Yale University

Sydney Weather Forecasting
-A Multivariate Time Series Perspective

Chang Shen

S&DS 583 Class at Yale

December 5, 2019

Contents

List of Figures

List of Tables

1 Introduction

1.1 Background

Weather forecasting is the application of science and technology to predict the conditions of the atmosphere for a given location and time [8], it relates closely to Production activities, social activities and daily activities. Forecasting weather requires interdisciplinary knowledge. Traditionally, the weather forecasting models were made by collecting as much data as possible about the current state of the atmosphere (temperature, humidity, rainfall meteorology understanding and wind), this processing can be complicated and only valid to short term forecasting. We want to build a model making this process more simple and explainable. Furthermore, we hope to summarize the macro features of the weather, like the trend, seasonality, etc.

This dataset contains daily weather observations from numerous Australian weather stations [1]. These observations have been taken from the Bureau of Meteorology's real-time system including the weather records of all states from 2013 to 2017. We are more interested in how the weather evolves in one specific area, thus we conducted research on forecasting weather of Sydney by applying multiple methods and hopefully to generalize the algorithms to other locations' weather change.

1.2 Objective

Time series models consider the white noise which can be used to simulate the chaotic nature of the atmosphere in real world. Univariate time series analysis methods(AR/ARMA/ARIMA/SARIMA) only captures the historical pattern of the time series. The interaction effect of multiple time series will be introduced to model time series more accurately by implementing the **Multivariate Time Series Models** (VAR/VARMA/LSTM).

A Multivariate Time Series is n time series within the same time frame,that is for any time t, $Y_t = (y_{1,t}, \cdots, y_{n,t})$. The analysis of Multivariate can be challenging since when modeling $y_{1,t}$, we need to include not only prior $\{y_{1,t-k}, \forall k\}$, but also interactions and latent information within the group of variables $\{y_{2,t-k}, \cdots, y_{n,t-k}, \forall k\}$.

The Primary objectives of the report are

- Initialize a VAR model(model selection/fitting/diagnose/forecating)
- Choose order, Diagnose, discuss over-fitting
- Introduce more models to make comparison(LSTM)
- Build an interactive shiny app with prediction model nested.

2 Exploratory Data Analysis

2.1 Original Data Statistics and Manipulation

2.1.1 Pre processing

The original data include 72588 observations of 46 different weather stations recordings, which was extracted from bureau of the Australia meteorology with date from 2013-03-01 to 2017-06-25. Extract the recordings of the Sydney weather station and drop the factor variable **rainfall** which is an indication for raining(since VAR can only be applied to numerical variables). After processing the data we use the descriptive statistics to check the distribution of the variable and detect the missing value.

2.1.2 Describe Statistics

As we can see from the describe statistics table below, the maximal temperature ranges from 11.7 °C to 40.9 °C with minor skew, minimal temperature ranges from 5 °C to 27.1 °C have 2 peaks in Sydney. Notice that the temperatures are exact to 1 decimal and the humidity, pressure and wind speed are integers. Also, notice there are some missing value in Multivariate Time Series, use the time series imputation to fill in the missing value.

weather recording of Sydney
9 Variables 1578 Observations

Date

n	missing	distinct
1578	0	1578

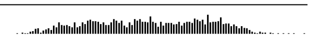
```
lowest : 2013-03-01 2013-03-02 2013-03-03 2013-03-04 2013-03-05
highest: 2017-06-21 2017-06-22 2017-06-23 2017-06-24 2017-06-25
```

MaxTemp

n	missing	distinct	Info	Mean	Gmd	.05	.10	.25	.50	.75	.90	.95
1574	4	226	1	23.53	5.017	16.7	17.9	20.3	23.4	26.4	28.8	31.1

```
lowest : 11.7 13.1 13.4 13.5 13.6, highest: 37.8 38.1 39.2 39.4 40.9
```

MinTemp



n	missing	distinct	Info	Mean	Gmd	.05	.10	.25	.50	.75	.90	.95
1574	4	192	1	15.13	5.179	7.90	8.90	11.40	15.20	19.00	20.97	21.90

```
lowest :  5.0  5.4  5.5  5.6  5.8, highest: 25.1 25.4 25.8 26.2 27.1
```

WindSpeed9am

n	missing	distinct	Info	Mean	Gmd	.05	.10	.25	.50	.75	.90	.95
1572	6	25	0.993	15.13	7.755	4	6	11	15	20	24	28

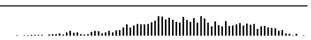
```
lowest :  0  2  4  6  7, highest: 37 39 41 44 54
```

WindSpeed3pm

n	missing	distinct	Info	Mean	Gmd	.05	.10	.25	.50	.75	.90	.95
1573	5	28	0.994	19.47	8.395	7	9	15	19	24	28	31

```
lowest :  0  2  4  6  7, highest: 43 44 46 48 57
```

Humidity9am



n	missing	distinct	Info	Mean	Gmd	.05	.10	.25	.50	.75	.90	.95
1574	4	79	1	66.69	17.32	40	48	57	67	78	87	91

```
lowest :  19  21  22  23  24, highest:  95  96  97  98 100
```

Humidity3pm

n	missing	distinct	Info	Mean	Gmd	.05	.10	.25	.50	.75	.90	.95
1574	4	85	1	53.27	17.94	27	32	43	54	63	73	83

```
lowest : 10 11 13 14 16, highest: 92 93 94 95 96
```

Pressure9am

n	missing	distinct	Info	Mean	Gmd	.05	.10	.25	.50	.75	.90	.95
1573	5	325	1	1019	8.065	1006	1009	1014	1019	1024	1028	1030

```
lowest :  996.5  996.7  998.3  998.6  999.1, highest: 1035.8 1036.2 1036.8 1038.8 1039.0
```

Pressure3pm

n	missing	distinct	Info	Mean	Gmd	.05	.10	.25	.50	.75	.90	.95
1573	5	323	1	1016	8.095	1004	1007	1012	1016	1021	1025	1028

```
lowest :  994.0  994.2  994.3  994.8  995.4, highest: 1033.7 1034.4 1034.9 1035.5 1036.0
```

Table 1: *describe statistics*

2.1.3 Missing Data Inspection and Imputation

Implement the linear Interpolation Imputation algorithms from to impute the missing value[4], which is for each univariate $y_{i,t}$ time series in the multivariate time series $Y_t = (y_{1,t}, \cdots, y_{n,t})$, use time variable t as predictor variable and $y_{i,t}$ as the response variable.

$$S(t) = \begin{cases} C_1(t), & t_0 \le t \le t_1 \\ \cdots & \\ C_i(t), & t_{i-1} < t \le t_i \\ \cdots & \\ C_n(t), & t_{n-1} < t \le t_n \end{cases}$$
$$= y_t$$

where there each $C_i = a_i + b_i t + c_i t^2 + d_i t^3 \, (d_i \neq 0)$
This consecutive processing can guarantee the interpolation value most reasonable.

2.1.4 Outlier Removal

Use the time series Outlier Detection to remove the potential outlier for miss recording, impute the outlier with the replacement generated by the algorithm.

2.2 Visualization and Smoothing

Visualize the final multivariate time series as follows, which is also available in interactive html shiny.changshen we can see from the figure below that there is no obvious trend in the weather time series of Sydney indicating the weather system is relative stable in Sydney, more methods and hypothesis test need to be included to prove whether the time series is stationary, there might be a seasonality in the time series data since it's a daily data, rigorous statistical analysis will be included in next section.

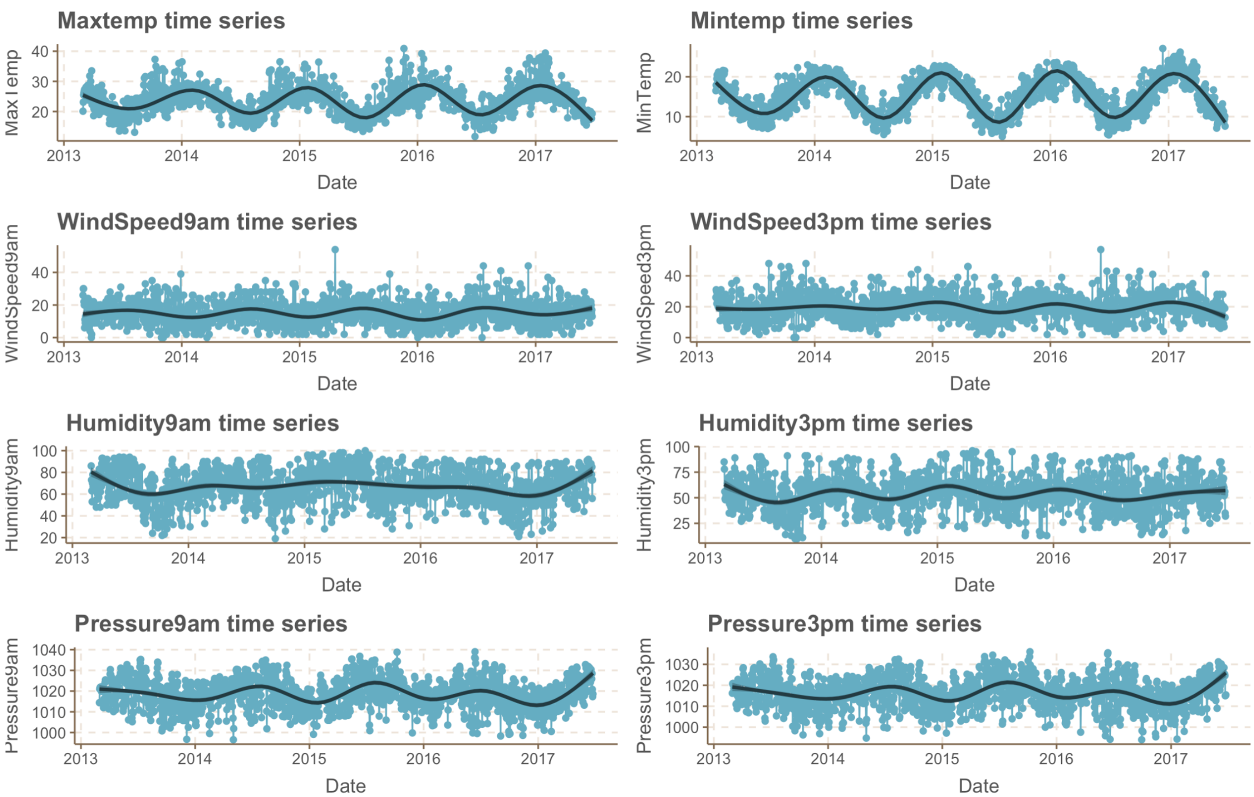


Figure 1: *Time series Visualization of variables*

2.3 Stationarity

2.3.1 Definition and methodology

Stationary means the mean of the time series wouldn't change with time and the covariance of y_t and y_s only relates to the difference between them.i.e,

$$
\begin{aligned}
&E_Y(t) = E_Y(t+\tau) && \text{for all } \tau \in R \\
&Cov_Y(t,s) = Cov_Y(t-s,0) && \text{for all } t,s \in R \\
&E\left[|X(t)|^2\right] < \infty && \text{for all } t \in R
\end{aligned}
$$

It's a presumption of most time series models including the main model we use in this report, **VAR** model, normally we use Augmented Dickey–Fuller test to test whether a unit root is present in a time series sample.The alternative hypothesis is that the time series is non-stationary[7].

2.3.2 ADF test and Transformation

We conduct ADF tests to each univatriate with in the multivariate time series separately and acquire the result as Table 2.

From the result of table 2, we reject the null hypothesis and we can conclude the weather data of Sydney is stationary in general except for the MinTemp time series. By implementing the common transformation methods: taking one order difference of lag one, we achieved the multivariate time series with all stationary variable.

Variable	Dickey-Fuller	P-value
MaxTemp	-3.9263	0.01267
MinTemp	2.7144	0.2759
WindSpeed9am	-8.327	$<< 0.01$
WindSpeed3pm	-7.4952	$<< 0.01$
Humidity9am	-7.0115	$<< 0.01$
Humidity3pm	-8.4925	$<< 0.01$
Pressure9am	-7.6682	$<< 0.01$
Pressure3pm	7.9157	$<< 0.01$

Table 2: *the ADF test result*

2.4 Seasonality

2.4.1 Visualization

This part is a calendar heatmap visualizations to show the possible seasonality for Max Temperature, We can see the June/July of a year are the months with lowest temperature and December and January of a year are months with highest temperature .There is not obvious pattern for weekends and weekdays, thus we can exclude the effect of human activities(commuting/holidays/etc) on temperature. [2],see figure 3.

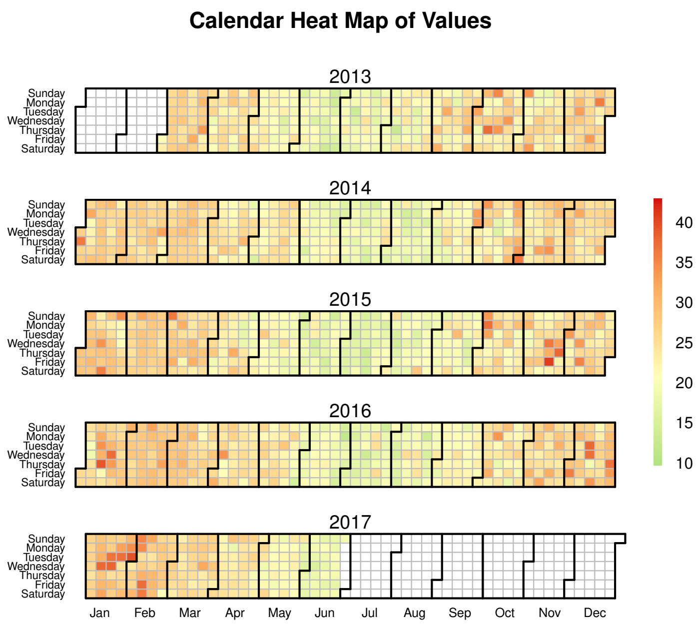


Figure 2: *Heat Map of Max Temperature*

2.4.2 Confirm the existence of Seasonality and the periodicity

To achieve statistical validity of seasonality, we conduct wo-test using the **isSeasonal** function, the result show that each variable has a seasonal pattern.
To determine whether there is multi seasonality, we also attempted the simple spectral analysis, the result shows that the dominant seasonality source from the year, with the frequency of **365.25**.

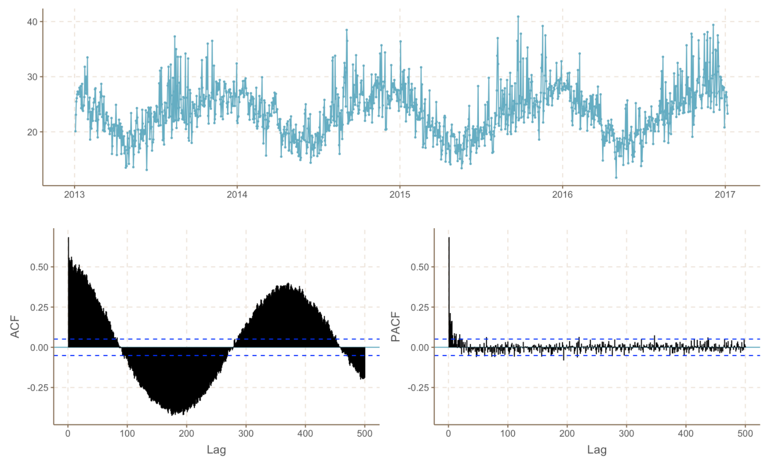


Figure 3: *Acf and Pacf of Max Temperature*

2.5 Correlation Between the the time series

Without consider the lag k correlation between two time series, i,e,$Corr(y_{i,t}, yj, t+k)$, we implement the general Pearson correlation to qualify whether the linear correlation between the multiple time series. We can see from the figure 4 below, there is a relatively strong correlation among Max Temperature and other variables, in next part, we will use the multivariate time series methods to forecast the max temperature.

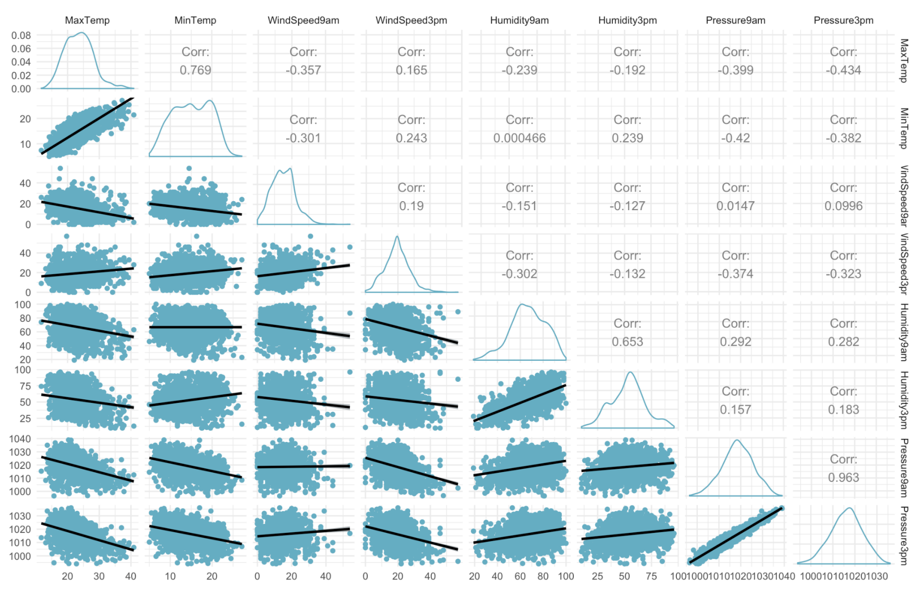


Figure 4: *Correlation between variables*

3 VAR model

3.1 Methodology

VAR(Vector Autoregression) is a linear model designed for multivariate time series. It's a generalization of the AR(Autoregression) model by including the capture of the linear interdependencies among multiple time series[6]. VAR has been applied successfully especially for describing the dynamic behavior of economic and financial time series and for forecasting[5] due to its flexibility and understandably.

The VAR model for a k dimension time series with order p has the definition as follows:

$$y_t = c + A_1 y_{t-1} + A_2 y_{t-2} + \cdots + A_p y_{t-p} + \varepsilon_t$$

where the y_{t-p} denotes ith lag of y, y_t is the vector $\{y_{1,t}, \cdots, y_{k,t}\}, \varepsilon$ is a zero-mean error term has a variance σ^2 and $cov(\varepsilon_i, \varepsilon_j) = 0, \forall i \neq j, A_i$ is a time-invariant $(k \times k)$-matrix.Similar to the **AR** model, VAR require stationary time series and the seasonality factor can be include by clarify the frequency or include dummy variables.
The model can be visualize as below

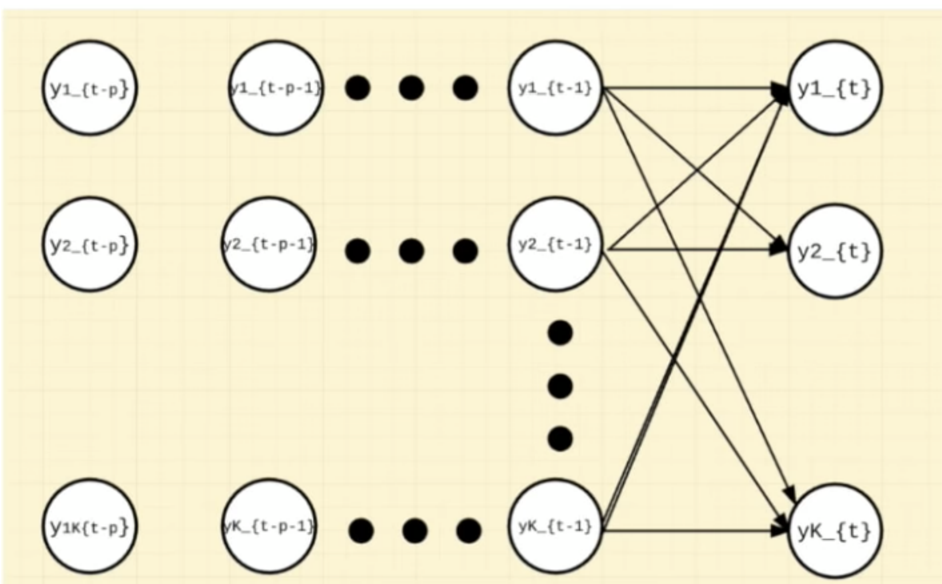


Figure 5: *the diagram of VAR[9]*

According to the meteorology knowledge and the simple correlation analysis, It's appropriate to initialize a VAR model to predict the weather change.

3.2 Include the seasonal dummy

Since in weather time series the seasonality is one of the most important factors, we want to include the seasonal predictor in the model, the R function provide the parameter **season** to realize our objective, however the frequency of the daily data 365 means there will be 365 more predictor variable in the final model, which will not only make the fitting processing longer but also will generate the high dimension problem and the over-fitting problem.

Thus, we simplify the frequency by manually creating 11 dummy variables indicating each month, by doing so, we consider the seasonality and maintain the model as simple as possible. The multivariate time

series after the dummy coding has 20 variables in total.
The final variables are
{**MaxTemp, MinTemp, WindSpeed9am, WindSpeed3pm, Humidity9am, Humidity3pm, Pressure9am, Pressure3pm, Month_03, Month_04, Month_05, Month_06, Month_07, Month_08, Month_09, Month_10, Month_11, Month_12, Month_02**}

3.3 Determine the order

Employ multiple metrics to determine the order for VAR model, the result list as Table 3, AIC and FPE tend to choose the model VAR(2), HQ and SC tend to choose model VAR(1), I fitted both models to see which one has a better prediction result.

metrics	1	2	3	4	5
AIC(n)	27.91	27.90	27.93	27.96	27.98
HQ(n)	28.03	28.11	28.25	28.38	28.51
SC(n)	28.22	28.47	28.77	29.07	29.37
FPE(n)	1.35×10^{12}	1.33×10^{12}	1.38×10^{12}	1.41×10^{12}	1.45×10^{12}

Table 3: *The order selection using different metrics*

Note:

$$AIC(n) = \ln\det\left(\tilde{\Sigma}_u(n)\right) + \frac{2}{T} nK^2$$
$$HQ(n) = \ln\det\left(\tilde{\Sigma}_u(n)\right) + \frac{2\ln(\ln(T))}{T} nK^2$$
$$SC(n) = \ln\det\left(\tilde{\Sigma}_u(n)\right) + \frac{\ln(T)}{T} nK^2$$
$$FPE(n) = \left(\frac{T+n^*}{T-n^*}\right)^K \det\left(\tilde{\Sigma}_u(n)\right)$$

With $\tilde{\Sigma}_u(n) = T^{-1}\sum_{t=1}^{T} \hat{u}_t \hat{u}_t'$ and n is the total number of the parameters in each equation[3].

3.4 Modelling

Use the first 1300 observations of each univariate time series as training set the last 278 univariate time series as the test set. For simplicity, here we only list the model for Max Temp.

3.4.1 Fitting Results

Two potential best VAR models have been listed as follows, all parameters but the wind speed are significant(***), as appendix 78.

In next part, we'll conduct the model diagnose to prove its validity and furthermore introduce another model frame, use the common metric to compare which one has a better performance in Sydney weather forecasting.

Variable	lag1
MaxTemp.l1	0.320
MinTemp.l1	0.111
WindSpeed9am.l1	−0.021
WindSpeed3pm.l1	−0.002
Humidity9am.l1	−0.017
Humidity3pm.l1	0.022
Pressure9am.l1	0.395
Pressure3pm.l1	−0.333
Month_03.l1	−0.585
Month_04.l1	−2.307
Month_05.l1	−3.295
Month_06.l1	−5.018
Month_07.l1	−4.984
Month_08.l1	−4.530
Month_09.l1	−3.115
Month_10.l1	−1.280
Month_11.l1	−1.332
Month_12.l1	−0.731
Month_02.l1	0.004
const	−47.537

Table 4: *fitting result for VAR(1)*

Variable	lag1	lag2
MaxTemp.l1	0.370	−0.099
MinTemp.l1	0.146	0.007
WindSpeed9am.l1	−0.019	−0.021
WindSpeed3pm.l1	−0.002	0.002
Humidity9am.l1	−0.007	−0.003
Humidity3pm.l1	0.026	−0.014
Pressure9am.l1	0.498	0.000
Pressure3pm.l1	−0.379	−0.075
Month_03.l1	−2.715	2.290
Month_04.l1	−1.366	−0.892
Month_05.l1	−3.466	0.338
Month_06.l1	−5.295	0.441
Month_07.l1	−4.453	−0.434
Month_08.l1	−3.508	−0.871
Month_09.l1	−4.941	2.043
Month_10.l1	−2.682	1.530
Month_11.l1	−2.735	1.466
Month_12.l1	−3.237	2.621
Month_02.l1	−0.250	0.347

Table 5: *fitting result for VAR(2)*

3.5 Model Diagnose

The presumption of the **VAR** model is that the residuals are White Noises with same variance, which means the Auto correlations function should be constantly lower than the cutoff values, here we only show the diagnose result for MaxTemp.

- **Test For Correlation**
  To test for serial correlation we applied a Portmanteau-test(Chi-squared = 2492.5, df = 3249, p-value = 1). The acf and pacf can also justify the residuals are $WN(0,\sigma^2)$ see the figure 6.
- **Test For Heteroscedasticity**
  To test whether there is heteroscedasticity in residuals, we performed a multivariate ARCH Lagrange-Multiplier test, result in Chi-squared = 244530, df = 433200, p-value = 1, proved the variances can be seen as constant in residuals.
- **Test For Normality**
  Conduct the Shapiro–Wilk test get a p-value ¿¿.99

4 More model comparsion

4.1 LSTM

Long Short Term Memory (LSTM) networks are special kind of Recurrent Neural Network (RNN) that are capable of learning long-term dependencies. In regular RNN small weights are multiplied over and over

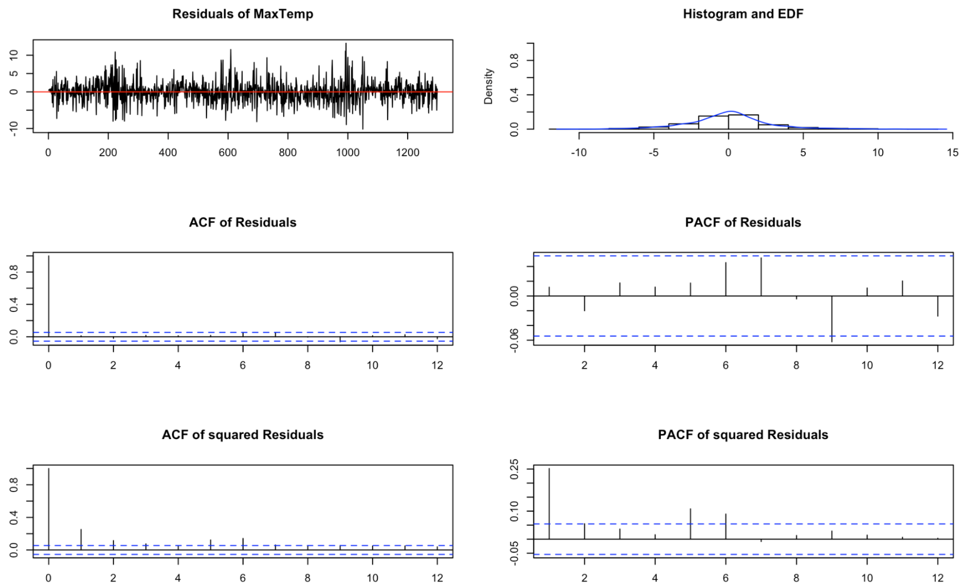


Figure 6: *The diagnose plot*

through several time steps and the gradients diminish asymptotically to zero-a condition known as vanishing gradient problem.

LSTM netowrk typically consists of memory blocks, referred to as cells, connected through layers. The information in the cells is contained in cell state Ct and hidden state ht and it is regulated by mechanisms, known as gates, through sigmoid and tanh activation functions.

Steps for fitting LSTM model

- Introduce Lag Variables(as predictor variables)
- Separate training and testing set
- Normalize the variables
- Model(define lookback/epoch/batch size,etc)

4.2 Model Performance and Forecasting

4.2.1 Fitting Criterion

With all the models we fitted in below sections, we attempt to predict the Max Temperature with historical multivariate time series on test set, we can see overall the VAR(2) has a slightly better performance compared to VAR(1), LSTM model has the worst performance this might cause by the relatively small dataset.

accuracy	rmse	mae	mape	mase
VAR(1)	3.856553	2.767140	0.1126411	100.536400
VAR(2)	3.810186	2.735609	0.1112156	99.171780
LSTM(epoch = 50)	4.473483	3.292713	0.1219252	1.230381

Table 6: *The performance of 3 models on test set*

4.2.2 Visualization of Forecasting

Visualization result of the three model forecasting, shows the VAR model can perfectly capture the changing pattern of the time series, while LSTM model might has a problem of lack of fitting we may need to conduct more parameter tuning/more epochs to get better performance, while LSTM is excellent in abnormal detection.

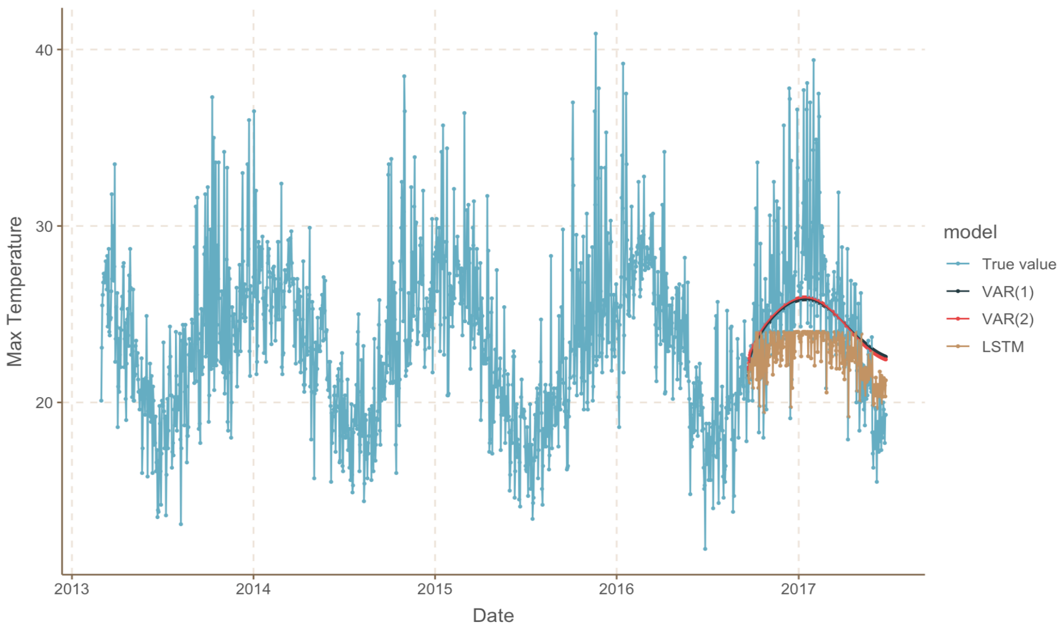


Figure 7: *Prediction of three models*

5 Discussion

5.1 VAR(Vector AutoRgression)

- Linearity character make it simple in format and can be explained.
- VAR is a powerful algorithm but it has a limitation since it is only applicable to numeric variables.
- Have to manually choose the order p
- Require the stationary presumption (may need transformation)
- Require the residuals be white noise

5.2 LSTM(Long Short Time Memory)

- A black box model can't be written as VAR
- Parameter tuning/ take long training time and more CPU memory
- More suitable for large scale data
- Can be applied to multiple scenarios like forecasting, classification, anomalously detection

References

[1] *Australian Government Bureau of Meteorology*. URL: http://www.bom.gov.au/climate/dwo/IDCJDW0000.shtml (visited on 09/30/2019).

[2] Sarang Gupta. *Time-Series Calendar Heatmaps*. en. Jan. 2019. URL: https://towardsdatascience.com/time-series-calendar-heatmaps-9f576578fcfe (visited on 11/29/2019).

[3] Edward J Hannan and Barry G Quinn. "The determination of the order of an autoregression". In: *Journal of the Royal Statistical Society: Series B (Methodological)* 41.2 (1979), pp. 190–195.

[4] Steffen Moritz and Thomas Bartz-Beielstein. "imputeTS: Time Series Missing Value Imputation in R". en. In: *The R Journal* 9.1 (2017), pp. 207–218. ISSN: 2073-4859. URL: https://journal.r-project.org/archive/2017/RJ-2017-009/index.html (visited on 11/28/2019).

[5] *multivariatetimeseries*. URL: https://faculty.washington.edu/ezivot/econ584/notes/multivariatetimeseries.pdf (visited on 09/30/2019).

[6] Peter CB Phillips et al. "Fully modified least squares and vector autoregression". In: *ECONOMETRICA-EVANSTON ILL-* 63 (1995), pp. 1023–1023.

[7] SAID E. SAID and DAVID A. DICKEY. "Testing for unit roots in autoregressive-moving average models of unknown order". In: *Biometrika* 71.3 (Dec. 1984), pp. 599–607. ISSN: 0006-3444. DOI: 10.1093/biomet/71.3.599. eprint: http://oup.prod.sis.lan/biomet/article-pdf/71/3/599/719376/71-3-599.pdf. URL: https://doi.org/10.1093/biomet/71.3.599.

[8] Wikipedia. *Weather_forecasting*. 2019. URL: https://www.wikiwand.com/en/Weather_forecasting (visited on 09/30/2010).

[9] Jeffrey Yau. *Time Series Forecasting using Statistical and Machine Learning Models — PyData NYC 2017*. Youtube. 2015. URL: https://www.youtube.com/watch?v=_vQ0W_qXMxk&t=10s.

Appendices

An Appendix of Some Kind

Weather Data in Australia
11 Variables 72588 Observations

Date

n	missing	distinct
72588	0	1578

```
lowest : 2013-03-01 2013-03-02 2013-03-03 2013-03-04 2013-03-05
highest: 2017-06-21 2017-06-22 2017-06-23 2017-06-24 2017-06-25
```

Location

n	missing	distinct
72588	0	46

```
lowest : Adelaide      Albany        Albury        AliceSprings  BadgerysCreek
highest: Watsonia      Williamtown   Witchcliffe   Wollongong    Woomera
```

MaxTemp

[2]	n	missing	distinct	Info	Mean	Gmd	.05	.10	.25	.50	.75	.90	.95
	70505	2083	493	1	23.45	8.201	12.90	14.54	18.00	23.00	28.60	33.30	35.70

`lowest : -4.8 -4.1 -3.8 -3.7 -3.2, highest: 46.6 46.8 46.9 47.0 47.3`

MinTemp

n	missing	distinct	Info	Mean	Gmd	.05	.10	.25	.50	.75	.90	.95
70385	2203	378	1	12.25	7.285	1.9	4.1	7.7	12.0	16.8	20.9	23.3

`lowest : -8.2 -7.8 -7.6 -7.5 -7.3, highest: 30.7 31.2 31.4 31.8 31.9`

Rainfall

[2]	n	missing	distinct	Info	Mean	Gmd	.05	.10	.25	.50	.75	.90	.95
	69942	2646	532	0.727	2.241	4.006	0.00	0.00	0.00	0.00	0.60	5.80	12.59

`lowest :   0.0   0.1   0.2   0.3   0.4, highest: 216.3 219.6 225.0 240.0 247.2`

WindSpeed9am

n	missing	distinct	Info	Mean	Gmd	.05	.10	.25	.50	.75	.90	.95
70158	2430	41	0.995	13.6	9.69	0	4	7	13	19	26	30

`lowest :   0   2   4   6   7, highest:  67  69  74  87 130`

WindSpeed3pm

n	missing	distinct	Info	Mean	Gmd	.05	.10	.25	.50	.75	.90	.95
68887	3701	40	0.995	18.37	9.746	6	7	11	17	24	30	33

`lowest :  0  2  4  6  7, highest: 65 69 72 74 83`

Humidity9am

n	missing	distinct	Info	Mean	Gmd	.05	.10	.25	.50	.75	.90	.95
69940	2648	101	1	69.18	21.59	34	44	57	70	84	95	98

`lowest :   0   1   2   3   4, highest:  96  97  98  99 100`

Humidity3pm

n	missing	distinct	Info	Mean	Gmd	.05	.10	.25	.50	.75	.90	.95
67976	4612	101	1	51.43	23.7	17	23	37	52	65	79	88

`lowest :   0   1   2   3   4, highest:  96  97  98  99 100`

Pressure9am

n	missing	distinct	Info	Mean	Gmd	.05	.10	.25	.50	.75	.90	.95
62836	9752	517	1	1018	7.997	1006	1009	1013	1018	1023	1027	1030

`lowest :  982.0  982.2  982.3  982.9  983.9, highest: 1039.2 1039.3 1039.9 1040.1 1040.3`

Pressure3pm

n	missing	distinct	Info	Mean	Gmd	.05	.10	.25	.50	.75	.90	.95
62842	9746	503	1	1015	7.97	1004	1007	1011	1015	1020	1025	1027

`lowest :  977.1  978.2  981.4  981.9  982.2, highest: 1036.8 1036.9 1037.0 1037.1 1037.3`

---

round	Estimate	Std. Error	t value	Pr(>| t |)
MaxTemp.l1	0.32	0.04	8.21	0.00
MinTemp.l1	0.11	0.05	2.36	0.02
WindSpeed9am.l1	−0.02	0.01	−1.62	0.11
WindSpeed3pm.l1	0.00	0.01	−0.19	0.85
Humidity9am.l1	−0.02	0.01	−2.21	0.03
Humidity3pm.l1	0.02	0.01	2.68	0.01
Pressure9am.l1	0.40	0.05	8.20	0.00
Pressure3pm.l1	−0.33	0.05	−7.03	0.00
$\text{Month}_0 3.l1$	−0.58	0.38	−1.52	0.13
$\text{Month}_0 4.l1$	−2.31	0.42	−5.55	0.00
$\text{Month}_0 5.l1$	−3.29	0.47	−7.05	0.00
$\text{Month}_0 6.l1$	−5.02	0.52	−9.66	0.00
$\text{Month}_0 7.l1$	−4.98	0.55	−9.05	0.00
$\text{Month}_0 8.l1$	−4.53	0.53	−8.58	0.00
$\text{Month}_0 9.l1$	−3.12	0.47	−6.58	0.00
$\text{Month}_1 0.l1$	−1.28	0.44	−2.92	0.00
$\text{Month}_1 1.l1$	−1.33	0.42	−3.15	0.00
$\text{Month}_1 2.l1$	−0.73	0.41	−1.78	0.07
$\text{Month}_0 2.l1$	0.00	0.42	0.01	0.99
const	−47.54	13.61	−3.49	0.00

Table 7: *VAR(1) summary*

round	Estimate	Std. Error	t value	Pr(>| t |)
MaxTemp.l1	0.37	0.04	8.99	0.00
MinTemp.l1	0.15	0.06	2.42	0.02
WindSpeed9am.l1	−0.02	0.01	−1.42	0.16
WindSpeed3pm.l1	0.00	0.01	−0.16	0.87
Humidity9am.l1	−0.01	0.01	−0.90	0.37
Humidity3pm.l1	0.03	0.01	3.17	0.00
Pressure9am.l1	0.50	0.07	7.62	0.00
Pressure3pm.l1	−0.38	0.05	−7.35	0.00
Month$_0$3.l1	−2.71	2.07	−1.31	0.19
Month$_0$4.l1	−1.37	2.32	−0.59	0.56
Month$_0$5.l1	−3.47	2.50	−1.39	0.17
Month$_0$6.l1	−5.30	2.60	−2.04	0.04
Month$_0$7.l1	−4.45	2.66	−1.68	0.09
Month$_0$8.l1	−3.51	2.64	−1.33	0.18
Month$_0$9.l1	−4.94	2.57	−1.92	0.05
Month$_1$0.l1	−2.68	2.38	−1.13	0.26
Month$_1$1.l1	−2.73	2.07	−1.32	0.19
Month$_1$2.l1	−3.24	1.54	−2.10	0.04
Month$_0$2.l1	−0.25	1.55	−0.16	0.87
MaxTemp.l2	−0.10	0.05	−2.19	0.03
MinTemp.l2	0.01	0.05	0.13	0.90
WindSpeed9am.l2	−0.02	0.01	−1.58	0.11
WindSpeed3pm.l2	0.00	0.01	0.20	0.84
Humidity9am.l2	0.00	0.01	−0.36	0.72
Humidity3pm.l2	−0.01	0.01	−1.55	0.12
Pressure9am.l2	0.00	0.06	0.01	0.99
Pressure3pm.l2	−0.07	0.07	−1.06	0.29
Month$_0$3.l2	2.29	2.07	1.11	0.27
Month$_0$4.l2	−0.89	2.32	−0.38	0.70
Month$_0$5.l2	0.34	2.50	0.14	0.89
Month$_0$6.l2	0.44	2.60	0.17	0.87
Month$_0$7.l2	−0.43	2.64	−0.16	0.87
Month$_0$8.l2	−0.87	2.64	−0.33	0.74
Month$_0$9.l2	2.04	2.57	0.79	0.43
Month$_1$0.l2	1.53	2.39	0.64	0.52
Month$_1$1.l2	1.47	2.08	0.71	0.48
Month$_1$2.l2	2.62	1.55	1.70	0.09
Month$_0$2.l2	0.35	1.55	0.22	0.82
const	−28.54	15.43	−1.85	0.06

Table 8: *VAR(2) summary*