Exploratory Analysis on EXPRESS ATHENS for FOC estimation

Dimitris Kaklis

DMC

Pipeline

Definition

- Extract LAROS data
- Filtering - prepossessing
- Mapping with TELEGRAMS
- Mapping with Weather Service
- Extract Vessel Profile from mapped Dataset (Statistical Analysis)
- Train Consumption Models
- Compare Consumption Models
- Select most suitable consumption Model

Extract LAROS RAW Data ∼ 8 Months

Consists of :

- 280.000 observations
- 420 features

	ttime	VesselHeading	CommandedRudderLimit	Latitude	Longitude	TrackDegreesMagnetic	RudderAngle	SpeedOverGround	TrackDegreesTrue	WindAngle	WindSpeed	InclinometerXmax
0	2019-09-10 00:00:00	316.760000	35.000000	5404.023300	752.608200	357.900000	0.100000	0.000000	0.500000	341.500000	20.375500	-0.670000
1	2019-09-10 00:01:00	316.760000	35.000000	5404.023300	752.608200	357.900000	0.100000	0.100000	0.500000	336.200000	19.750500	-0.680000
2	2019-09-10 00:02:00	315.300000	35.000000	5404.028300	752.616800	46.900000	-0.100000	0.100000	49.500000	349.100000	17.375500	-0.680000
3	2019-09-10 00:03:00	313.170000	35.000000	5404.031700	752.620700	45.800000	-0.100000	0.100000	48.400000	358.500000	17.000500	-0.680000
4	2019-09-10 00:04:00	312.130000	35.000000	5404.033300	752.623200	45.800000	-0.100000	0.100000	48.400000	358.500000	17.000500	-0.670000

Figure: Snapshot LAROS data

Feature selection (cont)

- Keep features relevant for our task → FOC estimation
  - ✓ stw
  - ✓ draft
  - ✓ ws
  - ✓ wd
- 'Map' our dataset with TELEGRAMS & Weather Service data
- Calculate correlations(dependency) between our new features
- Find co-linear features and remove them (circled next slide)

Feature selection (cont)

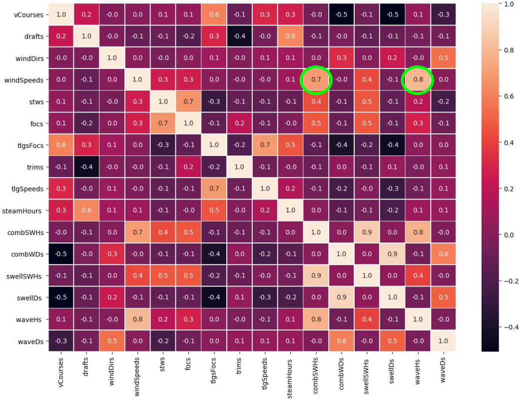


Figure: Pearson correlation coeff

Dataset after pre-processing

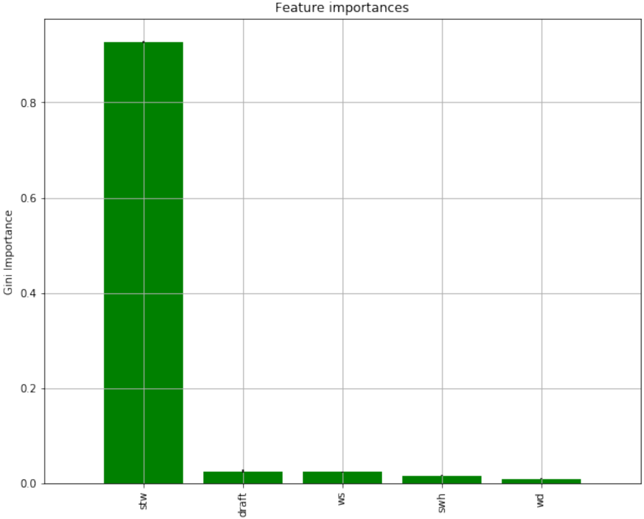
Dataset

	draft	wd	ws	stw	swh	foc	lat	lon	timestamp
0	7.640000	188.960000	20.450000	7.140000	1.190000	40.320000	54.036817	7.524675	2019-09-11 15:44
1	7.480000	186.850000	20.450000	8.140000	1.190000	38.880000	54.034250	7.523795	2019-09-11 15:45
2	7.490000	174.550000	23.020000	8.140000	1.190000	70.560000	54.034250	7.523795	2019-09-11 15:46
3	7.520000	161.510000	26.620000	9.510000	1.190000	72.000000	54.031233	7.522318	2019-09-11 15:47
4	7.810000	139.120000	23.920000	10.110000	1.190000	70.560000	54.028183	7.519653	2019-09-11 15:48

Figure: Snapshot of reduced dimensions dataset

Feature importance

	stw	draft	ws	swh	wd
0	0.928062	0.024027	0.023406	0.015132	0.009373



Consumption profiling for vessel

- On same round trip for different periods (event recognition)
- On same round trip for:
  - ✓ Different **Drafts**
  - ✓ Different **Speed ranges**
  - ✓ Different **SWH ranges**
  - ✓ Different **WD**

Consumption profiling for vessel (cont)

Trajectories

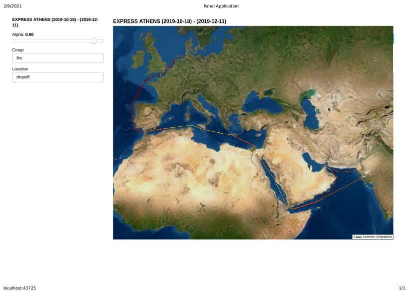


Figure: (2019-10-18) - (2019-12-11)

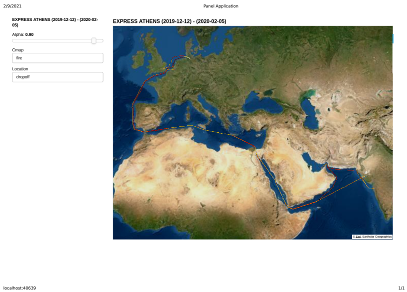
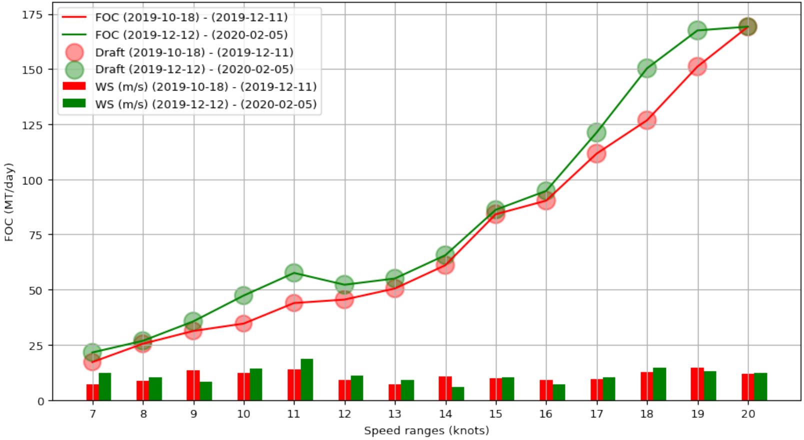


Figure: (2019-12-12) - (2020-02-05)

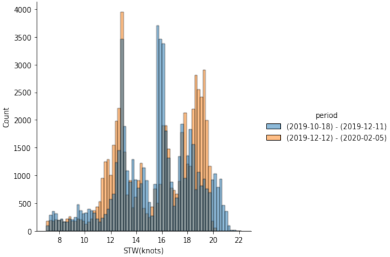
Consumption profiling for vessel (cont)

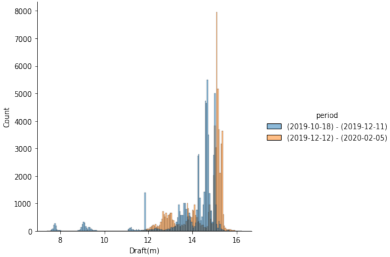
Consumption profile for the two periods



Consumption profiling for vessel (cont)

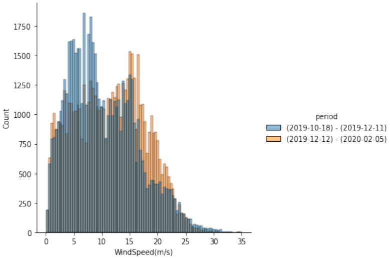
Distributions

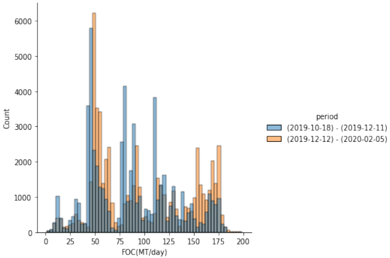




Consumption profiling for vessel (cont)

Distributions (cont)



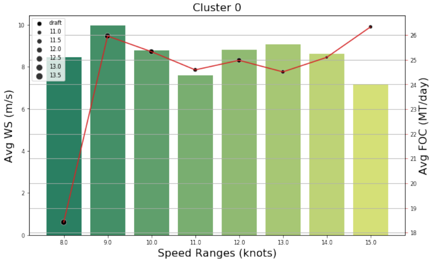


Consumption profiling for vessel (cont)

Separate vessel states

Cluster data based on most dominant features (stw,draft,ws) $\rightarrow$ FOC

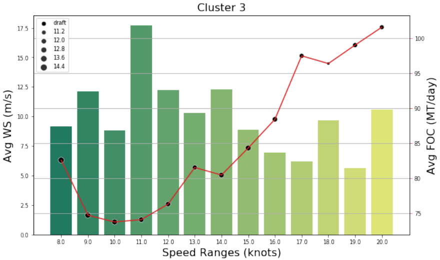
	draft	ws	stw	foc
0	14.500911	11.017526	17.669761	122.034095
1	11.947451	9.716434	13.179975	49.011670
2	14.011226	8.626996	15.761700	86.149118
3	14.399122	15.884578	19.267752	165.779715
4	13.025901	9.017101	9.326219	21.680181



Consumption profiling for vessel (cont)

Separate vessel states

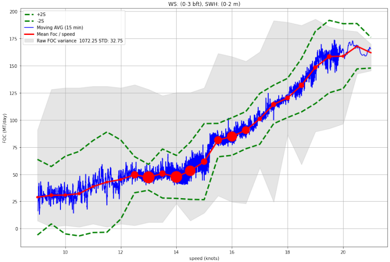
Cluster data based on most dominant features (stw,draft,ws) → FOC



Consumption profiling for vessel (cont)

Separate vessel states (per day)

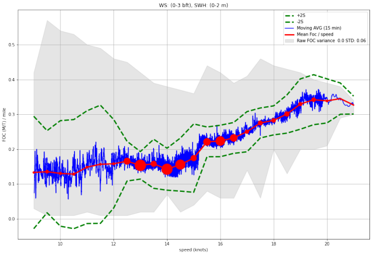
- ✓ Good weather conditions WS: (0-3 bft) SWH (0-2 m)
- ✓ Bad weather conditions WS: (3-8 bft) SWH (2-8 m)

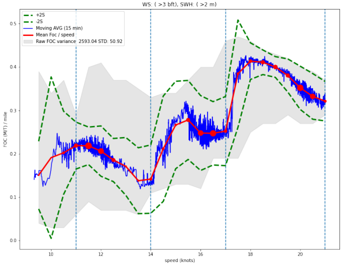


Consumption profiling for vessel (cont)

Separate vessel states (per mile)

- ✓ Good weather conditions WS: (0-3 bft) SWH (0-2 m)
- ✓ Bad weather conditions WS: (3-8 bft) SWH (2-8 m)



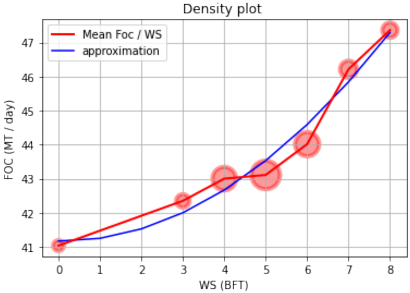


Consumption profiling for vessel (cont)

Separate vessel states

In order to observe better the impact of weather in FOC we calculate separately the percentage of change in FOC for WS, SWH, and WD keeping the rest of the features namely draft and speed constant.

WS impact



ws (bft)	perc change (%)
0-1	0.000000
1-2	0.570000
2-3	1.070000
3-4	1.670000
4-5	2.280000
5-6	2.650000
6-7	3.030000
7-8	3.740000

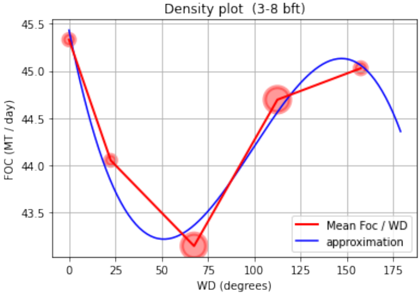
Consumption profiling for vessel (cont)

SWH impact

swh (m)	perc change (%)
0-1	0.000000
1-2	0.410000
2-3	0.970000
3-4	1.300000
4-5	1.840000
5-6	2.240000
6-7	2.710000
7-8	2.870000

Consumption profiling for vessel (cont)

WD impact



wd (deg)	perc change (%)
0-22.5	0.000000
22.5-67.5	2.820000
67.5-112.5	1.300000
112.5-157.5	2.390000
157.5-180	0.500000

Model Training - data preparation

Task description

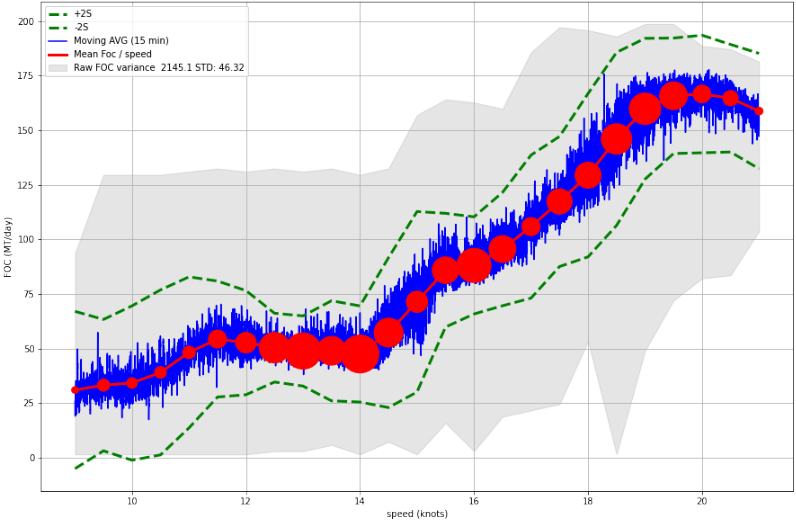
- Our task is to estimate the **mean FOC** value given a certain vector of features that corresponds to stw, ws, wd and draft
- Given the nature of our task and the volatility of FOC in raw data it is vital to transform our dataset in a way that resembles the average FOC value per timestep

Dataset Transformation

- Transform our dataset in 15 min rolling window averages
- Rolling window averages gives us the advantage to "clear" our dataset from "noise" and to train our models with a dataset that corresponds to the average value of FOC at every timestep

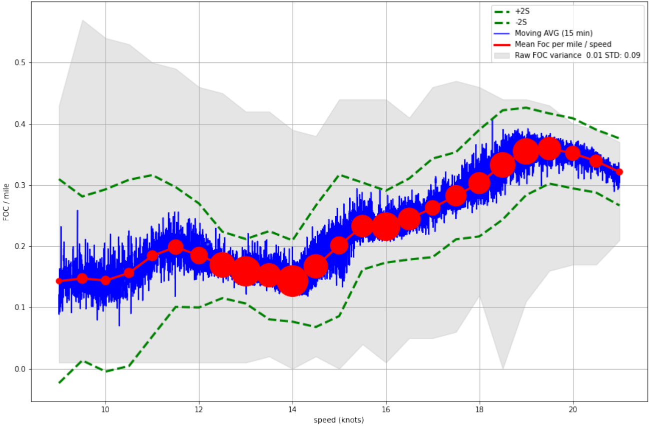
Model Training - data preparation (cont)

Visualize deviation between **raw data** - **rolling window averages** - **mean value** of FOC (MT/day) per speed range.



Model Training - data preparation (cont)

Visualize deviation between **raw data - rolling window averages - mean value** of FOC / mile per speed range.



Model Training - Decision Tree

Create "buckets" for DT

- Different vessel states (Different Speed Groups, Different WS, Different SWH, Different WD)
- Fill the buckets with the avg FOC value for corresponing to each state
- Interpolate between the buckets

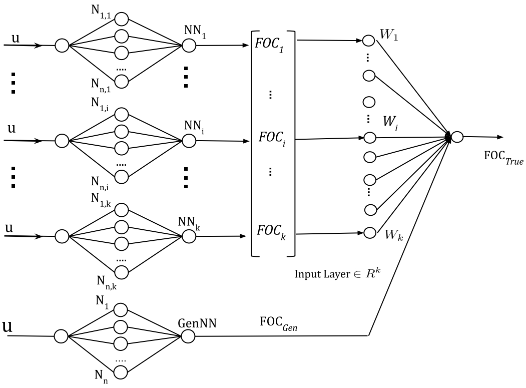
CONSUMPTION WHEN LADDEN								
Draft (m)				12.18				
Trim (m)								
MIN SPEED GROUP	9							
CONSUMPTION (MT/day)	Wind Speed (Beaufort)				Swell wave height (m)			
	0-1 Beaufort							
Relative Direction (Degrees)	(0-1)m	(1-2)m	(2-3)m	(3-4)m	(4-5)m	(5-6)m	(6-7)m	(7-8)m
Against (0-22.5)	30.78	30.81	30.83	30.86	30.88	30.91	30.94	30.97
Against side (22.5-67.5)	30.26	30.29	30.31	30.34	30.36	30.39	30.42	30.44
Side (67.5-112.5)	29.92	29.95	29.97	30.0	30.02	30.05	30.08	30.1
Side with (112.5-157.5)	29.83	29.86	29.88	29.91	29.93	29.96	29.99	30.01
With (157.5 - 180)	29.58	29.61	29.63	29.66	29.68	29.71	29.74	29.76
	1-2 Beaufort							
Relative Direction (Degrees)	(0-1)m	(1-2)m	(2-3)m	(3-4)m	(4-5)m	(5-6)m	(6-7)m	(7-8)m
Against (0-22.5)	**31.1**	**31.13**	**31.15**	**31.18**	**31.2**	**31.23**	**31.26**	**31.29**
Against side (22.5-67.5)	30.58	30.61	30.63	30.66	30.68	30.71	30.74	30.76
Side (67.5-112.5)	30.24	30.27	30.29	30.32	30.34	30.37	30.4	30.42
Side with (112.5-157.5)	30.14	30.17	30.19	30.22	30.24	30.27	30.3	30.32
With (157.5 - 180)	29.89	29.92	29.94	29.97	29.99	30.02	30.05	30.07
	2-3 Beaufort							
Relative Direction (Degrees)	(0-1)m	(1-2)m	(2-3)m	(3-4)m	(4-5)m	(5-6)m	(6-7)m	(7-8)m
Against (0-22.5)	**31.41**	**31.44**	**31.46**	**31.49**	**31.51**	**31.54**	**31.57**	**31.6**
Against side (22.5-67.5)	30.88	30.91	30.93	30.96	30.98	31.01	31.04	31.07
Side (67.5-112.5)	30.54	30.57	30.59	30.62	30.64	30.67	30.7	30.72
Side with (112.5-157.5)	30.44	30.47	30.49	30.52	30.54	30.57	30.6	30.62
With (157.5 - 180)	30.19	30.22	30.24	30.27	30.29	30.32	30.35	30.37
⋮			⋮				⋮	
SPEED GROUP 1	10							
CONSUMPTION (MT/day)	Wind Speed (Beaufort)				Swell wave height (m)			
	0-1 Beaufort							
Relative Direction (Degrees)	(0-1)m	(1-2)m	(2-3)m	(3-4)m	(4-5)m	(5-6)m	(6-7)m	(7-8)m
Against (0-22.5)	**42.03**	**42.07**	**42.1**	**42.14**	**42.17**	**42.21**	**42.25**	**42.28**
Against side (22.5-67.5)	41.32	41.36	41.39	41.43	41.46	41.5	41.54	41.57
Side (67.5-112.5)	40.85	40.88	40.91	40.95	40.98	41.02	41.06	41.09
Side with (112.5-157.5)	40.73	40.76	40.79	40.83	40.86	40.9	40.94	40.97

Vessel Basic Info | Cons profile for Draft (10.0 - 14.0) | Cons profile for Draft (6.0 - 10.0) | Max Acceptable Sea State | Statistics (TLG) | Statistics (RAW)

Model Training - Neural Ensemble - (Stacking)

Train different models for each cluster of data (slides 13-14)

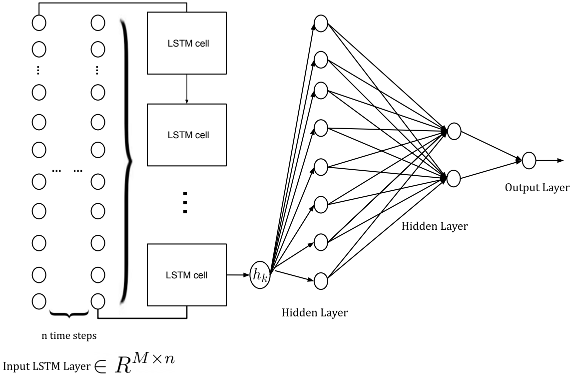
- Different estimators for different vessel states (Different Speed Groups, Different WS, Different SWH, Different WD)
- Take weighted average prediction of all models by training a meta-model that learns to combine them



Model Training - Neural (LSTM)

Train LSTM neural

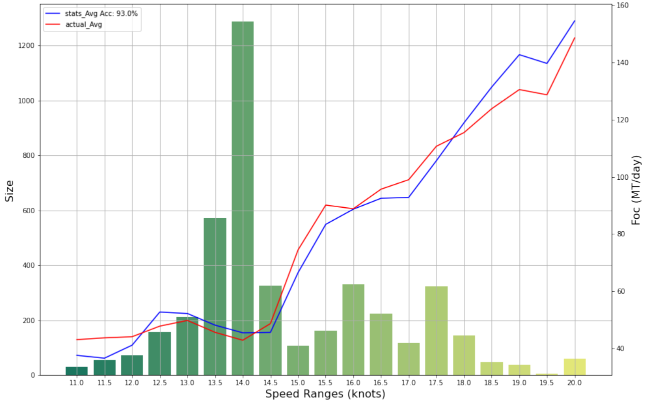
- LSTM architecture introduces the notion of "memory" to our models
- Transform dataset in vectors of $(N \times M \times n)$ dimensions
  - ✓ N: number of observations
  - ✓ M: number of features
  - ✓ n: "depth" of memory of our input vectors (timesteps to "look back")



Models Performance

Stats Decision Tree Performance

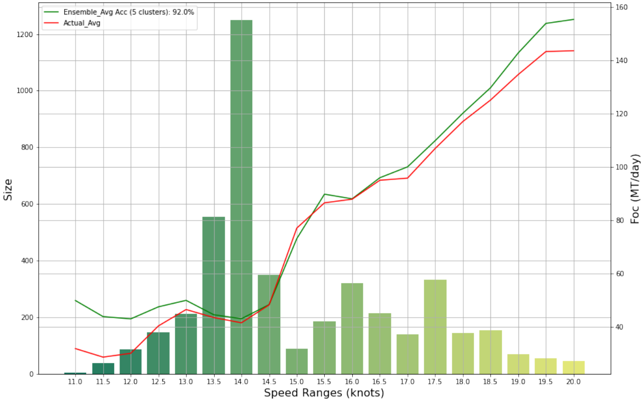
- Trained on $\simeq 4 * 10^4$ observations $\simeq$ 1 month of data
- Tested on $4 * 10^3 \simeq 10\%$ of trained data



Models Performance (cont)

Ensemble Neural Performance (5 clusters)

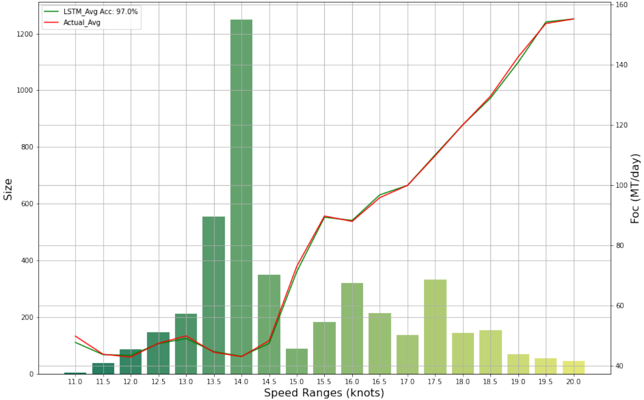
- Trained on $\simeq 4 * 10^4$ observations $\simeq$ 1 month of data
- Tested on $4 * 10^3$ unseen data $\simeq 10\%$ of trained data



Models Performance (cont)

LSTM Neural Performance

- Trained on $\simeq 4 * 10^4$ observations $\simeq$ 1 month of data
- Tested on $4 * 10^3$ unseen data $\simeq 10\%$ of trained data



Conclusions