

Monash Time-Series Forecasting Archive Replication

Fernando Urbano* Aben Carrington[†] Shrey Jain[‡]
Mukund Maheshwari[§]

March 10, 2024

Abstract

In this project we attempt to replicate results from a 2021 paper on the motivation and creation of the Monash Time Series Forecasting Archive, a project spearheaded by a group time series researchers from Monash University and the University of Sydney.

*fernandourbano@uchicago.edu

[†]acarrington@uchicago.edu

[‡]shreyjain@uchicago.edu

[§]mukundmaheshwari@uchicago.edu

1 Paper Importance

The Monash Archive is an essential asset to time series researchers as it provides a comprehensive benchmark time series forecasting archive to evaluate the performance of new global and multivariate forecasting algorithms. Specifically, as researchers branch further and further into the machine learning space, the Monash Archive allows them to test the generalized performance of their models against well-tested benchmark models, which is beneficial in addressing the questions of model overfitting and performance.

The archive contains datasets spanning multiple domains (industries) as well as 13 forecasting models, 6 of which are canonical univariate models, and 7 of which are global models that have shown positive results in recent years. In the following sections we provide a brief description of each of the datasets used, as well as an overview of the important aspects of the models used.

2 Data Sources

The archive repository contains 25 datasets where each data point is an individual time series, with most being of variable length. The datasets indicated as multivariate in Table 1 are constrained to have time series that are all of the same length so that multivariate forecasting models may be run on them without error. Additionally there are 5 datasets which contain a single very long time series. From these 30 primary datasets the authors have created 58 total datasets, where some are split according to differing model frequencies, and datasets with missing values are split into two datasets: one with and one without the missing values. It should be noted that global univariate and local univariate can be applied to all datasets in the repository.

We now go into further depth concerning each of the primary datasets:

2.1 Collections of multiple time series

This section describes the benchmark datasets that have a sufficient number of series from a particular frequency. The datasets may contain different categories in terms of domain and frequency.

2.1.1 M1 dataset

This dataset is from the M1 forecasting competition held in the year 1982. It contains 1001 time series with 3 different frequencies (monthly, quarterly, yearly) and is one of few belonging to multiple domains.

2.1.2 M3 dataset

This dataset is from the M3 forecasting competition held in the year 2000. It contains 3003 time series with 4 different frequencies (monthly, quarterly, yearly, other) and is one of few belonging to multiple domains.

2.1.3 M4 dataset

This dataset is from the M4 forecasting competition held in the year 2020. It contains 100,000 time series with 6 different frequencies (hourly, daily, weekly, monthly, quarterly, yearly) and is one of few belonging to multiple domains.

2.1.4 Tourism dataset

This dataset originates from a kaggle competition. It contains 1311 tourism related time series with 3 different frequencies (monthly, quarterly, yearly).

2.1.5 NN5 dataset

This dataset is from the NN5 neural forecasting competition held in the year 2008. It contains 111 daily time series of cash withdrawals from ATMs in the U.K. The original dataset contained missing values that were filled in by the authors of the paper using a median substitution method.

2.1.6 CIF 2016 dataset

This dataset is from the Computational Intelligence in Forecasting (CIF) competition held in 2016. It contains 72 monthly time series, where 24 originate from the banking sector and the other 48 are artificially generated. There are two datasets corresponding to different forecast horizons: 6 and 12.

2.1.7 Kaggle web traffic dataset

This dataset is from the Kaggle web traffic forecasting competition held in the year 2017. It contains 145063 daily time series representing the number of hits for a set of

pages on Wikipedia from 01/07/2015 to 10/09/2017. The authors also include their own aggregated version of weekly time series.

2.1.8 Solar dataset

This dataset corresponds to the solar power production in the state of Alabama throughout the year 2006. It contains 137 time series representing the amount of solar power produced every 10 minutes. The authors also include their own aggregated version of weekly time series.

2.1.9 Electricity dataset

This dataset corresponds to the amount of electricity consumed per hour by 321 clients, collected from 2012 to 2014 by UC Irvine. The authors also include their own aggregated version of weekly time series.

2.1.10 London smart meters dataset

This dataset corresponds to the energy consumption readings of London households in kWh from November 2011 to February 2014. It contains 5560 half-hourly time series. Two versions are included: one with missing values, and one where the missing values are filled in using the last observation carried forward (LOCF) method.

2.1.11 Australian electricity demand dataset

This dataset corresponds to the electricity demand of 5 Australian states: Victoria, New South Wales, Queensland, Tasmania and South Australia. It contains 5 half-hourly time series.

2.1.12 Wind farms dataset

This dataset contains very long minutely time series representing the wind power production of 339 wind farms in Australia. It is curated by the authors and is not available elsewhere. Two versions are included: one with missing values, and one where the missing values have been set to zero.

2.1.13 Car parts dataset

This dataset contains 2674 intermittent monthly time series showing car parts sales from January 1998 to March 2002. Two versions are included: one with missing

values, and one where the missing values have been set to zero.

2.1.14 Dominick dataset

This dataset corresponds to the profit of individual stock keeping units (SKUs) from a retailer collected from the online platform of the University of Chicago Booth School of Business Kilts Center. It contains 115704 weekly time series.

2.1.15 FRED-MD dataset

This dataset was extracted from the FRED-MD database and corresponds to a set of macro-economic indicators from the Federal Reserve Bank. It contains 107 monthly time series starting from 01/01/1959.

2.1.16 Bitcoin dataset

This dataset shows the potential factors influencing bitcoin price (such as transaction values and hash rate). It contains 18 daily time series, 2 of which show the public opinion of bitcoins in the form of tweets and google searches mentioning the keyword, bitcoin. It is curated by the authors and is not available elsewhere. Two versions are included: one with missing values, and one where the missing values are filled in using the LOCF method.

2.1.17 San Francisco Traffic dataset

This dataset corresponds to the road occupancy rates on San Francisco Bay area freeways. It contains 862 hourly time series taken from 2015 to 2016. The authors also include their own aggregated version of weekly time series.

2.1.18 Melbourne pedestrian counts dataset

This dataset contains hourly time series of pedestrian counts captured from 66 sensors in Melbourne from May 2009 to April 30, 2020.

2.1.19 Rideshare dataset

This dataset corresponds to attributes related to Uber and Lyft rideshare services (such as price and distance) for different locations in New York from 26/11/2018 to 18/12/2018. It contains 2304 hourly time series. Two versions are included: one with missing values, and one where the missing values have been set to zero.

2.1.20 Vehicle trips dataset

This dataset corresponds to the number of trips and vehicles belonging to a set of for-hire vehicle (FHV) companies in 2015, extracted from fivethirtyeight. It contains 329 daily time series. Two versions are included: one with missing values, and one where the missing values are filled in using the LOCF method.

2.1.21 Hospital dataset

This dataset corresponds to e patient counts related to medical products from January 2000 to December 2006. It contains 767 monthly time series.

2.1.22 COVID deaths dataset

This dataset represents the total COVID-19 deaths in a set of countries and states from 22/01/2020 to 20/08/2020, extracted from the Johns Hopkins repository. It contains 266 daily time series.

2.1.23 KDD cup 2018 dataset

This dataset originates from a 2018 competition. It contains 270 long hourly time series representing the air quality levels in 59 stations from 2 cities, Beijing (35 stations) and London (24 stations) from 01/01/2017 to 31/03/2018. It represents the air quality across multiple measurements.

2.1.24 Weather dataset

This dataset contains 3010 daily time series of four weather variables: rain, minimum temperature, maximum temperature, and solar radiation, measured at weather stations in Australia.

2.1.25 Temperature rain dataset

This dataset corresponds to the temperature/rainfall observations and forecasts, gathered by the Australian Bureau of Meteorology for 422 weather stations across Australia, between 02/05/2015 and 26/04/2017. It contains 32072 daily time series. Two versions are included: one with missing values, and one where the missing values have been set to zero.

2.2 Single long time series datasets

This section describes the benchmark datasets which have single time series with a large amount of data points.

2.2.1 Sunspot dataset

This dataset contains the single daily time series representing the sunspot numbers from 08/01/1818 to 31/05/2020. Two versions are included: one with missing values, and one where the missing values are filled in using the LOCF method.

2.2.2 Saugeen river flow dataset

This dataset contains a single very long time series representing the daily mean flow of the Saugeen River at Walkerton in cubic meters per second from 01/01/1915 to 31/12/1979. The length of the time series is 23,741.

2.2.3 US Births dataset

This dataset contains a single very long daily time series representing the number of births in the US from 01/01/1969 to 31/12/1988. The length of the time series is 7,305.

2.2.4 Solar power dataset

This dataset contains a single very long time series representing the solar power production of an Australian wind farm recorded every 4 seconds starting from 01/08/2019. The length of the time series is 7,397,222.

2.2.5 Wind power dataset

This dataset contains a single very long time series representing the wind power production of an Australian wind farm recorded every 4 seconds starting from 01/08/2019. The length of the time series is 7,397,147.

Table 1: Datasets in the current time series forecasting archive

	Dataset	Domain	No: of Series	Min. Length	Max. Length	No: of Freq	Missing	Competition	Multivariate
0	M1	Multiple	1023	18	150	3	No	Yes	No
1	M3	Multiple	3003	20	144	4	No	Yes	No
2	M4	Multiple	100000	19	9933	6	No	Yes	No
3	Tourism	Tourism	1311	11	333	3	No	Yes	No
4	CIF 2016	Banking	72	28	120	1	No	Yes	No
5	London Smart Meters	Energy	5560	288	39648	1	Yes	No	No
6	Aus. Electricity Demand	Energy	5	230736	232272	1	No	No	No
7	Wind Farms	Energy	339	6345	527040	1	Yes	No	No
8	Dominick	Sales	115704	28	393	1	No	No	No
9	Bitcoin	Economic	18	4581	4581	1	Yes	No	No
10	Pedestrian Counts	Transport	66	576	96424	1	No	No	No
11	Vehicle Trips	Transport	329	70	243	1	Yes	No	No
12	KDD Cup 2018	Transport	270	9504	10920	1	Yes	Yes	No
13	Weather	Weather	3010	1332	65981	1	No	No	No
14	NN5	Banking	111	791	791	2	Yes	Yes	Yes
15	Web Traffic	Web	145063	803	803	1	Yes	Yes	Yes
16	Solar	Energy	137	52560	52560	2	No	No	Yes
17	Electricity	Energy	321	26304	26304	2	No	No	Yes
18	Car Parts	Sales	2674	51	51	1	Yes	No	Yes
19	FRED-MD	Economics	107	728	728	1	No	No	Yes
20	San Francisco Traffic	Transport	862	17544	17544	2	No	No	Yes
21	Rideshare	Transport	2304	541	541	1	Yes	No	Yes
22	Hospital	Health	767	84	84	1	No	No	Yes
23	COVID Deaths	Nature	266	212	212	1	No	No	Yes
24	Temperature Rain	Nature	32072	725	725	1	Yes	No	Yes
25	Sunspot	Nature	1	73924	73924	1	Yes	No	No
26	Saugeen River Flow	Nature	1	23741	23741	1	No	No	No
27	US Births	Nature	1	7305	7305	1	No	No	No
28	Solar Power	Energy	1	7397222	7397222	1	No	No	No
29	Wind Power	Energy	1	7397147	7397147	1	No	No	No

3 Models/Evaluation

3.1 Models

As previously mentioned this project uses 6 traditional univariate models, and 7 global models, covering a representative set of state-of-the-art forecasting models from statistical, machine learning, and deep learning domains, for a total of 13 models.

The 6 traditional models used are Exponential Smoothing (ETS), Auto-Regressive Integrated Moving Average (ARIMA), Simple Exponential Smoothing (SES), Theta, Trigonometric Box-Cox ARMA Trend Seasonal (TBATS), and Dynamic Harmonic Regression ARIMA (DHR-ARIMA). As there is extensive literature on each of these models we forego any description here.

Table 2: Mean MASE results. The best model across each dataset is highlighted in boldface.

	Dataset	SES	Theta	ETS	(DHR-) ARIMA	PR	Cat Boost	ARIMA	TBATS
0	M1 Yearly	4.938	4.191	3.771	-	4.588	4.333	4.479	3.499
1	M1 Quarterly	1.929	1.702	1.658	-	1.892	2.040	1.787	1.694
2	M1 Monthly	1.379	1.091	1.074	-	1.123	1.220	1.165	1.118
3	M3 Quarterly	1.417	1.117	1.170	-	1.248	1.449	1.240	1.256
4	M3 Monthly	1.091	0.864	0.865	-	1.010	1.076	0.873	0.861
5	M4 Yearly	-	-	-	-	-	-	-	3.437
6	M4 Quarterly	-	-	-	-	-	-	-	1.186
7	M4 Weekly	-	-	-	-	-	-	-	0.505
8	Tourism Yearly	3.253	3.015	3.395	-	3.516	3.619	3.775	3.685
9	Tourism Quarterly	3.210	1.661	1.592	-	1.643	1.821	1.776	1.835
10	Tourism Monthly	3.306	1.649	1.526	-	1.678	1.712	1.587	1.751
11	Vehicle Trips	2.273	1.914	1.964	-	2.196	2.004	2.051	1.856
12	NN5 Daily	-	-	-	-	-	0.970	-	0.858
13	NN5 Weekly	0.903	0.885	0.911	0.887	0.854	0.854	-	0.872
14	Solar Weekly	1.215	1.224	1.134	0.848	1.053	1.477	-	0.916
15	Electricity Weekly	1.536	1.476	1.526	0.878	0.916	0.813	-	0.792
16	Traffic Weekly	1.116	1.121	1.125	1.191	1.122	1.122	-	1.148
17	Rideshare	4.040	4.872	-	-	-	-	-	4.384
18	Hospital	0.813	0.761	0.765	-	0.782	0.796	0.788	0.768
19	Sunspot	0.128	0.128	0.128	-	0.099	0.073	0.067	0.064
20	Bitcoin	5.289	5.223	4.538	-	4.616	5.653	5.498	4.602
21	CIF 2016	1.291	0.997	0.841	-	1.019	1.200	0.927	0.861
22	COVID Deaths	7.776	7.793	5.326	-	8.731	8.092	6.104	5.719
23	Car Parts	0.897	0.914	0.925	-	0.755	0.853	0.927	1.002
24	Fred Md	0.617	0.698	0.468	-	8.827	0.988	0.532	0.502
25	M3 Yearly	3.167	2.774	2.860	-	3.223	3.711	3.417	3.127
26	Saugeen River Flow	1.426	1.425	2.036	-	1.674	1.430	1.548	1.477
27	US Births	4.343	2.138	1.529	-	2.094	1.690	1.917	1.453

Table 3: Mean MAE results. The best model across each dataset is highlighted in boldface.

	Dataset	SES	Theta	ETS	(DHR-) ARIMA	PR	Cat Boost	ARIMA	TBATS
0	M1 Yearly	171.4K	152.8K	146.1K	-	134.2K	249.4K	145.6K	103.0K
1	M1 Quarterly	2.2K	2.0K	2.1K	-	1.6K	1.9K	2.2K	2.3K
2	M1 Monthly	2.3K	2.2K	1.9K	-	2.1K	2.1K	2.1K	2.2K
3	M3 Quarterly	572	486	513	-	519	594	559	562
4	M3 Monthly	743	624	626	-	693	736	655	631
5	M4 Yearly	-	-	-	-	-	-	-	960
6	M4 Quarterly	-	-	-	-	-	-	-	570
7	M4 Weekly	-	-	-	-	-	-	-	297
8	Tourism Yearly	95.6K	90.7K	94.8K	-	82.7K	81.3K	95.0K	94.1K
9	Tourism Quarterly	15.0K	7.7K	8.9K	-	9.1K	10.1K	10.4K	10.0K
10	Tourism Monthly	5.3K	2.1K	2.0K	-	2.2K	2.5K	2.5K	2.9K
11	Vehicle Trips	29.980	23.299	21.258	-	27.243	22.732	23.456	21.045
12	NN5 Daily	-	-	-	-	-	4.200	-	3.701
13	NN5 Weekly	15.665	15.305	15.698	15.383	14.937	15.359	-	14.985
14	Solar Weekly	1.2K	1.2K	1.1K	840	1.0K	1.5K	-	909
15	Electricity Weekly	74.1K	74.1K	67.7K	28.5K	44.9K	34.7K	-	24.4K
16	Traffic Weekly	1.125	1.131	1.144	1.222	1.125	1.181	-	1.166
17	Rideshare	6.293	7.620	-	-	-	-	-	6.877
18	Hospital	21.761	18.539	17.966	-	19.237	19.114	19.742	17.429
19	Sunspot	4.933	4.933	4.933	-	3.833	2.800	2.567	2.467
20	Bitcoin	1773.4Qi	1773.4Qi	1103.6Qi	-	666.4Qa	1921.1Qi	1047.2Qi	990.4Qa
21	CIF 2016	581.9K	714.8K	642.4K	-	563.2K	688.0K	469.1K	855.6K
22	COVID Deaths	354	321	85.591	-	348	486	85.768	96.288
23	Car Parts	0.548	0.530	0.564	-	0.407	0.531	0.561	0.583
24	Fred Md	2.8K	3.5K	2.0K	-	8.9K	2.6K	3.0K	2.0K
25	M3 Yearly	1.0K	957	1.0K	-	1.0K	1.1K	1.4K	1.2K
26	Saugeen River Flow	21.497	21.486	30.693	-	25.241	21.562	23.338	22.262
27	US Births	1.2K	587	420	-	575	464	526	399

Table 4: Mean RMSE results. The best model across each dataset is highlighted in boldface.

	Dataset	SES	Theta	ETS	(DHR-) ARIMA	PR	Cat Boost	ARIMA	TBATS
0	M1 Yearly	193.8K	171.5K	167.7K	-	152.0K	269.0K	175.3K	116.9K
1	M1 Quarterly	2.5K	2.3K	2.4K	-	1.9K	2.2K	2.5K	2.7K
2	M1 Monthly	2.7K	2.6K	2.3K	-	2.5K	2.5K	2.5K	2.6K
3	M3 Quarterly	671	568	599	-	606	698	651	654
4	M3 Monthly	894	754	755	-	830	879	791	765
5	M4 Yearly	-	-	-	-	-	-	-	1.1K
6	M4 Quarterly	-	-	-	-	-	-	-	673
7	M4 Weekly	-	-	-	-	-	-	-	358
8	Tourism Yearly	106.7K	99.9K	104.7K	-	89.6K	89.6K	106.1K	105.8K
9	Tourism Quarterly	17.3K	9.3K	10.8K	-	11.7K	12.6K	12.5K	12.0K
10	Tourism Monthly	7.0K	2.7K	2.5K	-	2.7K	3.1K	3.1K	3.7K
11	Vehicle Trips	36.525	27.814	26.153	-	31.692	27.348	28.535	25.503
12	NN5 Daily	-	-	-	-	-	5.715	-	5.204
13	NN5 Weekly	18.825	18.647	18.816	18.550	18.615	18.711	-	18.528
14	Solar Weekly	1.3K	1.3K	1.3K	968	1.2K	1.7K	-	1.0K
15	Electricity Weekly	77.1K	76.9K	70.4K	32.6K	47.8K	37.6K	-	28.0K
16	Traffic Weekly	1.514	1.529	1.534	1.545	1.503	1.511	-	1.528
17	Rideshare	7.174	8.604	-	-	-	-	-	8.096
18	Hospital	26.551	22.592	22.023	-	23.479	23.287	23.837	21.281
19	Sunspot	4.946	4.946	4.946	-	3.954	3.141	2.938	2.595
20	Bitcoin	1963.7Qi	1963.7Qi	1223.5Qi	-	829.2Qa	2002.0Qi	1198.1Qi	1164.3Qi
21	CIF 2016	657.1K	804.7K	722.4K	-	648.9K	760.0K	526.4K	940.1K
22	COVID Deaths	403	370	102	-	394	617	100	113
23	Car Parts	0.784	0.782	0.802	-	0.729	0.794	0.811	0.837
24	Fred Md	3.1K	3.9K	2.3K	-	9.7K	2.8K	3.3K	2.3K
25	M3 Yearly	1.2K	1.1K	1.2K	-	1.2K	1.3K	1.7K	1.4K
26	Saugeen River Flow	39.794	39.787	50.392	-	47.703	39.306	45.536	42.576
27	US Births	1.4K	736	607	-	732	635	706	607

Table 5: Mean sMAPE results. The best model across each dataset is highlighted in boldface.

	Dataset	SES	Theta	ETS	(DHR-) ARIMA	PR	Cat Boost	ARIMA	TBATS
0	M1 Yearly	0.231	0.202	0.186	-	0.188	0.200	0.195	0.174
1	M1 Quarterly	0.181	0.163	0.174	-	0.166	0.177	0.166	0.166
2	M1 Monthly	0.171	0.155	0.146	-	0.148	0.162	0.153	0.148
3	M3 Quarterly	0.109	0.092	0.097	-	0.098	0.112	0.102	0.102
4	M3 Monthly	0.162	0.139	0.141	-	0.152	0.165	0.143	0.138
5	M4 Yearly	-	-	-	-	-	-	-	0.149
6	M4 Quarterly	-	-	-	-	-	-	-	0.102
7	M4 Weekly	-	-	-	-	-	-	-	0.073
8	Tourism Yearly	0.341	0.319	0.365	-	0.469	0.328	0.334	0.339
9	Tourism Quarterly	0.274	0.154	0.151	-	0.159	0.167	0.165	0.172
10	Tourism Monthly	0.364	0.199	0.190	-	0.211	0.213	0.196	0.212
11	Vehicle Trips	0.362	0.301	0.313	-	0.350	0.308	0.308	0.291
12	NN5 Daily	-	-	-	-	-	0.239	-	0.211
13	NN5 Weekly	0.122	0.120	0.123	0.118	0.114	0.117	-	0.116
14	Solar Weekly	0.246	0.248	0.229	0.179	0.217	0.285	-	0.191
15	Electricity Weekly	0.142	0.146	0.141	0.108	0.100	0.097	-	0.085
16	Traffic Weekly	0.124	0.125	0.126	0.134	0.125	0.130	-	0.128
17	Rideshare	1.413	1.540	-	-	-	-	-	1.377
18	Hospital	0.179	0.173	0.175	-	0.176	0.179	0.178	0.176
19	Sunspot	1.924	1.924	1.924	-	1.901	1.858	1.730	1.860
20	Bitcoin	0.208	0.302	0.191	-	0.215	0.306	0.269	0.200
21	CIF 2016	0.149	0.130	0.122	-	0.123	0.151	0.114	0.122
22	COVID Deaths	0.153	0.156	0.086	-	0.183	0.158	0.092	0.087
23	Car Parts	0.649	0.593	0.658	-	0.432	0.655	0.657	0.659
24	Fred Md	0.087	0.097	0.084	-	0.308	0.093	0.080	0.080
25	M3 Yearly	0.178	0.168	0.170	-	0.171	0.197	0.188	0.174
26	Saugeen River Flow	0.360	0.360	0.675	-	0.453	0.362	0.398	0.373
27	US Births	0.118	0.058	0.041	-	0.058	0.045	0.052	0.038

Table 6: Median MAE results. The best model across each dataset is highlighted in boldface.

	Dataset	SES	Theta	ETS	(DHR-) ARIMA	PR	Cat Boost	ARIMA	TBATS
0	M1 Yearly	379	256	191	-	246	261	180	173
1	M1 Quarterly	22.296	19.554	19.588	-	19.195	19.804	16.228	18.871
2	M1 Monthly	45.333	38.230	38.508	-	37.365	39.924	40.538	35.776
3	M3 Quarterly	372	294	305	-	325	397	334	336
4	M3 Monthly	517	421	409	-	479	533	412	407
5	M4 Yearly	-	-	-	-	-	-	-	430
6	M4 Quarterly	-	-	-	-	-	-	-	256
7	M4 Weekly	-	-	-	-	-	-	-	164
8	Tourism Yearly	4.3K	4.1K	4.3K	-	4.3K	5.0K	4.6K	4.8K
9	Tourism Quarterly	1.9K	1.1K	1.0K	-	992	1.0K	1.0K	1.2K
10	Tourism Monthly	968	478	457	-	475	473	463	492
11	Vehicle Trips	6.033	4.667	4.667	-	6.967	5.367	4.967	4.433
12	NN5 Daily	-	-	-	-	-	3.684	-	3.458
13	NN5 Weekly	14.183	13.904	14.273	14.824	12.837	13.129	-	13.727
14	Solar Weekly	1.1K	1.1K	1.1K	761	942	1.3K	-	780
15	Electricity Weekly	11.0K	10.4K	11.0K	6.8K	7.1K	6.1K	-	6.1K
16	Traffic Weekly	0.918	0.924	0.918	0.976	0.930	0.948	-	0.942
17	Rideshare	1.652	1.975	-	-	-	-	-	1.795
18	Hospital	6.667	6.667	6.667	-	6.667	6.917	6.833	6.833
19	Sunspot	4.933	4.933	4.933	-	3.833	2.800	2.567	2.467
20	Bitcoin	23.2K	20.3K	19.4K	-	25.1K	20.9K	29.9K	27.3K
21	CIF 2016	107	103	70.431	-	95.132	111	80.656	67.118
22	COVID Deaths	2.233	4.417	1.650	-	6.767	3.217	1.783	1.800
23	Car Parts	0.333	0.250	0.333	-	0.250	0.417	0.333	0.417
24	Fred Md	1.894	1.940	2.350	-	41.359	4.114	2.732	1.992
25	M3 Yearly	703	660	641	-	712	860	701	638
26	Saugeen River Flow	21.497	21.486	30.693	-	25.241	21.562	23.338	22.262
27	US Births	1.2K	587	420	-	575	464	526	399

Table 7: Median MASE results. The best model across each dataset is highlighted in boldface.

	Dataset	SES	Theta	ETS	(DHR-) ARIMA	PR	Cat Boost	ARIMA	TBATS
0	M1 Yearly	3.772	3.155	2.324	-	2.847	2.912	2.127	2.215
1	M1 Quarterly	1.417	1.264	1.196	-	1.376	1.411	1.171	1.200
2	M1 Monthly	1.167	0.885	0.851	-	0.947	1.016	0.896	0.902
3	M3 Quarterly	1.073	0.831	0.855	-	0.902	1.126	0.917	0.914
4	M3 Monthly	0.861	0.721	0.712	-	0.825	0.900	0.704	0.699
5	M4 Yearly	-	-	-	-	-	-	-	2.402
6	M4 Quarterly	-	-	-	-	-	-	-	0.915
7	M4 Weekly	-	-	-	-	-	-	-	0.365
8	Tourism Yearly	2.442	2.360	2.373	-	2.356	3.000	2.719	2.518
9	Tourism Quarterly	2.309	1.348	1.275	-	1.361	1.368	1.388	1.478
10	Tourism Monthly	2.336	1.382	1.276	-	1.484	1.461	1.333	1.491
11	Vehicle Trips	1.402	0.999	0.964	-	1.429	1.129	1.020	0.963
12	NN5 Daily	-	-	-	-	-	0.902	-	0.834
13	NN5 Weekly	0.781	0.805	0.775	0.769	0.781	0.808	-	0.827
14	Solar Weekly	1.231	1.241	1.209	0.861	1.063	1.475	-	0.894
15	Electricity Weekly	1.341	1.303	1.337	0.798	0.842	0.732	-	0.705
16	Traffic Weekly	0.973	0.983	0.977	1.035	0.980	0.946	-	0.996
17	Rideshare	4.054	4.912	-	-	-	-	-	4.065
18	Hospital	0.745	0.723	0.731	-	0.740	0.754	0.736	0.734
19	Sunspot	0.128	0.128	0.128	-	0.099	0.073	0.067	0.064
20	Bitcoin	3.089	2.955	2.686	-	3.166	3.018	3.542	3.207
21	CIF 2016	0.862	0.662	0.532	-	0.746	0.861	0.559	0.537
22	COVID Deaths	1.554	2.192	0.614	-	5.313	2.052	0.982	0.605
23	Car Parts	0.562	0.482	0.562	-	0.375	0.562	0.600	0.596
24	Fred Md	0.430	0.407	0.385	-	8.458	0.618	0.355	0.370
25	M3 Yearly	2.261	1.985	1.907	-	2.267	2.726	2.003	1.900
26	Saugeen River Flow	1.426	1.425	2.036	-	1.674	1.430	1.548	1.477
27	US Births	4.343	2.138	1.529	-	2.094	1.690	1.917	1.453

Table 8: Median RMSE results. The best model across each dataset is highlighted in boldface.

	Dataset	SES	Theta	ETS	(DHR-) ARIMA	PR	Cat Boost	ARIMA	TBATS
0	M1 Yearly	416	323	230	-	305	298	208	204
1	M1 Quarterly	24.459	22.811	21.858	-	22.529	22.572	20.232	22.320
2	M1 Monthly	54.669	46.396	44.392	-	45.346	47.584	47.105	44.038
3	M3 Quarterly	436	356	369	-	378	480	406	400
4	M3 Monthly	634	517	496	-	582	634	500	493
5	M4 Yearly	-	-	-	-	-	-	-	495
6	M4 Quarterly	-	-	-	-	-	-	-	302
7	M4 Weekly	-	-	-	-	-	-	-	200
8	Tourism Yearly	4.7K	4.6K	4.6K	-	4.7K	5.5K	5.2K	5.2K
9	Tourism Quarterly	2.3K	1.4K	1.2K	-	1.2K	1.2K	1.2K	1.5K
10	Tourism Monthly	1.3K	675	599	-	596	620	606	671
11	Vehicle Trips	8.103	5.802	5.925	-	8.725	6.962	6.506	5.580
12	NN5 Daily	-	-	-	-	-	5.320	-	4.749
13	NN5 Weekly	17.524	16.816	17.523	17.487	16.263	16.060	-	16.990
14	Solar Weekly	1.2K	1.2K	1.2K	878	1.0K	1.4K	-	886
15	Electricity Weekly	12.5K	11.8K	12.5K	8.3K	8.2K	7.3K	-	7.3K
16	Traffic Weekly	1.201	1.215	1.210	1.211	1.195	1.159	-	1.214
17	Rideshare	1.841	2.190	-	-	-	-	-	2.021
18	Hospital	8.256	8.196	8.251	-	8.251	8.485	8.391	8.357
19	Sunspot	4.946	4.946	4.946	-	3.954	3.141	2.938	2.595
20	Bitcoin	30.3K	26.3K	24.3K	-	31.4K	28.3K	38.2K	33.0K
21	CIF 2016	129	118	85.771	-	109	131	103	79.025
22	COVID Deaths	3.087	5.290	2.205	-	8.283	3.941	2.164	2.129
23	Car Parts	0.707	0.645	0.707	-	0.577	0.707	0.707	0.707
24	Fred Md	2.306	2.362	2.702	-	45.182	4.512	3.490	2.515
25	M3 Yearly	804	740	759	-	825	969	815	753
26	Saugeen River Flow	39.794	39.787	50.392	-	47.703	39.306	45.536	42.576
27	US Births	1.4K	736	607	-	732	635	706	607

Table 9: Median RMSE results. The best model across each dataset is highlighted in boldface.

	Dataset	SES	Theta	ETS	(DHR-) ARIMA	PR	Cat Boost	ARIMA	TBATS
0	M1 Yearly	0.173	0.147	0.130	-	0.135	0.134	0.120	0.127
1	M1 Quarterly	0.112	0.086	0.084	-	0.101	0.116	0.097	0.086
2	M1 Monthly	0.143	0.112	0.108	-	0.119	0.125	0.115	0.113
3	M3 Quarterly	0.067	0.052	0.055	-	0.057	0.076	0.064	0.062
4	M3 Monthly	0.107	0.093	0.091	-	0.104	0.110	0.090	0.090
5	M4 Yearly	-	-	-	-	-	-	-	0.088
6	M4 Quarterly	-	-	-	-	-	-	-	0.058
7	M4 Weekly	-	-	-	-	-	-	-	0.048
8	Tourism Yearly	0.188	0.168	0.192	-	0.169	0.236	0.227	0.206
9	Tourism Quarterly	0.225	0.132	0.129	-	0.133	0.135	0.131	0.148
10	Tourism Monthly	0.302	0.174	0.172	-	0.185	0.189	0.180	0.190
11	Vehicle Trips	0.342	0.235	0.232	-	0.327	0.271	0.236	0.228
12	NN5 Daily	-	-	-	-	-	0.229	-	0.196
13	NN5 Weekly	0.109	0.110	0.108	0.111	0.105	0.104	-	0.110
14	Solar Weekly	0.248	0.249	0.244	0.176	0.218	0.282	-	0.184
15	Electricity Weekly	-	0.117	-	0.070	-	0.061	-	-
16	Traffic Weekly	0.097	0.098	0.098	0.105	0.098	0.102	-	0.101
17	Rideshare	2.000	2.000	-	-	-	-	-	1.976
18	Hospital	0.166	0.159	0.161	-	0.161	0.168	0.168	0.163
19	Sunspot	1.962	1.962	1.962	-	1.956	1.933	1.943	1.947
20	Bitcoin	0.182	0.187	0.188	-	0.172	0.187	0.192	0.175
21	CIF 2016	0.114	0.080	0.066	-	0.084	0.108	0.077	0.070
22	COVID Deaths	-	-	-	-	-	-	-	-
23	Car Parts	-	-	-	-	-	-	-	-
24	Fred Md	0.016	0.015	0.015	-	0.291	0.033	0.016	0.013
25	M3 Yearly	0.124	0.115	0.115	-	0.129	0.146	0.124	0.115
26	Saugeen River Flow	0.360	0.360	0.676	-	0.454	0.363	0.398	0.374
27	US Births	0.118	0.058	0.041	-	0.058	0.045	0.052	0.038