# Evaluation of a machine learning framework to forecast storm surge

Daryl Metters

Coastal Impacts Unit

Department of Environment and Science

Queensland Government

# The Coastal Impacts: what we do

- The primary Queensland Government agency involved in the management of extreme storm-tide events in Queensland

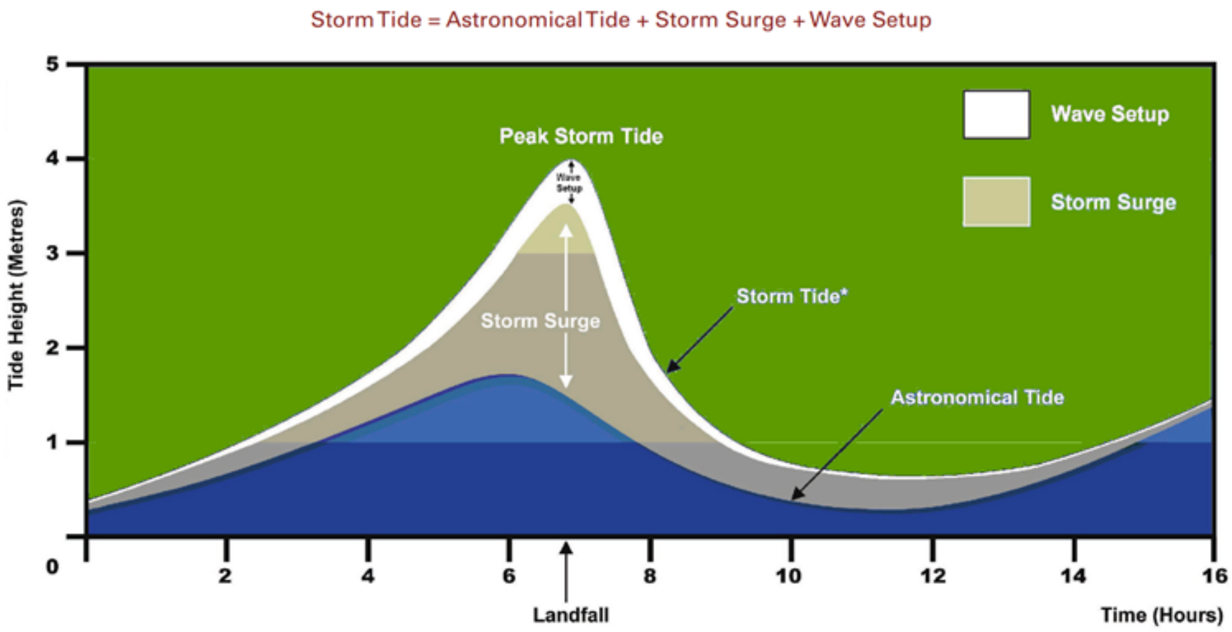
- Operate 36 storm-tide gauges and 14 tide gauges along the Queensland coast. These measure the magnitude of storm-tide during cyclonic events for the use by disaster management groups for evacuation purposes

- During severe events we Liaise with the Bureau of Meteorology to confirm information in storm-tide advice (warnings) and provide technical advice on storm-tide to local, district and state groups

# What is Storm Tide ?

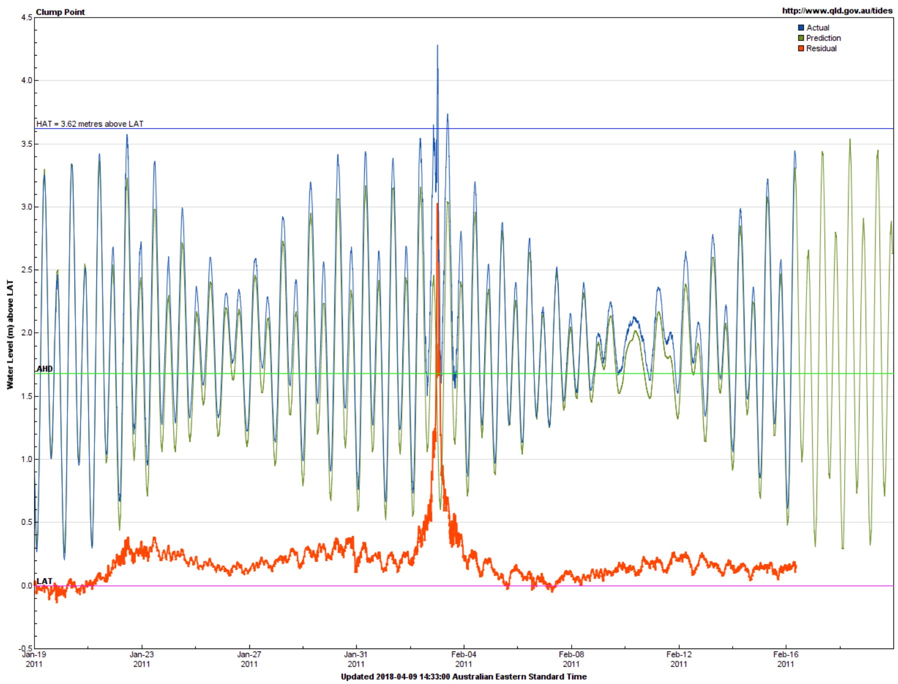
# What is Storm Tide ?

The **storm tide** is the total water level obtained by adding the **storm surge** and **wave setup** to the height of the **astronomical tide**



# What is Storm Surge?

Storm Surge = Actual - Tide Prediction



# Storm-tide data

- We present and distribute the actual storm-tide data along with the tide prediction and residual in near real time
- This process predicts or forecasts the tide component only and not the surge level
- The surge level is calculated in near real time as the levels are reported by each STG
- Storm Surge = actual (measured) level – tide prediction.

# Storm Surge forecasting

- Many methods developed over recent years to help forecast the storm surge level
- Most make use of the linear relationship between the driving parameter(s) and the actual water level recorded
- Successful methods used in modern times are based on numerical modelling of the physical driving forces responsible for the surge levels
- These modelling efforts are expensive to maintain due to the large computing power needed to operate the models

# Storm Surge forecast

Important for planning:

1. Evacuation during severe events
2. Recreational activities
3. Commercial transport
4. Scientific marine activities

# The Project

## Storm-tide forecasting using Machine Learning

A proof of concept

- Transfer of knowledge and understanding of machine learning principles to DES staff
- To establish and test various machine learning models, and use those machine learning models to forecast storm-tide levels
- Formulate an understanding of the effectiveness of machine learning in forecasting storm-tide

# Machine Learning

- A type of artificial intelligence
- Enables the ability to "learn" with data, without being explicitly programmed
- Explores the study and construction of algorithms that can learn from and make predictions on data
- This overcomes following strictly static program instructions by making data-driven predictions or decisions, through building a model from sample inputs
- Employed in a range of computing tasks where designing and programming explicit algorithms with good performance is difficult or unfeasible.

# Computing and code

- Python library: Scikit-learn
- Anaconda 3
- Used CPU and memory
- Amazon AWS High Performance Computing facilities

  72 cores of virtual CPUs and 144 GB memory

# Data preparation

- Storm tide, wind and pressure data checked for errors and gaps
- Filtered for out of range values: if out of range then removed
- Single missing points given the average of the two points before and after, larger gaps were considered missing data and removed.

# Model Training, Testing and Forecasting

- The prepared dataset was divided into training and testing datasets
- 2/3 training : to improve model accuracy
- 1/3 testing : to check model accuracy

- The ML model output is then used to forecast 72 hours beyond the training and testing datasets

# Input Data

- **Clump Point Storm Tide Gauge**
- Storm Tide: one minute
- Atmospheric pressure:  one minute
- Wind speed : 10 minute
- Wind direction: 10 minute

- **Tide predictions :** 10 minute

# Machine Learning

Two general types of machine learning models utilised

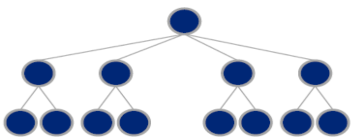
(1) Feature based models:

- Decision Tree
- Neural Networks
- Linear
- k-Nearest Neighbour (kNN)
- Support Vector Machine (SVM) and
- Random Forest

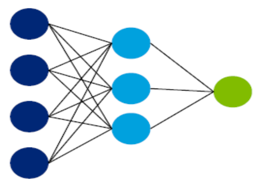
(2) Time series models: ARIMA and Prophet.

# Machine Learning Methods

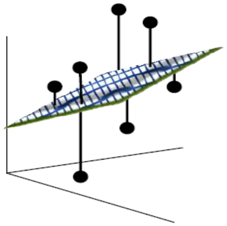
**As the eight models were developed, special effort was made to train DES staff on how to conceptualise and run the models in Python**



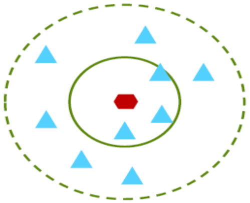
Decision Tree



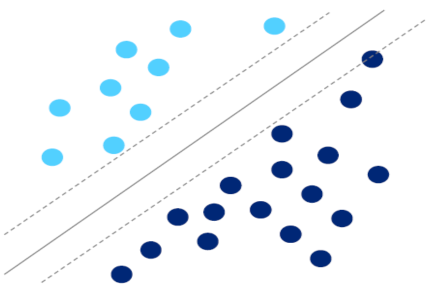
Neural Network



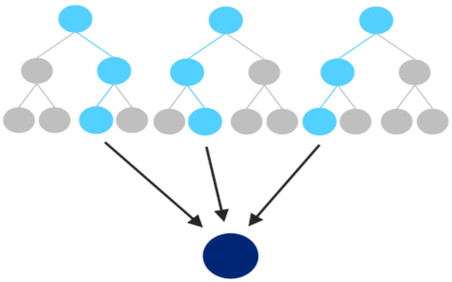
Linear



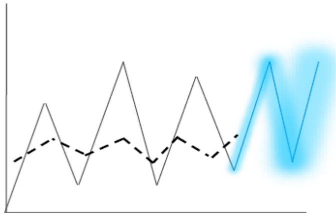
KNN



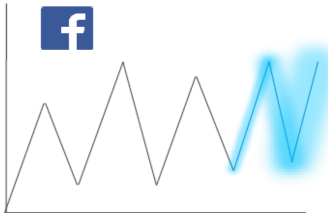
SVR



Random Forest



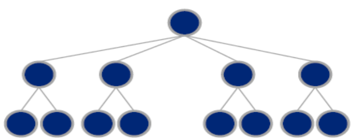
ARIMA



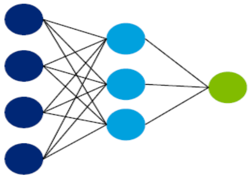
Prophet

# Machine Learning Methods

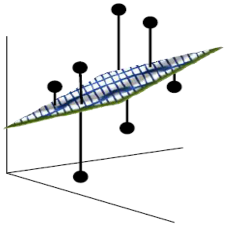
**As the eight models were developed, special effort was made to train DES staff on how to conceptualise and run the models in Python**



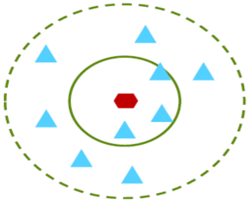
Decision Tree



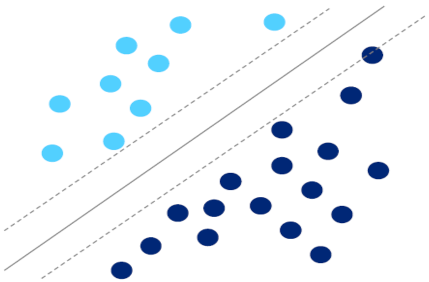
Neural Network



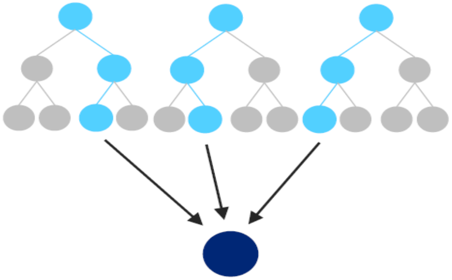
Linear



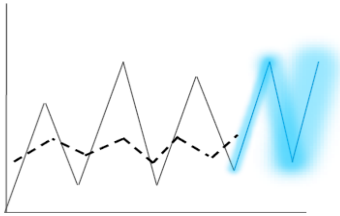
KNN



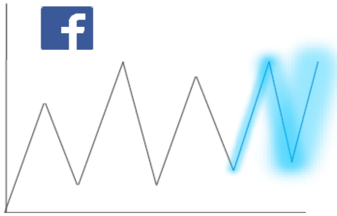
SVR



Random Forest



ARIMA



Prophet

# Model approach's

- **V1 Modelled weather approach**

  **Storm Tide**

  **Training/testing Input**: storm tide, atmospheric pressure, wind speed and direction, tide predictions

  **Forecast input**: BoM Access-G modelled weather forecast

  **Storm Surge**

  **Training/testing Input**: residual, atmospheric pressure, wind speed and direction

  **Forecast input**: BoM Access-G modelled weather forecast
- **V2 Dataset shift approach**
- **V3 Modelled weather forecast approach using the time series models**
- **V2 and V3 approach's not taken to forecast phase**

# Model performance

- Metrics used:
  1. Run time: Linux command 'time' used to generate real time: wall clock time - time from start to finish of the call.
  2. Mean Square error
  3. Correlation coefficients

# Model performance

## Model Run Time

| Forecast type | Model | Real time |
|---|---|---|
| Storm Tide | Decision Tree | 0m 9.5s |
| | Neural Network | 128m 7.9s |
| | Linear Model | 0m 8.7s |
| | KNN | 0m 25.8s |
| | Random Forest | 14m 35.8s |
| | SVR | 31m 23.9s |
| Residual | Decision Tree | 0m 12.0s |
| | Neural Network | 132m 45.7s |
| | Linear Model | 0m 8.6s |
| | KNN | 0m 56.8s |
| | Random Forest | 14m 50.4s |
| | SVR | 31m 29.1s |

# Model performance

## Model Run Time

| Forecast type | Model | Real time |
|---|---|---|
| Storm Tide | Decision Tree | 0m 9.5s |
| | Neural Network | 128m 7.9s |
| | Linear Model | 0m 8.7s |
| | KNN | 0m 25.8s |
| | Random Forest | 14m 35.8s |
| | SVR | 31m 23.9s |
| Residual | Decision Tree | 0m 12.0s |
| | Neural Network | 132m 45.7s |
| | Linear Model | 0m 8.6s |
| | KNN | 0m 56.8s |
| | Random Forest | 14m 50.4s |
| | SVR | 31m 29.1s |

# Model performance

## Model Run Time

| Forecast type | Model | Real time |
|---|---|---|
| Storm Tide | Decision Tree | 0m 9.5s |
| | Neural Network | 128m 7.9s |
| | Linear Model | 0m 8.7s |
| | KNN | 0m 25.8s |
| | Random Forest | 14m 35.8s |
| | SVR | 31m 23.9s |
| Residual | Decision Tree | 0m 12.0s |
| | Neural Network | 132m 45.7s |
| | Linear Model | 0m 8.6s |
| | KNN | 0m 56.8s |
| | Random Forest | 14m 50.4s |
| | SVR | 31m 29.1s |

# Model performance testing phase

Storm Tide with tide predictions

| Model | Mean Squared error | Correlation coefficient |
|---|---|---|
| KNN | 0.010 | 0.988 |
| SVR | 0.007 | 0.990 |
| Decision Tree | 0.008 | 0.990 |
| Random Forest | 0.007 | 0.991 |
| Linear Model | 0.007 | 0.990 |
| Neural Network | 0.007 | 0.990 |

High correlations

All models performed equally well

# Model performance testing phase

## Storm Tide without tide predictions

| Model | Mean Squared error | Correlation coefficient |
|---|---|---|
| KNN | 0.400 | 0.057 |
| SVR | 0.374 | 0.069 |
| Decision Tree | 0.374 | 0.111 |
| Random Forest | 0.374 | 0.106 |
| Linear Model | 0.377 | 0.019 |
| Neural Network | 0.378 | -0.014 |

Very low correlation

All models performed poorly

# Model performance testing phase

## Storm Surge (Residual)

| Model | 1 month | 3 months | 6 months | 12 months | average |
|---|---|---|---|---|---|
| KNN | 0.264 | 0.288 | 0.421 | 0.398 | 0.343 |
| SVR | 0.377 | 0.398 | 0.568 | 0.403 | 0.437 |
| Decision Tree | 0.307 | 0.326 | 0.400 | 0.425 | 0.365 |
| Random Forest | 0.375 | 0.383 | 0.610 | 0.543 | 0.478 |
| Linear Model | 0.400 | 0.397 | 0.383 | 0.343 | 0.381 |
| Neural Network | 0.400 | 0.393 | 0.514 | 0.512 | 0.455 |
| average | 0.354 | 0.364 | 0.483 | 0.437 | |

Moderate correlation

# Model performance testing phase

## Storm Surge (Residual)

| Model | 1 month | 3 months | 6 months | 12 months | average |
|---|---|---|---|---|---|
| KNN | 0.264 | 0.288 | 0.421 | 0.398 | 0.343 |
| SVR | 0.377 | 0.398 | 0.568 | 0.403 | 0.437 |
| Decision Tree | 0.307 | 0.326 | 0.400 | 0.425 | 0.365 |
| Random Forest | 0.375 | 0.383 | 0.610 | 0.543 | 0.478 |
| Linear Model | 0.400 | 0.397 | 0.383 | 0.343 | 0.381 |
| Neural Network | 0.400 | 0.393 | 0.514 | 0.512 | 0.455 |
| average | 0.354 | 0.364 | 0.483 | 0.437 | |

Moderate correlation

Increasing correlation with increase in data length

# Model performance testing phase

## Storm Surge (Residual)

| Model | 1 month | 3 months | 6 months | 12 months | average |
|---|---|---|---|---|---|
| KNN | 0.264 | 0.288 | 0.421 | 0.398 | 0.343 |
| SVR | 0.377 | 0.398 | 0.568 | 0.403 | 0.437 |
| Decision Tree | 0.307 | 0.326 | 0.400 | 0.425 | 0.365 |
| Random Forest | 0.375 | 0.383 | 0.610 | 0.543 | 0.478 |
| Linear Model | 0.400 | 0.397 | 0.383 | 0.343 | 0.381 |
| Neural Network | 0.400 | 0.393 | 0.514 | 0.512 | 0.455 |
| average | 0.354 | 0.364 | 0.483 | 0.437 | |

Moderate correlation

Increasing correlation with increase in data length

Random Forest best performing model

# Model performance testing phase

## Storm Surge (Residual)

| Model | 1 month | 3 months | 6 months | 12 months | average |
|---|---|---|---|---|---|
| KNN | 0.264 | 0.288 | 0.421 | 0.398 | 0.343 |
| SVR | 0.377 | 0.398 | 0.568 | 0.403 | 0.437 |
| Decision Tree | 0.307 | 0.326 | 0.400 | 0.425 | 0.365 |
| Random Forest | 0.375 | 0.383 | 0.610 | 0.543 | 0.478 |
| Linear Model | 0.400 | 0.397 | 0.383 | 0.343 | 0.381 |
| Neural Network | 0.400 | 0.393 | 0.514 | 0.512 | 0.455 |
| average | 0.354 | 0.364 | 0.483 | 0.437 | |

Moderate correlation

Increasing correlation with increase in data length

Random Forest and Neural Network best performing

# Model performance testing phase

## Time series models

| ARIMA model | Mean Squared error | Correlation Coefficient |
|---|---|---|
| Storm Tide | 0.438 | -0.02 |
| Wind Speed | 36.739 | 0.09 |
| Wind Direction | 6438.297 | 0.13 |
| Air Pressure | 67.119 | 0.51 |

| Prophet model | Mean Squared error | Correlation Coefficient |
|---|---|---|
| Storm Tide | 0.014 | 0.34 |
| Wind Speed | 86.645 | 0.13 |
| Wind Direction | 8958.065 | 0.41 |
| Air Pressure | 54.353 | 0.59 |

Very low correlation for Storm Tide, and wind speed and direction

# Model performance testing phase

## Time series models

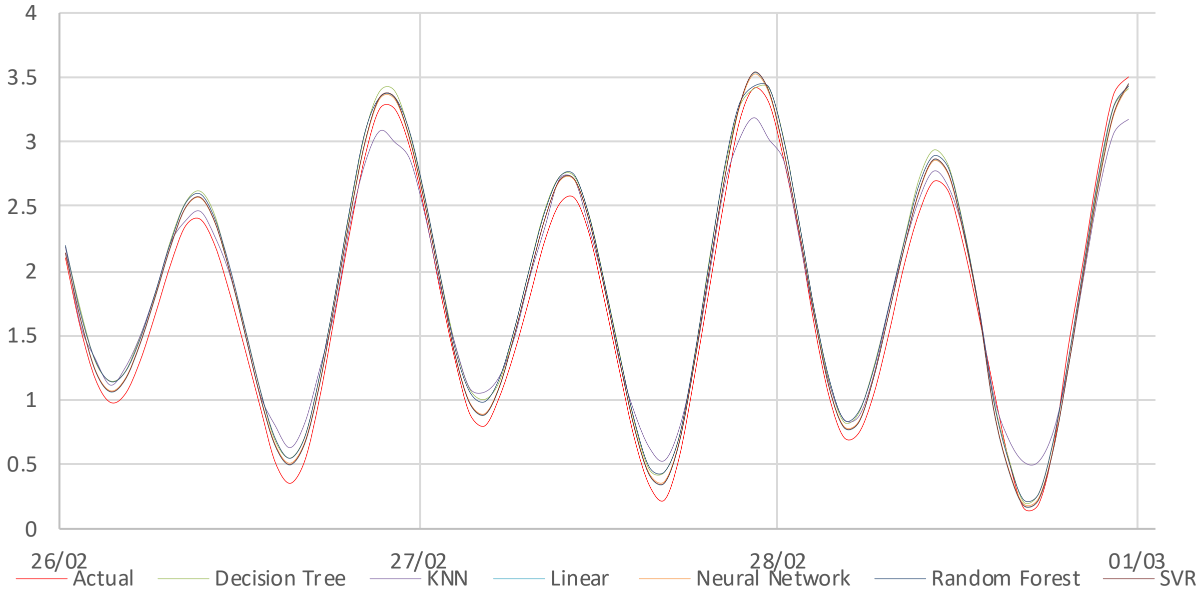
| ARIMA model | Mean Squared error | Correlation Coefficient |
|---|---|---|
| Storm Tide | 0.438 | -0.02 |
| Wind Speed | 36.739 | 0.09 |
| Wind Direction | 6438.297 | 0.13 |
| Air Pressure | 67.119 | 0.51 |

| Prophet model | Mean Squared error | Correlation Coefficient |
|---|---|---|
| Storm Tide | 0.014 | 0.34 |
| Wind Speed | 86.645 | 0.13 |
| Wind Direction | 8958.065 | 0.41 |
| Air Pressure | 54.353 | 0.59 |

Very low correlation for Storm Tide, and wind speed and direction but moderate correlation for pressure

# Model performance forecasting phase

## Storm Tide




Queensland Government

# Model performance forecasting phase

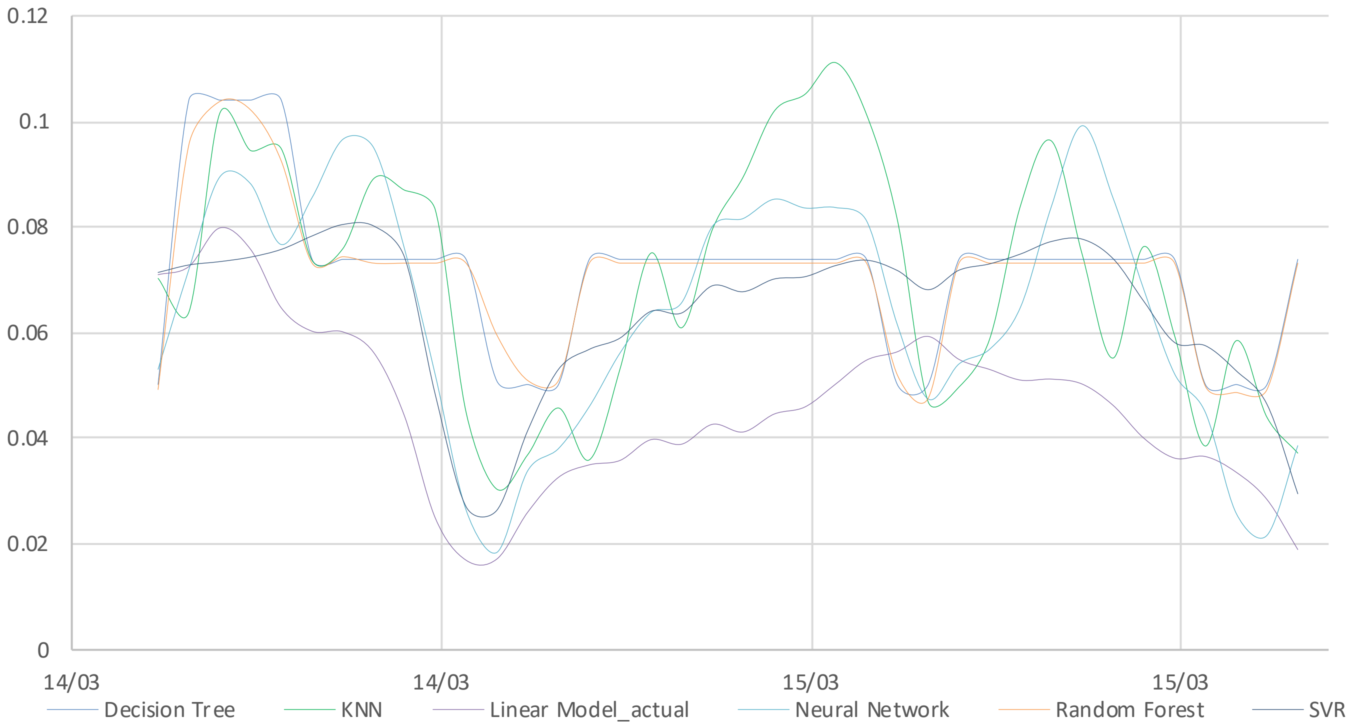
## Storm Tide

| Model | Forecast Storm Tide vs actual Storm |
|---|---|
| KNN | 0.996 |
| SVR | 0.999 |
| Decision Tree | 0.999 |
| Random Forest | 0.999 |
| Linear Model | 0.999 |
| Neural Network | 0.999 |

High correlations: better than testing phase

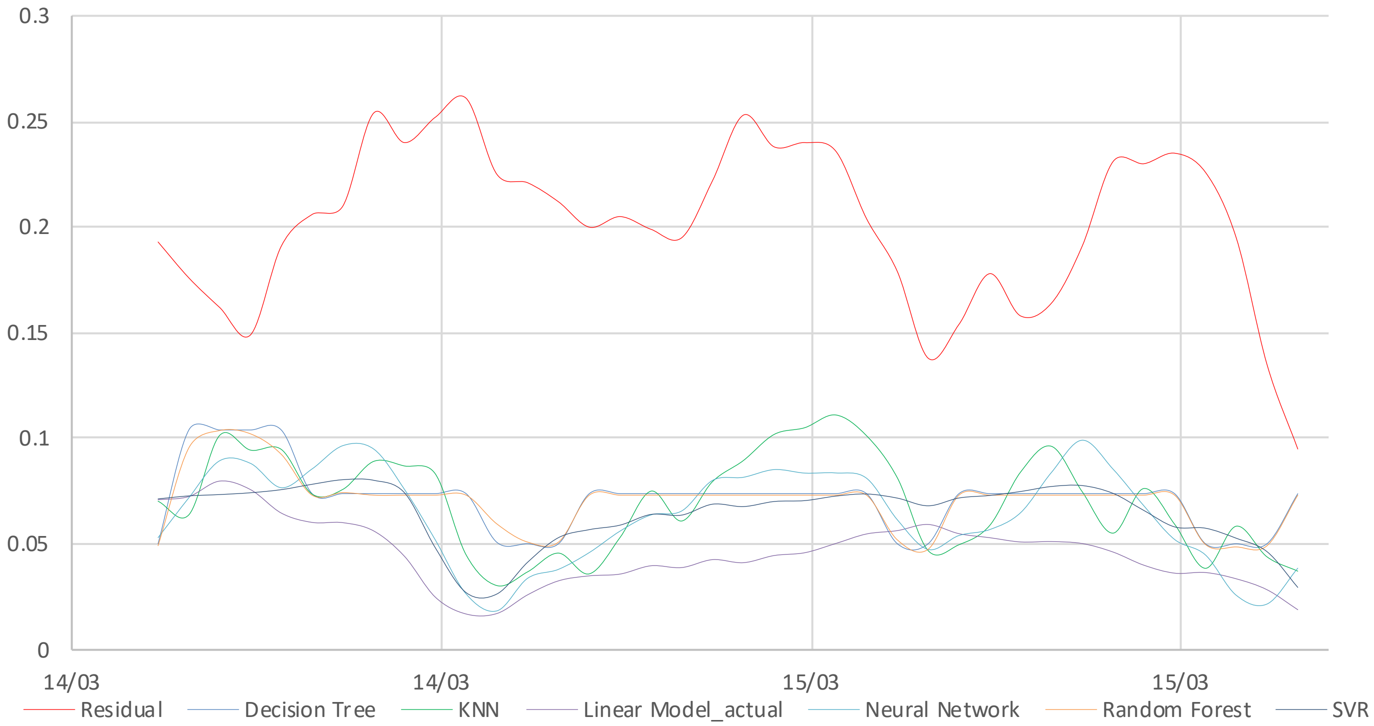
# Model performance forecasting phase

## Storm Surge (residual)



# Model performance forecasting phase

## Storm Surge (residual)




Queensland Government

# Model performance forecasting phase

## Storm Surge (residual)

| Model | Forecast Storm Tide vs actual Storm | Forecast residual vs residual |
|---|---|---|
| KNN | 0.996 | 0.199 |
| SVR | 0.999 | -0.021 |
| Decision Tree | 0.999 | -0.081 |
| Random Forest | 0.999 | -0.052 |
| Linear Model | 0.999 | -0.278 |
| Neural Network | 0.999 | 0.154 |

# Model performance Summary

- Modelling Storm Tide gave best testing phase performance over all models
  - Due to inclusion of tide predictions
- Storm Tide forecast gave best performance over all models
  - Due to inclusion of tide predictions
- Modelling of Storm Surge (residual) gave moderate testing phase correlations
- The Storm Surge forecast gave poor results

# Next phase

- Funding to continue with the project has ceased
- We are setting up a cluster of high-end PC's to take the project to the next phase
- Will be looking at the issues with the Storm Surge forecast
- Possible underlying problems with the Storm Surge forecast are:
  - Short learning data-set
  - Mismatch between the frequency of the input data and the BoM ACCESS-G model. Inputs= 10 minute, ACCESS-G = one hour
  - Offset between forecast and actual Storm Surge
  - Low occurrence of Storm Surge in the input data

Thank you

Daryl Metters
Coastal Impacts Unit
Queensland Department of Environment and Science

Queensland Government