|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Method** | **sMAPE** | | | | | | |
|  | **Yearly**  (23,000) | **Quarterly**  (24,000) | **Monthly**  (48,000) | **Weekly**  (359) | **Daily**  (4227) | **Hourly**  (414) | **All**  (100,000) |
| Naïve | 16.3422 | 11.6103 | 15.2565 | 9.1613 | 3.0453 | 43.003 | 14.2079 |
| sNaïve | 16.3422 | 12.5214 | 15.9883 | 9.1613 | 3.0453 | 13.9123 | 14.6574 |
| Naïve2 | 16.3422 | 11.0116 | 14.4268 | 9.1613 | 3.0453 | 18.3829 | 13.5641 |
| SES | 16.3982 | 10.6004 | 13.6182 | 9.0121 | 3.045 | **18.0939** | 13.0884 |
| Holt | 16.5347 | 10.9548 | 14.8282 | 9.7065 | 3.07 | 29.4737 | 13.8363 |
| Damped | 15.1626 | 10.2429 | 13.4726 | 8.8671 | 3.0631 | 19.2767 | 12.6537 |
| Theta | **14.5636** | 10.3132 | **13.0118** | 9.0889 | 3.0526 | 18.1381 | **12.3072** |
| Comb | 14.8743 | **10.1966** | 13.4339 | 8.947 | **2.9854** | 22.114 | 12.5664 |
|  |  | **Quarterly’** (2,400) | **Monthly’** (4,800) |  |  | **Hourly’** (42) | **Average** (weighted) |
| Naïve2’ |  | 10.8665 | 14.4595 |  |  | 21.2282 |  |
| Comb’ |  | 10.0628 | 13.4611 |  |  | 22.6851 |  |
| ARIMA | 15.1561 | 10.1525 | 13.5356 | 8.5913 | 3.186 | 16.8975 | 12.6551 |
| ETS | 15.3558 | 10.1163 | 13.5351 | 8.7266 | 3.046 | 19.5307 | 12.6975 |
| ETSARIMA | 14.8145 | **9.8515** | **12.9769** | **8.444** | 3.0814 | **17.3744** | **12.2331** |
| MLP\* |  |  |  | 9.3951 | 3.2726 | 16.7037 |  |
| DeepAR\* |  |  |  | 9.0204 | 3.5452 | 13.4565 |  |
| MQ-CNN\* |  |  |  | 9.5051 | 3.8066 | 19.359 |  |
| DeepAR\*\* |  |  |  |  |  | 12.0585 |  |
| Bold numbers highlight the best performing method within respective dataset  ‘ indicates subset | | | | | | | |

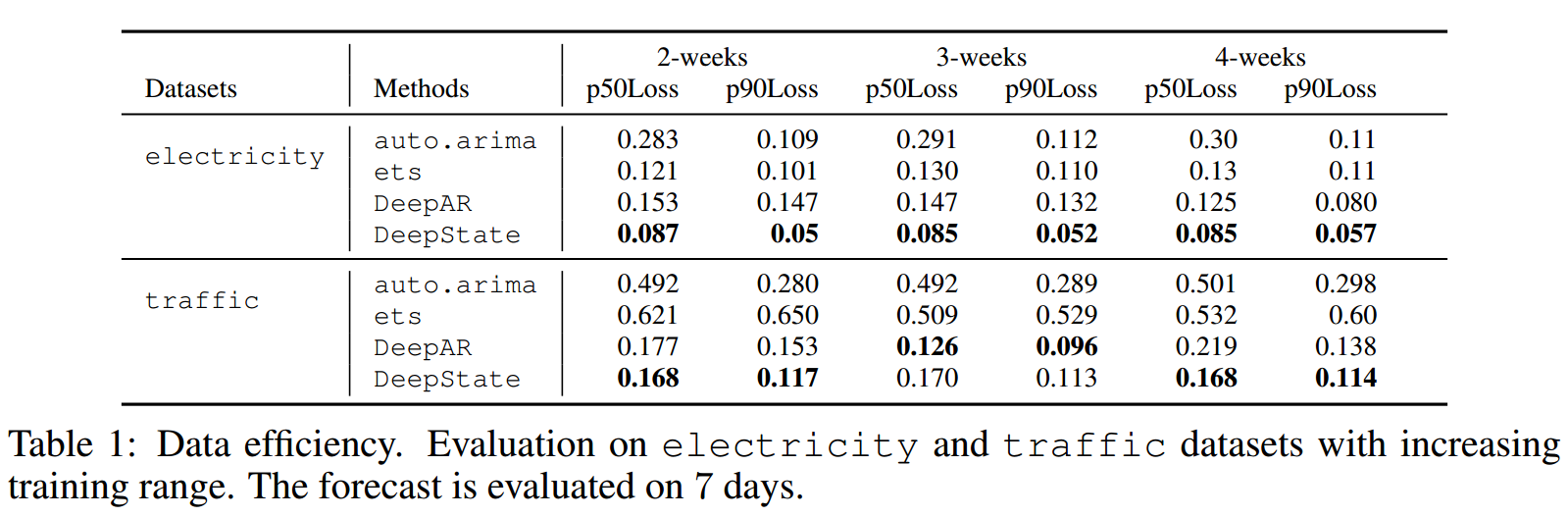
|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Method** | **MASE** | | | | | | |
|  | **Yearly**  (23,000) | **Quarterly**  (24,000) | **Monthly**  (48,000) | **Weekly**  (359) | **Daily**  (4227) | **Hourly**  (414) | **All**  (100,000) |
| Naïve | 3.9744 | 1.477 | 1.2051 | 2.7773 | 3.2784 | 11.6077 | 2.0436 |
| sNaïve | 3.9744 | 1.6022 | 1.2597 | 2.7773 | 3.2784 | 1.1932 | 2.0568 |
| Naïve2 | 3.9744 | 1.3714 | 1.0634 | 2.7773 | 3.2784 | **2.395** | 1.9121 |
| SES | 3.9811 | 1.3398 | 1.0196 | 2.6845 | 3.2809 | 2.3847 | 1.8848 |
| Holt | 3.5764 | 1.1988 | 1.0098 | 2.4132 | 3.2313 | 9.3799 | 1.7791 |
| Damped | 3.3721 | 1.1747 | 0.9719 | **2.4001** | 3.2343 | 2.9475 | 1.6816 |
| Theta | 3.3746 | 1.2313 | 0.9702 | 2.6387 | 3.2623 | 2.4545 | 1.6949 |
| Comb | **3.2823** | **1.1741** | **0.9663** | 2.4295 | **3.2051** | 4.5845 | **1.6637** |
|  |  | **Quarterly’** (2,400) | **Monthly’** (4,800) |  |  | **Hourly’** (42) | **Average** (weighted) |
| Naïve2’ |  | 1.3579 | 1.0614 |  |  | 2.4679 |  |
| Comb’ |  | 1.152 | 0.9641 |  |  | 4.4873 |  |
| ARIMA | 3.4018 | 1.1437 | 0.9399 | 2.5423 | 3.3993 | **0.9379** | 1.6648 |
| ETS | 3.4444 | 1.1457 | 0.945 | 2.527 | 3.2526 | 1.5777 | 1.6739 |
| ETSARIMA | 3.3408 | **1.1123** | **0.9092** | 2.4711 | 3.2707 | 1.1427 | **1.6236** |
| MLP\* |  |  |  | 2.8217 | 3.2726 | 1.5627 |  |
| DeepAR\* |  |  |  | 3.2609 | 3.5452 | 1.6072 |  |
| MQ-CNN |  |  |  | 2.7788 | 3.8066 | 3.2726 |  |
| DeepAR\*\* |  |  |  |  |  | 2.204 |  |
|  |  |  |  |  |  |  |  |
| Bold numbers highlight the best performing method within respective dataset  ‘ indicates subset | | | | | | | |

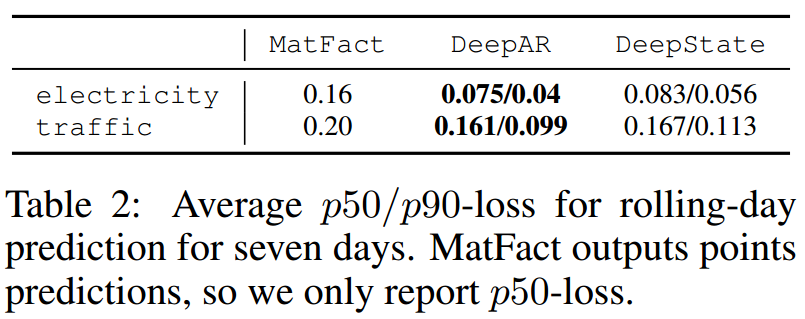
|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Method** | **OWA** | | | | | | |
|  | **Yearly**  (23,000) | **Quarterly**  (24,000) | **Monthly**  (48,000) | **Weekly**  (359) | **Daily**  (4227) | **Hourly**  (414) | **All**  (100,000) |
| Naïve | 1 | 1.0657 | 1.0954 | 1 | 1 | 3.5929 | 1.0723 |
| sNaïve | 1 | 1.1527 | 1.1464 | 1 | 1 | 0.6275 | 1.1054 |
| Naïve2 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| SES | 1.0026 | 0.9698 | 0.9514 | 0.9752 | 1.0003 | **0.99** | 0.9699 |
| Holt | 0.9558 | 0.9345 | 0.9887 | 0.9642 | 0.9969 | 2.7599 | 0.9757 |
| Damped | 0.8881 | 0.8934 | 0.9239 | 0.916 | 0.9962 | 1.1396 | 0.9123 |
| Theta | 0.8701 | 0.9172 | **0.9072** | 0.9711 | 0.9988 | 1.0058 | 0.9056 |
| Comb | **0.868** | **0.8911** | 0.92 | 0.9257 | **0.979** | 1.5586 | 0.9063 |
|  |  | **Quarterly**‘ (2,400) | **Monthly’** (4,800) |  |  | **Hourly’** (42) | **Average** (weighted) |
| Naïve2’ |  | 1 | 1 |  |  | 1 |  |
| Comb’ |  | 0.8872 | 0.9197 |  |  | 1.4434 |  |
| ARIMA | 0.8917 | 0.8883 | 0.9108 | 0.9266 | 1.0416 | **0.588** | 0.9053 |
| ETS | 0.9032 | 0.8874 | 0.9132 | 0.9312 | 0.9962 | 0.7797 | 0.9077 |
| ETSARIMA | 0.8736 | **0.8629** | **0.8771** | **0.9057** | 1.0048 | 0.6407 | **0.8774** |
| MLP\* |  |  |  | 1.0215 | 1.0364 | 0.7806 |  |
| DeepAR\* |  |  |  | 1.0802 | 1.1228 | 0.7015 |  |
| MQ-CNN |  |  |  | 1.0198 | 1.2056 | 1.2098 |  |
| DeepAR\*\* |  |  |  |  |  | 0.7881 |  |
|  |  |  |  |  |  |  |  |
| Bold numbers highlight the best performing method  ‘ indicates subset | | | | | | | |

\* default values for algorithms, epochs=100, num\_batches\_per\_epoch=50, context\_length=prediction\_length

\*\* context\_length=168 (7 days), epochs=100, num\_batches\_per\_epoch=100

**Replication study: RSG+18**





Own results:

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Data** | **Method** | **seed** | **Epochs / batches** | **Context length** | **freq** | **MASE** | **sMAPE** | **wQ50L** | **wQ90L** |
| electricity | DeepAR | 42 | 100/50 | 2-weeks | D | 1.0591 | 0.1204 | 0.0642 | 0.0309 |
| electricity | DeepAR | 43 | 100/50 | 2-weeks | D | 0.8695 | 0.1103 | 0.0896 | 0.0736 |
| electricity | DeepAR | 44 | 100/50 | 2-weeks | D | 0.8679 | 0.1100 | 0.0870 | 0.0592 |
| electricity | DeepState | 42 | 25/50 | 2-weeks | D | 1.2391 | 0.1447 | 0.0910 | 0.1171 |
| electricity | DeepState | 43 | 25/50 | 2-weeks | D | 1.2247 | 0.1420 | 0.0936 | 0.1258 |
| electricity | DeepState | 44 | 25/50 | 2-weeks | D | 1.1617 | 0.1377 | 0.0803 | 0.1082 |

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Data** | **Method** | **seed** | **Epochs / batches** | **Context length** | **freq** | **MASE** | **sMAPE** | **wQ50L** | **wQ90L** |
| electricity | DeepAR | 42 | 100/50 | 4-weeks | D | 0.8004 | 0.1053 | 0.0631 | 0.0430 |
| electricity | DeepAR | 43 | 100/50 | 4-weeks | D | 0.7933 | 0.1050 | 0.0590 | 0.0379 |
| electricity | DeepAR | 44 | 100/50 | 4-weeks | D | 0.7755 | 0.1053 | 0.0604 | 0.0369 |
| electricity | DeepState | 42 | 25/50 | 4-weeks | D | 1.4042 | 0.1559 | 0.0929 | 0.1054 |
| electricity | DeepState | 43 | 25/50 | 4-weeks | D | 1.3968 | 0.1527 | 0.1087 | 0.1411 |
| electricity | DeepState | 44 | 25/50 | 4-weeks | D | 1.5307 | 0.1744 | 0.1188 | 0.1488 |
| electricity | DeepState | 42 | 25/50 | default | D | 0.9637 | 0.1312 | 0.0736 | 0.0493 |
| electricity | DeepState | 42 | 25/50 | default | D | 0.9326 | 0.1295 | 0.0706 | 0.0497 |
| electricity | DeepState | 42 | 25/50 | default | D |  |  |  |  |

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Data** | **Method** | **seed** | **Epochs / batches** | **Context / prediction** | **freq** | **MASE** | **sMAPE** | **wQ50L** | **wQ90L** |
| traffic | DeepAR | 42 | 100/50 | 28/7 | D | 0.7978 | 0.1288 | 0.1522 | 0.0984 |
| traffic | DeepAR | 43 | 100/50 | 28/7 | D | 0.7830 | 0.1272 | 0.1494 | 0.0968 |
| traffic | DeepAR | 44 | 100/50 | 28/7 | D | 0.8162 | 0.1319 | 0.1561 | 0.0984 |
| traffic | DeepState | 42 | 25/50 | 28/7 | D |  |  |  |  |
| traffic | DeepState | 43 | 25/50 | 28/7 | D |  |  |  |  |
| traffic | DeepState | 44 | 25/50 | 28/7 | D |  |  |  |  |

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Data** | **Method** | **seed** | **Epochs / batches** | **Context / prediction** | **freq** | **MASE** | **sMAPE** | **wQ50L** | **wQ90L** |
| M4-hourly | DeepAR | 42 | 100/50 | 48/48 | H | 1.3395 | 0.1140 | 0.0483 | 0.0319 |
| M4-hourly | DeepAR | 43 | 100/50 | 48/48 | H | 1.4938 | 0.1258 | 0.0641 | 1.0240 |
| M4-hourly | DeepAR | 44 | 100/50 | 48/48 | H | 2.1627 | 0.1421 | 0.1348 | 0.1822 |
| M4-hourly | DeepState | 42 | 25/50 |  |  |  |  |  |  |
| M4-hourly | DeepState | 43 | 25/50 |  |  |  |  |  |  |
| M4-hourly | DeepState | 44 | 25/50 |  |  |  |  |  |  |

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Data** | **Method** | **seed** | **Epochs / batches** | **Context / prediction** | **freq** | **MASE** | **sMAPE** | **wQ50L** | **wQ90L** |
| M4-weekly | DeepAR | 42 | 100/50 | 13/13 | W | 2.6305 | 0.0952 | 0.0668 | 0.0310 |
| M4-weekly | DeepAR | 43 | 100/50 | 13/13 | W | 2.5731 | 0.0891 | 0.0643 | 0.0282 |
| M4-weekly | DeepAR | 44 | 100/50 | 13/13 | W | 2.5440 | 0.0895 | 0.0634 | 0.0299 |
| M4-weekly | DeepState | 42 | 25/50 |  |  |  |  |  |  |
| M4-weekly | DeepState | 43 | 25/50 |  |  |  |  |  |  |
| M4-weekly | DeepState | 44 | 25/50 |  |  |  |  |  |  |

MLND Project

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Data** | **Method** | **seed** | **Epochs / batches** | **Context / prediction** | **freq** | **MASE** | **sMAPE** | **wQ50L** | **wQ90L** |
| M4-daily | DeepAR | 42 | 100/50 | 14/14 | D | 4.9019 | 0.0444 | 0.0401 | 0.0224 |
| M4-daily | DeepAR | 43 | 100/50 | 14/14 | D | 4.0038 | 0.0367 | 0.0308 | 0.0141 |
| M4-daily | DeepAR | 44 | 100/50 | 14/14 | D | 3.5157 | 0.0334 | 0.0299 | 0.0159 |
| M4-daily | DeepAR | 42 | 100/100 | 14/14 | D | 4.6888 | 0.0424 | 0.0381 | 0.0194 |
| M4-daily | DeepAR | 43 | 100/100 | 14/14 | D | 3.7620 | 0.0351 | 0.0318 | 0.0150 |
| M4-daily | DeepAR | 44 | 100/100 | 14/14 | D | 3.7491 | 0.0347 | 0.0316 | 0.0141 |
| M4-daily | DeepAR | 42 | 100/200 | 14/14 | D | 5.4063 | 0.0475 | 0.0460 | 0.0263 |
| M4-daily | DeepAR | 43 | 100/200 | 14/14 | D | 4.3779 | 0.0392 | 0.0377 | 0.0177 |
| M4-daily | DeepAR | 44 | 100/200 | 14/14 | D | 3.6072 | 0.0338 | 0.0307 | 0.0141 |