VIX Forecasting

Amrita,Ben,Preethi,Tyler,Vinit

Table of Contents

- **Data Exploration**
- **Roadmap**
- **Forecasting - VECM**
- **Forecasting - RNN**
- **Forecasting - Naive**
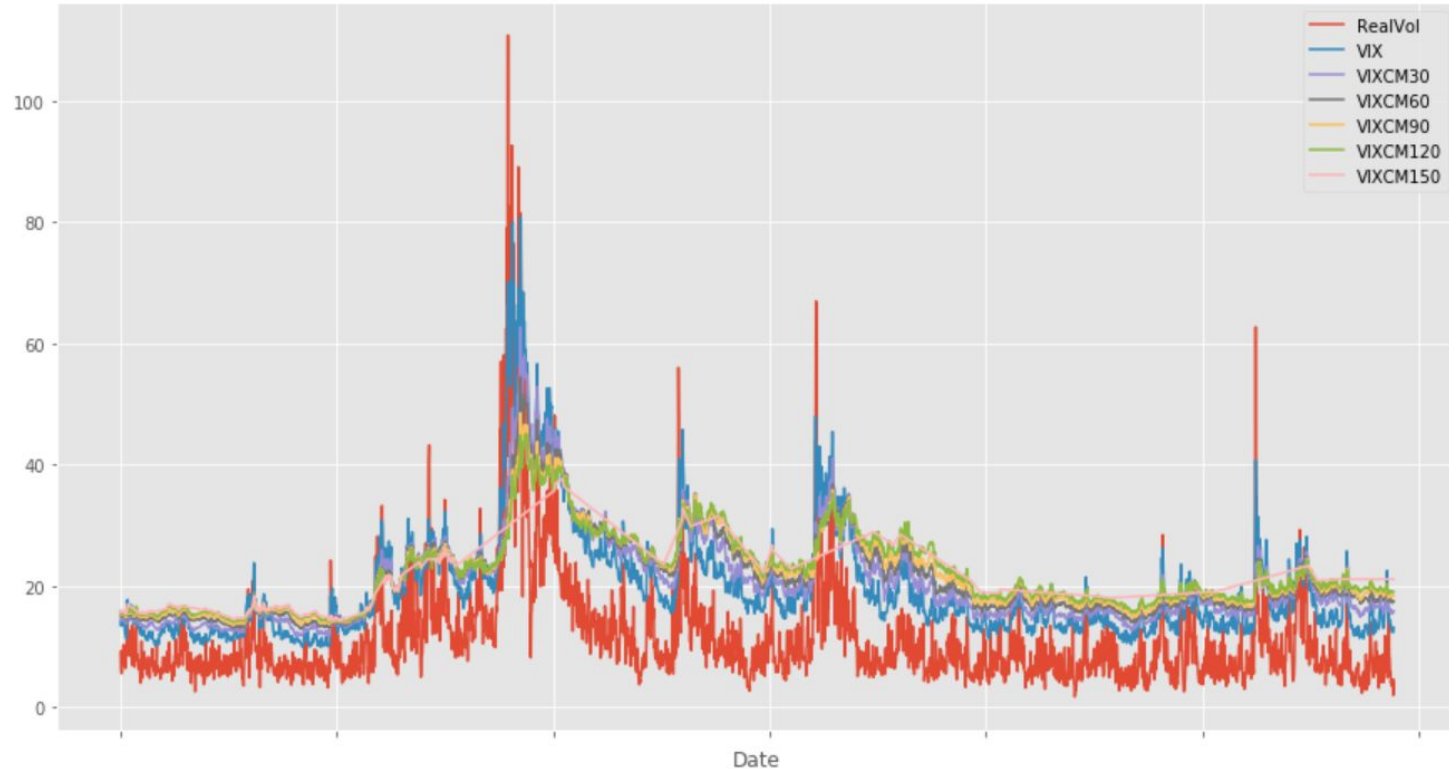- **Model Comparison**

Correlation

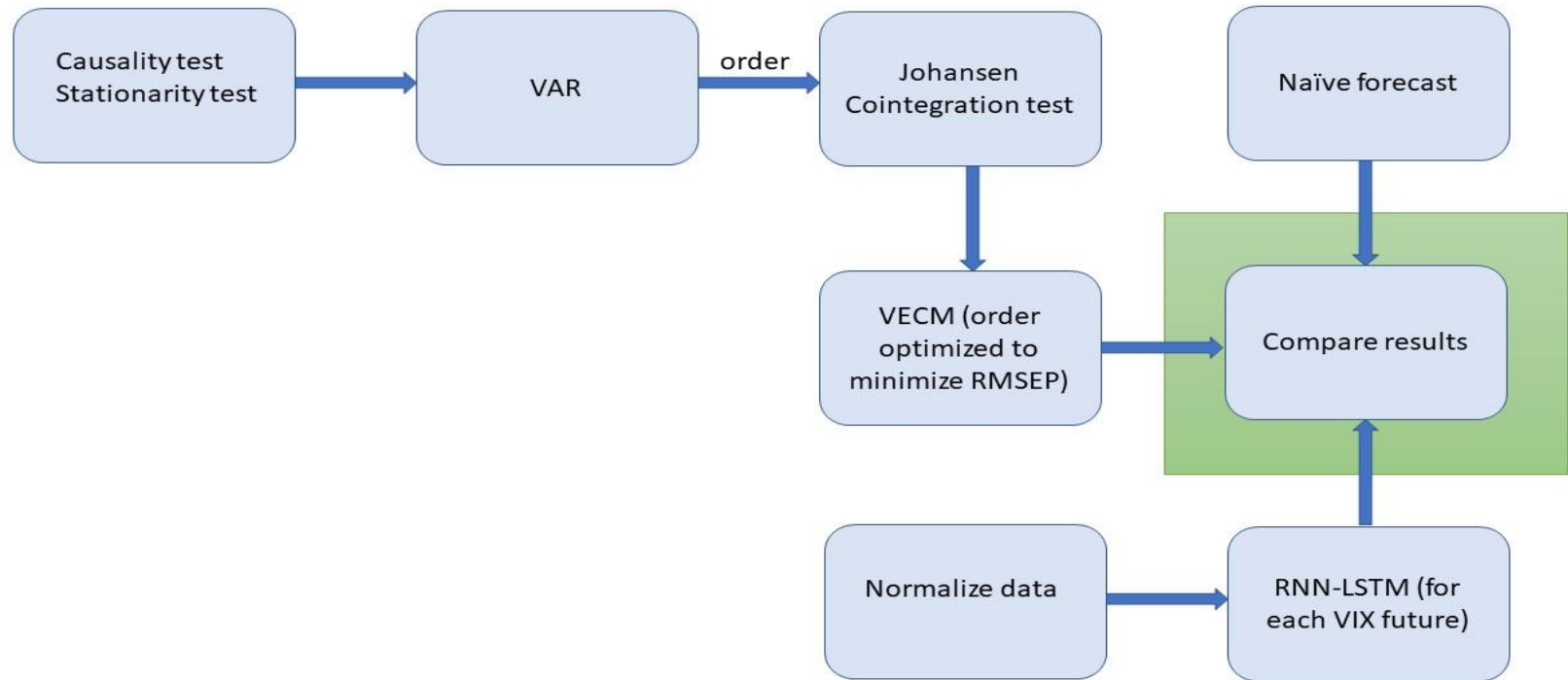|	SPY	RealVol	VIX	VIXCM30	VIXCM60	VIXCM90	VIXCM120	VIXCM150	SPY_ret
SPY	1.000000	-0.370416	-0.448691	-0.474916	-0.478567	-0.470011	-0.455977	-0.408777	0.019941
RealVol	-0.370416	1.000000	0.862610	0.732905	0.669239	0.624888	0.589693	0.490956	-0.137838
VIX	-0.448691	0.862610	1.000000	0.949276	0.904740	0.869575	0.839492	0.722753	-0.128789
VIXCM30	-0.474916	0.732905	0.949276	1.000000	0.988666	0.970015	0.950190	0.853157	-0.063839
VIXCM60	-0.478567	0.669239	0.904740	0.988666	1.000000	0.993938	0.982756	0.906139	-0.046847
VIXCM90	-0.470011	0.624888	0.869575	0.970015	0.993938	1.000000	0.996061	0.932606	-0.038479
VIXCM120	-0.455977	0.589693	0.839492	0.950190	0.982756	0.996061	1.000000	0.943444	-0.033899
VIXCM150	-0.408777	0.490956	0.722753	0.853157	0.906139	0.932606	0.943444	1.000000	-0.008257
SPY_ret	0.019941	-0.137838	-0.128789	-0.063839	-0.046847	-0.038479	-0.033899	-0.008257	1.000000

Missing Values

- VIXCM120 - 28 Null out of 2943 values
- VIXCM150 - 2434 Null out of 2943 values
- Solution - Linear Interpolation
- VIXCM150 Plots - Before & After Interpolation
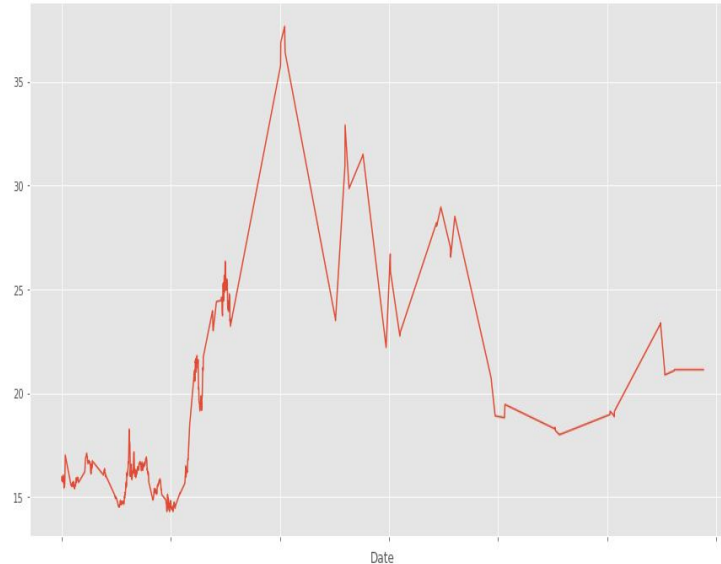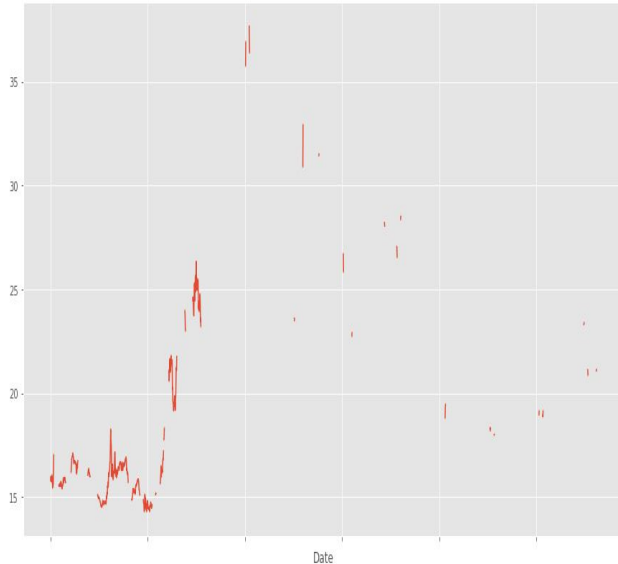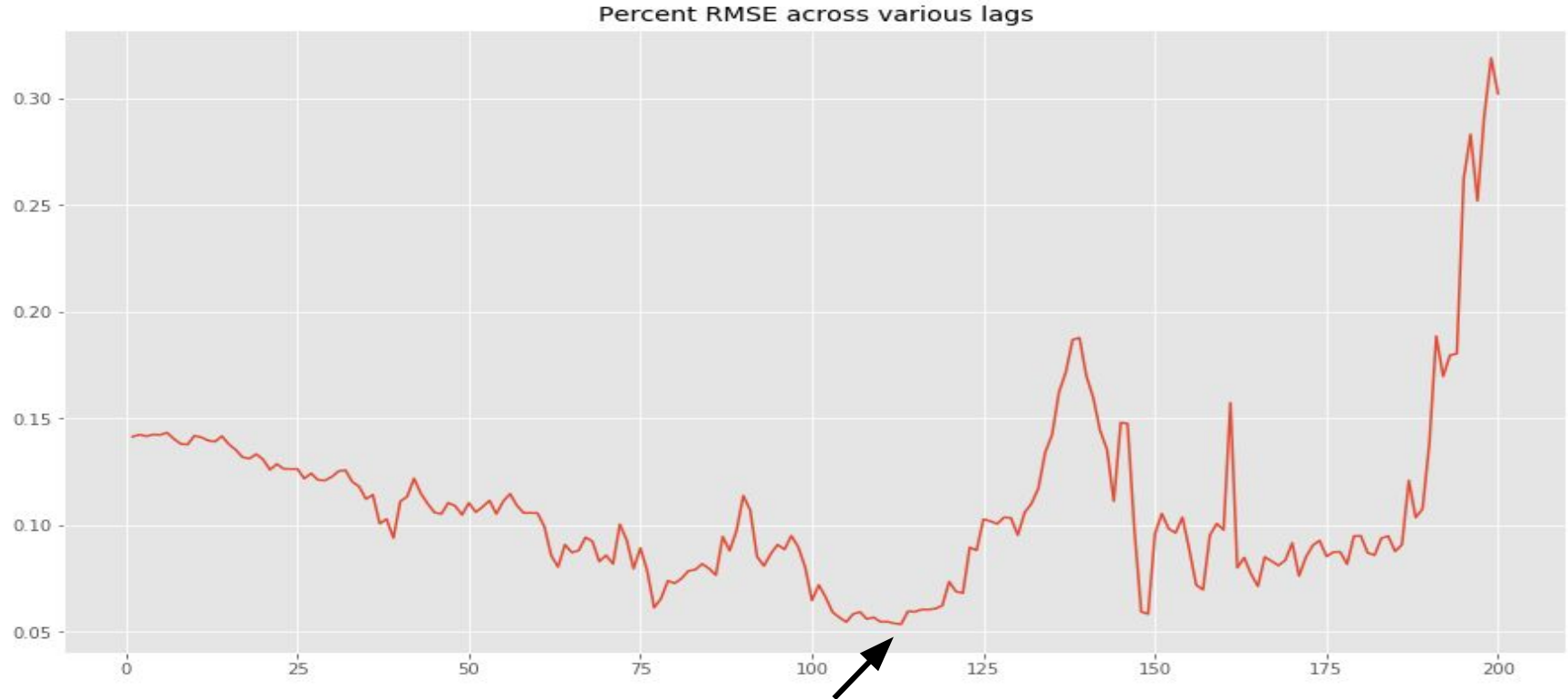
RealVol, VIX, VIX Futures Plot
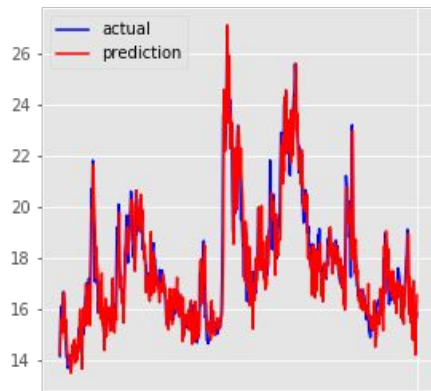
Roadmap

VECM Model

- Features - SPY_Ret, RealVol, VIX, VIXCM30, VIXCM60, VIXCM90, VIXCM120, VIXCM150
- Train-Test split: 80/20
- Optimized VECM model for lowest RMSEP
- **Metrics**: RMSEP
  - Penalizes extreme deviations more
  - Better comparison metric than raw error scores

Optimized VECM Model - lag 113
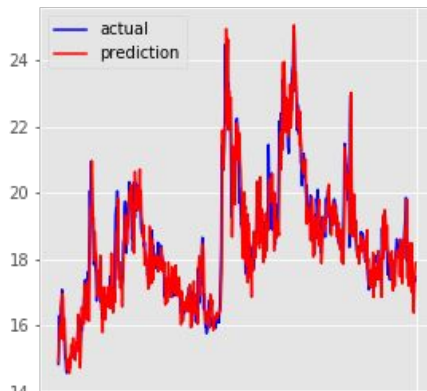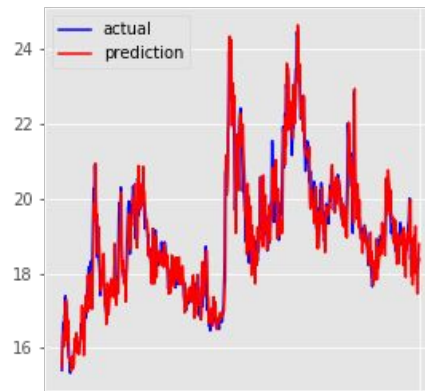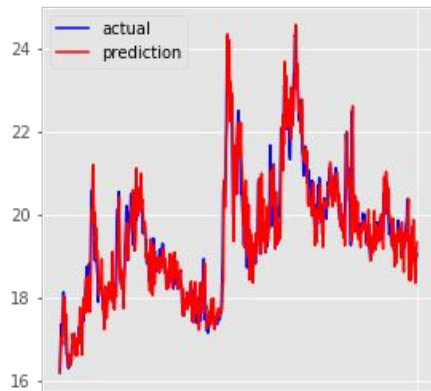
Percent RMSE across various lags
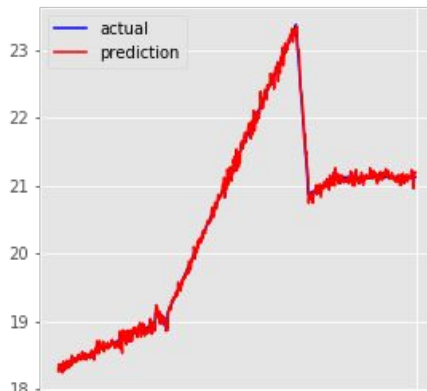
VECM 1- Day Forecast

VECM Multi-Day Forecast
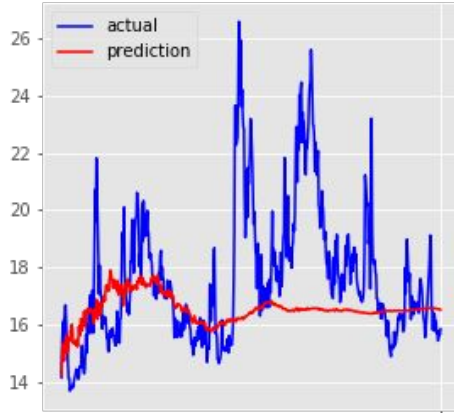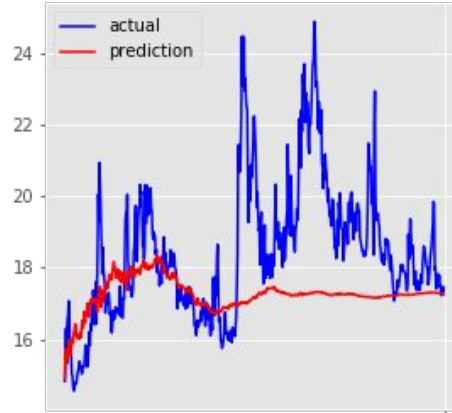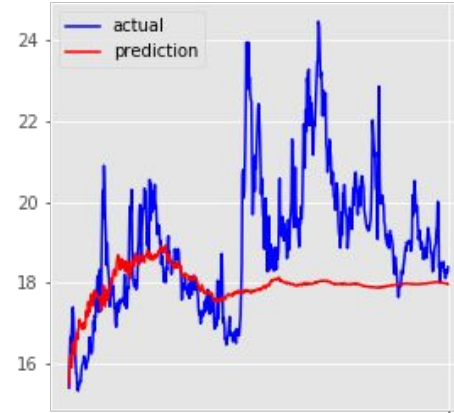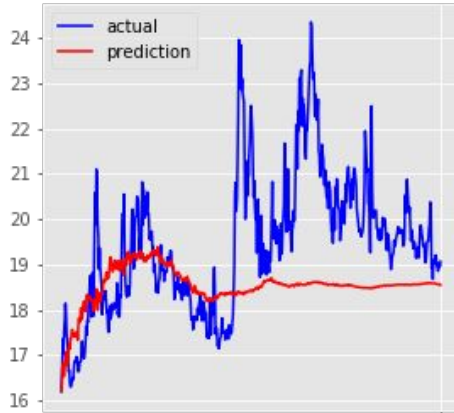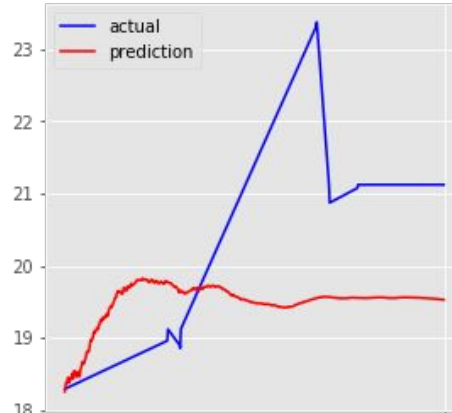
Naive Forecast - Rolling mean 1-Day Forecast
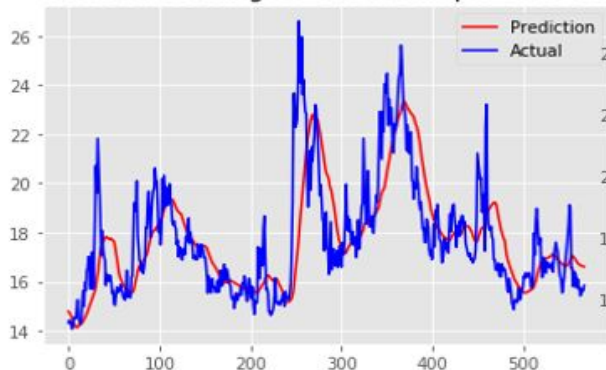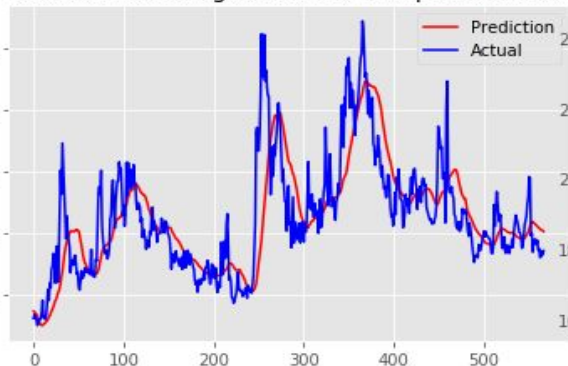
Naive Forecast - Rolling mean Multi-Day Forecast
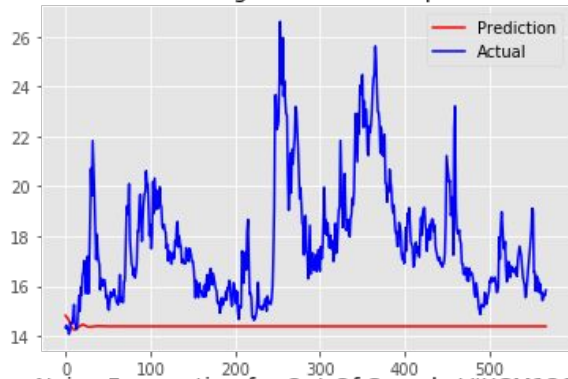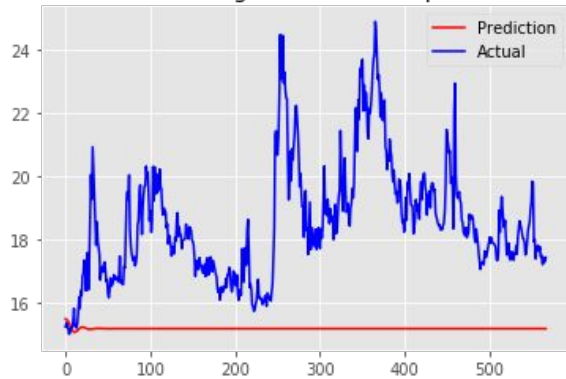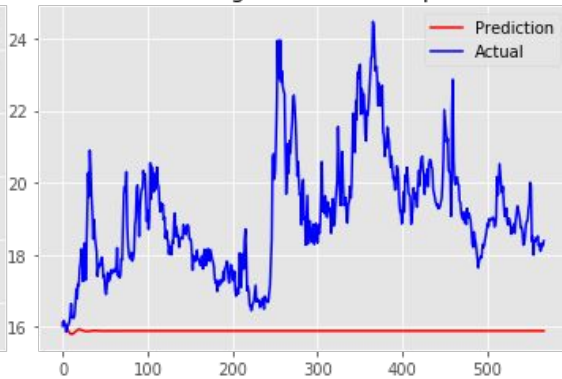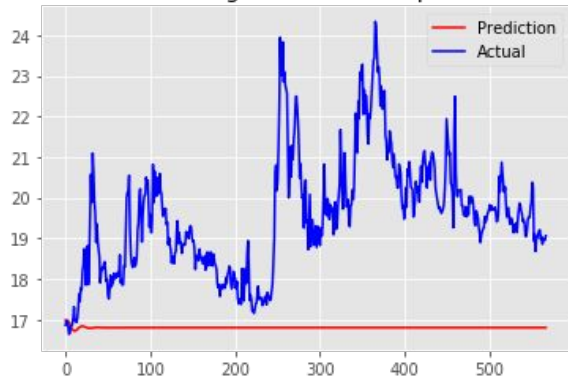
RNN Forecast

- Features - Lagged SPY, RealVol, VIX, VIX Futures
- Train-Test split: 80/20
- Normalization
- 1 Layer - 50 nodes
- LSTM
- Adam Optimizer, learning rate - 0.00035, Epochs - 35

RNN VIX Futures Forecast (Out of Sample)

VIXCM30

VIXCM60

VIXCM90

VIXCM120
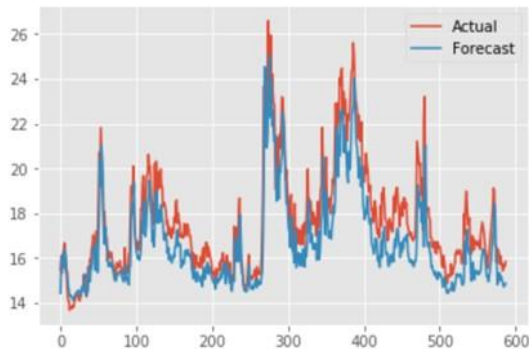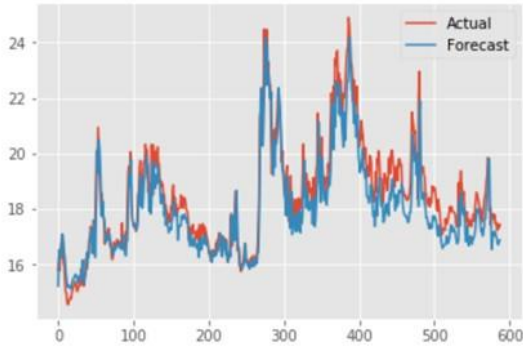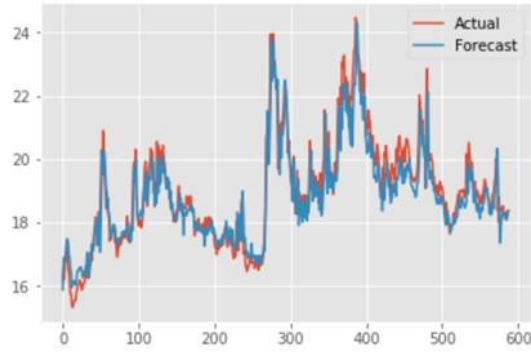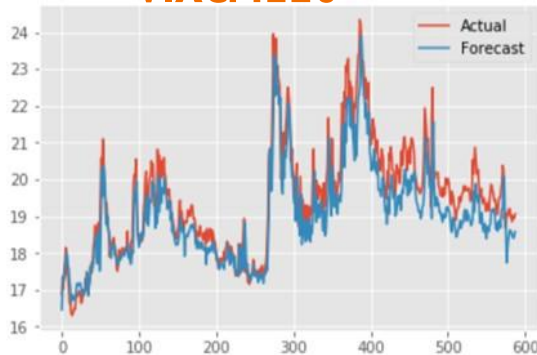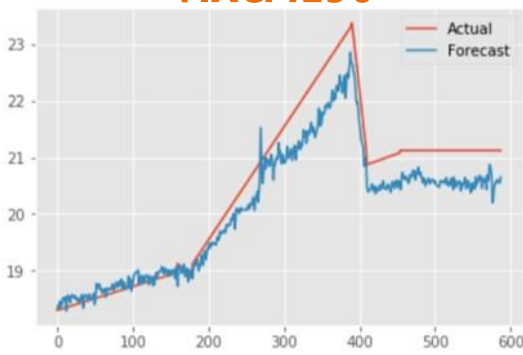

VIXCM150

Model Comparison

Model	RMSEP (VIXCM30)	RMSEP (VIXCM60)	RMSEP (VIXCM90)	RMSEP (VIXCM120)	RMSEP (VIXCM150)
LSTM	0.0570	0.0420	0.0350	0.0300	0.0300
VECM model (order = 113)	0.0545	0.0387	0.0314	0.0273	0.0035
Naive model	0.1841	0.1717	0.1691	0.1709	0.1698

Results & Limitations

- 1- Day forecast models are accurate to the value
- Multi-day forecast models provide accurate trends not values through the first 200 forecasted days
- VECM model and RNN model performs better than Naive model
- More complex LSTM model might yield better results
- Interpolation is a useful way to handle sporadic null values such as VIXCM120 but not a useful way to handle consecutive missing values such as VIXCM150

Thank You!