Multi-Output Chaining for Time Series Forecasting to Estimate Impact of Supply Chain Disruption: Case of Russia's Invasion of Ukraine on US Import and Export

Hyeonsup Lim, Majbah Uddin, Yuandong Liu, Shih-Miao Chin, and Ho-Ling Hwang
Oak Ridge National Laboratory (ORNL)

Introduction & Data Source

Motivation

- On top of "already-happening" supply chain disruptions due to the pandemic, the Russia's invasion of Ukraine brought an unprecedented challenge to predict and mitigate its impact to supply chain disruptions globally.
- How can we measure the overall impact to U.S. import & export?
- Some impacts may be noticeable and visible, but the key challenge is "we do not know what we do not know" happening in further downstream industries of supply chains.

Data Source and Scope

- USA Trade Online data (Jan 2002 – Sep 2022)
- Total 1,148 and 1,147 commodity types for imports and exports respectively (by 4-digit HS codes)
  - The covered HS codes account over 99% of US import and export value
- Independent/dependent variables: total monthly value of goods imported from and exported to U.S. by HS code

Proposed Model Approach

What is Multi-Output Chaining? And Why?

- Multiple target variables (i.e., value of import and export for each commodity) are estimated sequentially by a defined upstream-downstream relation.
- Cross-correlations, with lags in months, are used to account for the relationship between upstream and downstream in forecasting the supply chain disruption impacts. The main purpose is to capture both directly and indirectly impacted commodities for U.S. imports/exports.

Seasonal Auto-Regressive Integrated Moving Average with eXogenous factors (SARIMAX)

- Consideration of Russia's invasion as one exogenous variable (+ Great Recession, Covid, Suez Canal, Big Freeze)
- In addition, each HS code's model may utilize multiple other HS codes' prior months' value (sequentially estimated) as exogenous variables.

Scenarios

- Depending on "when" the invasion is going to be ended
  - Base scenario – Jan 2023
  - Four additional scenarios – Mar, Jun, Sep, and Dec 2023

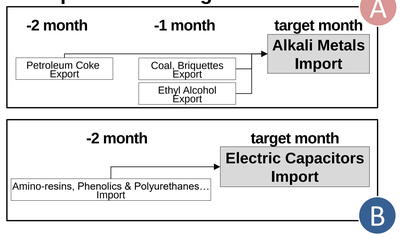
Overall Flow Chart

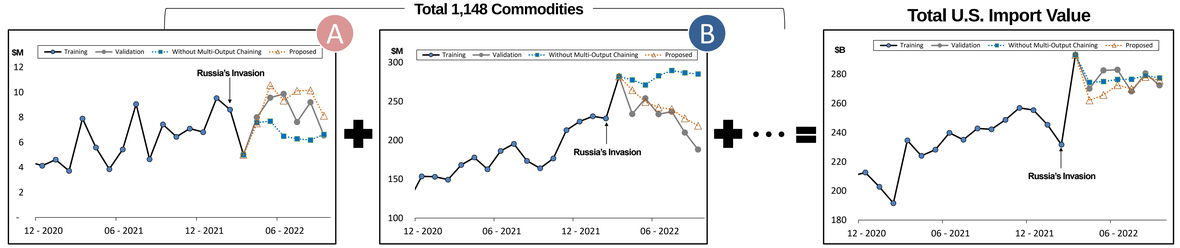
1. Make each variable stationary
2. Calculate cross-correlation between all variables – only select variables where highest correlation is with negative lags (past)
3. Split training/validation sets
4. Initial model fit with the training set
   i. SARIMA with one target variable
   ii. Add another variable with highest cross-correlation
   iii. Stop if no significant improvement in the validation sets
5. Final model fit with the entire dataset
6. Apply the final model in Step 5
   i. Predict +1 month import/export values for all HS codes
   ii. Update input variables for each HS codes based on the previous month estimates
7. Repeat Step 6 until the end (Dec 2023)

Model Result

Compared to the SARIMAX model without consideration of multi-output chaining, the proposed method reduced the mean absolute error by 36% and the root mean square error by 26%.

Examples of Chaining

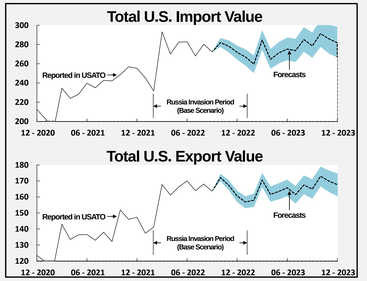




2023 Total Estimated US. Imports/Exports Value Forecasts by Scenarios

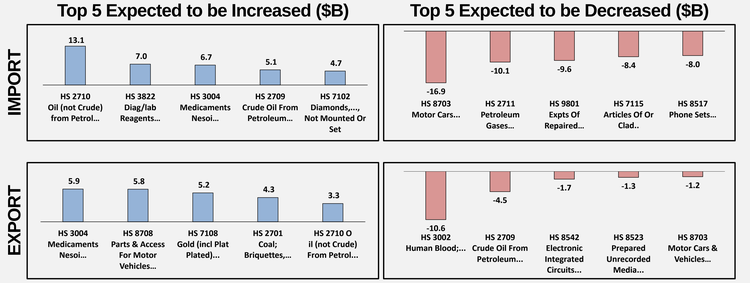
Scenario Invasion Ending on	Import Value	Export Value
(0) Jan 2023 (Base)	$3,319B (Δ=0)	$1,980B (Δ=0)
(1) Mar 2023	$3,325B (+$5B, 0.2%)	$1,998B (+$17B, 0.9%)
(2) Jun 2023	$3,330B (+$10B, 0.3%)	$2,015B (+$34B, 1.7%)
(3) Sep 2023	$3,334B (+$15B, 0.5%)	$2,031B (+$51B, 2.6%)
(4) Dec 2023	$3,340B (+$21B, 0.6%)	$2,048B (+$68B, 3.4%)

Estimated Impacts by Scenarios
(compared to the base scenario)



Base Scenario (Invasion Ending on Jan 2023)

Most Impacted Commodities (by 4-digit HS Codes) – Base vs Scenario 4



Estimated Impact = (4) Invasion Ending in Dec 2023 – (0) Invasion Ending in Jan 2023

Conclusions

Key Contributions:

- Employed the multi-output chaining for U.S. import/export value forecasting.
- Proposed a framework that can be sequentially applied where there are many (over thousand) target variables present.
- For certain commodities, the proposed method outperforms the alternative method, which do not consider cross-correlation with lags.
- The proposed method can potentially reveal upstream-downstream relations for selected commodities (one used as early indicator of others)

Limitations and Suggestions for Future Studies:

- The target variable is limited to the total value of imports and exports.
  - The model performance could have been better if other variables, such as quantity or price, were considered as well. Note that the USATO data provides quantity and price information for 10-digit HS codes only.
- Consideration of different levels of commodity details (2-digit, 6-digit, 10-digit), foreign countries, and other external data sources (GDP, economic data, sanctions, port-level information, etc.) can be included to further improve the model performance and its applicability in the future.

Contact Information

Hyeonsup Lim, ORNL
limh@ornl.gov