Pollen MSE

May 31, 2019

```
In [31]: import warnings
         warnings.filterwarnings('ignore')
         import pandas as pd
         import matplotlib.pyplot as plt
         import seaborn as sns
         sns.set(font='IPAGothic')
         import numpy as np
         import statsmodels.api as sm
         from sklearn.metrics import mean_squared_error

In [7]: data = pd.read_excel('cleaned_data/Pollen_full_data_Eve.xlsx', parse_dates=['date_time']

In [9]: data = data['pollenGlobalIndex']

In [23]: print(data.index.min(), data.index.max())

2017-09-06 14:00:00 2019-05-20 11:00:00


In [89]: a = data['2017-09-06 15:00:00':]

In [99]: print(data.index.min(), data.index.max())

2017-09-06 15:00:00 2019-05-20 11:00:00


In [90]: a.head()

Out[90]: date_time
         2017-09-06 15:00:00    9
         2017-09-06 16:00:00    9
         2017-09-06 17:00:00    9
         2017-09-06 18:00:00    9
         2017-09-06 19:00:00    9
         Name: pollenGlobalIndex, dtype: int64

In [91]: a.count()

Out[91]: 33234
```

```
In [85]: data.count()

Out[85]: 33235

In [92]: data = a

In [93]: data.count()

Out[93]: 33234

In [98]: data[16617:]

Out[98]: date_time
         2018-03-05 17:00:00     9
         2018-03-05 18:00:00     9
         2018-03-05 19:00:00    10
         2018-03-05 20:00:00    10
         2018-03-05 21:00:00    10
         2018-03-05 22:00:00    10
         2018-03-05 23:00:00    10
         2018-03-06 00:00:00     9
         2018-03-06 01:00:00     9
         2018-03-06 02:00:00     9
         2018-03-06 03:00:00    10
         2018-03-06 04:00:00     9
         2018-03-06 05:00:00     9
         2018-03-06 06:00:00     9
         2018-03-06 07:00:00    10
         2018-03-06 08:00:00     9
         2018-03-06 09:00:00     9
         2018-03-06 10:00:00     9
         2018-03-06 11:00:00     9
         2018-03-06 12:00:00     9
         2018-03-06 13:00:00     9
         2018-03-06 14:00:00     9
         2018-03-06 15:00:00    10
         2018-03-06 16:00:00     9
         2018-03-06 17:00:00    10
         2018-03-06 18:00:00     9
         2018-03-06 19:00:00     9
         2018-03-06 20:00:00     9
         2018-03-06 21:00:00     9
         2018-03-06 22:00:00     9
                                ..
         2017-10-26 12:00:00     1
         2017-11-06 16:00:00     1
         2017-11-07 12:00:00     1
         2017-11-07 13:00:00     1
         2017-11-07 14:00:00     1
```

```
2017-11-16 15:00:00     1
2017-11-16 16:00:00     1
2017-11-16 17:00:00     1
2017-11-16 18:00:00     3
2017-11-16 19:00:00     3
2017-11-16 20:00:00     3
2017-11-16 21:00:00     3
2017-11-16 22:00:00     3
2017-11-16 23:00:00     3
2017-11-17 00:00:00     1
2017-11-17 01:00:00     1
2017-11-17 02:00:00     1
2017-11-17 03:00:00     1
2017-11-17 04:00:00     1
2017-11-17 05:00:00     1
2017-11-17 06:00:00     1
2017-11-17 07:00:00     1
2017-11-17 08:00:00     1
2017-11-17 09:00:00     1
2017-11-17 10:00:00     1
2017-11-17 11:00:00     1
2017-11-17 13:00:00     1
2017-11-17 14:00:00     1
2017-11-17 15:00:00     1
2017-11-17 16:00:00     1
Name: pollenGlobalIndex, Length: 16617, dtype: int64
```

```
In [94]: 33234/2
```

```
Out[94]: 16617.0
```

```
In [70]: data['2019-05-18 15:00:00']
```

```
Out[70]: date_time
         2019-05-18 15:00:00    9
         Name: pollenGlobalIndex, dtype: int64
```

```
In [100]: tr_start, tr_end = '2017-09-06 15:00:00', '2018-03-05 17:00:00'
          te_start, te_end = '2018-03-05 18:00:00', '2019-05-20 11:00:00'
```

```
In [101]: tra = data[tr_start:tr_end].dropna()
          tes = data[te_start:te_end].dropna()
```

```
In [103]: tra.count()
```

```
Out[103]: 22411
```

```
In [104]: tes.count()
```

```
Out[104]: 22411
```

```
In [28]: resDiff = sm.tsa.arma_order_select_ic(tra, max_ar=7, max_ma=7, ic='aic', trend='c')
         print('ARMA(p,q) =',resDiff['aic_min_order'],'is the best.')
```

```
/home/omar/.local/lib/python3.5/site-packages/statsmodels/tsa/base/tsa_model.py:225: ValueWarnin
  ' ignored when e.g. forecasting.', ValueWarning)
/home/omar/.local/lib/python3.5/site-packages/statsmodels/tsa/base/tsa_model.py:225: ValueWarnin
  ' ignored when e.g. forecasting.', ValueWarning)
/home/omar/.local/lib/python3.5/site-packages/statsmodels/tsa/base/tsa_model.py:225: ValueWarnin
  ' ignored when e.g. forecasting.', ValueWarning)
/home/omar/.local/lib/python3.5/site-packages/statsmodels/tsa/base/tsa_model.py:225: ValueWarnin
  ' ignored when e.g. forecasting.', ValueWarning)
/home/omar/.local/lib/python3.5/site-packages/statsmodels/tsa/base/tsa_model.py:225: ValueWarnin
  ' ignored when e.g. forecasting.', ValueWarning)
/home/omar/.local/lib/python3.5/site-packages/statsmodels/tsa/base/tsa_model.py:225: ValueWarnin
  ' ignored when e.g. forecasting.', ValueWarning)
/home/omar/.local/lib/python3.5/site-packages/statsmodels/tsa/base/tsa_model.py:225: ValueWarnin
  ' ignored when e.g. forecasting.', ValueWarning)
/home/omar/.local/lib/python3.5/site-packages/statsmodels/tsa/base/tsa_model.py:225: ValueWarnin
  ' ignored when e.g. forecasting.', ValueWarning)
/home/omar/.local/lib/python3.5/site-packages/statsmodels/tsa/base/tsa_model.py:225: ValueWarnin
  ' ignored when e.g. forecasting.', ValueWarning)
/home/omar/.local/lib/python3.5/site-packages/statsmodels/tsa/base/tsa_model.py:225: ValueWarnin
  ' ignored when e.g. forecasting.', ValueWarning)
/home/omar/.local/lib/python3.5/site-packages/statsmodels/tsa/base/tsa_model.py:225: ValueWarnin
  ' ignored when e.g. forecasting.', ValueWarning)
/home/omar/.local/lib/python3.5/site-packages/statsmodels/tsa/base/tsa_model.py:225: ValueWarnin
  ' ignored when e.g. forecasting.', ValueWarning)
/home/omar/.local/lib/python3.5/site-packages/statsmodels/tsa/base/tsa_model.py:225: ValueWarnin
  ' ignored when e.g. forecasting.', ValueWarning)
/home/omar/.local/lib/python3.5/site-packages/statsmodels/tsa/base/tsa_model.py:225: ValueWarnin
  ' ignored when e.g. forecasting.', ValueWarning)
/home/omar/.local/lib/python3.5/site-packages/statsmodels/tsa/base/tsa_model.py:225: ValueWarnin
  ' ignored when e.g. forecasting.', ValueWarning)
/home/omar/.local/lib/python3.5/site-packages/statsmodels/tsa/base/tsa_model.py:225: ValueWarnin
  ' ignored when e.g. forecasting.', ValueWarning)
/home/omar/.local/lib/python3.5/site-packages/statsmodels/tsa/base/tsa_model.py:225: ValueWarnin
  ' ignored when e.g. forecasting.', ValueWarning)
/home/omar/.local/lib/python3.5/site-packages/statsmodels/tsa/base/tsa_model.py:225: ValueWarnin
  ' ignored when e.g. forecasting.', ValueWarning)
/home/omar/.local/lib/python3.5/site-packages/statsmodels/tsa/base/tsa_model.py:225: ValueWarnin
  ' ignored when e.g. forecasting.', ValueWarning)
/home/omar/.local/lib/python3.5/site-packages/statsmodels/tsa/base/tsa_model.py:225: ValueWarnin
  ' ignored when e.g. forecasting.', ValueWarning)
/home/omar/.local/lib/python3.5/site-packages/statsmodels/tsa/base/tsa_model.py:225: ValueWarnin
  ' ignored when e.g. forecasting.', ValueWarning)
/home/omar/.local/lib/python3.5/site-packages/statsmodels/tsa/base/tsa_model.py:225: ValueWarnin
```

```
  ' ignored when e.g. forecasting.', ValueWarning)
/home/omar/.local/lib/python3.5/site-packages/statsmodels/tsa/base/tsa_model.py:225: ValueWarnin
  ' ignored when e.g. forecasting.', ValueWarning)
/home/omar/.local/lib/python3.5/site-packages/statsmodels/tsa/base/tsa_model.py:225: ValueWarnin
  ' ignored when e.g. forecasting.', ValueWarning)
/home/omar/.local/lib/python3.5/site-packages/statsmodels/tsa/base/tsa_model.py:225: ValueWarnin
  ' ignored when e.g. forecasting.', ValueWarning)
/home/omar/.local/lib/python3.5/site-packages/statsmodels/tsa/base/tsa_model.py:225: ValueWarnin
  ' ignored when e.g. forecasting.', ValueWarning)
/home/omar/.local/lib/python3.5/site-packages/statsmodels/tsa/base/tsa_model.py:225: ValueWarnin
  ' ignored when e.g. forecasting.', ValueWarning)
/home/omar/.local/lib/python3.5/site-packages/statsmodels/base/model.py:488: HessianInversionWar
  'available', HessianInversionWarning)
/home/omar/.local/lib/python3.5/site-packages/statsmodels/base/model.py:508: ConvergenceWarning:
  "Check mle_retvals", ConvergenceWarning)
/home/omar/.local/lib/python3.5/site-packages/statsmodels/tsa/base/tsa_model.py:225: ValueWarnin
  ' ignored when e.g. forecasting.', ValueWarning)
/home/omar/.local/lib/python3.5/site-packages/statsmodels/base/model.py:488: HessianInversionWar
  'available', HessianInversionWarning)
/home/omar/.local/lib/python3.5/site-packages/statsmodels/base/model.py:508: ConvergenceWarning:
  "Check mle_retvals", ConvergenceWarning)
/home/omar/.local/lib/python3.5/site-packages/statsmodels/tsa/base/tsa_model.py:225: ValueWarnin
  ' ignored when e.g. forecasting.', ValueWarning)
/home/omar/.local/lib/python3.5/site-packages/statsmodels/tsa/base/tsa_model.py:225: ValueWarnin
  ' ignored when e.g. forecasting.', ValueWarning)
/home/omar/.local/lib/python3.5/site-packages/statsmodels/tsa/base/tsa_model.py:225: ValueWarnin
  ' ignored when e.g. forecasting.', ValueWarning)
/home/omar/.local/lib/python3.5/site-packages/statsmodels/tsa/base/tsa_model.py:225: ValueWarnin
  ' ignored when e.g. forecasting.', ValueWarning)
/home/omar/.local/lib/python3.5/site-packages/statsmodels/tsa/base/tsa_model.py:225: ValueWarnin
  ' ignored when e.g. forecasting.', ValueWarning)
/home/omar/.local/lib/python3.5/site-packages/statsmodels/tsa/base/tsa_model.py:225: ValueWarnin
  ' ignored when e.g. forecasting.', ValueWarning)
/home/omar/.local/lib/python3.5/site-packages/statsmodels/tsa/base/tsa_model.py:225: ValueWarnin
  ' ignored when e.g. forecasting.', ValueWarning)
/home/omar/.local/lib/python3.5/site-packages/statsmodels/tsa/base/tsa_model.py:225: ValueWarnin
  ' ignored when e.g. forecasting.', ValueWarning)
/home/omar/.local/lib/python3.5/site-packages/statsmodels/tsa/base/tsa_model.py:225: ValueWarnin
  ' ignored when e.g. forecasting.', ValueWarning)
/home/omar/.local/lib/python3.5/site-packages/statsmodels/tsa/base/tsa_model.py:225: ValueWarnin
  ' ignored when e.g. forecasting.', ValueWarning)
/home/omar/.local/lib/python3.5/site-packages/statsmodels/tsa/base/tsa_model.py:225: ValueWarnin
  ' ignored when e.g. forecasting.', ValueWarning)
/home/omar/.local/lib/python3.5/site-packages/statsmodels/tsa/base/tsa_model.py:225: ValueWarnin
  ' ignored when e.g. forecasting.', ValueWarning)
/home/omar/.local/lib/python3.5/site-packages/statsmodels/tsa/base/tsa_model.py:225: ValueWarnin
```

```
  ' ignored when e.g. forecasting.', ValueWarning)
/home/omar/.local/lib/python3.5/site-packages/statsmodels/tsa/base/tsa_model.py:225: ValueWarnin
  ' ignored when e.g. forecasting.', ValueWarning)
/home/omar/.local/lib/python3.5/site-packages/statsmodels/tsa/base/tsa_model.py:225: ValueWarnin
  ' ignored when e.g. forecasting.', ValueWarning)
/home/omar/.local/lib/python3.5/site-packages/statsmodels/tsa/base/tsa_model.py:225: ValueWarnin
  ' ignored when e.g. forecasting.', ValueWarning)
/home/omar/.local/lib/python3.5/site-packages/statsmodels/tsa/base/tsa_model.py:225: ValueWarnin
  ' ignored when e.g. forecasting.', ValueWarning)
/home/omar/.local/lib/python3.5/site-packages/statsmodels/base/model.py:488: HessianInversionWar
  'available', HessianInversionWarning)
/home/omar/.local/lib/python3.5/site-packages/statsmodels/tsa/base/tsa_model.py:225: ValueWarnin
  ' ignored when e.g. forecasting.', ValueWarning)
/home/omar/.local/lib/python3.5/site-packages/statsmodels/base/model.py:488: HessianInversionWar
  'available', HessianInversionWarning)
/home/omar/.local/lib/python3.5/site-packages/statsmodels/tsa/base/tsa_model.py:225: ValueWarnin
  ' ignored when e.g. forecasting.', ValueWarning)
/home/omar/.local/lib/python3.5/site-packages/statsmodels/tsa/base/tsa_model.py:225: ValueWarnin
  ' ignored when e.g. forecasting.', ValueWarning)
/home/omar/.local/lib/python3.5/site-packages/statsmodels/tsa/base/tsa_model.py:225: ValueWarnin
  ' ignored when e.g. forecasting.', ValueWarning)
/home/omar/.local/lib/python3.5/site-packages/statsmodels/tsa/base/tsa_model.py:225: ValueWarnin
  ' ignored when e.g. forecasting.', ValueWarning)
/home/omar/.local/lib/python3.5/site-packages/statsmodels/tsa/base/tsa_model.py:225: ValueWarnin
  ' ignored when e.g. forecasting.', ValueWarning)
/home/omar/.local/lib/python3.5/site-packages/statsmodels/base/model.py:508: ConvergenceWarning:
  "Check mle_retvals", ConvergenceWarning)
/home/omar/.local/lib/python3.5/site-packages/statsmodels/tsa/base/tsa_model.py:225: ValueWarnin
  ' ignored when e.g. forecasting.', ValueWarning)
/home/omar/.local/lib/python3.5/site-packages/statsmodels/base/model.py:508: ConvergenceWarning:
  "Check mle_retvals", ConvergenceWarning)
/home/omar/.local/lib/python3.5/site-packages/statsmodels/tsa/base/tsa_model.py:225: ValueWarnin
  ' ignored when e.g. forecasting.', ValueWarning)
/home/omar/.local/lib/python3.5/site-packages/statsmodels/tsa/base/tsa_model.py:225: ValueWarnin
  ' ignored when e.g. forecasting.', ValueWarning)
/home/omar/.local/lib/python3.5/site-packages/statsmodels/tsa/base/tsa_model.py:225: ValueWarnin
  ' ignored when e.g. forecasting.', ValueWarning)
/home/omar/.local/lib/python3.5/site-packages/statsmodels/tsa/base/tsa_model.py:225: ValueWarnin
  ' ignored when e.g. forecasting.', ValueWarning)
/home/omar/.local/lib/python3.5/site-packages/statsmodels/tsa/base/tsa_model.py:225: ValueWarnin
  ' ignored when e.g. forecasting.', ValueWarning)
/home/omar/.local/lib/python3.5/site-packages/statsmodels/tsa/base/tsa_model.py:225: ValueWarnin
  ' ignored when e.g. forecasting.', ValueWarning)
/home/omar/.local/lib/python3.5/site-packages/statsmodels/tsa/base/tsa_model.py:225: ValueWarnin
  ' ignored when e.g. forecasting.', ValueWarning)
/home/omar/.local/lib/python3.5/site-packages/statsmodels/tsa/base/tsa_model.py:225: ValueWarnin
  ' ignored when e.g. forecasting.', ValueWarning)
/home/omar/.local/lib/python3.5/site-packages/statsmodels/tsa/base/tsa_model.py:225: ValueWarnin
```

```
  ' ignored when e.g. forecasting.', ValueWarning)
/home/omar/.local/lib/python3.5/site-packages/statsmodels/tsa/base/tsa_model.py:225: ValueWarnin
  ' ignored when e.g. forecasting.', ValueWarning)
/home/omar/.local/lib/python3.5/site-packages/statsmodels/tsa/base/tsa_model.py:225: ValueWarnin
  ' ignored when e.g. forecasting.', ValueWarning)
/home/omar/.local/lib/python3.5/site-packages/statsmodels/tsa/base/tsa_model.py:225: ValueWarnin
  ' ignored when e.g. forecasting.', ValueWarning)


ARMA(p,q) = (7, 7) is the best.
```

Parameters

```
In [30]: arima = sm.tsa.statespace.SARIMAX(tra,order=(7,1,7),seasonal_order=(0,0,0,0),
                                           enforce_stationarity=False, enforce_invertibility=Fals
         arima.summary()

/home/omar/.local/lib/python3.5/site-packages/statsmodels/tsa/base/tsa_model.py:225: ValueWarnin
  ' ignored when e.g. forecasting.', ValueWarning)
/home/omar/.local/lib/python3.5/site-packages/statsmodels/base/model.py:508: ConvergenceWarning:
  "Check mle_retvals", ConvergenceWarning)


Out[30]: <class 'statsmodels.iolib.summary.Summary'>
         """
                                  Statespace Model Results
         ==============================================================================
         Dep. Variable:      pollenGlobalIndex   No. Observations:                30344
         Model:               SARIMAX(7, 1, 7)   Log Likelihood             -45555.644
         Date:                Thu, 30 May 2019   AIC                          91141.288
         Time:                        22:38:07   BIC                          91266.089
         Sample:                             0   HQIC                         91181.311
                                         - 30344
         Covariance Type:                  opg
         ==============================================================================
                          coef    std err          z      P>|z|      [0.025      0.975]
         ------------------------------------------------------------------------------
         ar.L1         -0.3179      0.011    -29.973      0.000      -0.339      -0.297
         ar.L2         -0.1826      0.009    -19.276      0.000      -0.201      -0.164
         ar.L3          0.1550      0.010     14.772      0.000       0.134       0.176
         ar.L4         -0.2773      0.011    -25.041      0.000      -0.299      -0.256
         ar.L5         -0.1478      0.008    -17.755      0.000      -0.164      -0.131
         ar.L6         -0.2519      0.007    -34.589      0.000      -0.266      -0.238
         ar.L7          0.6020      0.008     73.819      0.000       0.586       0.618
         ma.L1          0.1330      0.010     13.464      0.000       0.114       0.152
         ma.L2          0.0894      0.009     10.286      0.000       0.072       0.106
         ma.L3         -0.3877      0.009    -42.449      0.000      -0.406      -0.370
         ma.L4          0.1502      0.010     14.974      0.000       0.131       0.170
```

```
ma.L5          0.0361      0.008      4.311      0.000       0.020       0.052
ma.L6          0.0083      0.009      0.944      0.345      -0.009       0.026
ma.L7         -0.9067      0.008   -107.486      0.000      -0.923      -0.890
sigma2         1.0554      0.011     92.209      0.000       1.033       1.078
===================================================================================
Ljung-Box (Q):                     5985.72   Jarque-Bera (JB):            375630.19
Prob(Q):                              0.00   Prob(JB):                         0.00
Heteroskedasticity (H):               0.93   Skew:                             0.40
Prob(H) (two-sided):                  0.00   Kurtosis:                        20.22
===================================================================================

Warnings:
[1] Covariance matrix calculated using the outer product of gradients (complex-step).
"""
```

pred = arima.predict(tr_end,te_end, dynamic=True)[1:] print('ARIMA model MSE:{}'.format(mean_squared_error(tes,pred)))

```
In [115]: pred = arima.predict(10823,33234, dynamic=True)[1:]
          print('ARIMA model MSE:{}'.format(mean_squared_error(tes,pred)))
```

```
ARIMA model MSE:29.15488319030185
```