Proceedings of the 25th International Conference on
Port and Ocean Engineering under Arctic Conditions
June 9-13, 2019, Delft, The Netherlands

Machine learning for tactical iceberg drift forecasting

Renat Yulmetov and Freeman Ralph

C-CORE, St. John's, Canada

ABSTRACT

Iceberg drift forecasting on tactical scales remains challenging due to the lack of accurate and timely data on ocean currents, winds, iceberg geometry and mass. Current operational models are mechanistic and require a set of constants to be additionally determined for each iceberg. Alternatively, statistical methods and dead reckoning have demonstrated greater performance up to the first 36 hours of forecasting. On the other hand, purely statistical models may not be able to perform in rare outlying cases. This study describes a neural network applied to iceberg drift forecasting, that is essentially a statistical approach. The network is trained and tested using iceberg drift data recorded during exploratory drilling offshore Labrador in 1979. Initial drift track parts, ocean currents and winds are used to train the network, and then the model is used to forecast 24 hours ahead. The model performance is promising and potentially can be improved even further given more data such as more accurate currents and winds, or additional inputs, for example, information about waves or sea surface gradients.

KEY WORDS: Icebergs, drift forecasting, machine learning, time series, neural network

INTRODUCTION

Tactical iceberg drift forecasting is still challenging after almost 40 years of development offshore Newfoundland. Complex bathymetry and the mixture of North Atlantic and Labrador ocean currents result in very dynamic environmental conditions in the vicinity of major offshore facilities. This in turn complicates met/ocean forecasting that directly affects iceberg drift modelling, because winds and ocean currents are usually the input parameters. Roughly speaking, predicting iceberg drift becomes a forecast built upon another forecast, which carries additional uncertainties. There is also a lack of accurate drift and ocean measurements caused by the remoteness of location and complexity of data collection methods.

Current operational drift models are based on the momentum equation and use wind and ocean currents to calculate the drag forces (Kubat et al., 2005). The input for these models comes from large-scale gridded models, providing large uncertainties for the instantaneous current/wind velocities. Ocean currents and winds measured *in situ* are more accurate, however, rarely available. The currents can be estimated based on the observed iceberg drift as was done by Turnbull et al. (2018). Recent studies show that purely statistical ocean current estimation delivers lower forecasting errors than the dynamic models in case of uncertain ocean input (Andersson et al., 2018).

This study demonstrates a machine learning approach to iceberg drift forecasting. It is a purely statistical approach that builds a model that minimizes the hindcasting error for a large number of observations. Then the model can be used in the forecasting mode to predict the drift several time steps ahead.

The paper will briefly present the method and the data which are used to build the drift model. Then the forecasted trajectories and forecasting error will be presented and analyzed. It will be followed by a short discussion about the machine learning capabilities.

METHOD

Machine learning is a generalization for a wide range of algorithms and techniques aiming to classify data, predict them or recognize patterns in the data. A shallow neural network (Rojas, 1996) will be used in this study to perform time-series prediction. This method constructs a non-linear function that maps a given set of arguments onto the set of predictions.

The method starts with some initial guess of the function and iteratively refines the guess by comparing the calculated function values with the known output values. The larger the error, the more adjustments to be made to the function. This process is called training.

The neural network can be represented by a set of interconnected elements, (called neurons) grouped in input, hidden and output layers (Figure 1). The function of the input can be calculated in the following way. The input data samples, represented by vectors with components x_i, are used to calculate weighted sums for each element in the hidden layer and apply the transfer function, which is usually the hyperbolic tangent, in the form:

$$y_j = \tanh\left(\sum_i w_{ij} \cdot x_i + b_j\right) \tag{1}$$

where w_{ij} are the weights corresponding to i-th input component of the j-th element and b_j are the bias coefficients corresponding to j-th element in the hidden layer.

The output layer, in turn, calculates its own weighted sum of the values y_j from all the elements in the hidden layer, then calculates one more weighted sum to find the prediction, as follows:

$$z_k = \sum_j \acute{w}_{jk} \cdot y_j + \acute{b}_k \tag{2}$$

where $\acute{w}_{jk}$ and $\acute{b}_k$ are the weights and bias coefficients in the output layer, j and k are now related to the elements in the hidden layer and the outputs correspondingly.

Once the prediction is made for given input sample, the weight and bias coefficients are adjusted via gradient descent based on the prediction error. Note that the weight and bias coefficients have to be randomized when the network is initialized. This sequence is repeated iteratively for various input samples, eventually minimizing the prediction error. The number of times the algorithm goes through the whole training dataset is called the number of epochs.

Applied to the iceberg drift prediction, the neural network predicts the iceberg velocity based on a few past iceberg observations, wind and ocean current velocities. The predicted iceberg velocity is then integrated to find the trajectory.

It is assumed that the iceberg velocity has been recorded at least twice prior to the forecast. In addition, the wind and ocean current velocities are expected to be known at the moment of forecast and their 24-h forecast is available. The following input samples are constructed for the neural network:

$$[U_x^{i-1}, U_y^{i-1}, U_x^i, U_y^i, V_{w,x}^i, V_{w,y}^i, V_{w,x}^{i+1}, V_{w,y}^{i+1}, V_{a,x}^i, V_{a,y}^i, V_{a,x}^{i+1}, V_{a,y}^{i+1}] \quad (3)$$

where U corresponds to the iceberg velocity, V_w is the ocean current velocity, and V_a is the wind velocity. The upper index indicates the moment of time, where i corresponds to the current moment. The output samples are just iceberg velocity component at the time $i+1$.

$$[U_x^{i+1}, U_y^{i+1}] \quad (4)$$

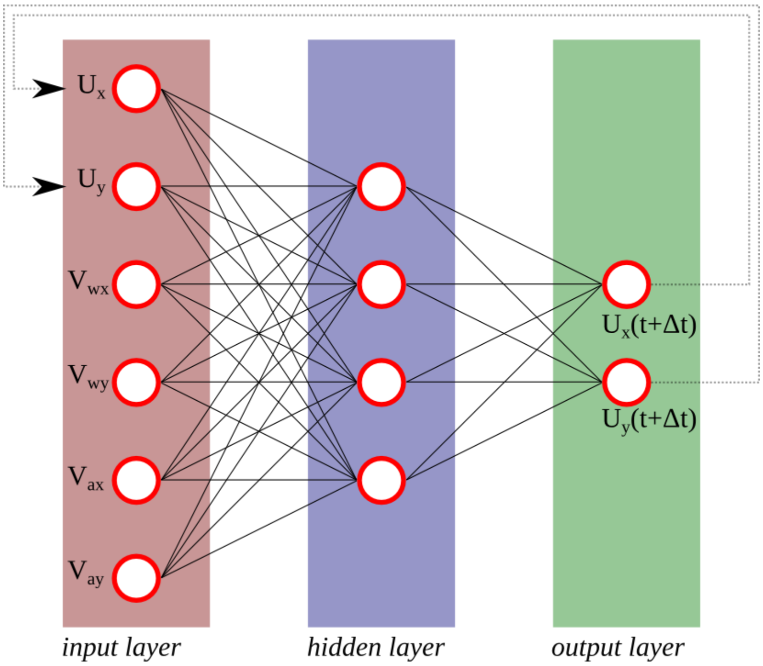

Figure 1. A schematic description of a neural network with six input values, four elements in the hidden layer and two output values.

These samples constructed for the observed iceberg tracks and recorded met/ocean data are used to train the model. The trained model is used to calculate the 24h velocity forecasts by performing consecutive single step predictions in a loop, where the input vector is updated each iteration using the newly-found velocity.

The number of epochs and the number of elements in the hidden layer are the parameters of the neural network. The optimal values of the parameters were selected by measuring performance of the network on the training dataset using their various combinations. The number of epochs and the number of elements in the hidden layer were found to be equal to 200 and 10 respectively. The method was implemented using an open-source Python module called *Scikit-learn* (Pedregosa et al., 2011).

DATA

The quality and the size of a training dataset determines the performance of the model. The data have to be accurate and representative of what can be expected when using the model to forecast. The data records from the exploratory drilling program in offshore Labrador have been used (Woodworth-Lynas et al., 1985).

The drilling was performed in 1979 using Pelerin and Neddril 2 drillships at the wellsite called Roberval K-92 (Figure 2). The well was located at 54°51'N, 55°44′W at a water depth of 269 m. The drilling was performed in the ice-free season, although the area is known to be covered by sea ice for a substantial portion of the year.

The iceberg records contain bearing and distance to icebergs from the drilling rig at hourly time intervals. The ocean currents measured on site were recorded hourly at 10, 25 and 50 m depths. Some of the current measurements were absent, therefore, the average current velocity for the whole water column was estimated based on the available measurements (Figure 3a). Note that some icebergs were drifting up to 45 km distance from the drilling rig, where the ocean currents could be different (Figure 4a) from those measured under the rig. In the absence of any other measurements, it was assumed that the measured currents could be applied to any of the icebergs recorded. Finally, the wind data is presented by hourly record of direction and speed (Figure 3b).

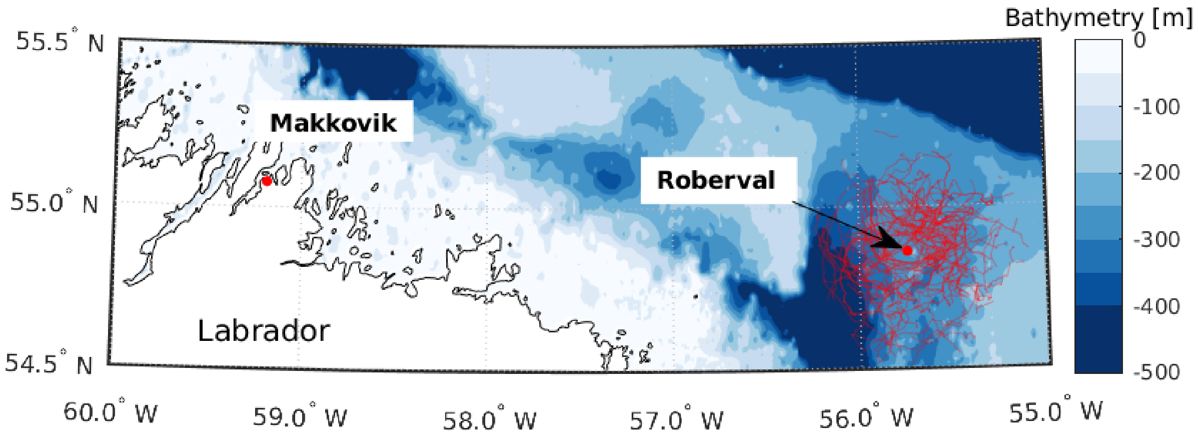

Figure 2. The map showing the drill site location, bathymetry and iceberg tracks (red lines).

It is possible to see that the ocean current measurements are likely to contain higher levels of noise than the measured wind velocities. Unfortunately, no information about the measuring equipment is available to assess the setup and data acquisition process.

Some icebergs were towed in order to prevent downtime and disconnection of the drilling rig. The corresponding tracks have been split into separate tracks by removing the parts of trajectories when these icebergs were towed.

The whole dataset consisted of 123 iceberg tracks with corresponding observed wind and ocean velocities. It resulted in 6809 individual data samples for training and testing. To achieve an unbiased performance estimate, the whole dataset had to be divided into the training and testing subsets. A model that is trained and tested using the same dataset may demonstrate unexpected performance when applied to new “unfamiliar” data. In comparison, a model that is poorly trained, but well-tested, is expected to underperform. This consideration resulted in the training data set consisting of 100 tracks (5634 samples) and the model was used in the forecasting mode for the remaining 23 “unfamiliar” tracks (1175 samples) to estimate its performance.

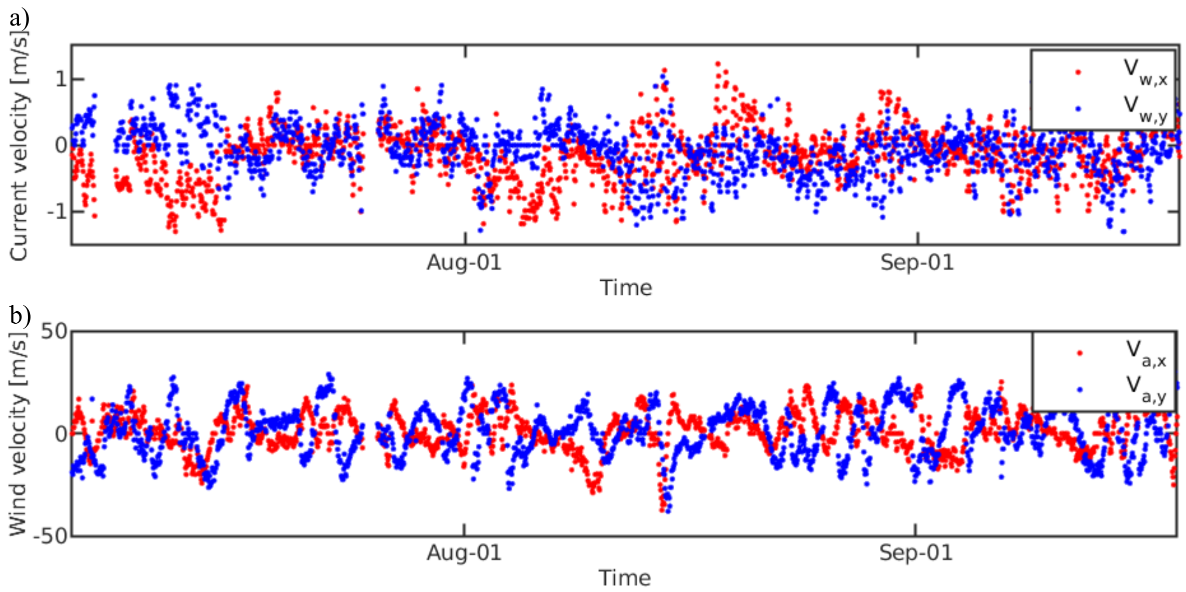

Figure 3. a) 50 m-depth-average ocean current velocity and b) wind velocity.

The iceberg drift velocity distribution is shown in Figure 4b. The average drift speed was found to be equal to 0.2 m/s, and the fastest iceberg drifted at 0.83 m/s. Most of the time, the drift speed did not exceed 0.5 m/s. A similar mean speed was found to the north-east of Greenland for icebergs that were not influenced by the sea ice (Yulmetov et al., 2016), while the maximum speed off Greenland reached a value twice as high during a storm event. This extreme value is also affected by the longer period of observation for the Greenland icebergs.

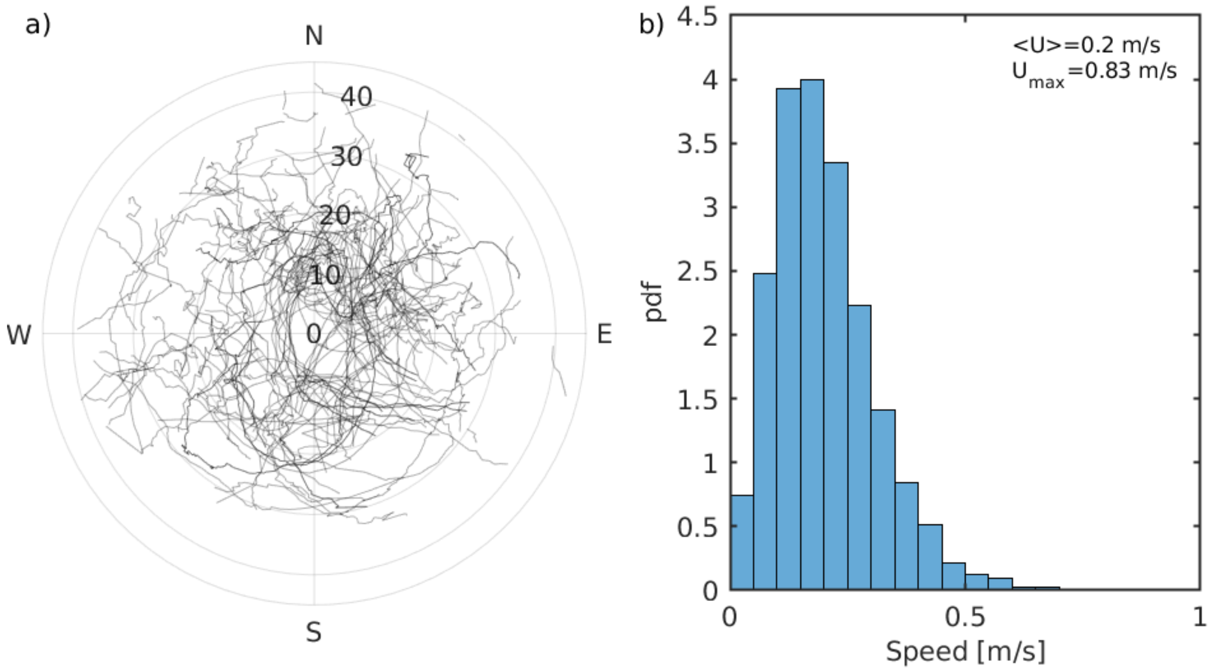

Figure 4. a) Iceberg tracks b) the probability density function (pdf) of iceberg drift speed derived from binning the drift speed observations.

RESULTS

After training the model was applied in forecasting mode to the icebergs in the testing dataset. Given only two hours of observations (three iceberg positions) and wind and weather forecasts for the next 24 hours, the iceberg velocity was forecasted and integrated by making 24 hourly predictions in the closed loop. Note, that the met/ocean forecasts were in fact observed winds and currents, which makes it is a hindcasting exercise.

Two example forecasts are shown in Fig 5 and compared to the observed trajectories. The first case represents the situation where the model delivers better performance. It is usually characterized by persistent currents and winds acting in same or similar directions. Even despite the high drift velocity, the forecast error remains low.

In the second case (Figure 5b), the iceberg underwent significant change in drift direction. Such processes are usually hard to capture accurately by any type of the drift model, whether dynamic or statistical primarily due to lack of accurate local data. Although the initial drift reversal from east to west was reproduced to certain degree of accuracy, the model eventually diverted the iceberg to the east. This resulted in the opposite drift direction leading to high forecast error. Note, that this was the case of a slow-drifting iceberg caused by a slow ocean current changing direction. Given a certain level of noise in the ocean current measurements, in cases of low drift and current speeds the relative uncertainty is large, which leads to significant forecasting errors. In this case, the iceberg travelled less than half the distance than in the first case, but the predicted trajectory error was unacceptably large.

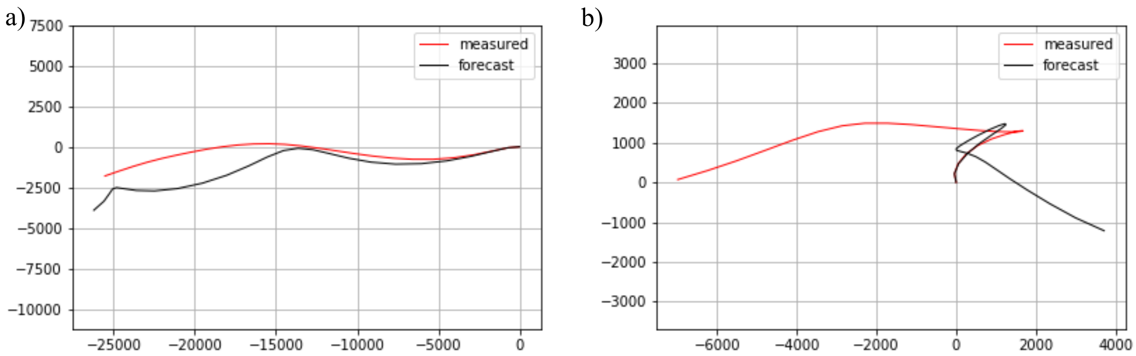

Figure 5. Examples of observed iceberg trajectories and their a) good and b) poor forecasts.

DISCUSSION

The model performance can be assessed based on the average forecasting error progression in time. The forecasting error was calculated as the distance between the predicted and observed iceberg locations.

$$\delta(t) = \left\|\vec{X}_o(t) - \vec{X}_f(t)\right\| \tag{5}$$

where $\vec{X}_o(t)$ is the observed iceberg posit ion and $\vec{X}_f(t)$ is forecasted iceberg position, $\delta(t)$ is essentially the forecasting error progression in time. Once averaged between multiple icebergs, it becomes a good estimate of the model accuracy.

Figure 6a shows the forecasting error issued for all of the forecasts in the testing dataset. The thick red line is the average error. Clearly, there is large variance in the individual forecast performance. Some forecasts are very accurate, but there are few forecasts that are unacceptable and provide more than 15 km error in just 12 hours. These errors appear because of the model misbehavior caused, in its turn, by inaccurate data used for training.

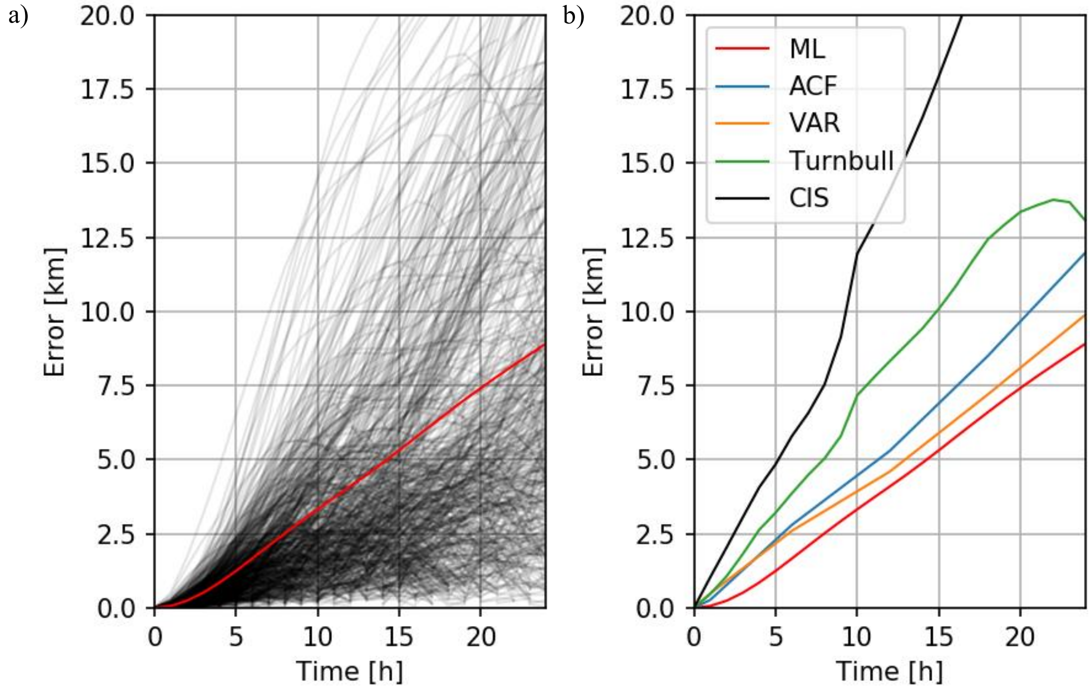

Figure 6. a) Forecasting error for all of the individual tracks (grey), average forecasting error (red). b) average performance for various drift prediction models.

It is likely that the ocean currents measured at the drilling rig location differ significantly from those 35 km away. Understanding the range limit would be valuable even for strategic locations of future subsea current measurements systems. The 24-h average individual forecast error, however, seems to demonstrate a negative trend based on the iceberg-facility distance (Figure 7). This occurs because the forecasting error is minimized for the whole dataset at once, without assuming that the currents at the distance are less certain. It might be noticed that the lowest levels of error correspond to the bins with the highest data density, i.e. at 10-20 km or at 30-35 km. It is suggested for future modeling to filter data samples based on the distance to the facility while training the model.

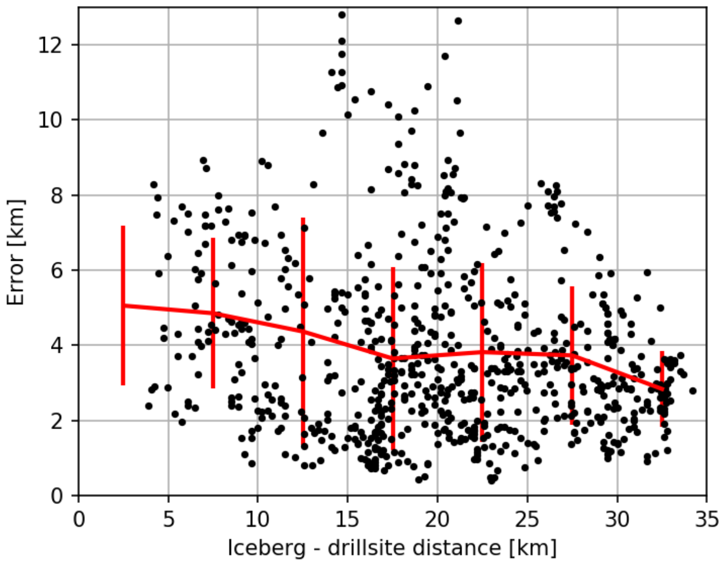

Figure 7. Individual forecast error seems to negatively depend on distance to the facility. Red errorbar shows mean and standard deviation for 5 km-binned data.

The performance of multiple models, including the one constructed using the machine learning approach, are compared in Figure 6b. The average forecasting errors presented below were calculated for the models applied to various datasets. All of them, however, were used to forecast iceberg drift in the offshore Newfoundland waters.

Andersson et al. (2018) managed to perform a comparative analysis of multiple models applied to the same dataset. One of their best-performing models, vector auto regression (VAR), was purely statistical and delivered less than 10 km error after 24 hours of forecast (Andersson et al., 2019). This model predicts ocean current velocity using vector auto regression large period of current data derived from gridded large-scale reanalysis. Then the predicted current is used to predict iceberg drift velocity assuming that it is 2% of the wind velocity relative to the ocean current.

A hybrid model (ACF), that calculates a correction (ancillary current) to the ocean current by using the moving horizon estimator and then integrates the equation of motion (Andersson et al., 2016), results in marginally larger error on the same dataset.

The dynamic model of Turnbull et al. (2018) integrates the equation of motion. Unless the measurements are available, the ocean currents are estimated locally using a similar linear relationship. The model also performs drag coefficients estimation based on the observed drift. It was applied to 14 iceberg tracks recorded using radar and resulted in 13 km error in 24 hours. The CIS model (Carrieres et al., 2001) forecasts for the same icebergs, but using winds and currents derived from large scale models, resulted in more than twice larger errors. This was a clear indication of the impact of inaccurate ocean current input for the model. Note, that actual forecast winds were used in Turnbull et al. (2018) in contrast to this study that employs only the measured data, which may explain the performance difference.

As seen from Figure 6b, the machine learning approach results in the lowest error levels overall. Although it might seem marginal, given the number of forecasts, the difference is statistically significant.

It is of large interest to understand for each model the level of uncertainty in the input data and the level of uncertainty in the model itself. Given perfectly accurate input data the dynamic and statistical models should demonstrate comparable level of performance.

CONCLUSIONS

This study presents a new statistical approach to iceberg drift forecasting. The approach is based on a shallow neural network implementation. The network uses two past iceberg velocity records, current wind and ocean velocities and their forecasts to predict the iceberg velocity. The model is trained using a large iceberg drift dataset recorded using marine radar during exploratory drilling offshore Labrador in 1979. The model performance is estimated using a part of the dataset that was not used for training.

Preliminary testing demonstrates that the model is capable of achieving on average less than 10 km error in 24 hours. However, the individual forecasts may significantly deviate from the observed tracks in some cases. Most often, those are the cases when the drift speed is low and ocean currents are slow. It occurs due to larger relative uncertainty in the input data.

It has also been shown that the forecasting error does not depend on the distance to the facility. This is the feature of the algorithm that minimizes error across the whole dataset independently of the distance to the current measurement point. It is suggested for the future development to penalize the data samples corresponding to the distant iceberg tracks during training.

In general, it is still hard to quantify uncertainty of the input data vs. uncertainty of the model. In order to do so, more accurate data have to be collected offshore Newfoundland. A small number of strategically-placed current profilers around the major offshore developments would enable more precious data collection and be beneficial to many players across various industries.

REFERENCES

Andersson, L.E., Scibilia, F., Copland, L., Imsland, L., 2018. Comparison of statistical iceberg forecast models. Cold Regions Science and Technology 155, 69–89. https://doi.org/10.1016/j.coldregions.2018.07.003

Andersson, L.E., Scibilia, F., Imsland, L., 2019. An iceberg forecast approach based on a statistical ocean current model. Cold Regions Science and Technology 158, 128–142. https://doi.org/10.1016/j.coldregions.2018.11.016

Andersson, L.E., Scibilia, F., Imsland, L., 2016. An estimation-forecast set-up for iceberg drift prediction. Cold Regions Science and Technology 131, 88–107. https://doi.org/10.1016/j.coldregions.2016.08.001

Carrieres, T., Sayed, M., Savage, S., Crocker, G., 2001. Preliminary Verification of an Operational Iceberg Drift Model, in: Proceedings of the 16th International Conference on Port and Ocean Engineering Under Arctic Conditions. Ottawa, Ontario, pp. 1107–1116.

Kubat, I., Sayed, M., Savage, S.B., Carrieres, T., 2005. An Operational Model of Iceberg Drift. Int J Offshore Polar 15, 125–131.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., Duchesnay, É., 2011. Scikit-learn: Machine Learning in Python. Journal of Machine Learning Research 12, 2825−2830.

Rojas, R., 1996. Neural Networks: A Systematic Introduction. Springer-Verlag, Berlin Heidelberg.

Turnbull, I.D., King, T., Ralph, F., 2018. Development of a New Operational Iceberg Drift Forecast Model for the Grand Banks of Newfoundland, in: OTC Arctic Technology Conference. Presented at the OTC Arctic Technology Conference, Offshore Technology Conference, Houston, Texas, USA. https://doi.org/10.4043/29109-MS

Woodworth-Lynas, C.M.T., Simms, A., Rendell, C.M., 1985. Iceberg grounding and scouring on the Labrador Continental Shelf. Cold Regions Science and Technology 10, 163–186. https://doi.org/10.1016/0165-232X(85)90028-X

Yulmetov, R., Marchenko, A., Løset, S., 2016. Iceberg and sea ice drift tracking and analysis off north-east Greenland. Ocean Engineering 123, 223–237. https://doi.org/10.1016/j.oceaneng.2016.07.012