UNIVERSITEIT·STELLENBOSCH·UNIVERSITY
jou kennisvennoot • your knowledge partner

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA
Denkleiers • Leading Minds • Dikgopolo tša Dihlalefi

INTERNATIONAL MARITIME INFORMATION SYSTEMS
IMIS GLOBAL

Model Selection: When is overkill - overkill?

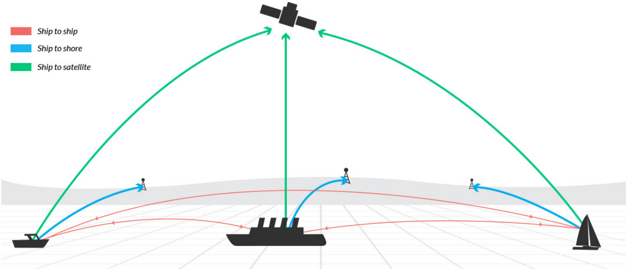
Creating prediction models in the maritime environment

Presenter: Neil Burger

Supervisors:
Dr T.L. Grobler & Prof W. Kleynhans

- **Research Focus:**
  - Improved vessel trajectory & coordinate prediction of Tanker and Cargo vessels.
- **Data - AIS:**
  - **A**utomatic **I**dentification **S**ystem Data
  - Each vessel uniquely identifiable - fitted with a transmitter
  - Recorded attributes we are interested in:
    - Longitude (LON), Latitude (LAT), Datetimestamp (t), Speed Over Ground (SOG), Course Over Ground (COG)

Figure: AIS transmittors: T-AIS & S-AIS[1]



[1] "What is AIS," VT Explorer. [Online]. Available: http://www.vtexplorer.com/what-is-ais/. [Accessed: 20-Oct-2020].

Spatial Distribution Map (SDM) - Celtic Sea , Bay of Biscay & English Channel

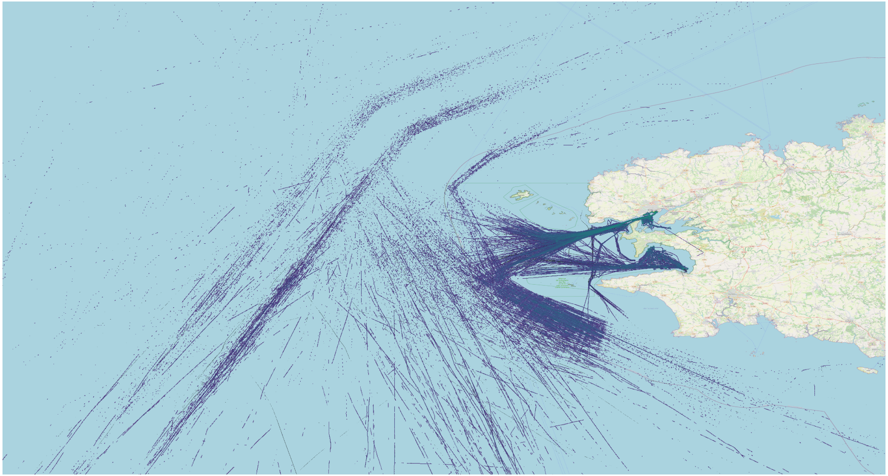


Figure: Signal Distribution from French Maritime dataset[1], T-AIS & S-AIS

[1] https://doi.org/10.5281/zenodo.1167594

Why prediction? & Vessel Nature

- Why are we interested in prediction?
  - Improved ETA prediction of vessels.
  - Illicit activity detection.
  - Trajectory reconstruction.
  - Data reduction/compression (520 Million observations per day - MarineTraffic).

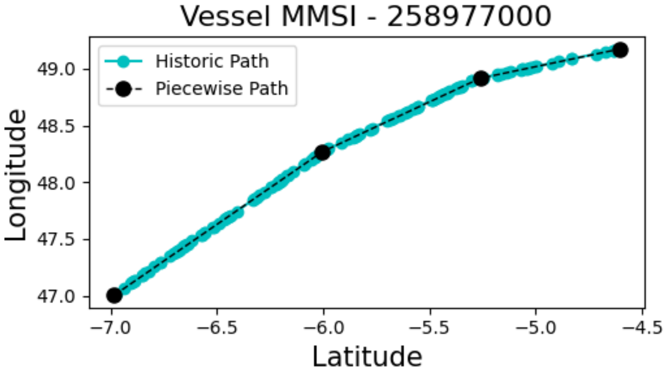


Figure: Piecewise linear nature of Tanker and Cargo vessels

Discrete Kalman Filter[2](DKF)

VS

Linear Regression Model[3](LRM)

- Journal paper - Uses a DKF to estimate coordinates.
- We think a DKF is **overkill** for this use case.

[2] K. Jaskolski, "Automatic Identification System (AIS) Dynamic Data Estimation Based on Discrete Kalman Filter (KF) Algorithm,"Zeszyty Naukowe Akademii Marynarki Wojennej, pp. 1–1, 12 2017

[3] C. N. Burger, T. L. Grobler and W. Kleynhans, "Discrete Kalman Filter and Linear Regression Comparison for Vessel Coordinate Prediction" 2020 21st IEEE International Conference on Mobile Data Management (MDM), Versailles, France, 2020, pp. 269-274

What is the Discrete Kalman Filter (DKF)?

- Estimates a state of a system by the recursive evaluation of two sets of equations, in the presence of noise
  - **Predictor equations (propagation):**
    - Estimate the current state and the associated uncertainty .
  - **Measurement update equations (correction):**
    - Updates parameters using the true current observed state.

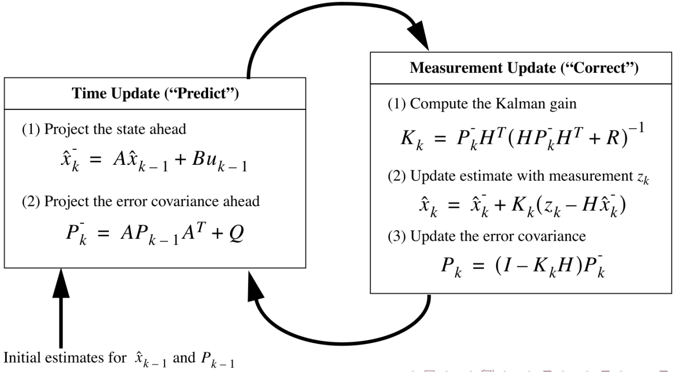


Figure: Complete DKF operation[3]

DKF:

- Estimates three variables simultaneously:
    - LON, LAT, SOG.
- **Q** & **R**
    - Used in the predict and measurement update equations respectively.
    - Sensitivity & Confidence of the DKF.
    - Updated each iteration.
    - Initialised to the values in the paper[2].

LRM:

- Setup similar to the DKF with two sets of equations.
- SOG estimated using a rolling window method.
- LAT and LON estimates are derived from the SOG.

[2] K. Jaskolski, "Automatic Identification System (AIS) Dynamic Data Estimation Based on Discrete Kalman Filter (KF) Algorithm,"Zeszyty Naukowe Akademii Marynarki Wojennej, pp. 1–1, 12 2017

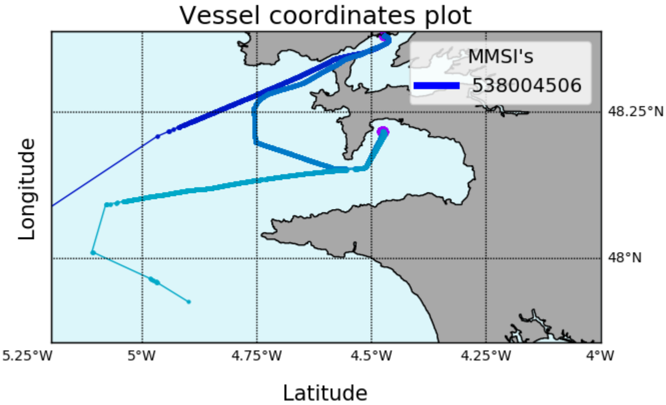
Experiment & Results

DKF vs LRM

Tested on 30 independent vessel trajectories,
each trajectory was undersampled
20 times with different interval sizes

Example - DKF vs LRM

- Vessel information:
  - MMSI: 538004506
  - #Recorded Observations: 339
  - Time observed: 3749s



Example - DKF vs LRM

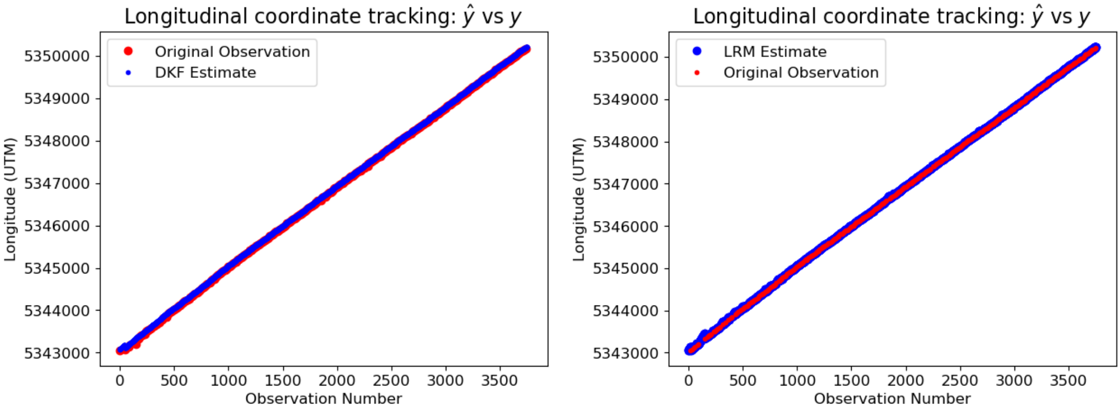


Figure: DKF vs LRM Longitude prediction

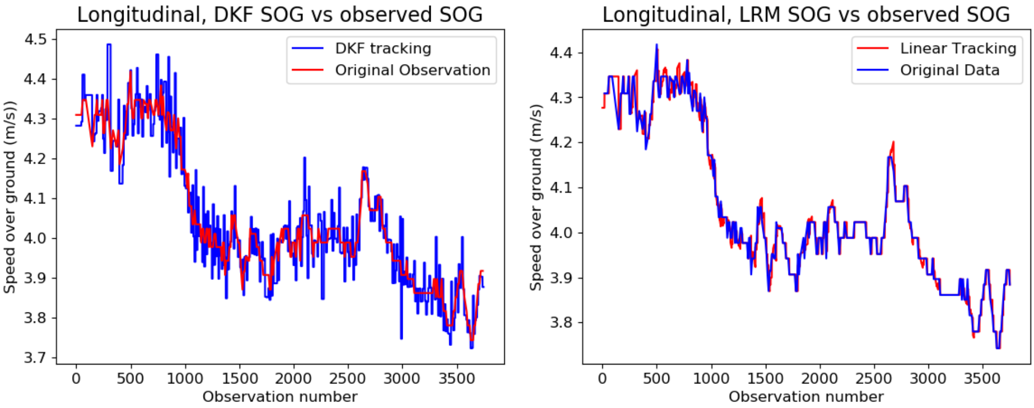


Figure: DKF vs LRM SOG prediction

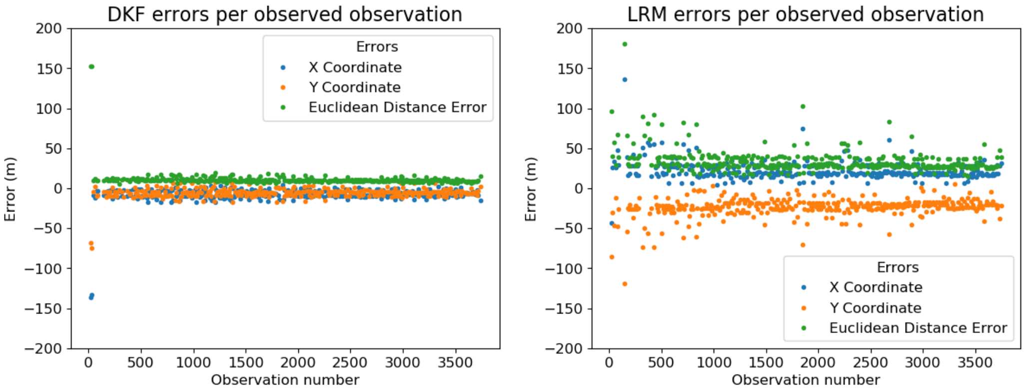


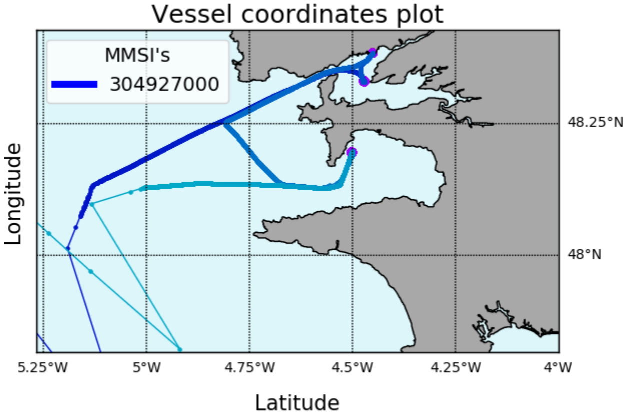
Figure: DKF vs LRM Euclidean Error

Error Comparison			
	DKF	**LRM**	**$\|\Delta MED\|$**
MED[4]	187	216	29

[3] Mean Euclidean Distance Error

Example - DKF vs LRM

- Extracted trajectory information:
  - MMSI: 304927000
  - #Recorded Observations: 699
  - Time observed: 7269s



Example - DKF vs LRM

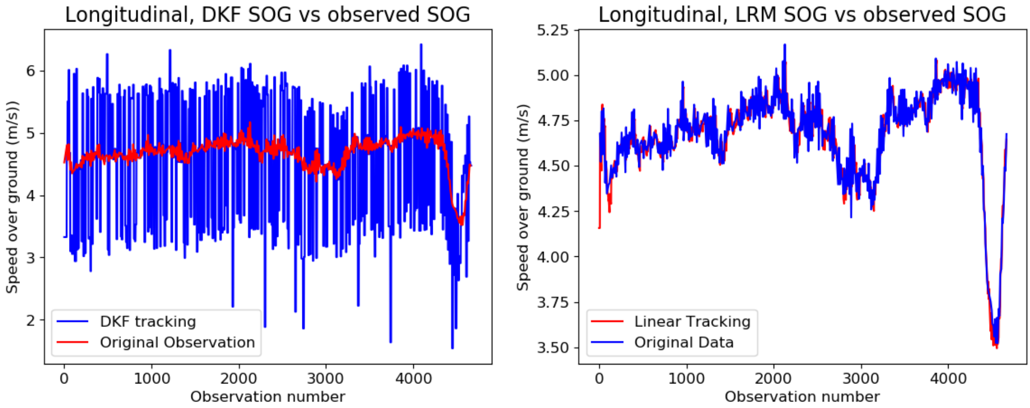


Figure: DKF vs LRM SOG prediction

Example - DKF vs LRM

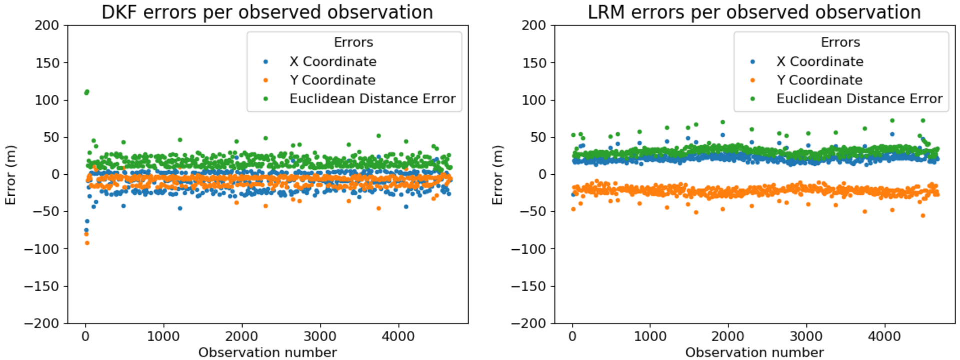
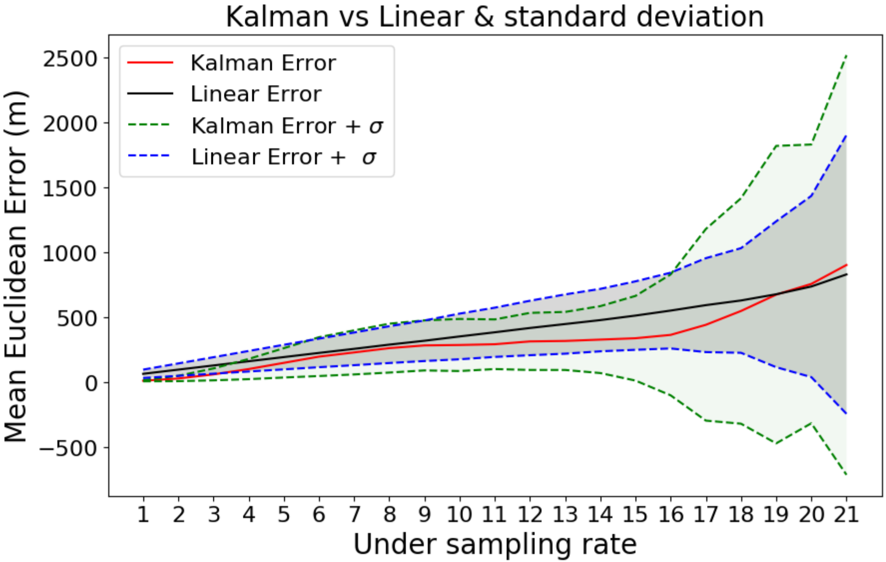


Figure: DKF vs LRM Euclidean Error

Error Comparison			
	DKF	**LRM**	**$\|\Delta$MED$\|$**
MED[3]	449	229	220

[3] Mean Euclidean Distance Error

Results - All vessels



Conclusion

- Model Complexity vs increase in model performance
- DKF requires the optimisation of both the **Q** and **R** matrices
- LRM requires the optimisation of a single variable - the window size
- LRM for the use on Tankers and Cargo vessels
    - Majority of the route segments are of a linear nature.

Figure: QR: Discrete Kalman Filter and Linear Regression Comparison for Vessel Coordinate Prediction

Current Work

- Extend LRM to make use of a priori information
  - Improved prediction performance
  - Improved trajectory reconstruction
- Reduce runtime complexity of prediction algorithms

Questions or Suggestions