1. Time Series Forecasting of Temperatures using SARIMA: An Example from Nanjing

<https://iopscience.iop.org/article/10.1088/1757-899X/394/5/052024/meta>

Monthly mean temperature of Nanjing from January 1951 to December 2017; highly seasonal; SARIMA

**Note**:

* MSE as the metric.
* AIC (AIC (p) = n\*ln (RSS / n) + 2K, n is the number of data points and RSS is the residual sums of squares) is used to determine the parameters. Model with the minimal AIC will be selected as the best forecasting model.
* Also use ACF and PACF to determine appropriate parameters
* Data is **rescaled** to stabilize the variance using formula:

A black line with black text

Description automatically generated

* Diagnostic test to check p-value for statistically significance.
* Plots for residual: residual over time; histogram plot; Q-Q plot; ACF.
* Check Kernel Density Estimation (KDE) for normal distribution

1. Traffic Volume Prediction using Memory-Based Recurrent Neural Networks: A comparative analysis of LSTM and GRU

<https://arxiv.org/abs/2303.12643>

LSTM, GRU

* The long-short-term memory (LSTM) and gated recurrent unit (GRU) is capable of holding a long sequence of past observations and making a correlation in sequence prediction, which makes them more suitable for time-series datasets.
* Hourly traffic volume from 2012-2018, considering weather features and holidays that impact traffic volume
* Use the MinMaxScaler technique to normalize the feature values between 0 and 1
* Use interquartile range technique to remove the outlier (paper)
* Considering other features as inputs
* The main evaluation metrics are mean squared error (MSE), mean absolute error (MAE), and mean absolute percentage error (MAPE)
* Empirically, the autoregressive integrated moving average (ARIMA) is not suitable when the dataset is more complex and has long sequences. So, the author directly chose LSTM and GRU.
* Two experiment setting, all feature vs. only 4 features

1. Air-quality prediction based on the ARIMA-CNN-LSTM combination model optimized by dung beetle optimizer

<https://www.nature.com/articles/s41598-023-36620-4#:~:text=D.,has%20the%20best%20aggregation%20effect>.

ARIMA-CNN-LSTM combination model

1. Deep learning-based forecasting of electricity consumption

<https://www.nature.com/articles/s41598-024-56602-4#:~:text=The%20primary%20goal%20of%20this,period%20of%20predicted%20electricity%20consumption>.

LSTM

* Time series data refers to anything that is measured over a period of time, and its primary use is to identify trends and patterns that can be used to anticipate and draw conclusions about future outcomes.
* Weekday-weekend patterns
* A diagram of a data processing process

  Description automatically generated
* LSTM algorithm is particularly built to avoid the problem of long-term dependency. Although some can also process time series data, but they cannot manage time series data with huge delay because of the disappearance of gradient.
* For the LSTM model to begin learning, we transformed the time series into X and Y matrices. Normalize the data as a pre-processing step in a range between 0 and 1 with mean 0 and variance 1.
* The data standardization is necessary to prevent the training from diverging and make the model better fit.
* Objective function: maximize the R-squared and conversely minimize the mean absolute loss.
* The training set was further split in training and validation set that was helpful and used as a preventive measure against over-fitting.
* Tried different optimizer, finally ‘Adam’