Monash Time-Series Forecasting Archive Replication

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

March 9, 2024

Abstract

The "Monash Time Series Forecasting Archive" significantly enhances research in time series forecasting, unveiling a meticulously selected assortment of 25 datasets across diverse sectors like energy, banking, and tourism.

This initiative tackles the pressing need for diverse, comprehensive datasets to benchmark a broad spectrum of forecasting models, spanning from conventional univariate to modern global and multivariate methodologies.

*fernandourbano@uchicago.edu
†shreyjain@uchicago.edu
‡acarrington@uchicago.edu
§mukundmaheshwari@uchicago.edu

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

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

	Dataset	Cat Boost	ARIMA	TBATS
0	M1 Yearly	-	4.479	**3.499**
1	M1 Quarterly	-	-	**1.694**
2	M1 Monthly	-	-	**1.118**
3	M3 Quarterly	-	-	**1.256**
4	M3 Monthly	-	-	**0.861**
5	M4 Yearly	-	-	**3.437**
6	M4 Quarterly	-	-	**1.186**
7	M4 Weekly	-	-	**0.505**
8	Tourism Yearly	-	-	**3.685**
9	Tourism Quarterly	-	-	**1.835**
10	Tourism Monthly	-	-	**1.751**
11	NN5 Daily	0.970	-	**0.858**
12	CIF 2016	-	-	**0.861**
13	M3 Yearly	-	-	**3.127**

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

	Dataset	Cat Boost	ARIMA	TBATS
0	M1 Yearly	-	145608.872	**103006.950**
1	M1 Quarterly	-	-	**2326.464**
2	M1 Monthly	-	-	**2237.507**
3	M3 Quarterly	-	-	**561.766**
4	M3 Monthly	-	-	**630.577**
5	M4 Yearly	-	-	**960.446**
6	M4 Quarterly	-	-	**570.217**
7	M4 Weekly	-	-	**296.808**
8	Tourism Yearly	-	-	**94121.085**
9	Tourism Quarterly	-	-	**9972.417**
10	Tourism Monthly	-	-	**2940.081**
11	NN5 Daily	4.200	-	**3.701**
12	CIF 2016	-	-	**855578.358**
13	M3 Yearly	-	-	**1192.847**

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

	Dataset	Cat Boost	ARIMA	TBATS
0	M1 Yearly	-	175343.759	**116850.923**
1	M1 Quarterly	-	-	**2673.911**
2	M1 Monthly	-	-	**2594.483**
3	M3 Quarterly	-	-	**653.614**
4	M3 Monthly	-	-	**765.240**
5	M4 Yearly	-	-	**1099.947**
6	M4 Quarterly	-	-	**672.697**
7	M4 Weekly	-	-	**357.536**
8	Tourism Yearly	-	-	**105799.355**
9	Tourism Quarterly	-	-	**12001.480**
10	Tourism Monthly	-	-	**3661.512**
11	NN5 Daily	5.715	-	**5.204**
12	CIF 2016	-	-	**940099.906**
13	M3 Yearly	-	-	**1386.329**

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

	Dataset	Cat Boost	ARIMA	TBATS
0	M1 Yearly	-	0.195	**0.174**
1	M1 Quarterly	-	-	**0.166**
2	M1 Monthly	-	-	**0.148**
3	M3 Quarterly	-	-	**0.102**
4	M3 Monthly	-	-	**0.138**
5	M4 Yearly	-	-	**0.149**
6	M4 Quarterly	-	-	**0.102**
7	M4 Weekly	-	-	**0.073**
8	Tourism Yearly	-	-	**0.339**
9	Tourism Quarterly	-	-	**0.172**
10	Tourism Monthly	-	-	**0.212**
11	NN5 Daily	0.239	-	**0.211**
12	CIF 2016	-	-	**0.122**
13	M3 Yearly	-	-	**0.174**

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

	Dataset	Cat Boost	ARIMA	TBATS
0	M1 Yearly	-	179.979	**173.359**
1	M1 Quarterly	-	-	**18.871**
2	M1 Monthly	-	-	**35.776**
3	M3 Quarterly	-	-	**335.693**
4	M3 Monthly	-	-	**406.592**
5	M4 Yearly	-	-	**429.689**
6	M4 Quarterly	-	-	**255.646**
7	M4 Weekly	-	-	**163.678**
8	Tourism Yearly	-	-	**4789.949**
9	Tourism Quarterly	-	-	**1176.187**
10	Tourism Monthly	-	-	**492.461**
11	NN5 Daily	3.684	-	**3.458**
12	CIF 2016	-	-	**67.118**
13	M3 Yearly	-	-	**637.810**

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

	Dataset	Cat Boost	ARIMA	TBATS
0	M1 Yearly	-	**2.127**	2.215
1	M1 Quarterly	-	-	**1.200**
2	M1 Monthly	-	-	**0.902**
3	M3 Quarterly	-	-	**0.914**
4	M3 Monthly	-	-	**0.699**
5	M4 Yearly	-	-	**2.402**
6	M4 Quarterly	-	-	**0.915**
7	M4 Weekly	-	-	**0.365**
8	Tourism Yearly	-	-	**2.518**
9	Tourism Quarterly	-	-	**1.478**
10	Tourism Monthly	-	-	**1.491**
11	NN5 Daily	0.902	-	**0.834**
12	CIF 2016	-	-	**0.537**
13	M3 Yearly	-	-	**1.900**

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

	Dataset	Cat Boost	ARIMA	TBATS
0	M1 Yearly	-	207.818	**204.193**
1	M1 Quarterly	-	-	**22.320**
2	M1 Monthly	-	-	**44.038**
3	M3 Quarterly	-	-	**400.010**
4	M3 Monthly	-	-	**493.189**
5	M4 Yearly	-	-	**495.022**
6	M4 Quarterly	-	-	**302.405**
7	M4 Weekly	-	-	**200.317**
8	Tourism Yearly	-	-	**5156.832**
9	Tourism Quarterly	-	-	**1470.608**
10	Tourism Monthly	-	-	**670.852**
11	NN5 Daily	5.320	-	**4.749**
12	CIF 2016	-	-	**79.025**
13	M3 Yearly	-	-	**752.691**

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

	Dataset	Cat Boost	ARIMA	TBATS
0	M1 Yearly	-	**0.120**	0.127
1	M1 Quarterly	-	-	**0.086**
2	M1 Monthly	-	-	**0.113**
3	M3 Quarterly	-	-	**0.062**
4	M3 Monthly	-	-	**0.090**
5	M4 Yearly	-	-	**0.088**
6	M4 Quarterly	-	-	**0.058**
7	M4 Weekly	-	-	**0.048**
8	Tourism Yearly	-	-	**0.206**
9	Tourism Quarterly	-	-	**0.148**
10	Tourism Monthly	-	-	**0.190**
11	NN5 Daily	0.229	-	**0.196**
12	CIF 2016	-	-	**0.070**
13	M3 Yearly	-	-	**0.115**