Table 1 contains the results of the Informer model on 13 experimental datasets across 10 evaluation metrics: mean MASE, median MASE, mean sMAPE, median sMAPE, mean msMAPE, median msMAPE, mean MAE, median MAE, mean RMSE and median RMSE.

For these experiments, we have selected the datasets with equal-length series. The "features" parameter of the informer model is set to "M" and thus, the model is executed as a Global Forecasting Model (GFM) across series.

For a fair comparison of benchmarking models, we use the same number of lagged values and forecast horizons for a given dataset during model evaluation. Thus, for each dataset, the informer model is also executed with the same number of lagged values and the forecast horizon that are explained in our paper. In particular, the "seq_len" and "label_len" parameters of the informer model are set to the number of lagged values. Both "itr" and "batch_size" parameters are set to 1. The number of time series in the dataset is used as the value of "enc_in", "dec_in" and "c_out" parameters. Furthermore, the default train-validation-test proportion used in the implementation, 7:1:2, is not applicable for Traffic Weekly and Hospital datasets due to their series lengths, considered number of lagged values, and forecast horizons. Thus, for Traffic Weekly and Hospital datasets, the above proportion is respectively set to 8:1:1 and 6:2:2.

Table 1: Informer Results

Dataset	Mean MASE	Median MASE	Mean sMAPE	Median sMAPE	Mean msMAPE	Median msMAPE	Mean MAE	Median MAE	Mean RMSE	Median RMSE
NN5 Daily	0.933	0.870	23.05	21.38	22.96	21.30	4.07	3.61	5.52	4.89
NN5 Weekly	1.079	0.946	14.85	12.49	14.84	12.49	19.45	16.24	23.03	19.68
Solar 10 mins	1.614	1.604	170.02	170.06	157.22	158.60	3.67	3.25	6.41	5.88
Solar Weekly	2.408	2.387	42.14	42.84	42.14	42.84	2360.71	2155.42	2623.95	2410.24
Electricity Hourly	2.682	2.424	28.87	21.90	28.83	21.90	441.77	127.28	629.88	172.59
Electricity Weekly	1.444	1.313	16.03	11.68	16.03	11.68	48773.67	11815.75	54022.60	13241.67
FRED-MD	17.839	14.329	72.32	66.10	70.59	65.50	32700.73	79.77	32867.61	82.46
Traffic Hourly	1.439	1.233	48.60	36.05	7.34	6.10	0.02	0.02	0.04	0.03
Traffic Weekly	1.323	1.173	15.66	11.79	15.55	11.72	1.42	1.09	1.76	1.37
Hospital	1.221	1.055	26.92	24.27	26.87	24.24	38.32	10.08	44.25	12.08
Sunspot	0.504	0.504	198.33	198.33	197.22	197.22	19.43	19.43	20.31	20.31
Saugeen	1.896	1.896	49.43	49.43	49.38	49.38	28.59	28.59	44.42	44.42
US Births	2.220	2.220	5.94	5.94	5.94	5.94	609.43	609.43	734.44	734.44

Most of the datasets in our archive contain unequal-length series where each series has a different start and end date. The informer model implementation takes a dataframe as an input where the first column contains the time variable and the remaining columns are corresponding with each series. Including unequal-length series in a dataframe incurs a large number of missing values and it prevents us from executing the model as a GFM for the corresponding datasets. It is possible to execute the informer as a local model for these datasets, where one model is built per each series. For that, the "features" parameter of the informer model can be set to "S". However, building a deep learning model individually on

each series is highly time consuming and we expect to publish the corresponding results in the future.

Furthermore, the intermittent datasets in our archive such as Carparts, Rideshare, Web Traffic, Covid Deaths and Temperature Rain are not considered for the experiments.