A neural network methodology for forecasting constant and dynamic demand rate for intermittent demand time series

The International Institute of Forecasters presents
The 30th Annual International Symposium on Forecasting

Nikolaos Kourentzes
Sven F. Crone
LUMS—Department of Management Science
Motivation
Intermittent demand

Forecasting intermittent time series

- Croston’s method widely used for intermittent demand forecasting
  - Based on exponential smoothing [Croston, 72]
  - Later corrected for positive bias [Syntetos & Boylan, 01, Leven Segerstedt, 04]

- Model assumptions → Incorrect! [Shenstone & Hyndman, 03]
  - Assume continuous data
  - Include negative values
  - Problematic → Intermittent demand is non-negative, integer valued
Artificial Neural Networks in Forecasting

► Nonparametric nonlinear data driven models → no data assumptions
  ➢ Learn from available information and generalise [Church and Curram, 96]
  ➢ Approximate any data generating process [Hornik & Stinchcombe, 89, Hornik, 91]
  ➢ Flexible models → minimise modelling biases [Zhang et al., 98]

► Promising performance 64% (out of 126) articles found ANNs outperforming benchmarks [Kourentzes, 10] (73% according to Adya & Collopy, 98)

► Use ANNs to avoid data assumptions
  ➢ Have been used before [Guitierrez et al., 08] → Not flexible for bias correction

Propose a more generic ANN formulation
The key concept is to split time series in nonzero demand & demand intervals.

**Demand**

Only 5 months of data!

**Forecast them with the same single exponential smoothing model**

Count every how many months there is demand
Intermittent Demand

**Intermittent Time Series**

- **Nonzero Demand Forecasting Model**
- **Time Interval Forecasting Model**
- **Use Neural Networks to forecast**

**Intermittent Demand**

Croston’s Method

The output of Croston’s method is not a demand forecast, but a demand rate forecast.

Use Neural Networks to forecast

The use of neural networks is not straightforward in this case.
Use a pair of NNs to forecast demand and intervals (Single [input] NN)

- Model separately the demand and interval time series → Fit appropriate NNs
- Advantage over Croston’s → No data assumptions

→ NNs allow more elegant formulations...
Use a single NN to forecast simultaneously demand and intervals (Dual [input] NN)

In addition ➔ exploit information of demand to forecast intervals and vice versa
Intermittent Demand
Neural Network Methods

NNs offer alternative forecasting procedures

**Croston’s method:**

→ Use the last estimated demand rate as a constant forecast
→ Equivalent NN forecast formulation

**Dynamic method:**

→ NNs are NAR models [Connor et al., 94]
→ Produce dynamic forecasts

**Constant forecast**

- Constrain to constant demand rate

**Dynamic forecast**

- Allow changing demand rate

→ ANNs typically require large samples for training → How to resolve?
In forecasting we typically use a training sample to parameterise our model and a test or hold-out sample to fit our model.

**NNs are powerful universal approximators** → fit any function to any arbitrary degree of accuracy (in-sample) [Hornik 91]

→ This may lead to overfitting and poor out-of-sample performance.

To resolve this we use an additional “validation set” to control the degree of overfitting.
NN Training

Early stopping training example

“Lost” training sample

This can be critical when there is limited data sample

→ Not enough training samples to train the network
NN Training

Regularisation example

![Graph showing the difference between normal training and regularised training](image)

**Normal training (Overfit)**

**Regularised training**

Cost Function:

\[ mse = \frac{1}{N} \sum_{i=1}^{N} (e_i)^2 \]

Forecast errors

Cost Function:

\[ msereg = \gamma mse + (1 - \gamma)msw \]

NN weights & biases

Can be used with small training samples!
Bias correction

- Croston’s method (divide non-zero demand with intervals) leads to positive forecasting bias [Syntetos & Boylan, 01]

- The proposed ANN approach suffers from same bias

- Data driven bias correction:

  \[
  \text{Final Forecast} = c \cdot \text{ANN forecast}
  \]

  \[
  c = \frac{\text{Observed in-sample demand rate}}{\text{ANN in-sample demand rate}}
  \]

 Jauneor method to de-bias ANN forecasts?
Experimental Design

Real dataset with 3000 time series

- Automotive spare parts [Syntetos & Boylan, 05]
- 24 observations history → Use 19 for training
- Forecast last 5 observation in each time series
- Error measures – Overcome intermittent demand problems
  - Mean Absolute Scaled Error - MASE [Hyndman et al., 06]
    \[
    MASE = \frac{\text{MAE of method}}{\text{MAE of in-sample naive}}
    \]
  - Geometric Mean Absolute Error - GMAE [Hyndman, 06]
    \[
    GMAE = \text{Geometric Mean}(|e_t|)
    \]
  - Identical to GRMSE suggested by Syntetos & Boylan [05]
Forecasting methods

- Time series methods: Naïve, Moving Average (MA) 3 & 5 periods, Single Exponential Smoothing (SES) with $\alpha = 0.05, 0.10, 0.15, 0.20$

- Croston’s original method + bias corrected variants with $\alpha = 0.05, 0.10, 0.15, 0.20$

- “Croston” moving averages (3&5 periods) + bias corrected variants

- ANNs (single/dual – constant/dynamic) + bias corrected variants

- 34 models in total
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean GMAE</td>
<td>2.26</td>
<td>2.91</td>
<td>3.06</td>
<td>3.42</td>
<td>3.20</td>
<td>3.02</td>
<td>3.00</td>
<td>3.31</td>
<td>3.12</td>
<td>2.99</td>
<td>3.53</td>
<td>2.57</td>
<td>3.16</td>
<td>3.25</td>
<td>3.75</td>
<td>2.72</td>
<td>3.56</td>
<td>2.84</td>
</tr>
<tr>
<td>Median GMAE</td>
<td>0.50</td>
<td>1.47</td>
<td>2.00</td>
<td>1.95</td>
<td>2.04</td>
<td>1.85</td>
<td>1.83</td>
<td>1.95</td>
<td>1.83</td>
<td>1.77</td>
<td>2.43</td>
<td>1.74</td>
<td>1.64</td>
<td>1.54</td>
<td>2.20</td>
<td>1.78</td>
<td>2.21</td>
<td>1.68</td>
</tr>
<tr>
<td>Mean MASE</td>
<td>1.06</td>
<td>0.82</td>
<td>0.92</td>
<td>1.01</td>
<td>0.92</td>
<td>0.90</td>
<td>0.90</td>
<td>0.96</td>
<td>0.92</td>
<td>0.91</td>
<td>0.98</td>
<td>0.89</td>
<td>0.97</td>
<td>0.98</td>
<td>1.04</td>
<td>0.92</td>
<td>1.00</td>
<td>0.91</td>
</tr>
<tr>
<td>Median MASE</td>
<td>0.85</td>
<td>0.65</td>
<td>0.76</td>
<td>0.77</td>
<td>0.77</td>
<td>0.74</td>
<td>0.75</td>
<td>0.77</td>
<td>0.74</td>
<td>0.73</td>
<td>0.79</td>
<td>0.70</td>
<td>0.72</td>
<td>0.76</td>
<td>0.79</td>
<td>0.73</td>
<td>0.83</td>
<td>0.73</td>
</tr>
</tbody>
</table>

- **SB-** De-bias according to Syntetos & Boylan, 01
- **SH-** De-bias according to Leven & Segerstedt, 04
- **OP-** De-biased neural networks

► Results counterintuitive → Time series methods better than Croston?

- It has been shown that this is an incorrect result → Croston is better for intermittent [Syntetos & Boylan, 01, Strijbosch et al., 00, Teunter & Sani, 08, Teunter et al., 09]

- Forecasting accuracy between **realised demand** (actuals) and **demand rate** (forecasts) → Is this correct?
Inventory Simulation Setup

- Simulate intermittent time series from empirical distributions

- 250 time series – 30 observations in-sample – 100 observations out-of-sample

- Order-up-to (T,S) policy $\rightarrow T = 1$ $\rightarrow$ Common in practice [Teunter & Sani, 08]

- Simulate service levels 80%, 90%, 95%, 99%

- Compare methods using:
  - Service levels $\alpha$ & $\beta$
  - Inventory – backlog efficiency [Teunter et. al., 09]
  - Test for statistically significant differences
Results

Service level $\alpha$ & $\beta$ differences significant at 1%, 5% and 10%
Conclusions

► ANNs provide promising results

➢ Higher service levels with equal or lower stock

► Forecast accuracy results are mixed → Are they representative of model performance?

► Direct comparisons of the models → Inventory simulation

Future research

➢ Best way to de-bias ANNs?

➢ Degree of intermittency and ANN applicability?
Thank you for your attention!

Questions?

Nikolaos Kourentzes
Lancaster University Management School
Centre for Forecasting
Lancaster, LA1 4YX, UK
Tel.  +44 (0) 7960271368
email  nikolaos@kourentzes.com