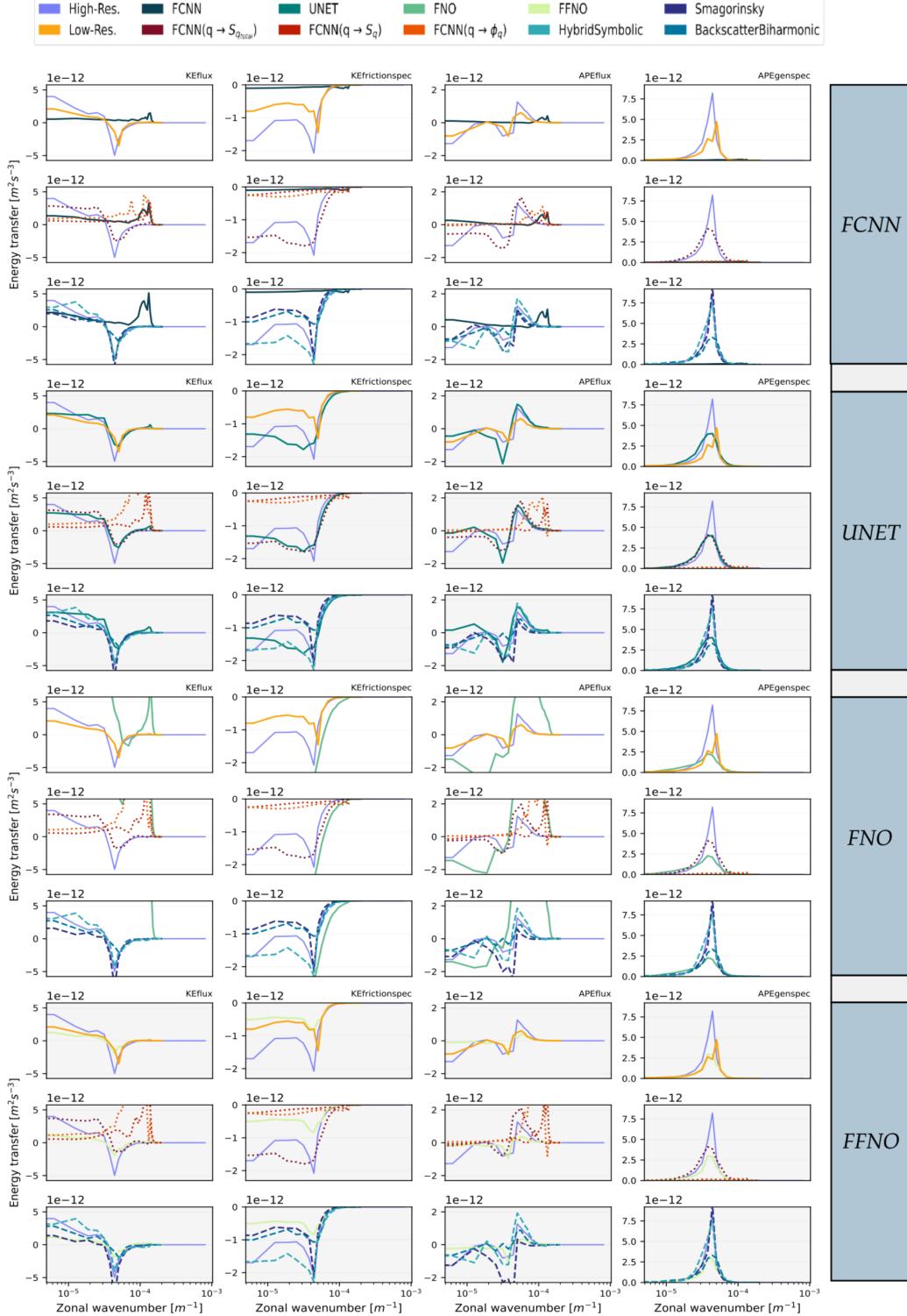


Figure 56: |Online - Phase 3 - Energy budget| This table displays energy spectra for KEflux, KEfrictionspec, APEflux, and APEgenspec using parameterizations grouped in Tab.1, these were trained on jets mixed (20000) and tested on eddies online. Each parameterization spectrum is compared against high-resolution and various low-resolution simulations, including neural networks from Ross et al., 2023 and analytical parameterizations from Smagorinsky, 1963; Jansen and Held, 2014; Zanna and Bolton, 2020.

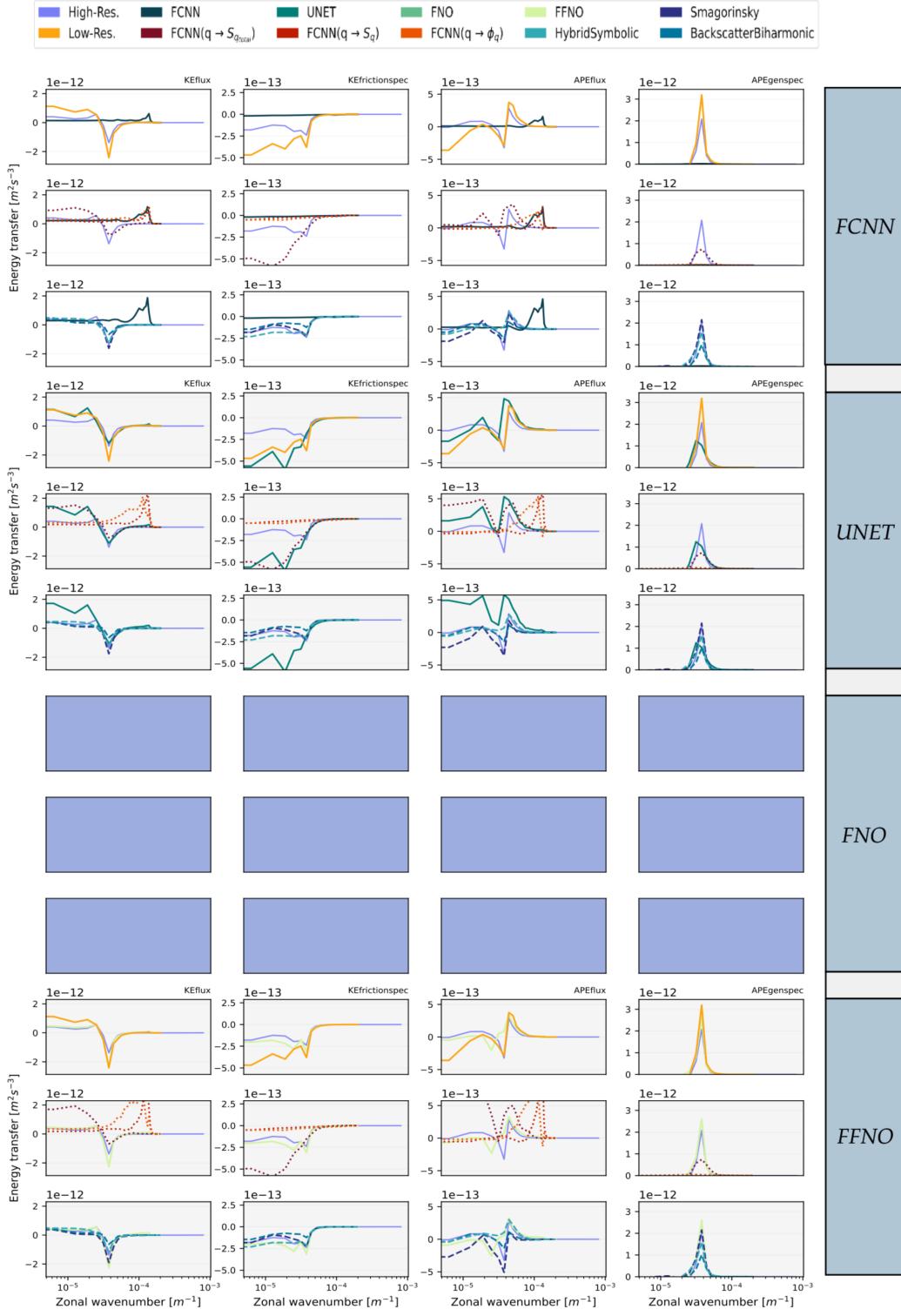


Figure 57: |Online - Phase 3 - Energy budget| This table displays energy spectra for KEflux, KEfrictionspec, APEflux, and APEgenspec using parameterizations grouped in Tab.1, these were trained on jets mixed (20000) and tested on jets online. Each parameterization spectrum is compared against high-resolution and various low-resolution simulations, including neural networks from Ross et al., 2023 and analytical parameterizations from Smagorinsky, 1963; Jansen and Held, 2014; Zanna and Bolton, 2020.

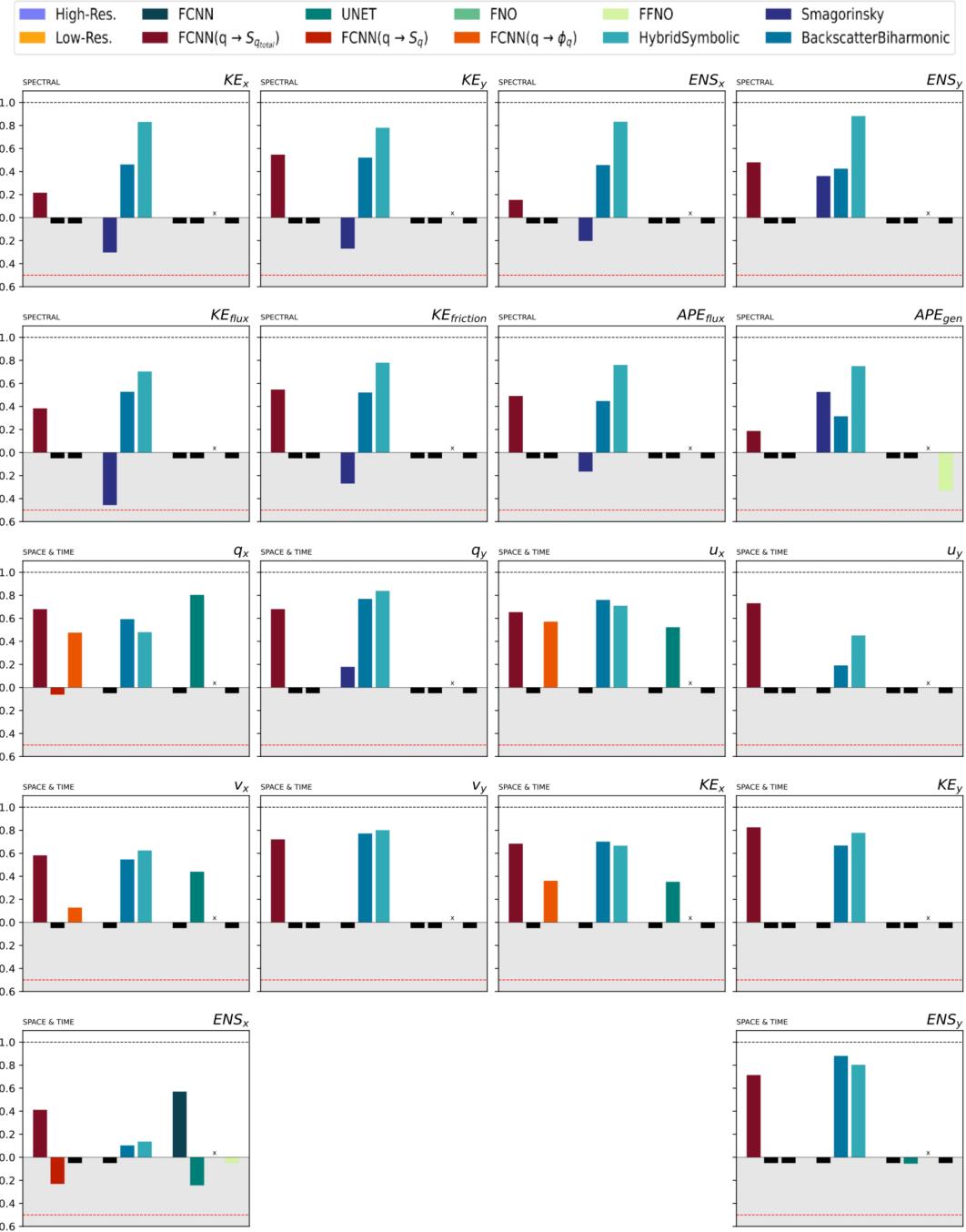


Figure 58: |Online - Phase 3 - Similarities| This table provides a summary of the Earth mover's distance, reformulated as a similarity metric for various flow quantities represented in either spectral or spatiotemporal domains. A value approaching 1 indicates strong agreement between the distribution obtained from high-resolution simulations and the current observations. Negative values are considered unfavorable, and values lower than -0.5 are disregarded. The tested parameterizations are grouped in Tab.1, they are trained on **eddies mixed (20000)** and tested on **eddies online**. For comparison, the results of neural networks (Ross et al., 2023) and analytical parameterizations are also presented (Smagorinsky, 1963; Jansen and Held, 2014; Zanna and Bolton, 2020).

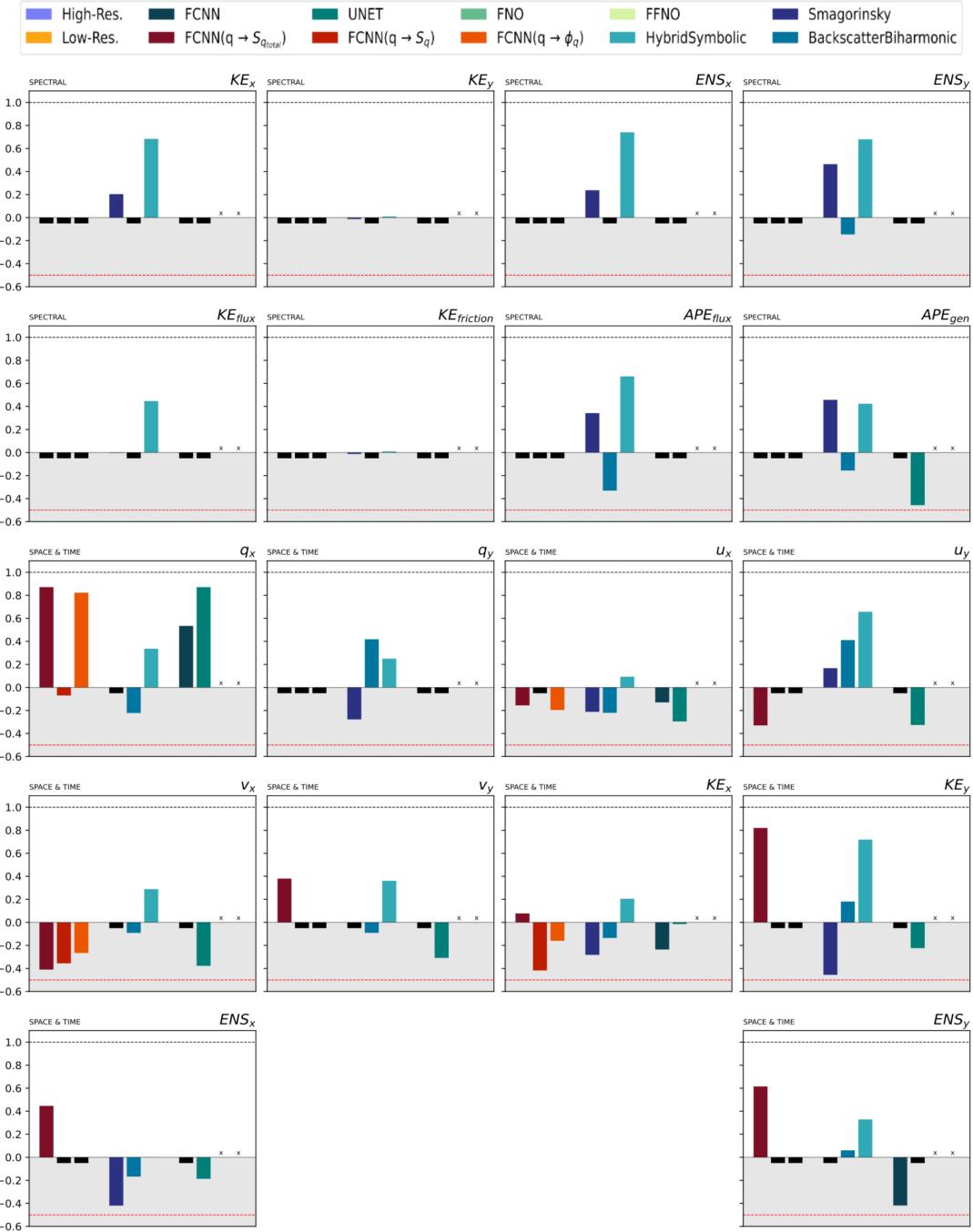


Figure 59: |Online - Phase 3 - Similarities| This table provides a summary of the Earth mover’s distance, reformulated as a similarity metric for various flow quantities represented in either spectral or spatiotemporal domains. A value approaching 1 indicates strong agreement between the distribution obtained from high-resolution simulations and the current observations. Negative values are considered unfavorable, and values lower than -0.5 are disregarded. The tested parameterizations are grouped in Tab.1, they are trained on **eddies mixed (20000)** and tested on **jets online**. For comparison, the results of neural networks (Ross et al., 2023) and analytical parameterizations are also presented (Smagorinsky, 1963; Jansen and Held, 2014; Zanna and Bolton, 2020).

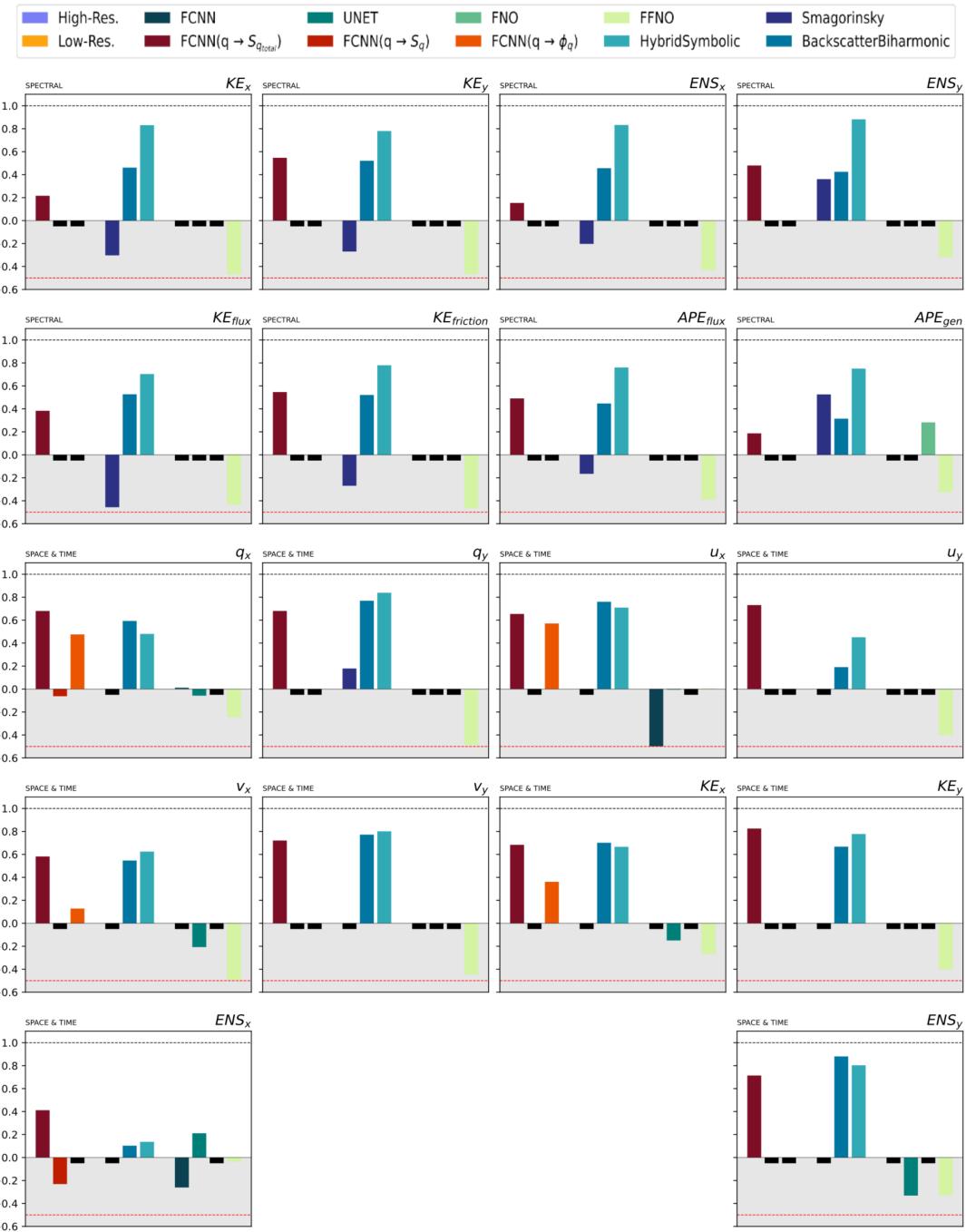


Figure 60: |Online - Phase 3 - Similarities| This table provides a summary of the Earth mover's distance, reformulated as a similarity metric for various flow quantities represented in either spectral or spatiotemporal domains. A value approaching 1 indicates strong agreement between the distribution obtained from high-resolution simulations and the current observations. Negative values are considered unfavorable, and values lower than -0.5 are disregarded. The tested parameterizations are grouped in Tab.1, they are trained on **jets mixed (20000)** and tested on **eddies online**. For comparison, the results of neural networks (Ross et al., 2023) and analytical parameterizations are also presented (Smagorinsky, 1963; Jansen and Held, 2014; Zanna and Bolton, 2020).

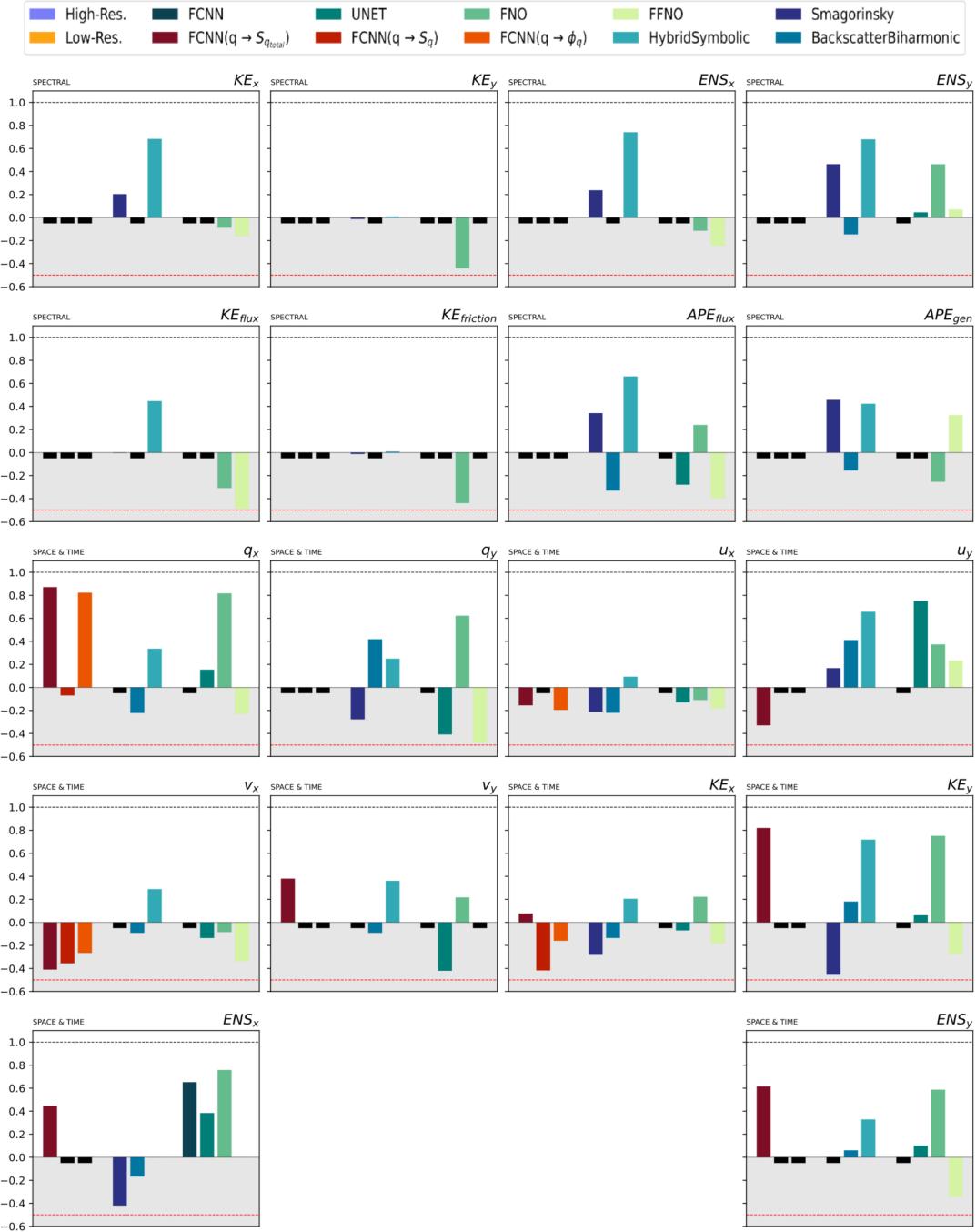


Figure 61: |Online - Phase 3 - Similarities| This table provides a summary of the Earth mover's distance, reformulated as a similarity metric for various flow quantities represented in either spectral or spatiotemporal domains. A value approaching 1 indicates strong agreement between the distribution obtained from high-resolution simulations and the current observations. Negative values are considered unfavorable, and values lower than -0.5 are disregarded. The tested parameterizations are grouped in Tab.1, they are trained on **jets mixed (20000)** and tested on **jets online**. For comparison, the results of neural networks (Ross et al., 2023) and analytical parameterizations are also presented (Smagorinsky, 1963; Jansen and Held, 2014; Zanna and Bolton, 2020).

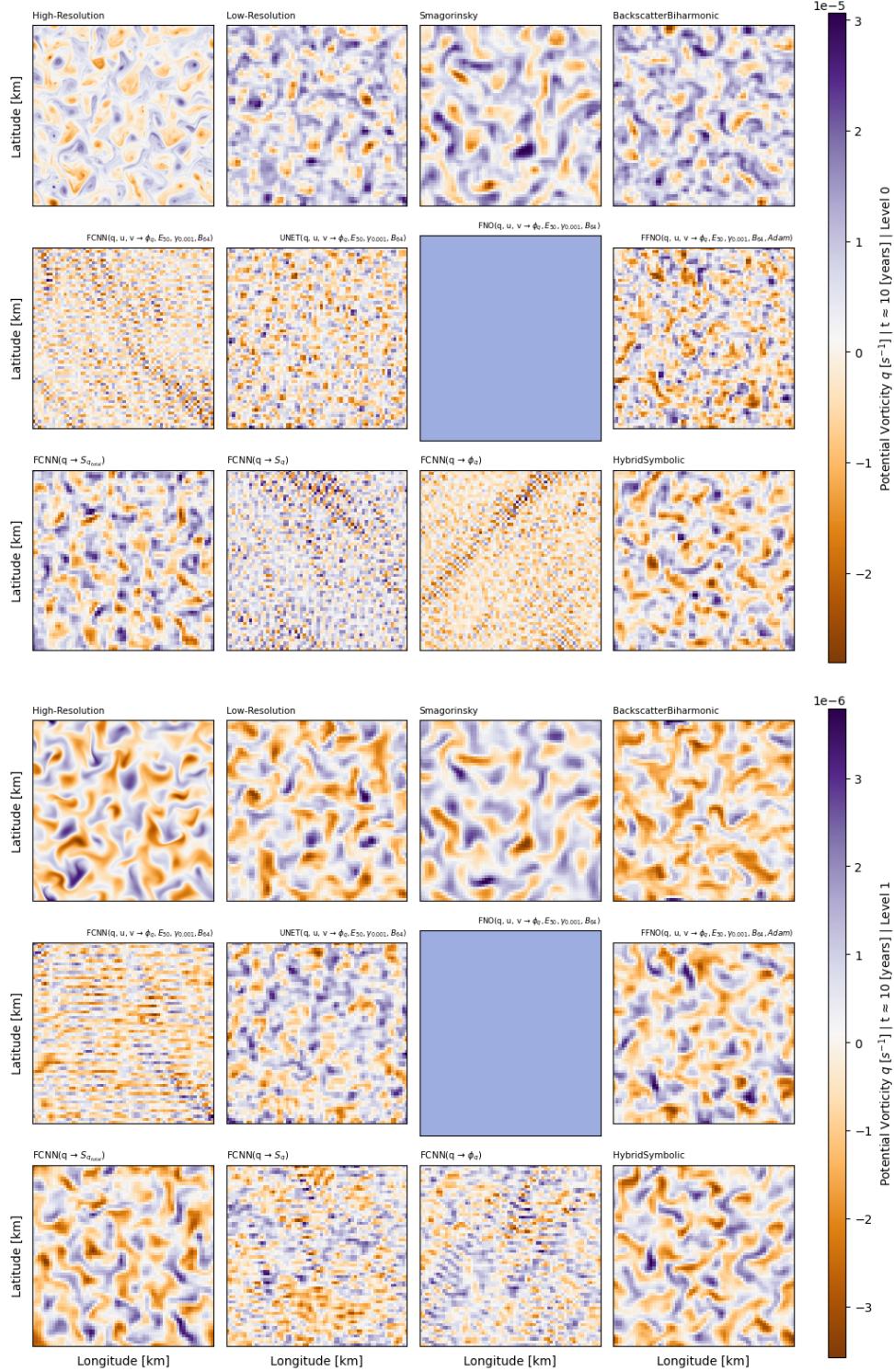
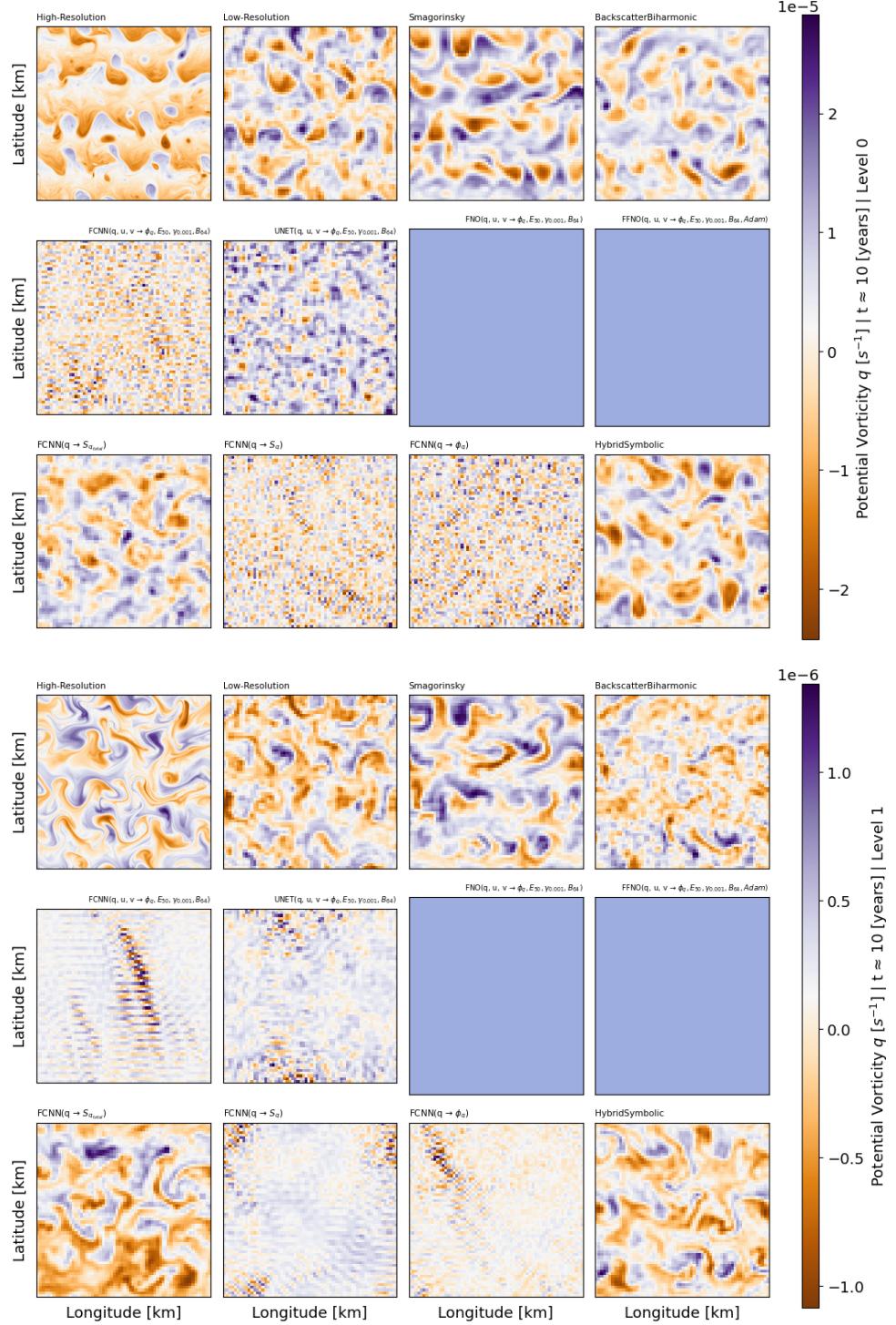
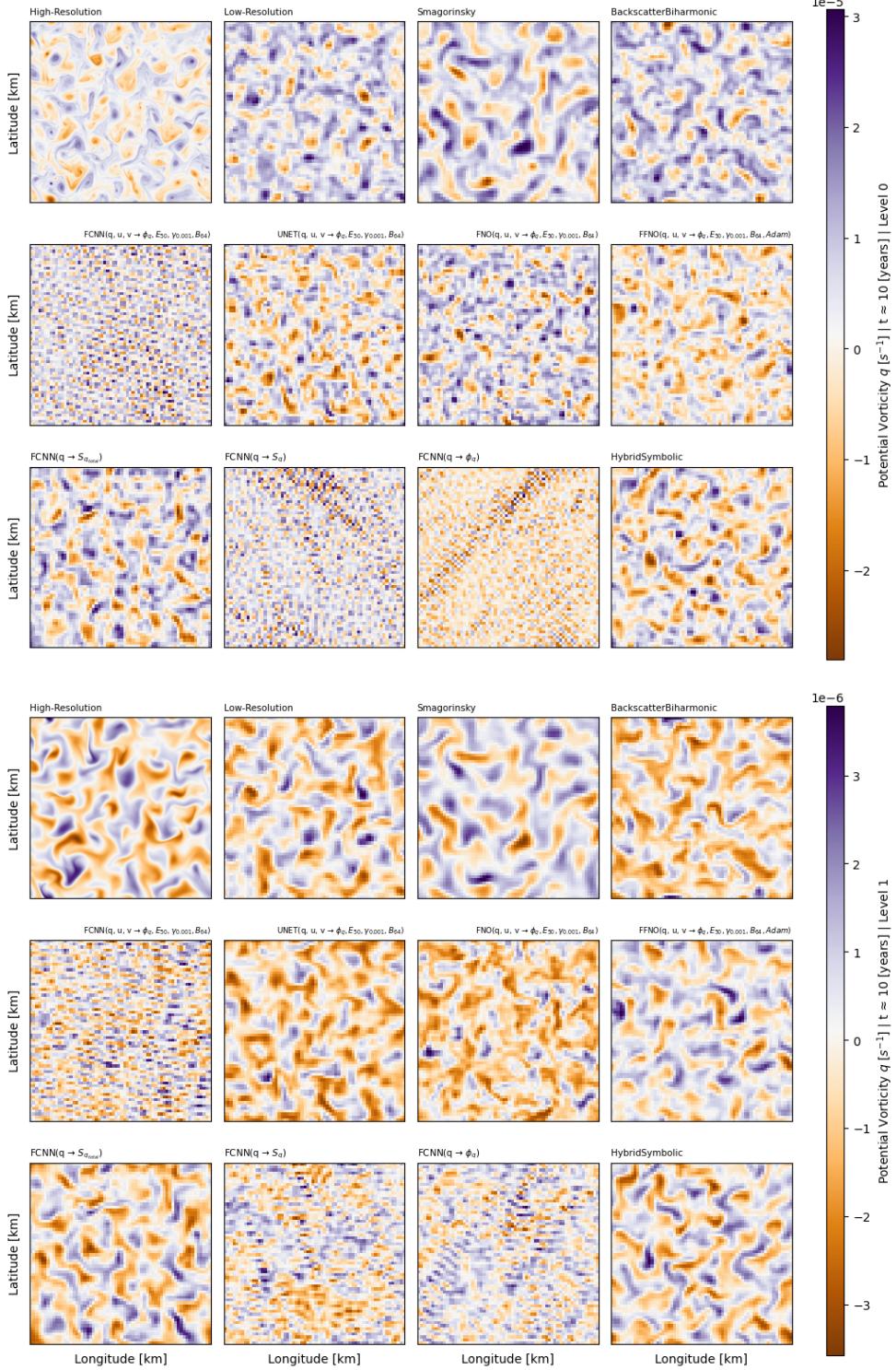


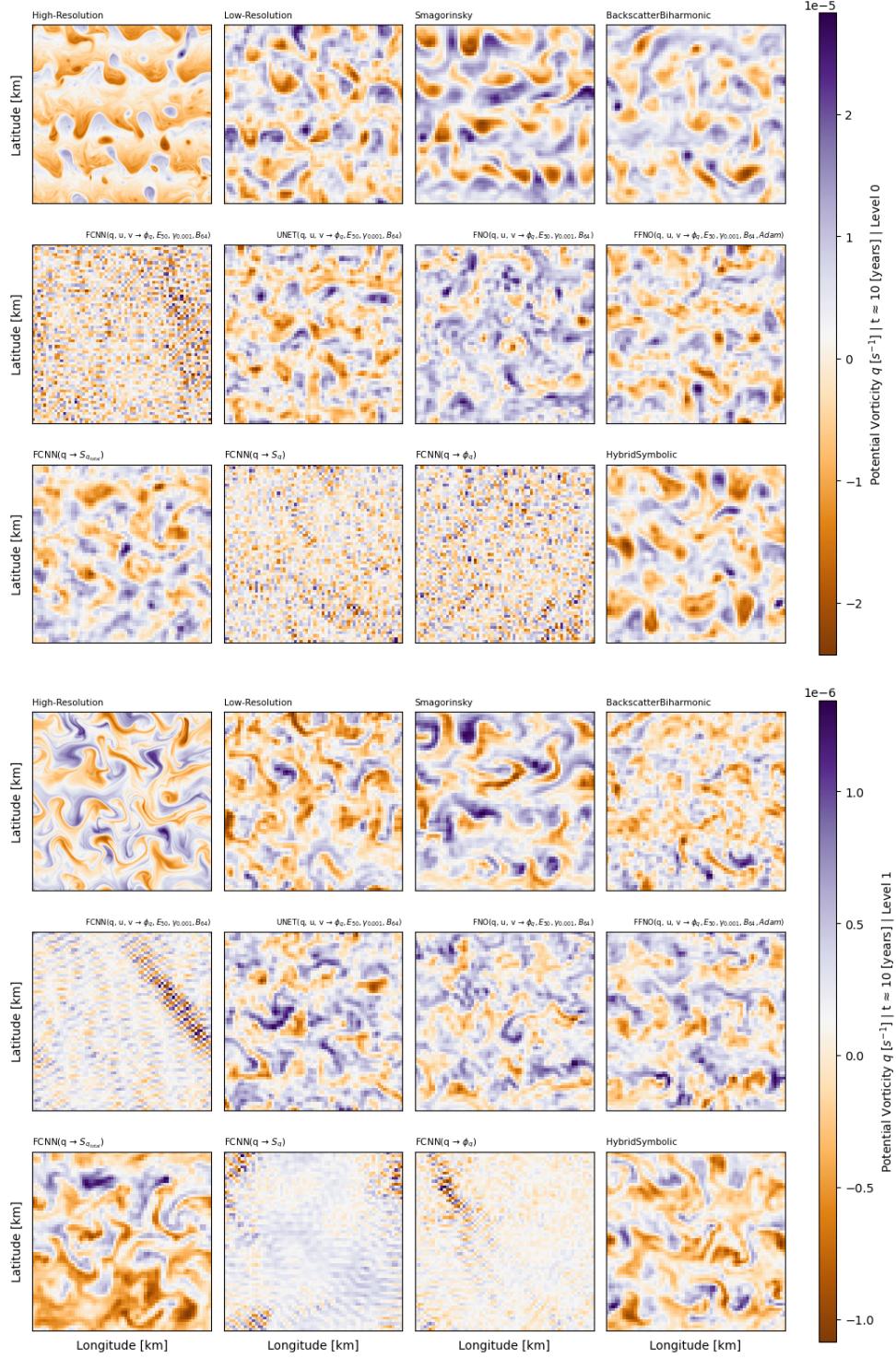
Figure 62: |Online - Phase 3 - Potential vorticity| Visualization of potential vorticity  $q$  for both upper (first three rows) and lower (last three rows) layers across different simulation types, indicated at the top of each image. Each image represents the  $q$  value spanning the entire computational domain after 10 years of simulations. The objective is to emphasize and visualize simulations that lose their physical relevance, becoming mere pixel grids, and to illustrate the divergence from the high-resolution simulation. Furthermore, the evaluated parameterizations are detailed in Tab.1, they were trained using **eddies mixed (20000)** and tested on **eddies online**.



**Figure 63: | Online - Phase 3 - Potential vorticity |** Visualization of potential vorticity  $q$  for both upper (first three rows) and lower (last three rows) layers across different simulation types, indicated at the top of each image. Each image represents the  $q$  value spanning the entire computational domain after 10 years of simulations. The objective is to emphasize and visualize simulations that lose their physical relevance, becoming mere pixel grids, and to illustrate the divergence from the high-resolution simulation. Furthermore, the evaluated parameterizations are detailed in Tab.1, they were trained using **eddies mixed (20000)** and tested on **jets online**.



**Figure 64: |Online - Phase 3 - Potential vorticity|** Visualization of potential vorticity  $q$  for both upper (first three rows) and lower (last three rows) layers across different simulation types, indicated at the top of each image. Each image represents the  $q$  value spanning the entire computational domain after 10 years of simulations. The objective is to emphasize and visualize simulations that lose their physical relevance, becoming mere pixel grids, and to illustrate the divergence from the high-resolution simulation. Furthermore, the evaluated parameterizations are detailed in Tab.1, they were trained using **jets mixed (20000)** and tested on **eddies online**.



**Figure 65: |Online - Phase 3 - Potential vorticity|** Visualization of potential vorticity  $q$  for both upper (first three rows) and lower (last three rows) layers across different simulation types, indicated at the top of each image. Each image represents the  $q$  value spanning the entire computational domain after 10 years of simulations. The objective is to emphasize and visualize simulations that lose their physical relevance, becoming mere pixel grids, and to illustrate the divergence from the high-resolution simulation. Furthermore, the evaluated parameterizations are detailed in Tab.1, they were trained using **jets mixed (20000)** and tested on **jets online**.

# PHASE IV

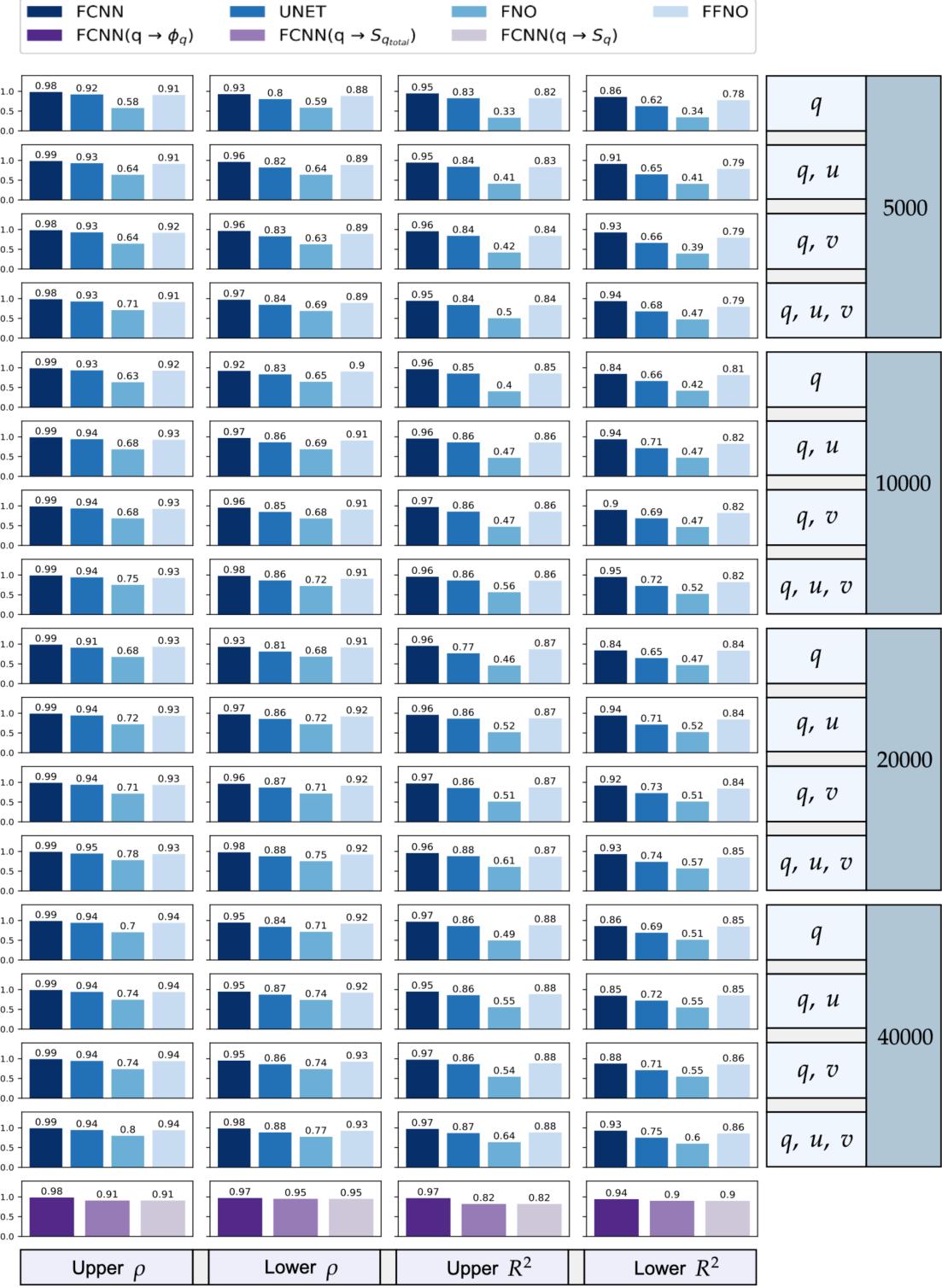


Figure 66: |Offline - Phase 4| This table summarizes offline results, including correlations (columns 1 and 2) and mean-squared errors (columns 3 and 4), for parameterizations trained on **full dataset**, evaluated on dataset **eddies offline** and predicting subgrid flux  $\Phi_q$  (see Eq. 23). On the right, input details for each parameterization are supplied, along with the number of samples used for training. The final row showcases outcomes from the three FCNN parameterizations introduced in Ross et al., 2023.

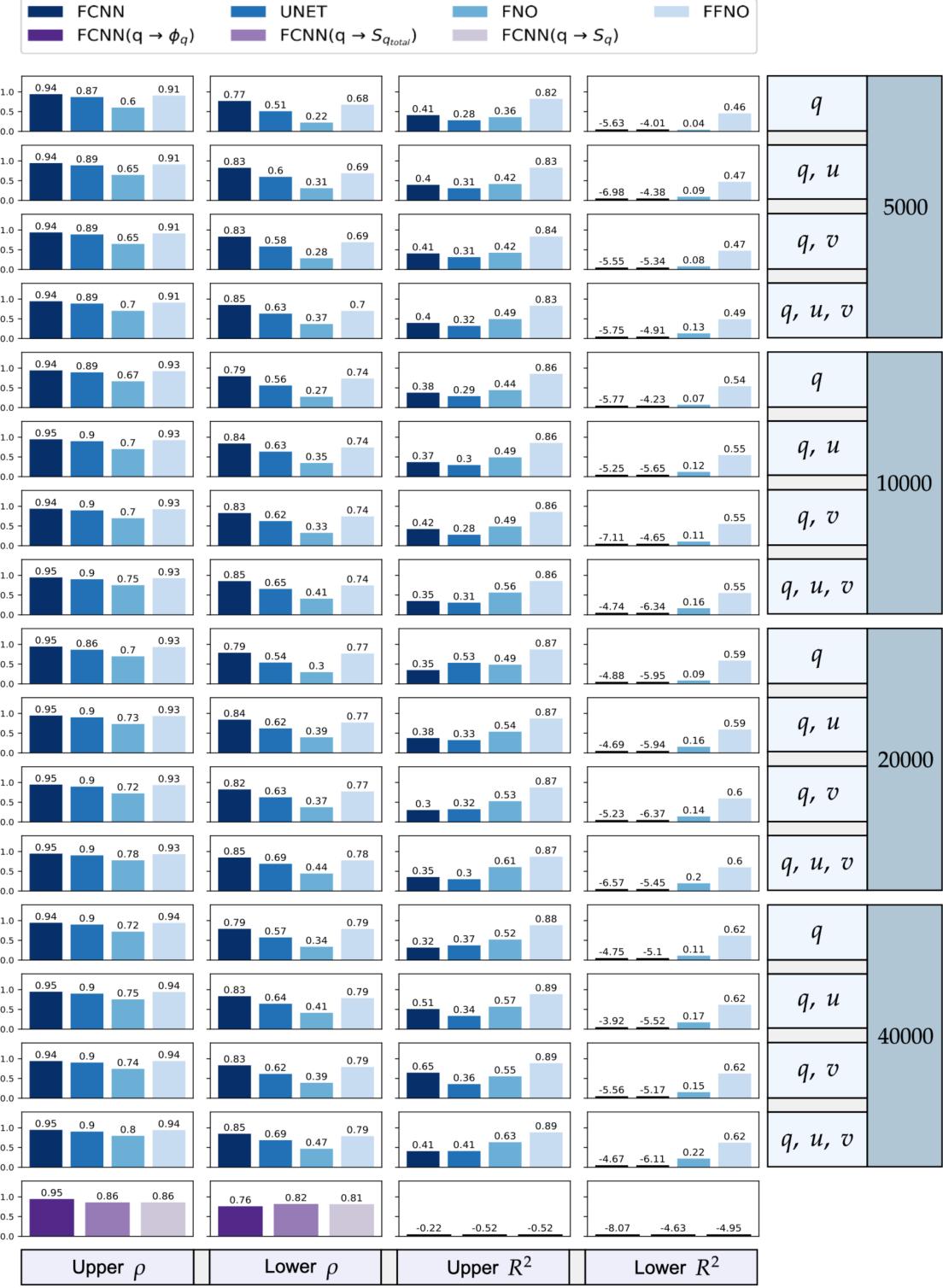


Figure 67: |Offline - Phase 4| This table summarizes offline results, including correlations (columns 1 and 2) and mean-squared errors (columns 3 and 4), for parameterizations trained on **full dataset**, evaluated on dataset **jets offline** and predicting subgrid flux  $\Phi_q$  (see Eq. 23). On the right, input details for each parameterization are supplied, along with the number of samples used for training. The final row showcases outcomes from the three FCNN parameterizations introduced in Ross et al., 2023.

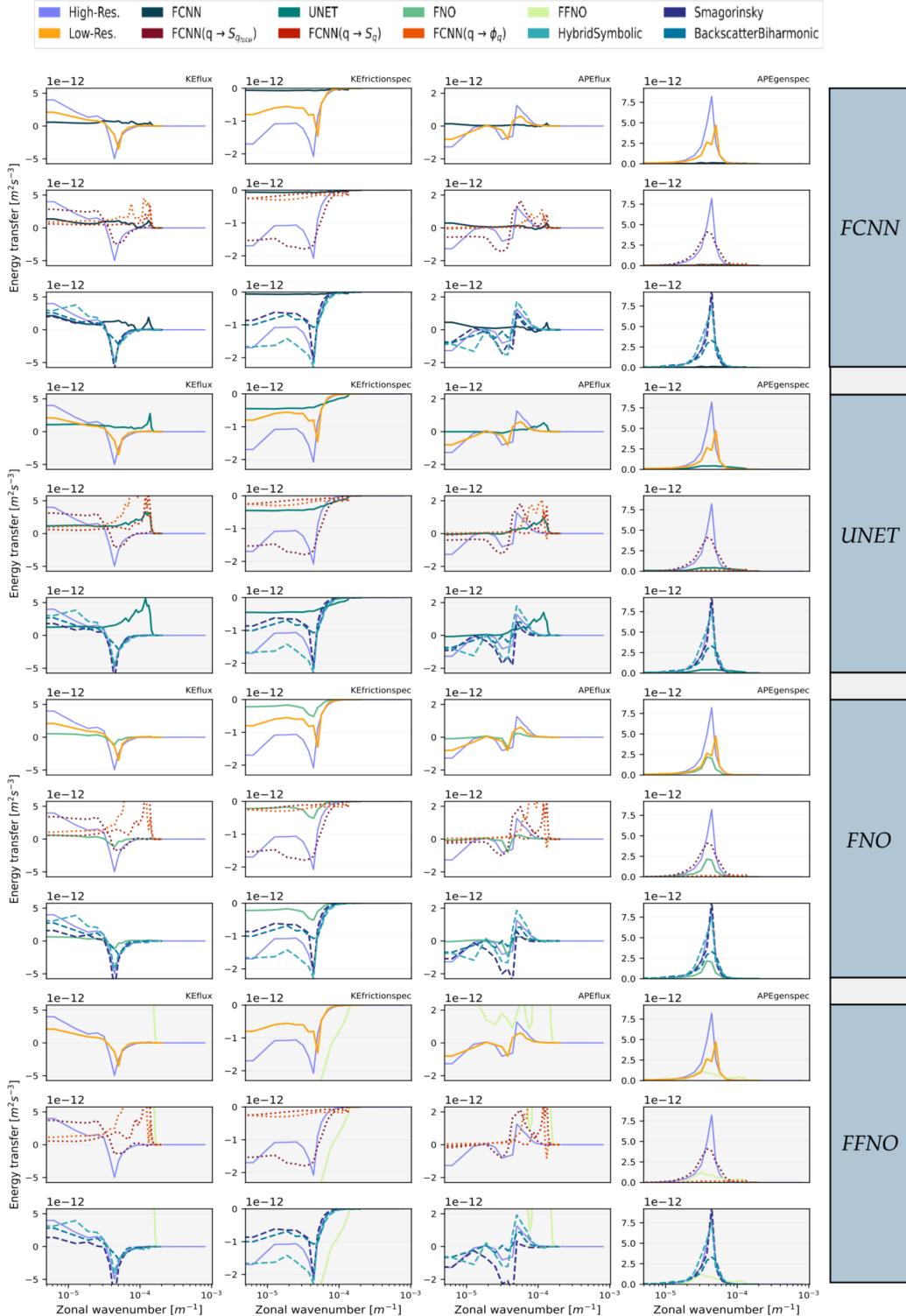


Figure 68: |Online - Phase 4 - Energy budget| This table displays energy spectra for **KEflux**, **KEfrictionspec**, **APEflux**, and **APEgenspec** using parameterizations grouped in Tab.??, these were trained on **full 40000** and tested on **eddies online**. Each parameterization spectrum is compared against high-resolution and various low-resolution simulations, including neural networks from Ross et al., 2023 and analytical parameterizations from Smagorinsky, 1963; Jansen and Held, 2014; Zanna and Bolton, 2020.

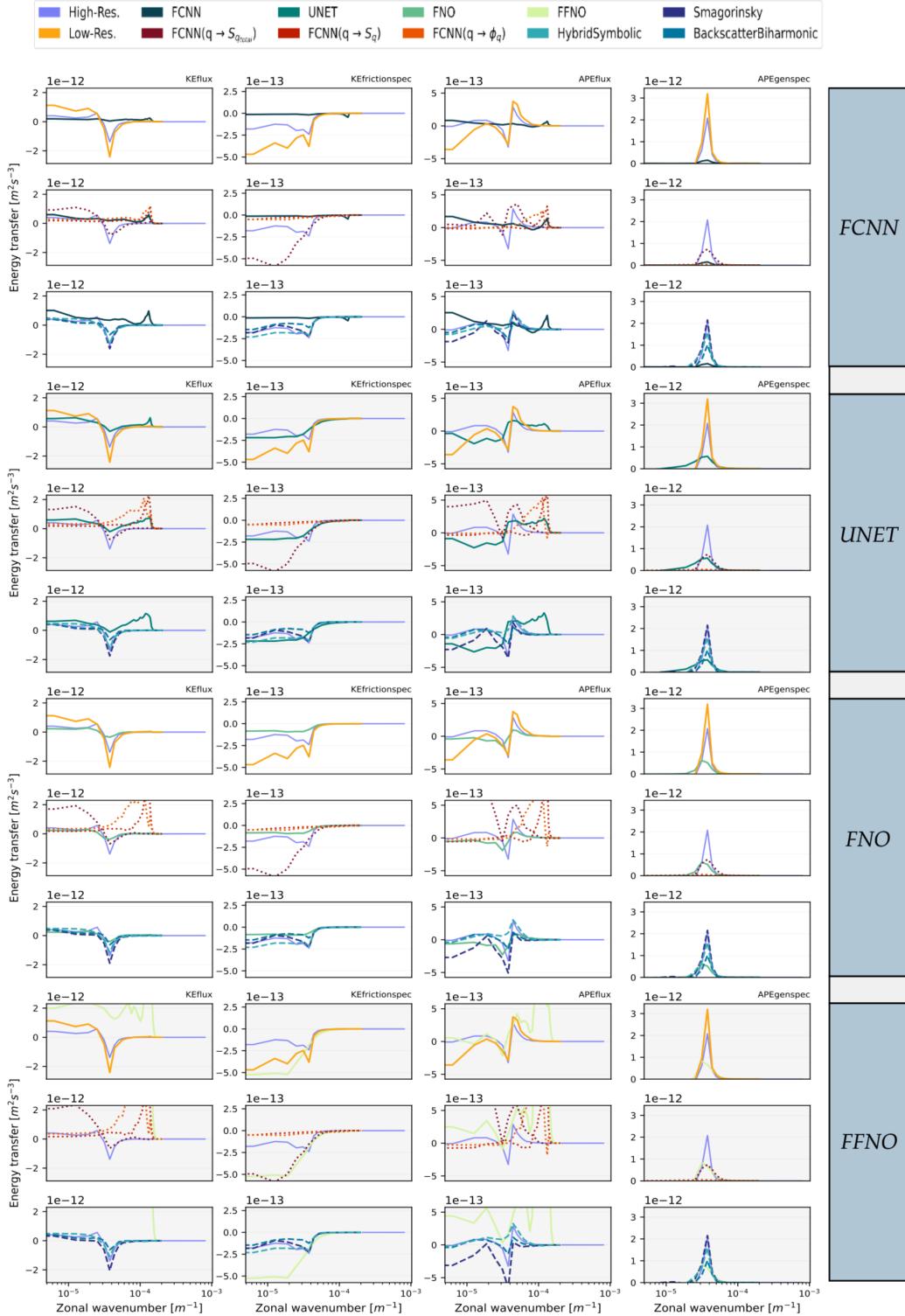


Figure 69: |Online - Phase 4 - Energy budget| This table displays energy spectra for **KEflux**, **KEfrictionspec**, **APEflux**, and **APEngenspec** using parameterizations grouped in Tab.??, these were trained on **full 40000** and tested on **jets online**. Each parameterization spectrum is compared against high-resolution and various low-resolution simulations, including neural networks from Ross et al., 2023 and analytical parameterizations from Smagorinsky, 1963; Jansen and Held, 2014; Zanna and Bolton, 2020.

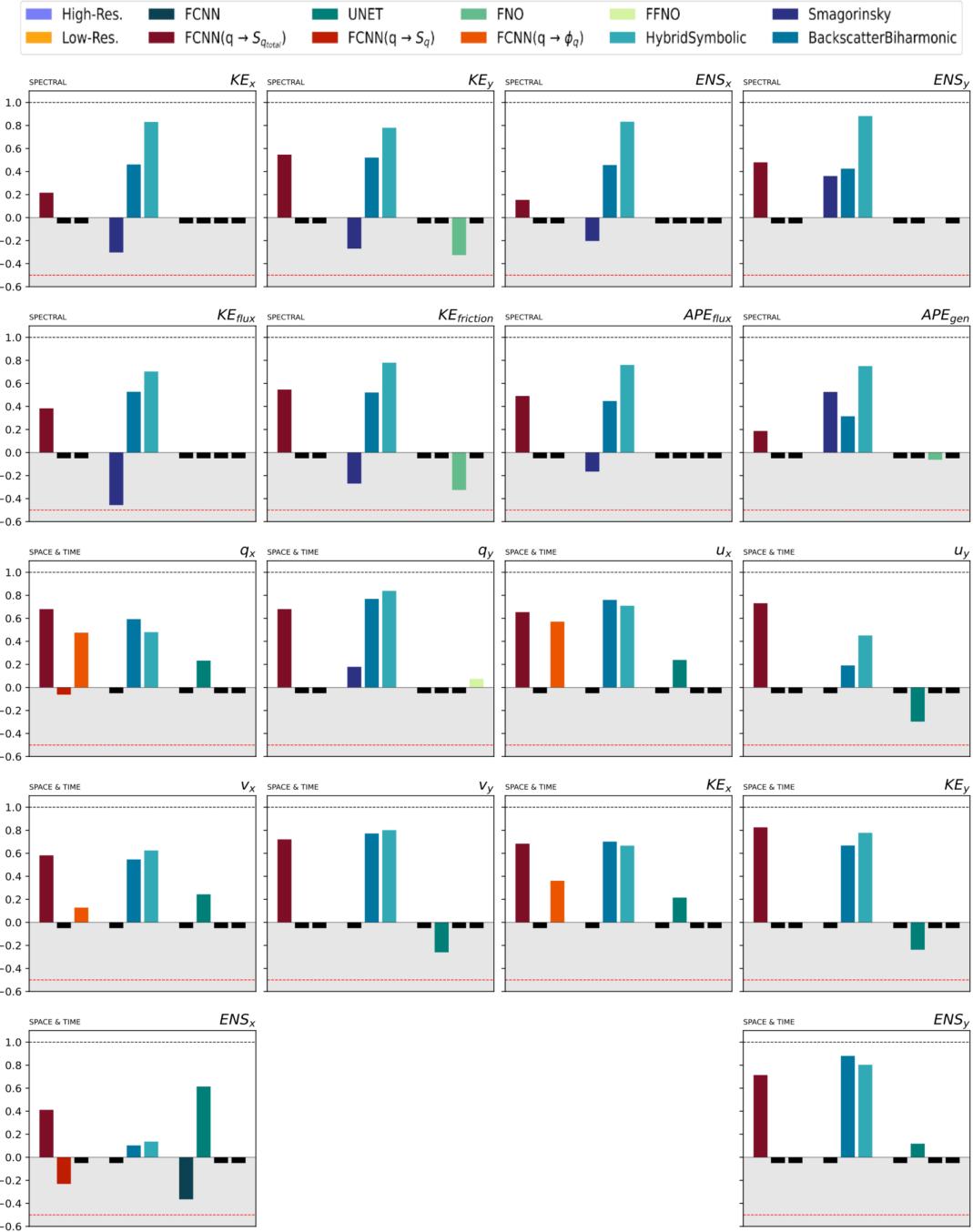


Figure 70: |Online - Phase 4 - Similarities| This table provides a summary of the Earth mover's distance, reformulated as a similarity metric for various flow quantities represented in either spectral or spatiotemporal domains. A value approaching 1 indicates strong agreement between the distribution obtained from high-resolution simulations and the current observations. Negative values are considered unfavorable, and values lower than -0.5 are disregarded. The tested parameterizations are grouped in Tab.??, they are trained on **full 40000** and tested on **eddies online**. For comparison, the results of neural networks (Ross et al., 2023) and analytical parameterizations are also presented (Smagorinsky, 1963; Jansen and Held, 2014; Zanna and Bolton, 2020).

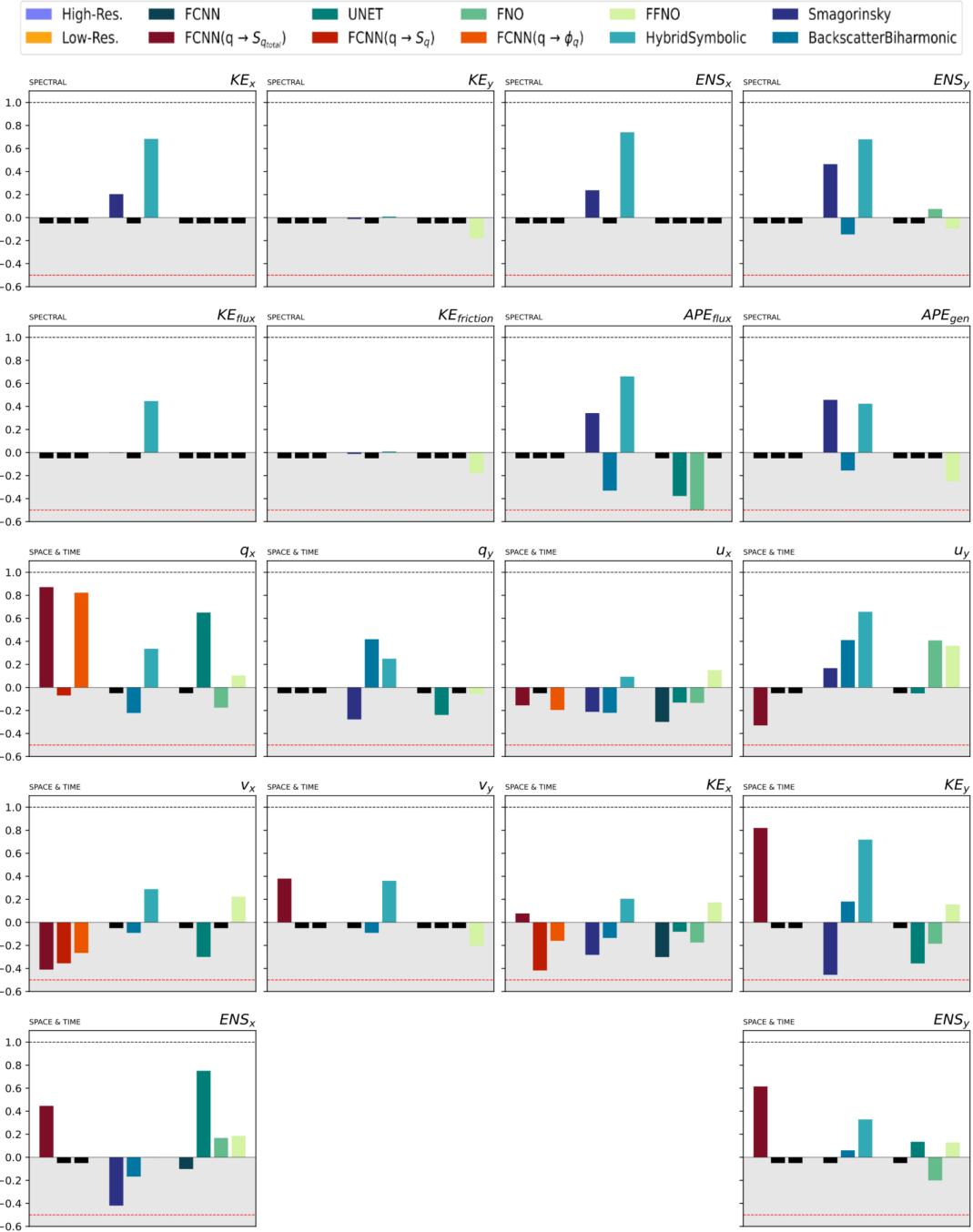


Figure 71: |Online - Phase 4 - Similarities| This table provides a summary of the Earth mover's distance, reformulated as a similarity metric for various flow quantities represented in either spectral or spatiotemporal domains. A value approaching 1 indicates strong agreement between the distribution obtained from high-resolution simulations and the current observations. Negative values are considered unfavorable, and values lower than -0.5 are disregarded. The tested parameterizations are grouped in Tab.??, they are trained on **full 40000** and tested on **jets online**. For comparison, the results of neural networks (Ross et al., 2023) and analytical parameterizations are also presented (Smagorinsky, 1963; Jansen and Held, 2014; Zanna and Bolton, 2020).

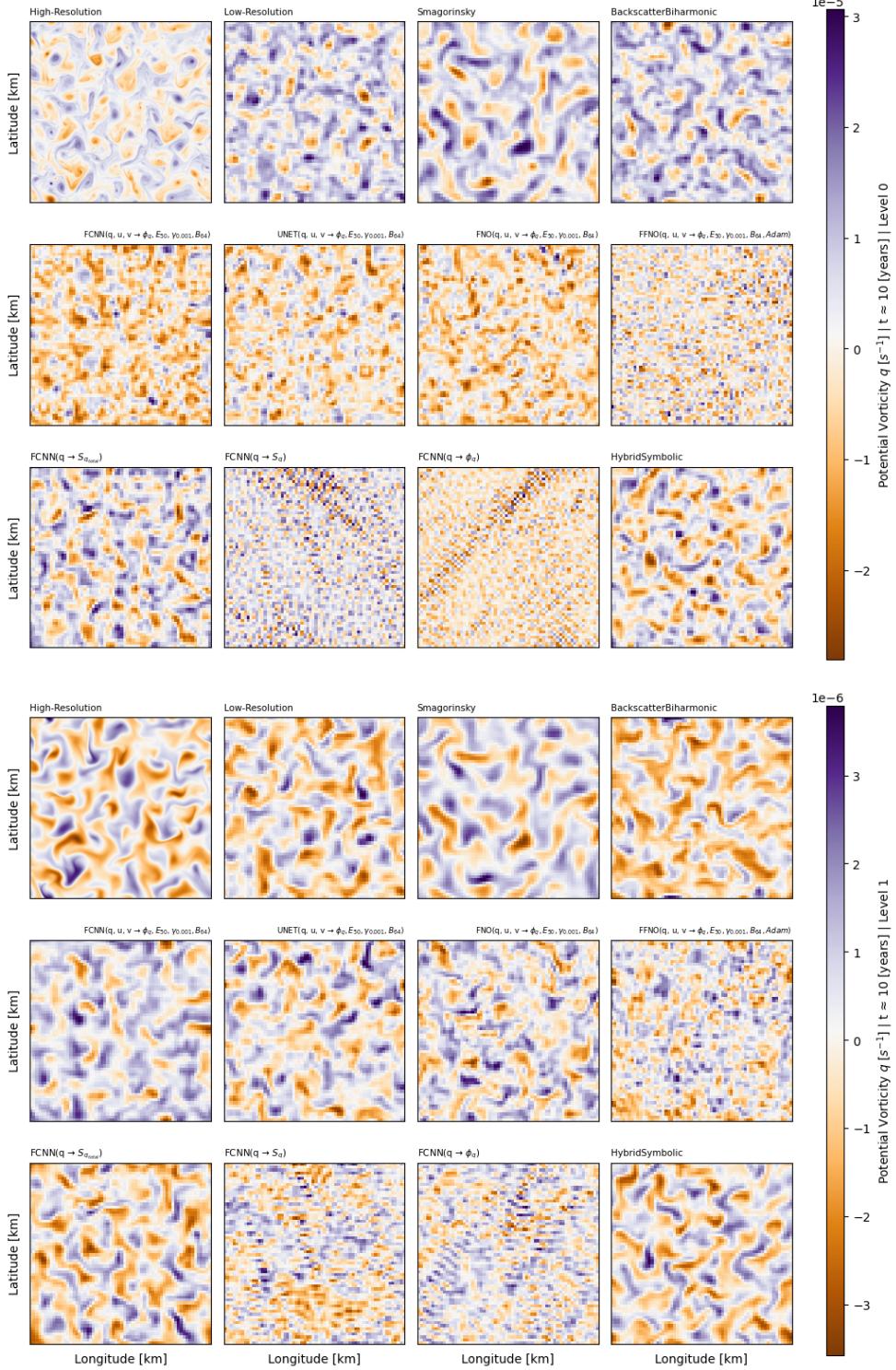
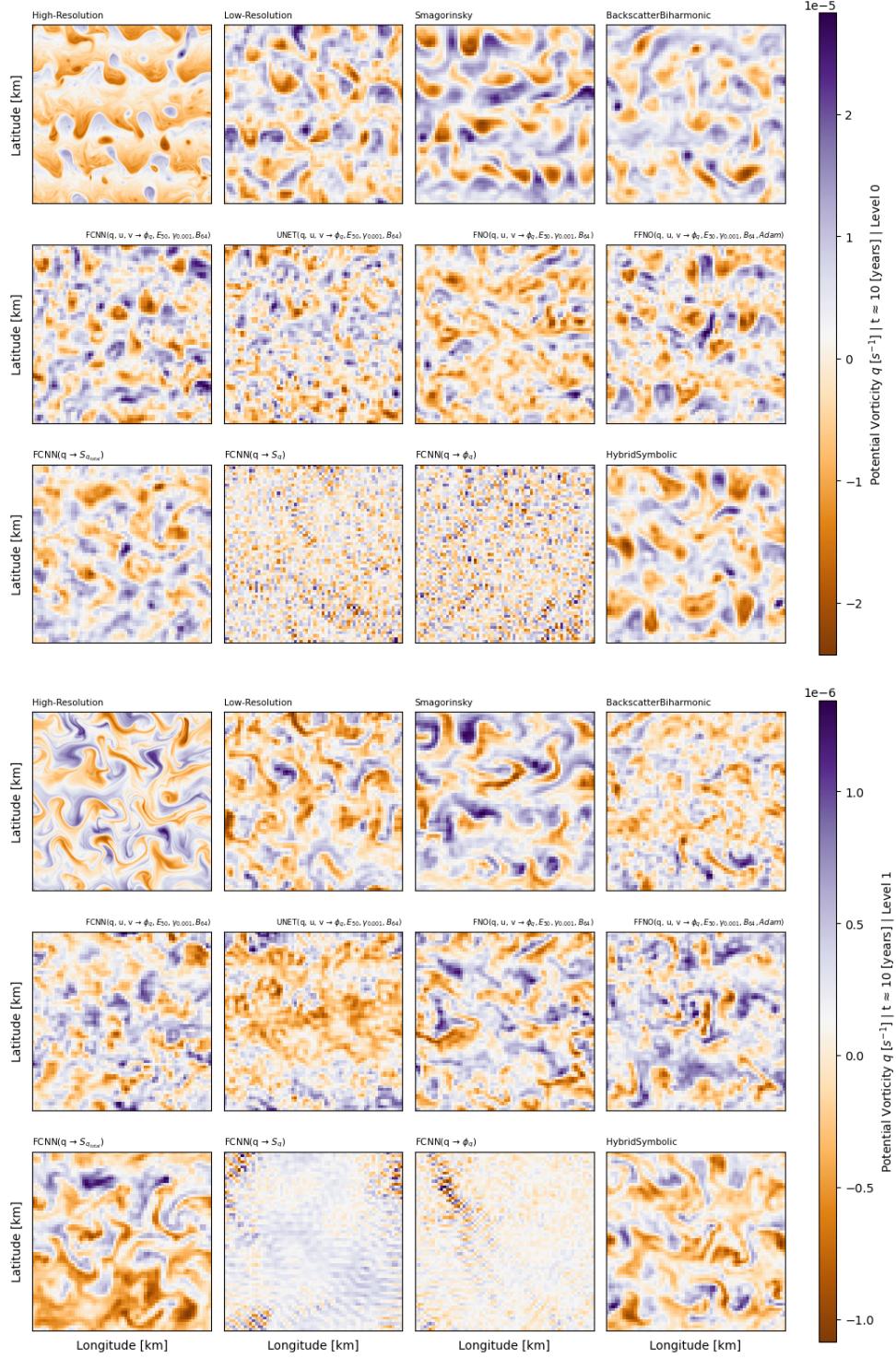


Figure 72: |Online - Phase 4 - Potential vorticity| Visualization of potential vorticity  $q$  for both upper (first three rows) and lower (last three rows) layers across different simulation types, indicated at the top of each image. Each image represents the  $q$  value spanning the entire computational domain after 10 years of simulations. The objective is to emphasize and visualize simulations that lose their physical relevance, becoming mere pixel grids, and to illustrate the divergence from the high-resolution simulation. Furthermore, the evaluated parameterizations are detailed in Tab.??, they were trained using **full 40000** and tested on **eddies online**.



**Figure 73: |Online - Phase 4 - Potential vorticity|** Visualization of potential vorticity  $q$  for both upper (first three rows) and lower (last three rows) layers across different simulation types, indicated at the top of each image. Each image represents the  $q$  value spanning the entire computational domain after 10 years of simulations. The objective is to emphasize and visualize simulations that lose their physical relevance, becoming mere pixel grids, and to illustrate the divergence from the high-resolution simulation. Furthermore, the evaluated parameterizations are detailed in Tab.??, they were trained using **full 40000** and tested on **jets online**.

# PHASE V

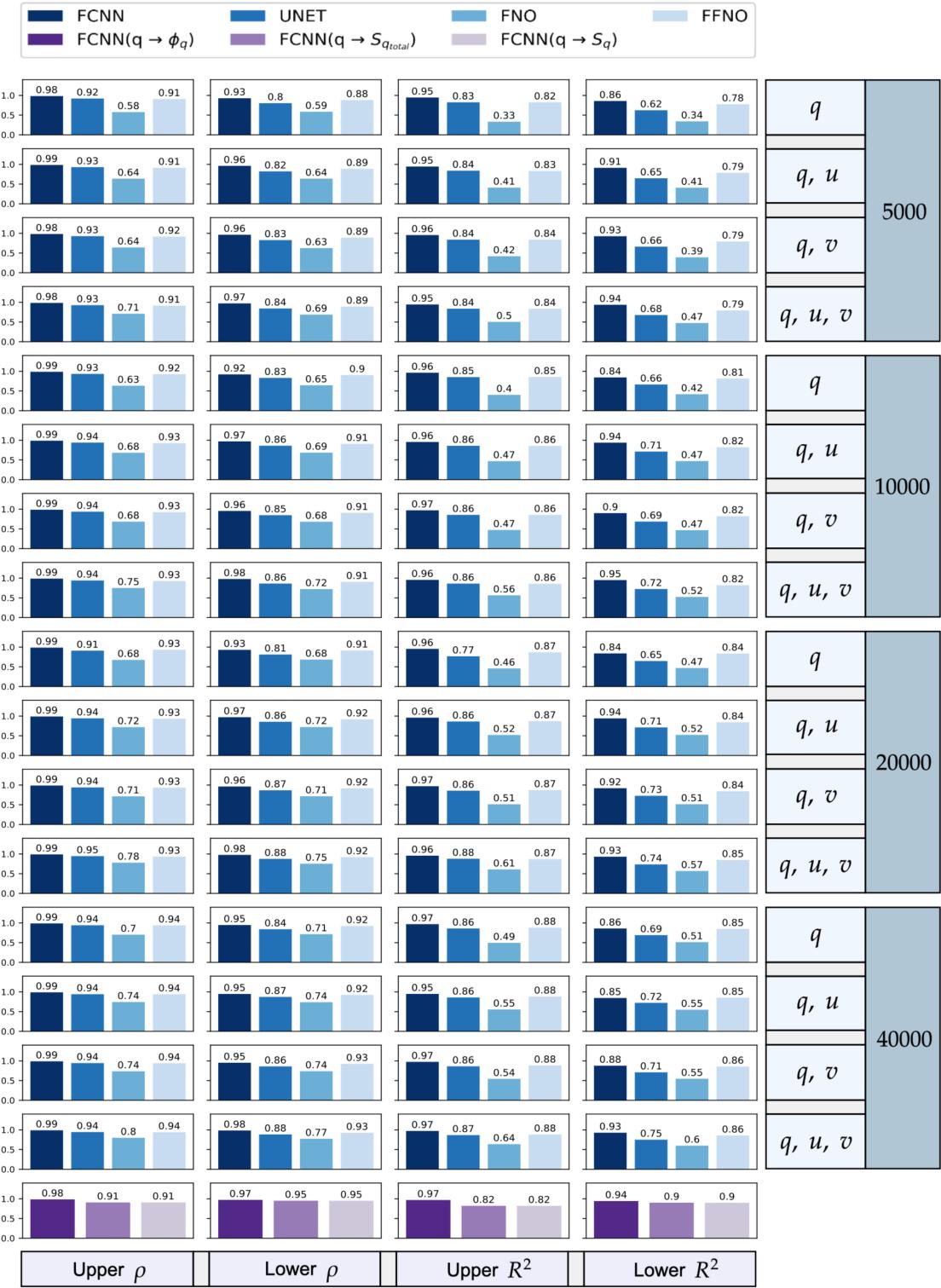


Figure 74: | Offline - Phase 5 | This table summarizes offline results, including correlations ( $\rho$ ) and mean-squared errors ( $R^2$ ), for parameterizations trained on **full dataset 5000**, evaluated on dataset **eddies offline** and predicting subgrid flux  $\Phi_q$  (see Eq. 23). On the right, details regarding the optimizer and scheduler employed for training are provided, along with the corresponding learning rates displayed in the legend. The bottom row presents results obtained using the three FCNN parameterizations introduced in Ross et al., 2023.

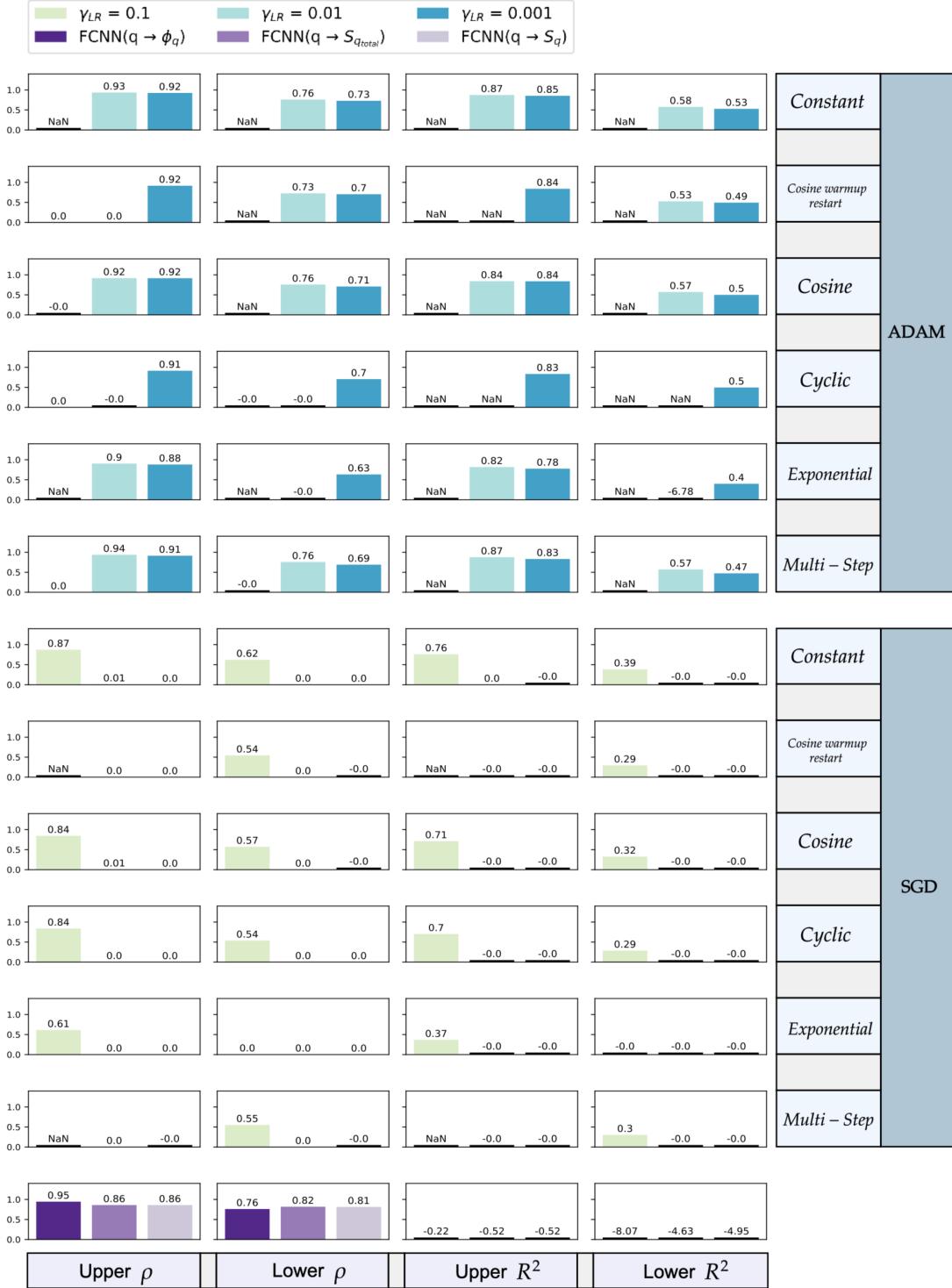


Figure 75: | Offline - Phase 5| This table summarize offline results, including correlations (columns 1 and 2) and mean-squared errors (columns 3 and 4), for parameterizations trained on **full dataset 5000**, evaluated on dataset **jets offline** and predicting subgrid flux  $\Phi_q$  (see Eq. 23). On the right, details regarding the optimizer and scheduler employed for training are provided, along with the corresponding learning rates displayed in the legend. The bottom row presents results obtained using the three FCNN parameterizations introduced in Ross et al., 2023.

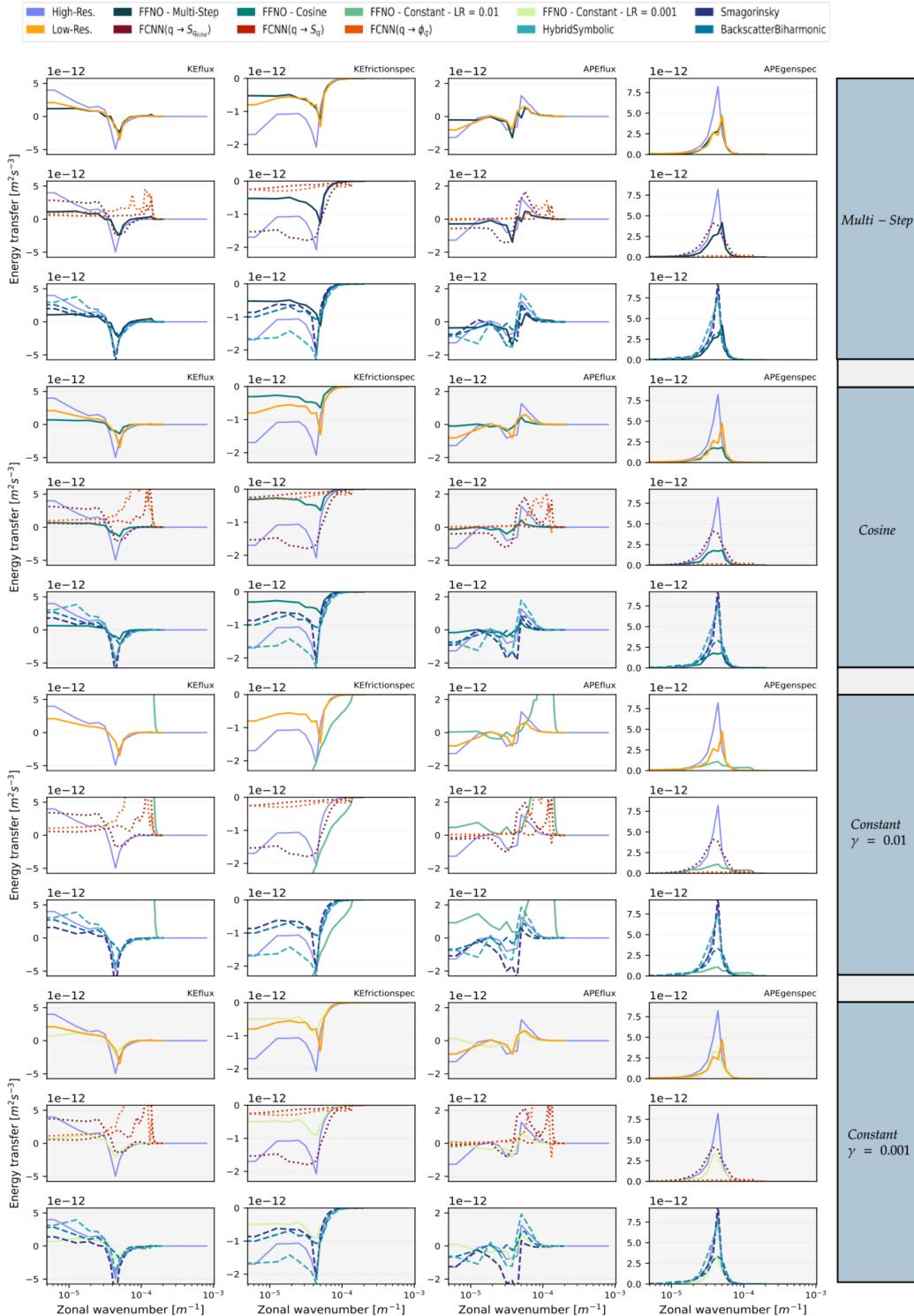


Figure 76: |Online - Phase 5 - Energy budget| This table displays energy spectra for **KEflux**, **KEfrictionspec**, **APEflux**, and **APEgenspec** using parameterizations of Tab.??, these were trained on full 5000 and tested on **eddies online**. Each parameterization spectrum is compared against high-resolution and various low-resolution simulations, including neural networks from Ross et al., 2023 and analytical parameterizations from Smagorinsky, 1963; Jansen and Held, 2014; Zanna and Bolton, 2020.

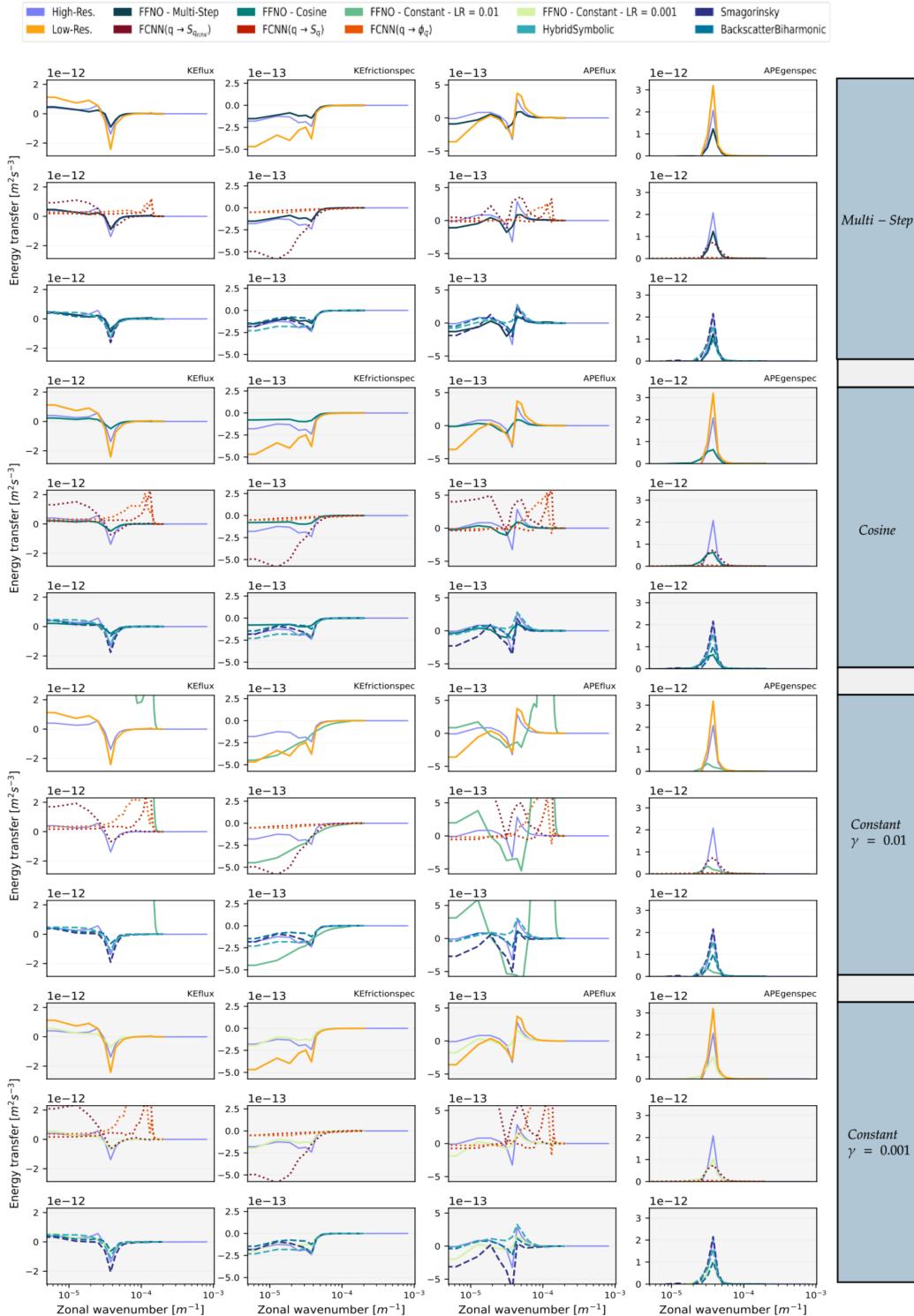


Figure 77: |Online - Phase 5 - Energy budget| This table displays energy spectra for KEflux, KEfrictionspec, APEflux, and APEgenspec using parameterizations of Tab.??, these were trained on **full 5000** and tested on **jets online**. Each parameterization spectrum is compared against high-resolution and various low-resolution simulations, including neural networks from Ross et al., 2023 and analytical parameterizations from Smagorinsky, 1963; Jansen and Held, 2014; Zanna and Bolton, 2020.

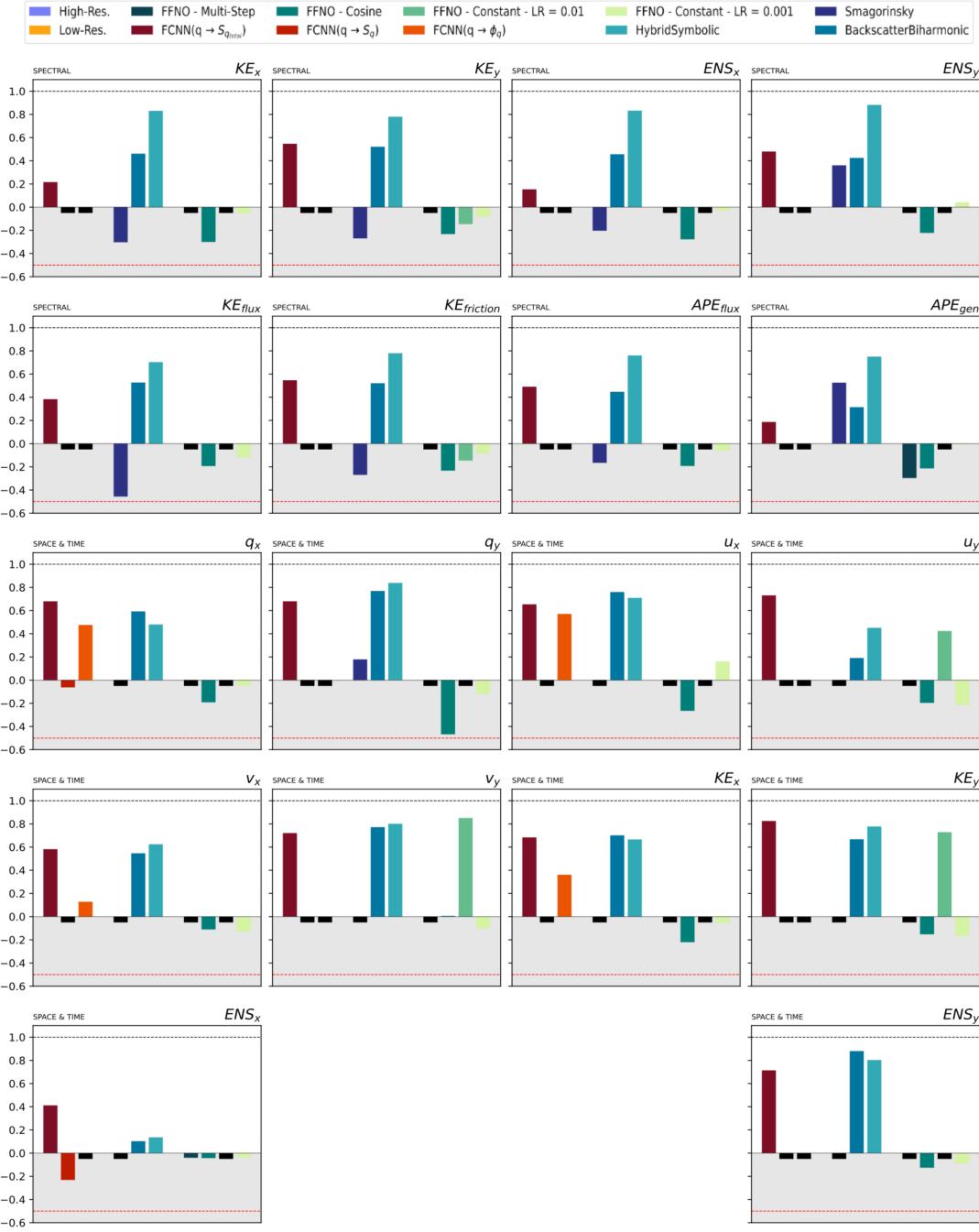


Figure 78: |Online - Phase 5 - Similarities| This table provides a summary of the Earth mover's distance, reformulated as a similarity metric for various flow quantities represented in either spectral or spatiotemporal domains. A value approaching 1 indicates strong agreement between the distribution obtained from high-resolution simulations and the current observations. Negative values are considered unfavorable, and values lower than -0.5 are disregarded. The tested parameterizations comes from Tab.??, they are trained on **full 5000** and tested on **eddies online**. For comparison, the results of neural networks (Ross et al., 2023) and analytical parameterizations are also presented (Smagorinsky, 1963; Jansen and Held, 2014; Zanna and Bolton, 2020).



Figure 79: |Online - Phase 5 - Similarities| This table provides a summary of the Earth mover's distance, reformulated as a similarity metric for various flow quantities represented in either spectral or spatiotemporal domains. A value approaching 1 indicates strong agreement between the distribution obtained from high-resolution simulations and the current observations. Negative values are considered unfavorable, and values lower than -0.5 are disregarded. The tested parameterizations comes from Tab.??, they are trained on **full 5000** and tested on **jets online**. For comparison, the results of neural networks (Ross et al., 2023) and analytical parameterizations are also presented (Smagorinsky, 1963; Jansen and Held, 2014; Zanna and Bolton, 2020).

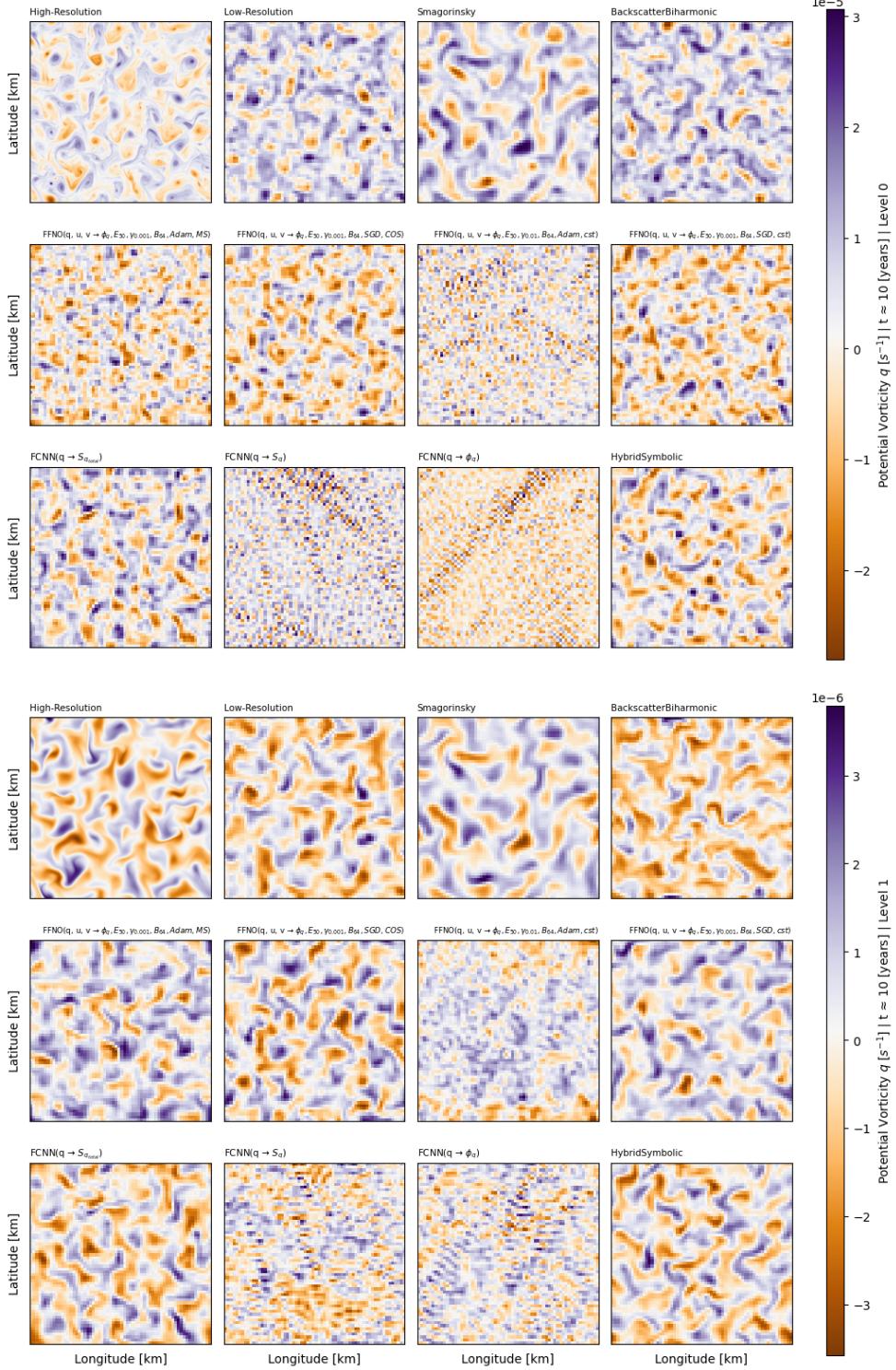


Figure 80: |Online - Phase 5 - Potential vorticity| Visualization of potential vorticity is presented for both upper (first three rows) and lower (last three rows) layers across different simulation types, indicated at the top of each image. Each image represents the  $q$  value spanning the entire computational domain after 10 years of simulations. The objective is to emphasize and visualize simulations that lose their physical relevance, becoming mere pixel grids, and to illustrate the divergence from the high-resolution simulation. Furthermore, the evaluated parameterizations are detailed in Tab.??, they were trained using **full 5000** and assessed against **eddies online**.

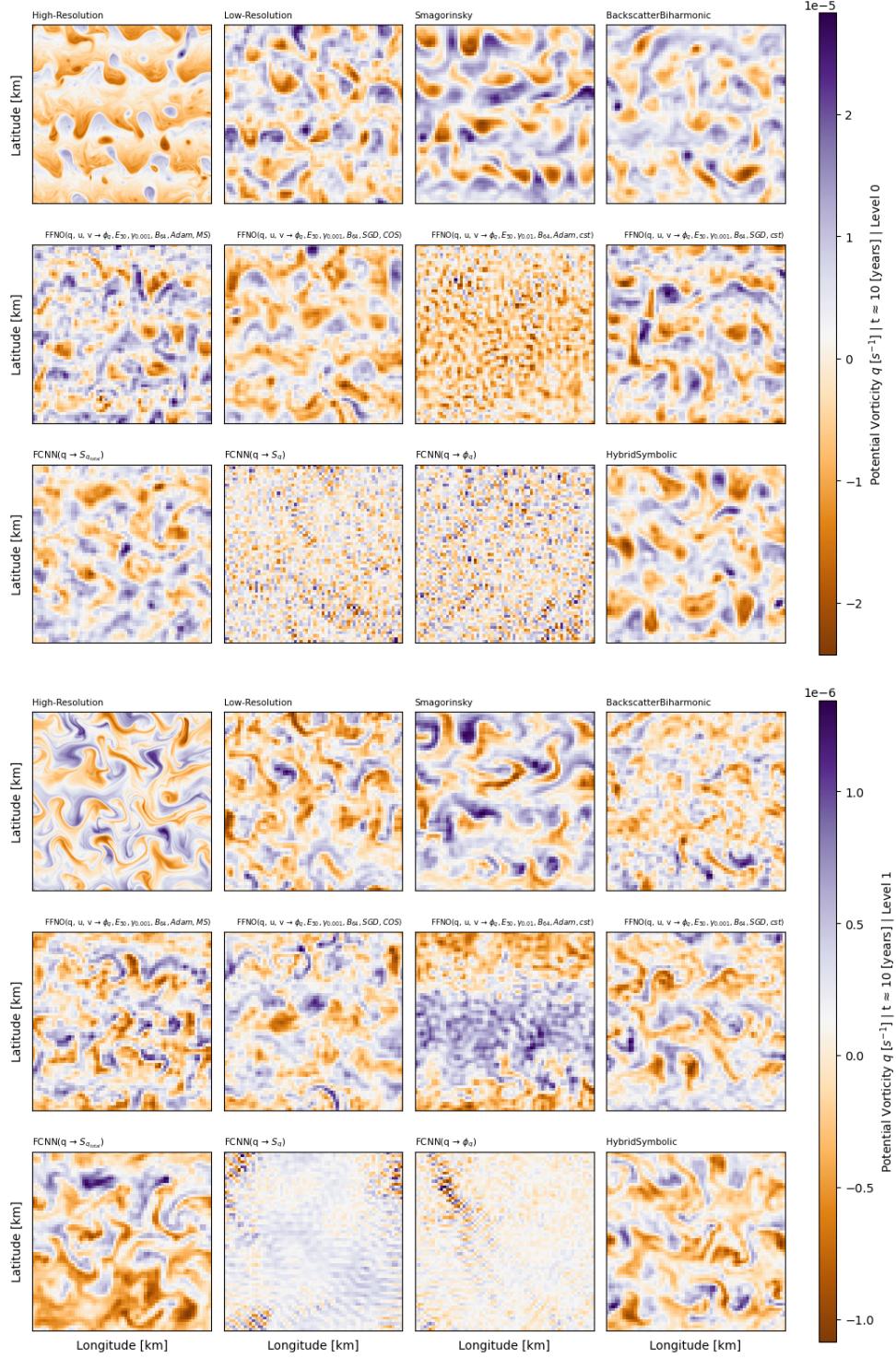


Figure 81: |Online - Phase 5 - Potential vorticity| Visualization of potential vorticity is presented for both upper (first three rows) and lower (last three rows) layers across different simulation types, indicated at the top of each image. Each image represents the  $q$  value spanning the entire computational domain after 10 years of simulations. The objective is to emphasize and visualize simulations that lose their physical relevance, becoming mere pixel grids, and to illustrate the divergence from the high-resolution simulation. Furthermore, the evaluated parameterizations are detailed in Tab.??, they were trained using full 5000 and assessed against jets online.

# PHASE VI

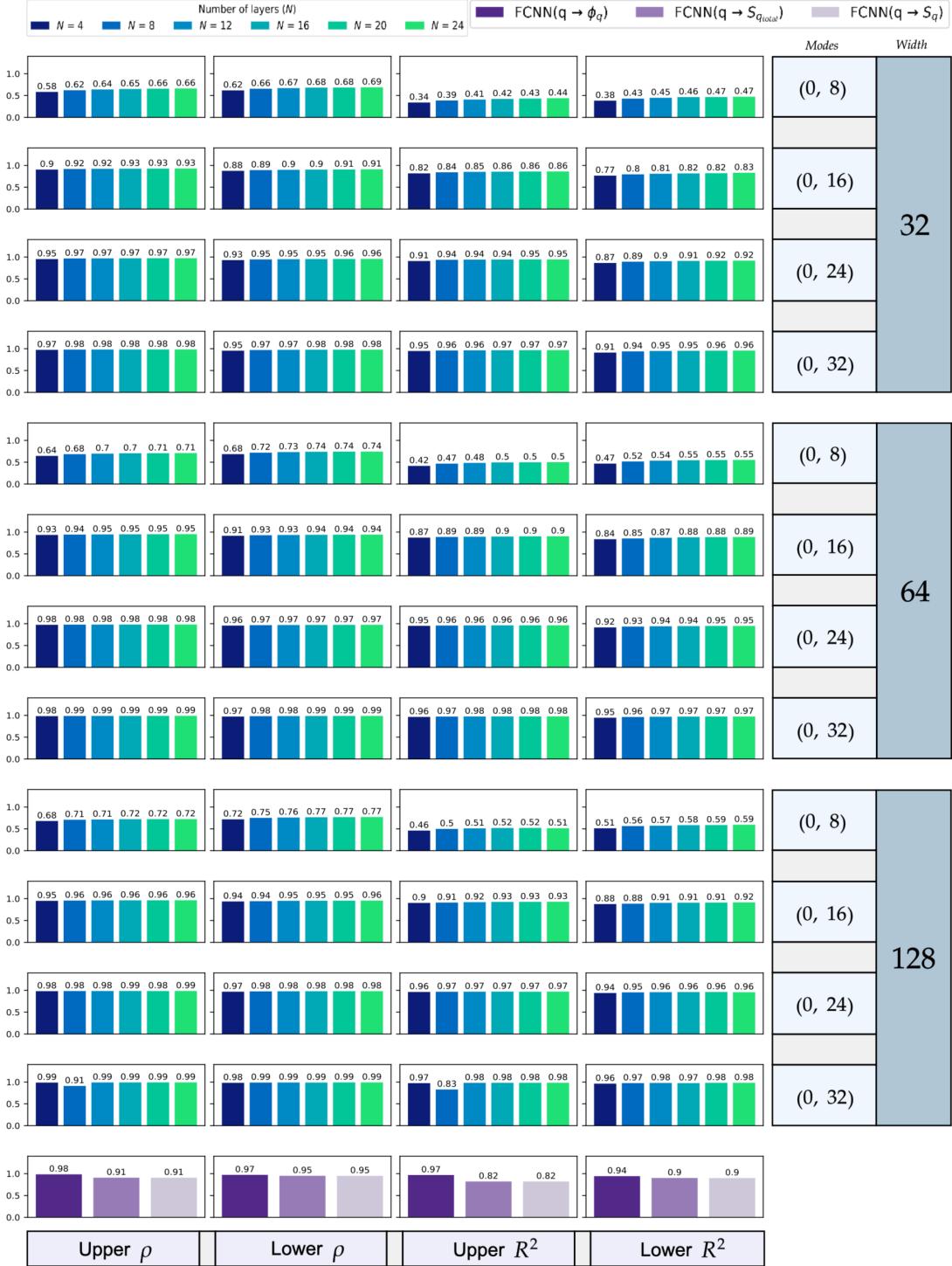


Figure 82: |Offline - Phase 6 - Part 1| This table summarize offline results, including correlations (columns 1 and 2) and mean-squared errors (columns 3 and 4), for parameterizations trained on **full dataset 5000**, evaluated on dataset **eddies offline** and predicting subgrid flux  $\Phi_q$  (see Eq. 23). On the right-hand side, the width value, retained Fourier modes, and total number of layers for the FFNO are provided in the legend. The bottom row presents results obtained using the three FCNN parameterizations introduced in Ross et al., 2023.

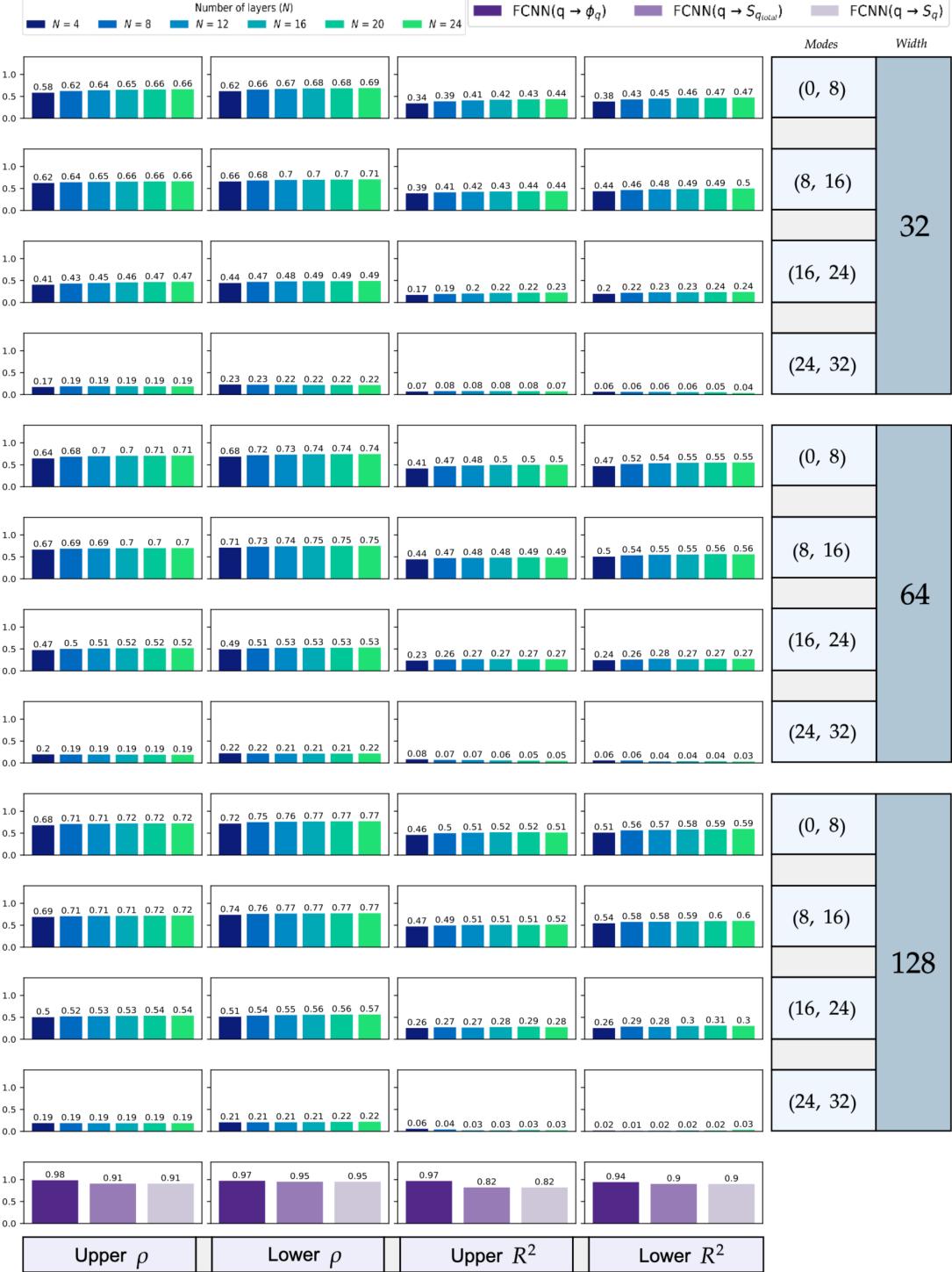


Figure 83: | Offline - Phase 6 - Part 2 | This table summarizes offline results, including correlations (columns 1 and 2) and mean-squared errors (columns 3 and 4), for parameterizations trained on **full dataset 5000**, evaluated on dataset **eddies offline** and predicting subgrid flux  $\Phi_q$  (see Eq. 23). On the right-hand side, the width value, retained Fourier modes, and total number of layers for the FFNO are provided in the legend. The bottom row presents results obtained using the three FCNN parameterizations introduced in Ross et al., 2023.

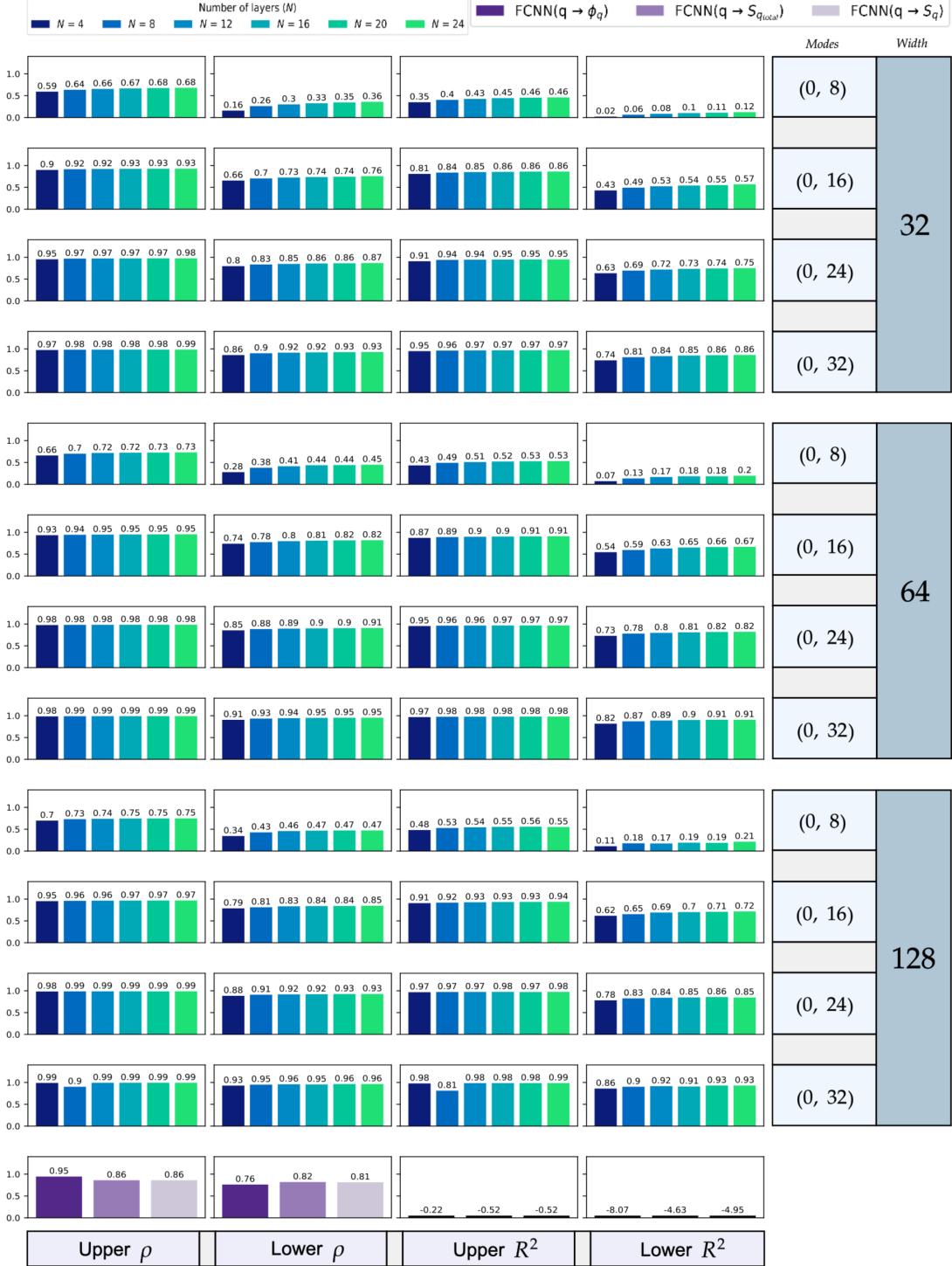


Figure 84: |Offline - Phase 6 - Part 1| This table summarizes offline results, including correlations (columns 1 and 2) and mean-squared errors (columns 3 and 4), for parameterizations trained on **full dataset 5000**, evaluated on dataset **jets offline** and predicting subgrid flux  $\Phi_q$  (see Eq. 23). On the right-hand side, the width value, retained Fourier modes, and total number of layers for the FFNO are provided in the legend. The bottom row presents results obtained using the three FCNN parameterizations introduced in Ross et al., 2023.

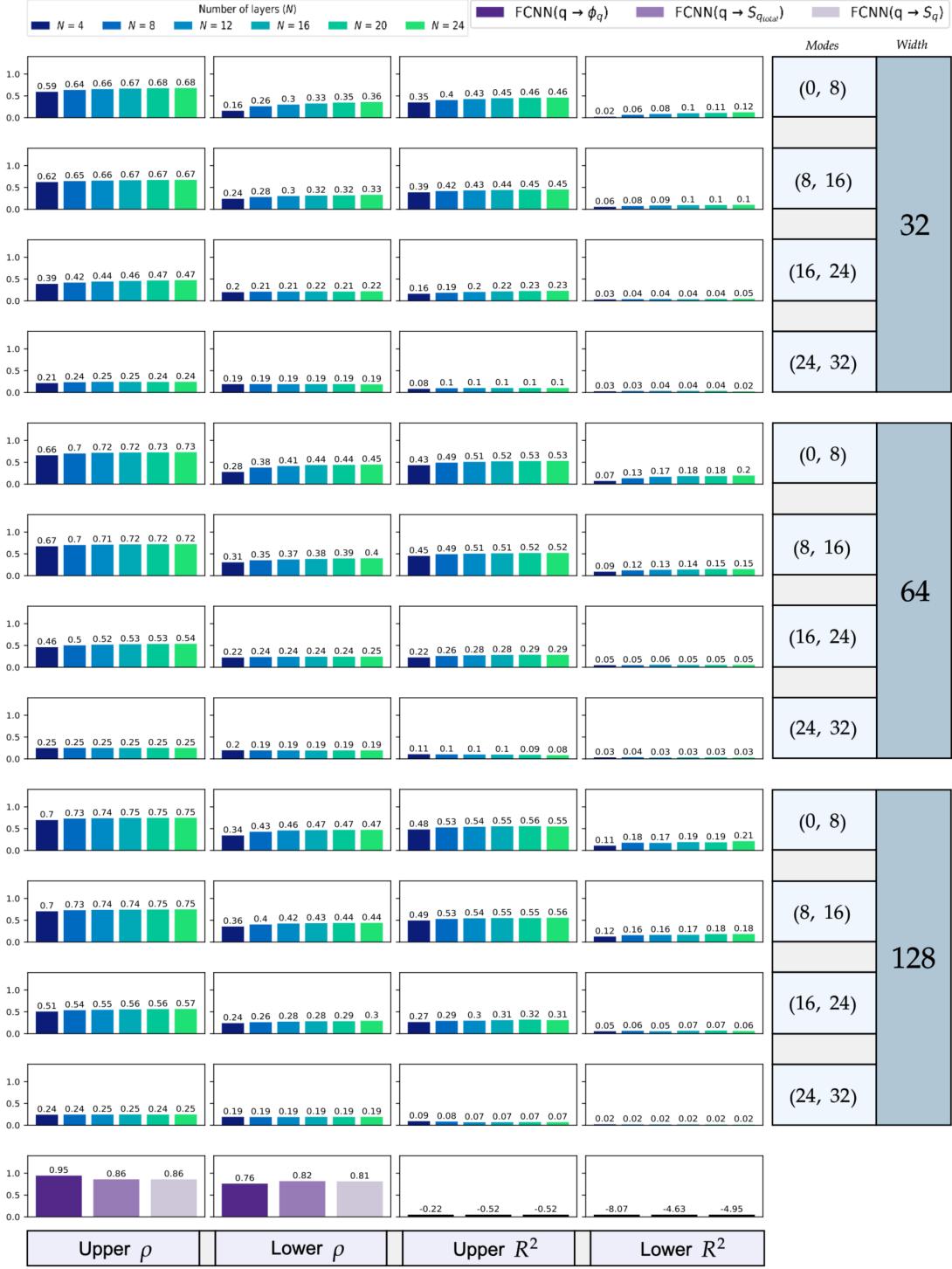


Figure 85: | Offline - Phase 6 - Part 2 | This table summarizes offline results, including correlations (columns 1 and 2) and mean-squared errors (columns 3 and 4), for parameterizations trained on **full dataset 5000**, evaluated on dataset **jets offline** and predicting subgrid flux  $\Phi_q$  (see Eq. 23). On the right-hand side, the width value, retained Fourier modes, and total number of layers for the FFNO are provided in the legend. The bottom row presents results obtained using the three FCNN parameterizations introduced in Ross et al., 2023.

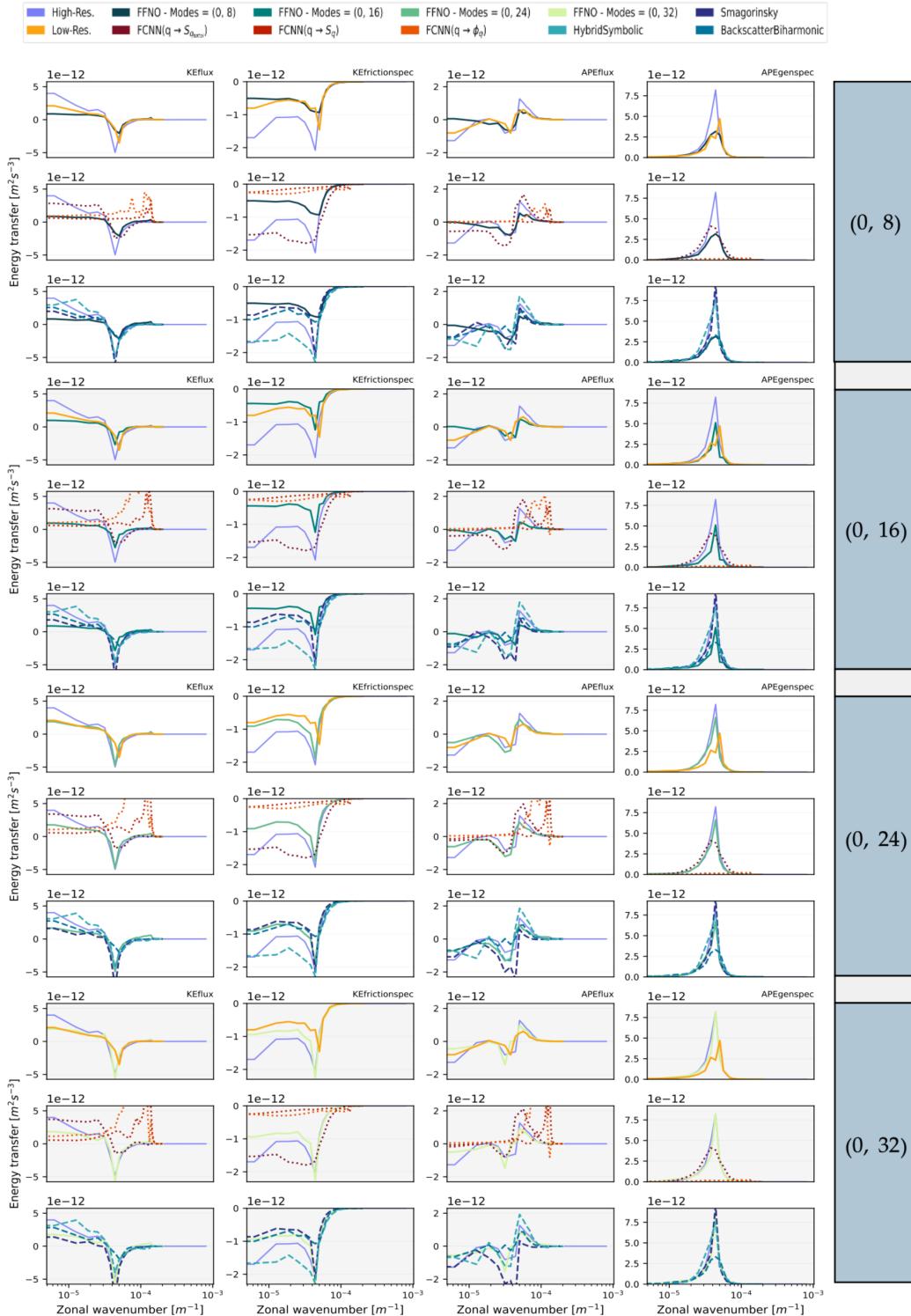


Figure 86: |Online - Phase 6 - Energy budget| This table displays energy spectra for KEflux, KEfrictionspec, APEflux, and APEgenspec using parameterizations of Tab.??, these were trained on full 5000 and tested on **eddies online**. Each parameterization spectrum is compared against high-resolution and various low-resolution simulations, including neural networks from Ross et al., 2023 and analytical parameterizations from Smagorinsky, 1963; Jansen and Held, 2014; Zanna and Bolton, 2020.

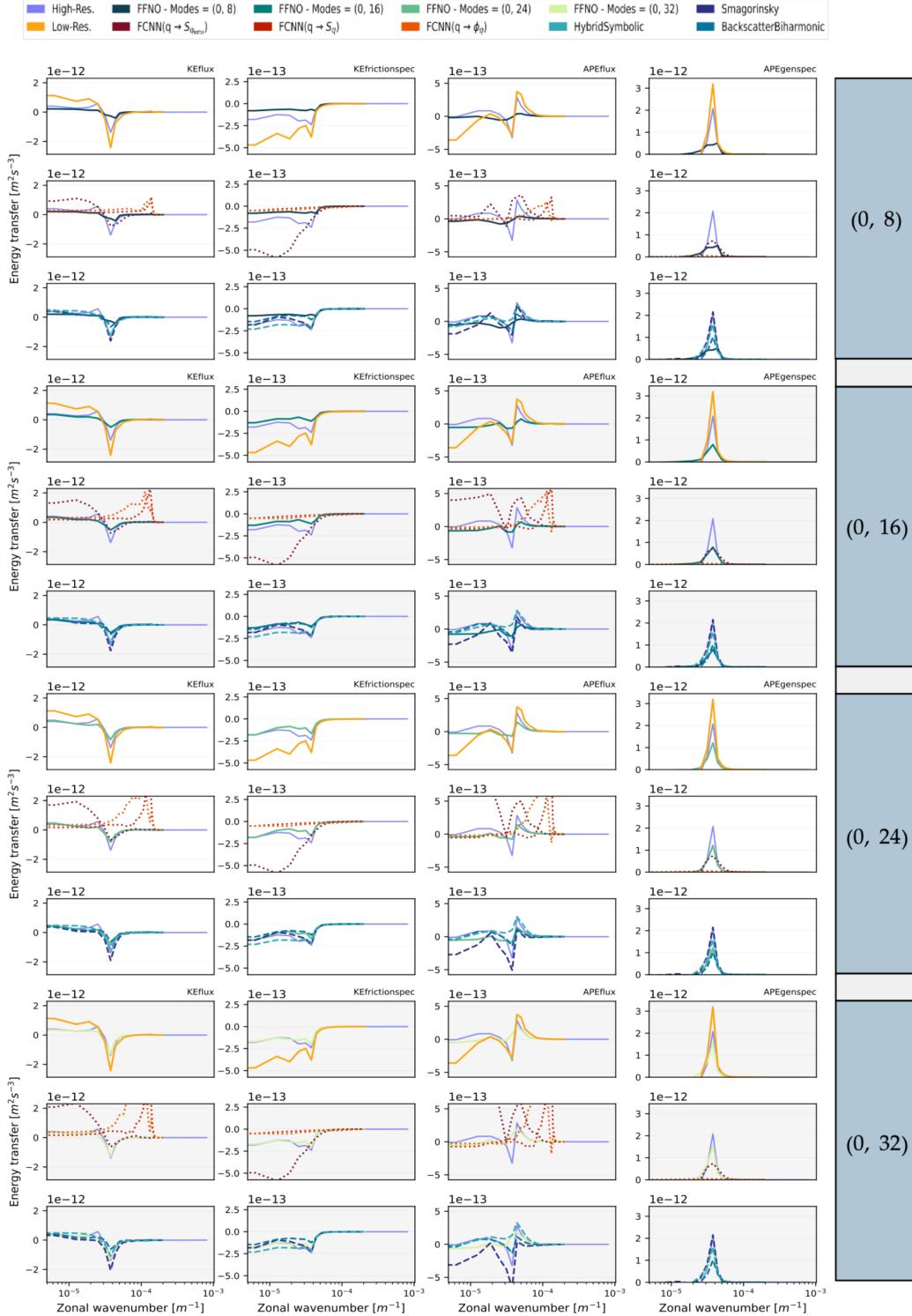


Figure 87: |Online - Phase 6 - Energy budget| This table displays energy spectra for **KEflux**, **KEfrictionspec**, **APEflux**, and **APEgenspec** using parameterizations of Tab.??, these were trained on **full 5000** and tested on **jets online**. Each parameterization spectrum is compared against high-resolution and various low-resolution simulations, including neural networks from Ross et al., 2023 and analytical parameterizations from Smagorinsky, 1963; Jansen and Held, 2014; Zanna and Bolton, 2020.

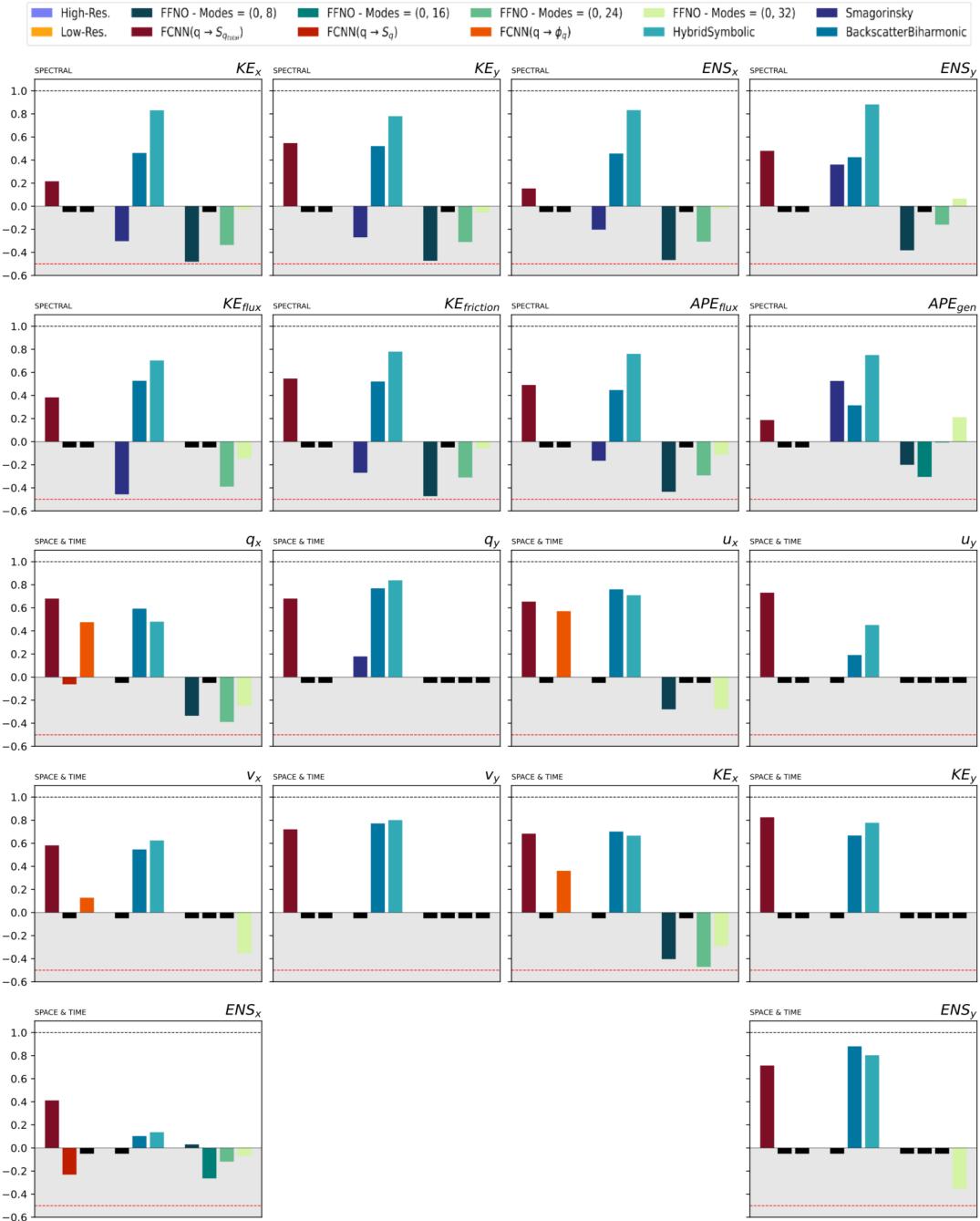


Figure 88: |Online - Phase 6 - Similarities| This table provides a summary of the Earth mover's distance, reformulated as a similarity metric for various flow quantities represented in either spectral or spatiotemporal domains. A value approaching 1 indicates strong agreement between the distribution obtained from high-resolution simulations and the current observations. Negative values are considered unfavorable, and values lower than -0.5 are disregarded. The tested parameterizations comes from Tab.??, they are trained on **full 5000** and tested on **eddies online**. For comparison, the results of neural networks (Ross et al., 2023) and analytical parameterizations are also presented (Smagorinsky, 1963; Jansen and Held, 2014; Zanna and Bolton, 2020).

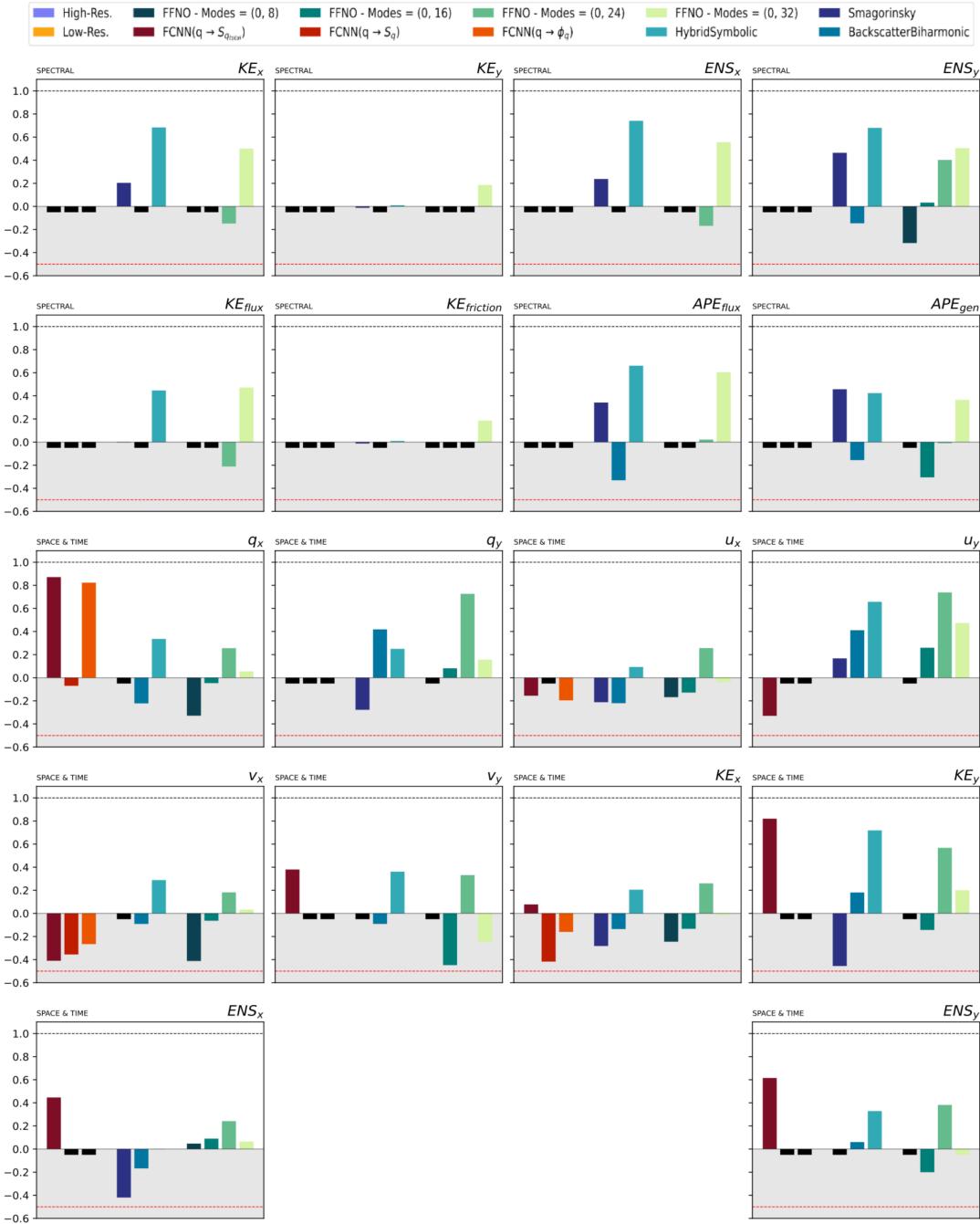


Figure 89: |Online - Phase 6 - Similarities| This table provides a summary of the Earth mover's distance, reformulated as a similarity metric for various flow quantities represented in either spectral or spatiotemporal domains. A value approaching 1 indicates strong agreement between the distribution obtained from high-resolution simulations and the current observations. Negative values are considered unfavorable, and values lower than -0.5 are disregarded. The tested parameterizations comes from Tab.??, they are trained on **full 5000** and tested on **jets online**. For comparison, the results of neural networks (Ross et al., 2023) and analytical parameterizations are also presented (Smagorinsky, 1963; Jansen and Held, 2014; Zanna and Bolton, 2020).

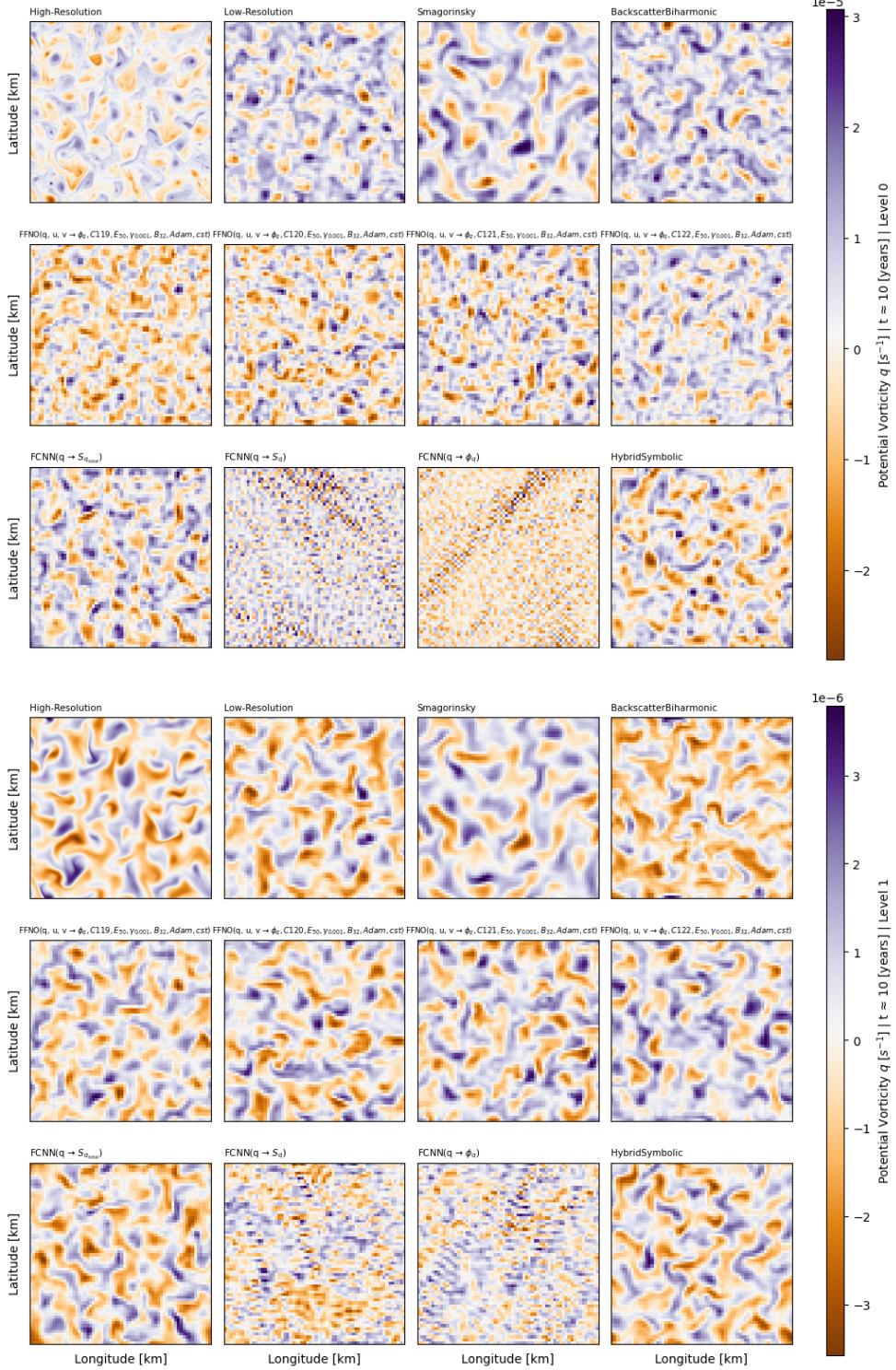


Figure 90: |Online - Phase 6 - Potential vorticity| Visualization of potential vorticity is presented for both upper (first three rows) and lower (last three rows) layers across different simulation types, indicated at the top of each image. Each image represents the  $q$  value spanning the entire computational domain after 10 years of simulations. The objective is to emphasize and visualize simulations that lose their physical relevance, becoming mere pixel grids, and to illustrate the divergence from the high-resolution simulation. Furthermore, the evaluated parameterizations are detailed in Tab.??, they were trained using **full 5000** and assessed against **eddies online**.

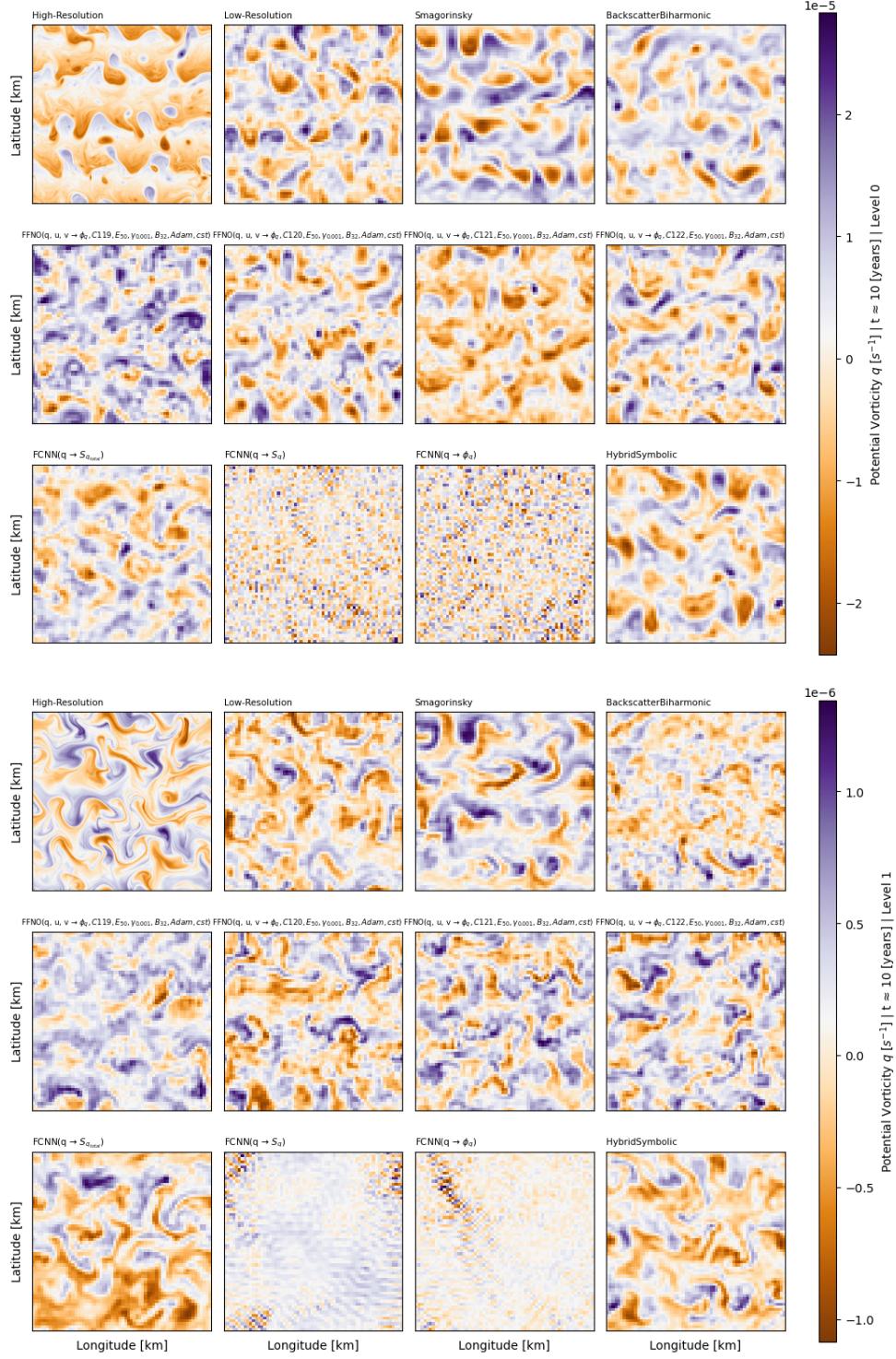


Figure 91: | Online - Phase 6 - Potential vorticity | Visualization of potential vorticity is presented for both upper (first three rows) and lower (last three rows) layers across different simulation types, indicated at the top of each image. Each image represents the  $q$  value spanning the entire computational domain after 10 years of simulations. The objective is to emphasize and visualize simulations that lose their physical relevance, becoming mere pixel grids, and to illustrate the divergence from the high-resolution simulation. Furthermore, the evaluated parameterizations are detailed in Tab.??, they were trained using full 5000 and assessed against jets online.