Interactive Probing of Multivariate Time Series Prediction Models: A Case of Freight Rate Analysis

Xu Haonan[*], Li Haotian[‡*], Wang Yong[*]

[*]Singapore Management University, [‡]The Hong Kong University of Science and Technology

Motivation

Freight rate analysts actively ask the **"what if"** question: what happens to the freight rates when an unusual event takes place? This is because freight rates are highly volatile and large vessel sizes pose magnifying effects on the final fees.
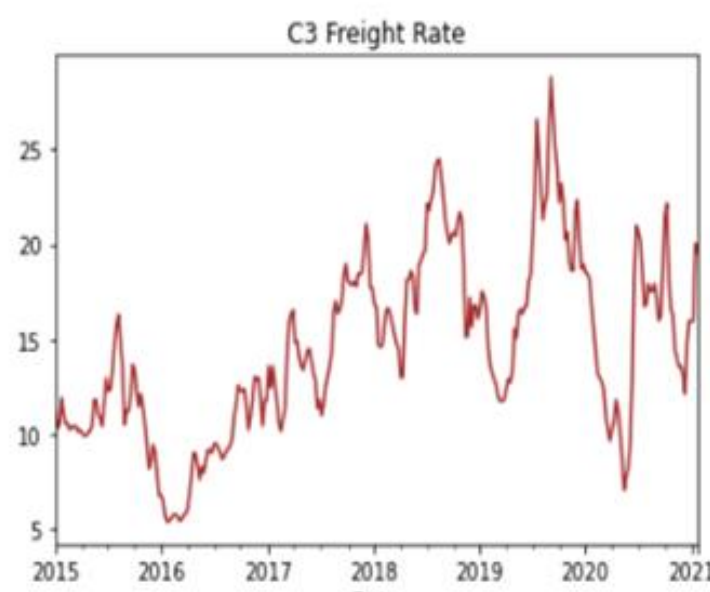

Figure 1: High volatility. Unit: $ per ton.

Figure 2: Huge vessel size. Image by Martin Farrer via Guardian.

Previous qualitative gauging methods increase risk exposure. A visualization solution is needed to enable **intuitive** variable manipulation and create market scenarios.

Introduction

We introduce a visual solution to **interactively** generate market scenarios and **quantitatively** indicate the new predictions, below are the major contributions:

1. An interactive time series input and output view to intuitively show what-if results.
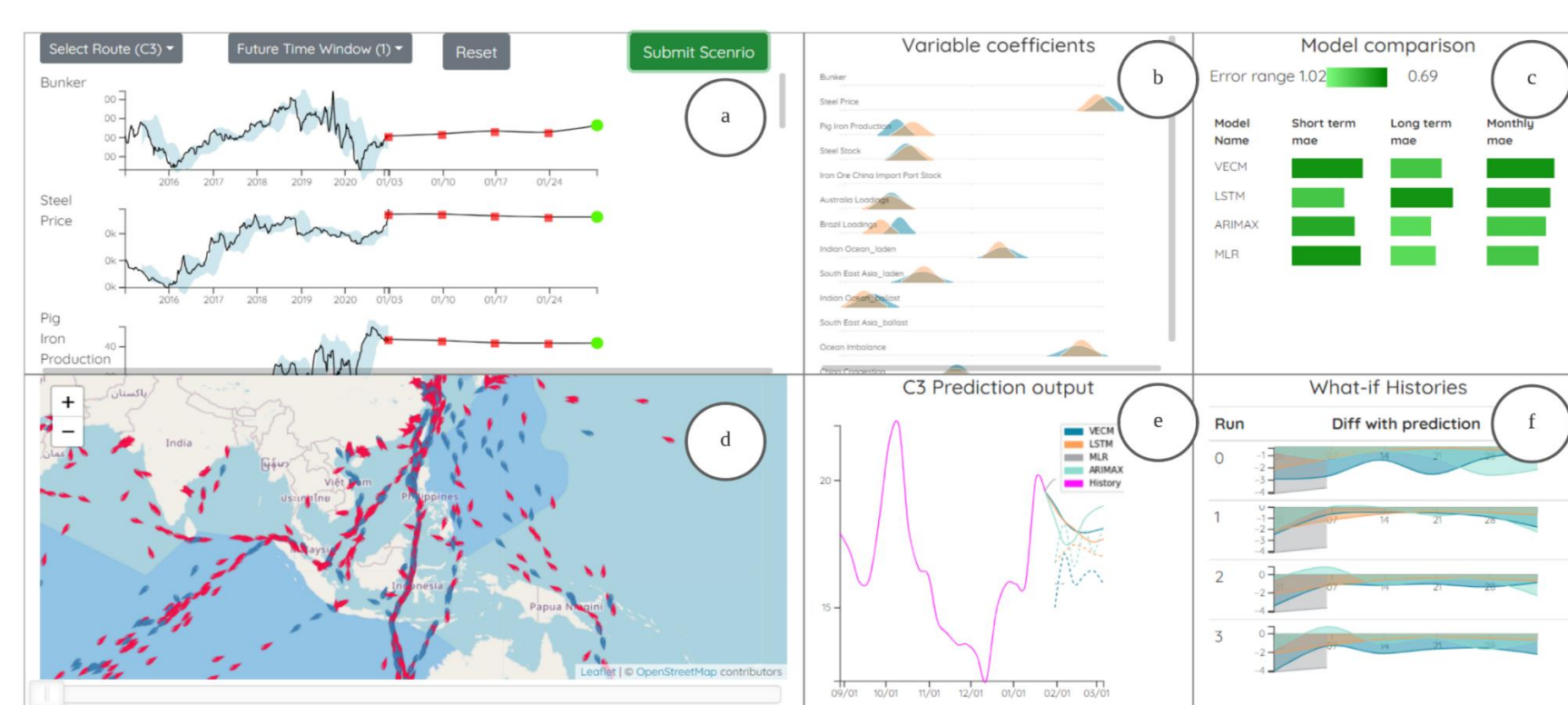2. A usability study with domain experts.

Figure 3: (a) The what-if input view allows users to detect past patterns and adjust what-if inputs. (b) Coefficient impact view facilitates indicator effect comprehension. (c) The model comparison view encodes model performances for model selection and ranking. (d) The Spatial data aggregation view supports spatial data summary, which helps users to anticipate possible future values. (e) The prediction and what-if view compares, and contracts outputs generated from existing and hypothetical data. (f) What if histories view logs hypothetical scenarios.

Visual Design of the Views

What-if Input View: It's a compound horizontally spliced line chart that allows users to *create and update market scenarios*. It supports zooming, Bollinger Band visualization and dragging to alter what-if values. Forward fill strategy is employed.
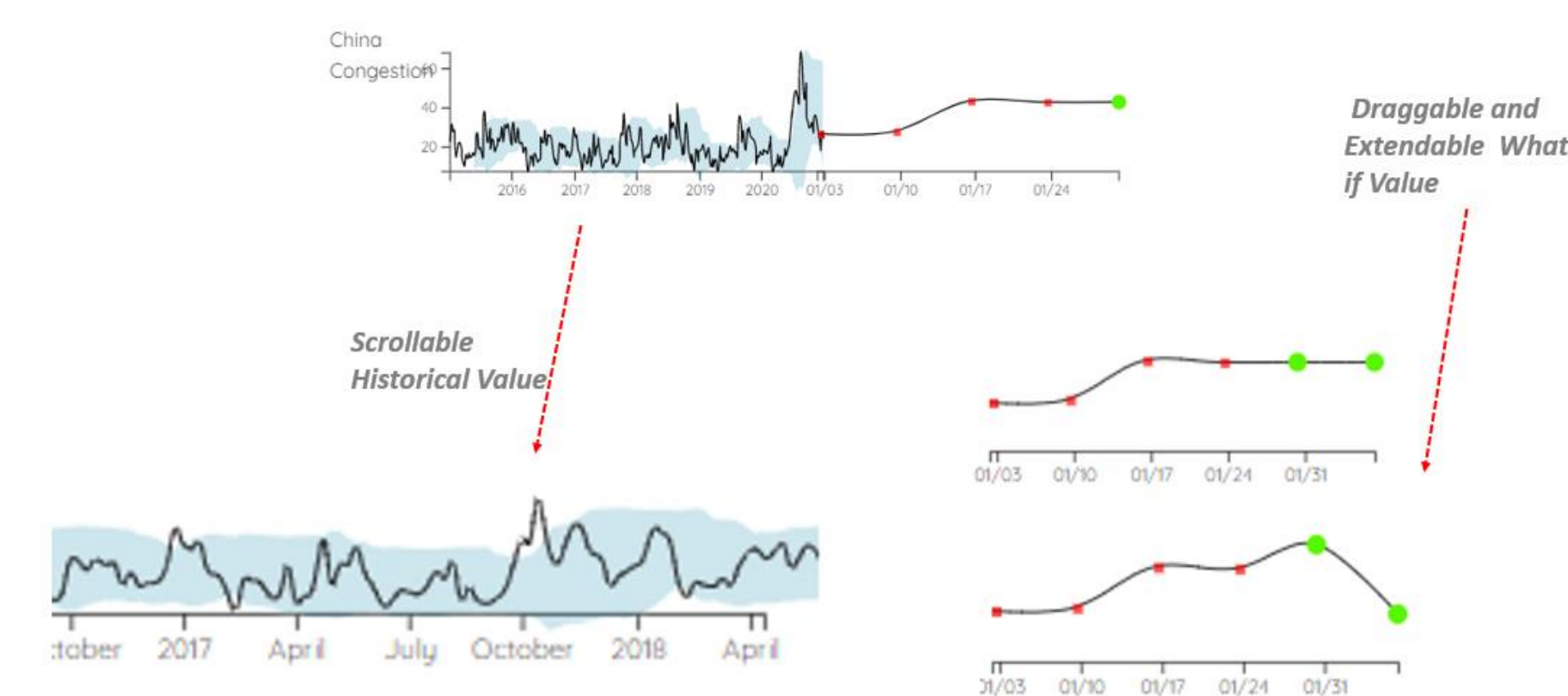

Figure 4: Design of the variable input view.

The **coefficient impact view** in Fig. 5(a) is vertically aligned area chart shaped in the normal distribution, variables' relative locations indicate their effects on the target freight rate.
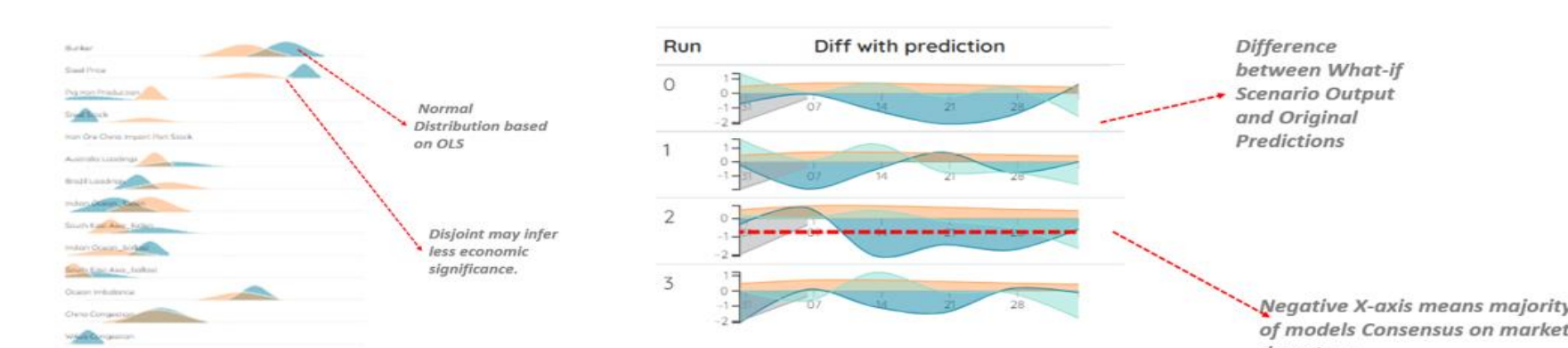

Figure 5: (a) Design of the coefficient Impact view. (b) Design of the what-If history view.

The **what-if history view** in Fig. 5(b) displays the what-if difference with the original predictions. The X-Axis is positioned at the middle of the Y-Axis (Difference).

The **spatial data aggregation view** in Fig. 6 is a map displaying the raw spatial data. The tip of the droplets indicates the vessel's heading, and the colour reflects cargo status.
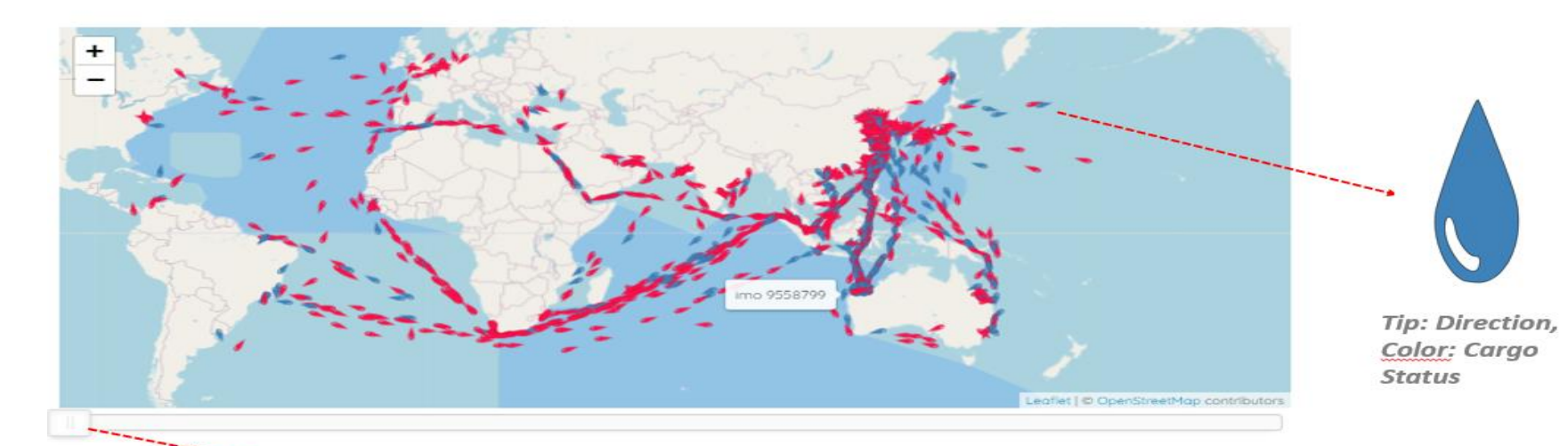

Figure 6: Design of the spatial data aggregation view.

Evaluation

We interviewed 3 freight traders (E1, E2, E3 respectively) and asked them to perform the same set of actions on the UI. The process explored the effect of a torrential flood around a Brazil iron ore mine on route C3's freight rate (Vergilio et al., 2020).

1. The user selects route C3 from Fig. 3(a) and waits for the pre-trained models' predictions.
2. The user locates Brazil Iron Loadings using the coefficient impact view (shown in Fig. 6(a)), then drags down the value for 2000 mt as shown in Fig. 6(c).
3. In Fig. 7(a), the user notices the middle line is y=-2, and all difference curves are below 0.
4. The readings suggest that the freight index is going to drop around 16% to around $15/mt.
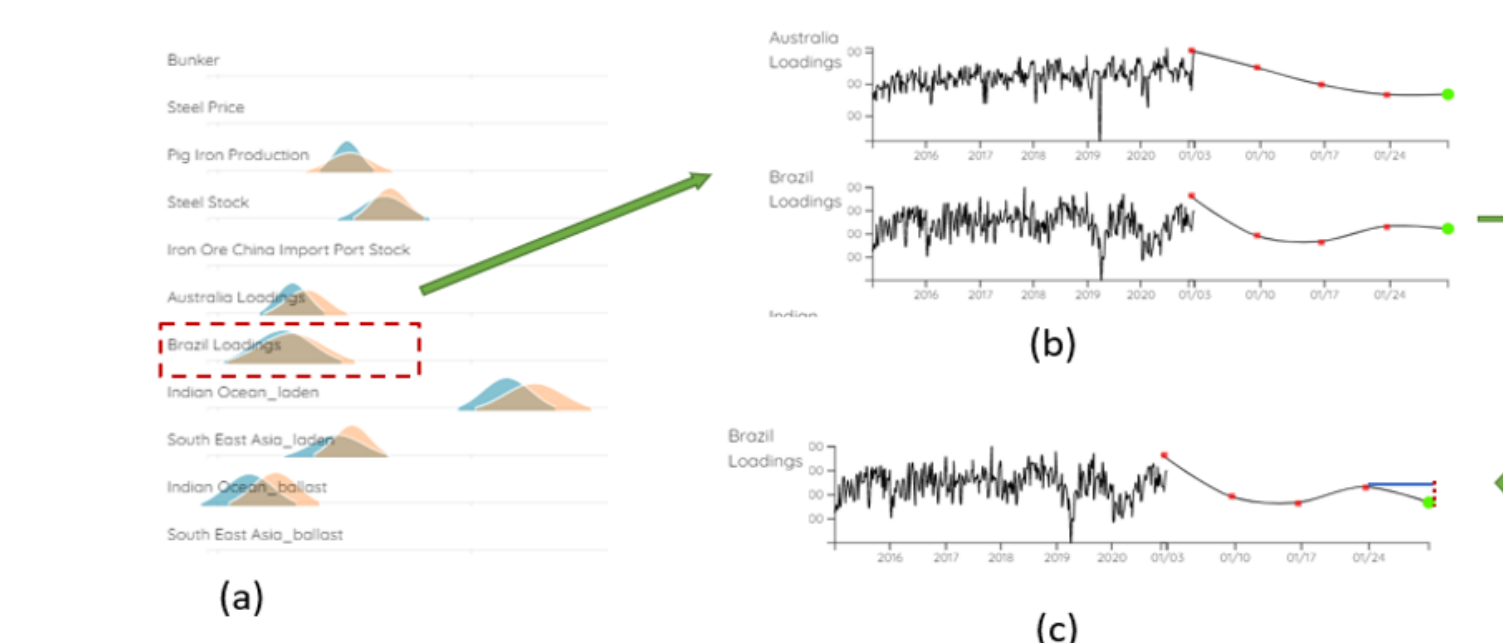

Figure 6: (a) and (b) Locating target variable. (c) The user changes the scenario value (b) by -2000mt.

5. The user looks for the empty vessels approaching Brazil in the map view Fig. 7(b) and takes down the IMOs. some unfortunate ship owners will be forced to charter out the vessels at a much lower price, that's an opportunity.
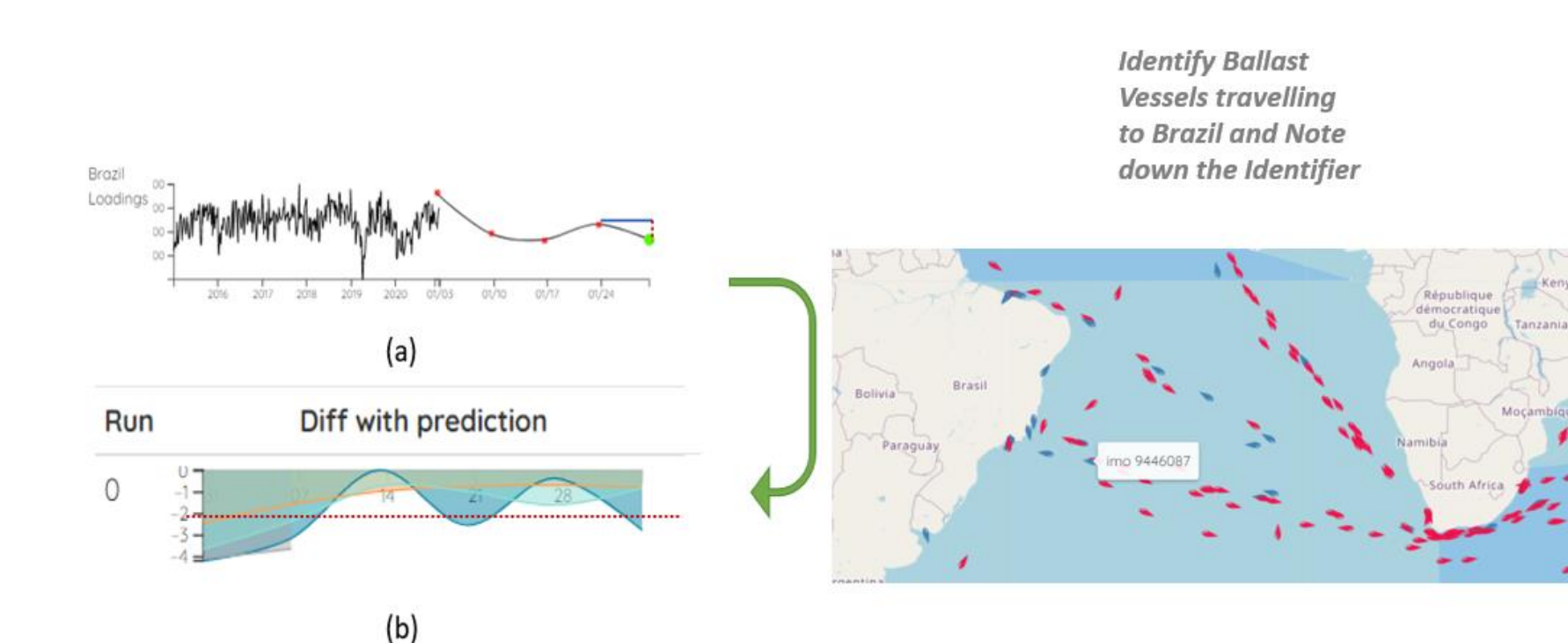

Figure 7: (a) User observes output of the what-if scenario. (b) User locates chartering targets in the map.

Evaluation Outcomes and Discussions

Positive feedbacks:

1. Intuitive: what-if values easy to interact with.
2. Explainable: variable view enables a better understanding of the models.

Limitations and plans:

1. Variable Cointegration: the tool assumes independence between exogenous variables. We plan to implement statistical learning on individual series (Kegel et al., 2017).
2. Assumption of normal distribution: the error range for coefficients assumed to be normal. we intend to add t-distribution support.
3. Forward fill strategy: the process accumulates error as terms increase. We will benchmark the simple forward fill strategy against the statistical learning approach and find the right trade-off.

CONCLUSION

We present a what-if tool for creating and probing market scenarios for freight rate prediction. We evaluated the solution with industry practitioners. The user study reflects that the solution enables users to accurately capture freight index changes and respond to market events proactively.

In the future, we plan to improve the accuracy of the machine-filled what-if data through learning statistical patterns of variables to resolve the effect of cointegrating variables and forward fill strategies.

References

Vergilio, Cristiane dos Santos, et al. 'Metal Concentrations and Biological Effects from One of the Largest Mining Disasters in the World (Brumadinho, Minas Gerais, Brazil)'. Scientific Reports, vol. 10, no. 1, Apr. 2020, p. 5936. www.nature.com

Kegel, Lars, et al. Generating What-If Scenarios for Time Series Data | Proceedings of the 29th International Conference on Scientific and Statistical Database Management. 17 June 2017