Time series forecasting with a statistical ensemble model:

An application for global goods trade volume prediction

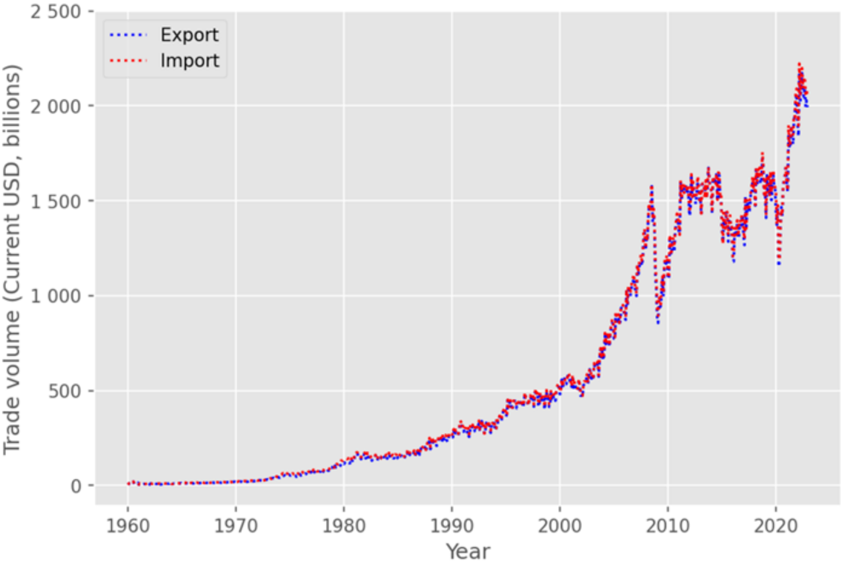
Fitahiana A.A. Razafimahenina
Madagascar AI Lab, Universitat Rovira i Virgili
Mihaela Simionescu
University of Bucharest, Institute for Economic Forecasting, Romanian Academy

20–21 February 2025
University of Pretoria, Economic Research Southern Africa Workshop

Introduction

- Several factors contribute to the complexity of predicting global trade volume for goods, including:
    - Global economic growth
    - Trade policies
    - Supply chain disruptions
    - Technological advancements
- The goal of this paper is to determine whether using a statistical ensemble model offers a significant advantage over relying on individual component models.

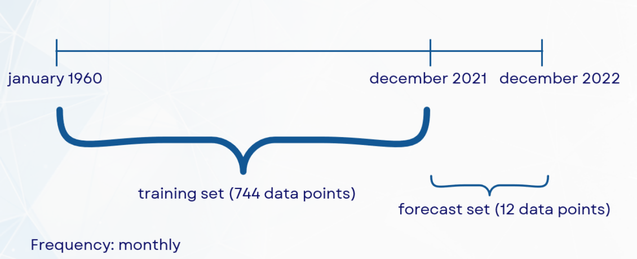
Dataset



Data source: IMF database

Temporal coverage

- Temporal coverage: January 1960 – December 2022
- Training set: 744 data points
- Forecast set: 12 data points
- Frequency: Monthly



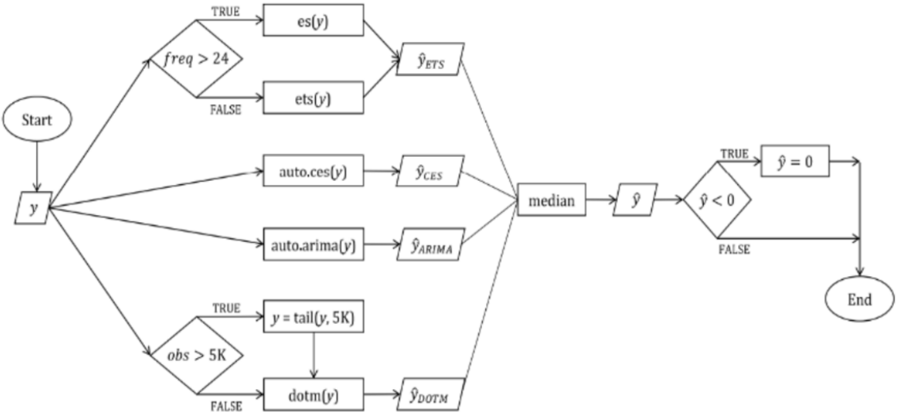
Data Processing

- Trade volume = mean(Export volume, Import volume)
- Values not adjusted for deflation
- Standardization: Standard Scaler, centering the distribution of features around a mean of zero and scaling them to possess a unit variance.

Models

- SCUM: Simple Combination of Univariate Models
- The Statistical Ensemble Model is a SCUM (Petropoulos and Svetunkov, 2020).
- Forecasts and prediction intervals are derived from four distinct univariate statistical models, using the median:
  - Exponential Triple Smoothing (ETS)
  - Complex Exponential Smoothing (CES)
  - AutoRegressive Integrated Moving Average (ARIMA)
  - Dynamic Optimized Theta Model (DOTM)

Model SEM architecture

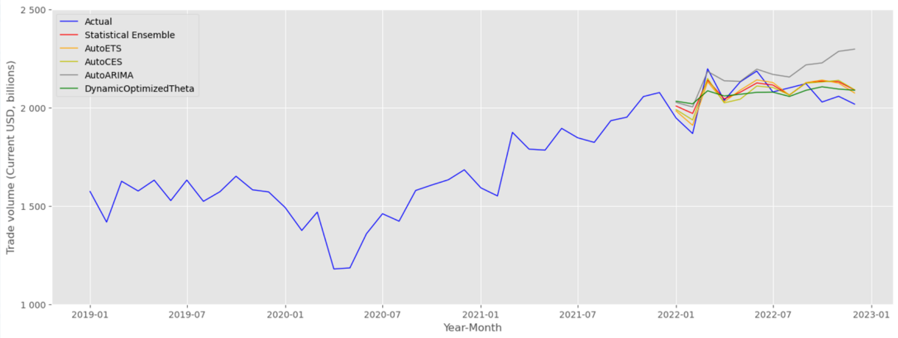


Source: Petropoulos and Svetunkov, 2020

Models

Model	Importance to our use case
Exponential Triple Smoothing	The ETS model accommodates both additive and multiplicative effects of the error, trend, and seasonality components. Enables the model to effectively address the seasonal patterns identified during our exploratory analysis.
Complex Exponential Smoothing	The CES model excels in capturing the complexities of long-term dependencies and non-linear tendencies inherent in time series data. This proficiency is especially relevant when forecasting the global goods trade volume, given that while a linear up ward trend is prevalent, discernible nonlinear trends emerge, particularly evident after the end of the 2000 years.
AutoRegressive Integrated Moving Average	The utilization of the ARIMA model is justified by its efficacy in capturing the linear dynamics and trends inherent in the global goods trade volume data. This model facilitates the representation of temporal correlation and non-stationarity, discerned during exploratory data analysis, owing to its capacity to encapsulate interdependencies within time series data.
Dynamic Optimized Theta Model	Particularly adept at capturing both short-term curvatures and long-term trends observed in the global goods trade volume dataset.

Results and discussion



Proximity of the output from the Statistical Ensemble Model to the actual data, particularly evident during the first months of the forecasting period, a characteristic also observed in the other models.

Results and discussion

Table 1: Performance metrics comparison

	ETS	CES	ARIMA	DOTM	SEM
MAPE	2.28	2.66	5.30	3.28	2.74
MASE	1.74	2.02	3.97	2.48	2.07
RMSE	54 096.38	62 460.44	136 916.02	78 605.72	63 651.85
U_1	0.01	0.02	0.03	0.02	0.02
U_2	0.42	0.48	1.06	0.60	0.49

Inferior performance in predicting the global goods trade volume when contrasted with both the CES and ETS models (but better than DOTM & ARIMA)

Results and discussion

- While the MASE value of 2.07 is considered acceptable, it signifies that the forecasting error of the Statistical Ensemble Model is approximately twice as large as that of the naive reference model.
- An RMSE of 63.65 billions USD indicates that, on average, the model's predictions deviate from the actual value by 63.65 billion USD. Moreover, given an average global goods trade volume of 556.21 billion USD, the relatively high RMSE implies that the average error accounts for approximately 11.44% of the mean value.

Results and discussion

- MAPE value of 2.74, suggesting that, on average, the predictions generated by the SEM deviate by 2.74% from the actual values. This indicates a relatively low level of error and suggests the model's precision.
- Additionally, the U1 coefficient value, which approaches 0, implies a higher level of accuracy for the Statistical model. Moreover, the U2 coefficient value, which is less than 1, contradicts the assertion made by MASE by affirming that the SEM achieves a higher degree of accuracy compared to the naive model.

Results and discussion

Table 2: Diebold-Mariano test results based on Statistical Ensemble Model

	DM Statistic	DM p-value
ETS	1.21	0.26
CES	-0.01	0.99
ARIMA	-2.16	0.06
DOTM	-1.19	0.26

- Only the ETS model outperforms the SEM.
- Notably, ARIMA performs slightly worse than SEM, while CES shows similar performance.
- **No statistically significant difference between the performance of the SEM and the other four models**

Conclusion

- The median effect within the Statistical Ensemble Model appears to mitigate extreme values in our case and maintains close alignment with the best-performing models, enabling it to visually and generally track the curves of real data, unlike the ARIMA model, for example.
- This ability to track curves is critical for economic forecasting, particularly in predicting global goods trade volume to anticipate international economic trends.
- SEM is designed to capture unique information that the others cannot, in theory SEM should be better (not verified for our use case).
- Future research: AI-based models

Thank you for your attention!