**Data:** **m4\_yearly, freq: Y**

Different order compared to M4 Yearly overall? Changes in composition.

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **M4 Yearly** | | | | | | | | | | | |
|  | **End: 1991** | | |  | **End: 2009** | | |  | **End: 2004** | | |
| **Method** | **OWA** | **sMAPE** | **MASE** |  | **OWA** | **sMAPE** | **MASE** |  | **OWA** | **sMAPE** | **MASE** |
| Naïve2 | 1.000 | 0.207 | 5.468 |  | 1.000 | 0.159 | 3.466 |  | 1.000 | 0.136 | 3.472 |
| Comb | 0.768 | 0.165 | 4.032 |  | 1.004 | 0.166 | 4.032 |  | 0.866 | 0.124 | 2.840 |
| ARIMA | 0.701 | 0.151 | 3.680 |  | 1.058 | 0.174 | 3.680 |  | 0.837 | 0.121 | 2.728 |
| ETS | 0.729 | 0.159 | 3.778 |  | 1.082 | 0.175 | 3.778 |  | 0.833 | 0.119 | 2.727 |
| ETSARIMA | 0.704 | 0.152 | 3.678 |  | 1.034 | 0.168 | 3.678 |  | 0.812 | 0.117 | 2.654 |
| L&K | 0.649 | 0.139 | 3.418 |  | 0.830 | 0.137 | 3.418 |  | 0.880 | 0.125 | 2.911 |
| DeepAR\* | 0.634 | 0.136 | 3.336 |  | 0.903 | 0.147 | 3.336 |  | 0.866 | 0.124 | 2.837 |
| Values for DeepAR indicates best median of three settings. Each setting is repeated for ten trials.  \* using series-specific dummy variables | | | | | | | | | | | |

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **M4 Quarterly** | | | | | | | | | | | |
|  | **End: Q4/2003** | | |  | **End: Q2/2013** | | |  | **End: Q2/2015** | | |
| **Method** | **OWA** | **sMAPE** | **MASE** |  | **OWA** | **sMAPE** | **MASE** |  | **OWA** | **sMAPE** | **MASE** |
| Naïve2 | 1.000 | 0.047 | 1.219 |  | 1.000 | 0.104 | 1.324 |  | 1.000 | 0.107 | 1.328 |
| Comb | 0.755 | 0.035 | 0.925 |  | 0.838 | 0.087 | 1.107 |  | 0.919 | 0.102 | 1.180 |
| ARIMA | 0.748 | 0.036 | 0.904 |  | 0.830 | 0.087 | 1.089 |  | 0.911 | 0.102 | 1.149 |
| ETS | 0.786 | 0.037 | 0.954 |  | 0.827 | 0.086 | 1.089 |  | 0.915 | 0.102 | 1.167 |
| ETSARIMA | 0.754 | 0.036 | 0.913 |  | 0.813 | 0.085 | 1.068 |  | 0.887 | 0.099 | 1.127 |
| L&K | 0.754 | 0.035 | 0.934 |  | 0.851 | 0.088 | 1.125 |  | 0.886 | 0.097 | 1.144 |
| DeepAR\* | 0.870 | 0.041 | 1.057 |  | 0.838 | 0.087 | 1.108 |  | 0.885 | 0.098 | 1.136 |
| Values for DeepAR indicates the median of the best of three settings, each setting uses consists of ten trials.  \* using series-specific dummy variables. | | | | | | | | | | | |

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **M4 Monthly** | | | | | | | | | | | |
|  | **End: 04/2014** | | |  | **End: 06/2014** | | |  | **End: 10/2015** | | |
| **Method** | **OWA** | **sMAPE** | **MASE** |  | **OWA** | **sMAPE** | **MASE** |  | **OWA** | **sMAPE** | **MASE** |
| Naïve2 | 1.000 | 0.184 | 1.087 |  | 1.000 | 0.074 | 1.048 |  | 1.000 | 1.290 | 0.306 |
| Comb | 0.969 | 0.182 | 1.035 |  | 0.906 | 0.067 | 0.938 |  | 0.890 | 1.145 | 0.273 |
| ARIMA | 0.982 | 0.184 | 1.047 |  | 0.885 | 0.068 | 0.888 |  | 0.889 | 1.148 | 0.271 |
| ETS | 0.987 | 0.184 | 1.059 |  | 0.891 | 0.068 | 0.897 |  | 0.914 | 1.180 | 0.279 |
| ETSARIMA | 0.955 | 0.178 | 1.027 |  | 0.860 | 0.066 | 0.865 |  | 0.884 | 1.139 | 0.271 |
| L&K | 0.975 | 0.181 | 1.050 |  | 0.921 | 0.068 | 0.960 |  | 0.847 | 1.105 | 0.256 |
| DeepAR\* | 1.003 | 0.184 | 1.098 |  | 0.974 | 0.073 | 0.998 |  | 0.921 | 1.209 | 0.276 |
| Values for DeepAR indicates the median of the best of three settings, each setting uses consists of ten trials.  \* using series-specific dummy variables | | | | | | | | | | | |