CLS
trusted market solutions

The Value of trading volume in FX: a trader's advantage?

Not long ago, decentralization of the FX market hindered real time access to global trade volumes. Things have changed...

Trading volume, the number of units or contracts traded in a financial instrument during a given period of time, is the second most important indicator of market activity after price. Measured on this basis, the FX market is by far the largest market in the world.

The decentralized nature of the FX market had hindered real-time access to globally traded volumes. Advancements and innovations in technology, however, have led to the accessibility and availability of trade data sources in the market.

Capitalizing on this trend, CLS recently started to publish real-time, anonymized, and aggregated FX trade and volume data for a large number of currency pairs, aimed at supporting market participants in developing innovative trading and risk management solutions now on the rise due to regulation.

Technical analysts and traders, followers of the Dow theory, have long been using historical volumes of stocks as a way to confirm a price movement and identify the best entry and exit points for their trading strategies.

Positioning, liquidity, and speculation

Predictions of traded volumes also have a significant economic value. The impact of bank balance sheet pressures, the fragmentation of liquidity, and clients' demand for best execution – supported by the Markets in Financial Instruments Directive II (MiFID II) – has meant a greater focus on tools such as transaction cost analysis (TCA). Market impact, the effect on price due to an investor's trading activity, is dependent on the liquidity demands of the investor and the information content of its trades. It is one of the more costly components to trade execution, always resulting in a drag on performance, but can be mitigated with knowledge of future market liquidity. In pre-trade TCA, volume predictions can be used for estimating the movement in price caused by a particular trade or order. Traditional market impact models rely on historical averages of volumes; improving upon this benchmark with forecasted volumes will help deliver best execution.

Another aspect of pre-trade analysis is forecasting future price volatility. Volatility serves as a measure of uncertainty, which has implications for investment decisions and risk management. It is regarded as the most important variable in the pricing of derivative securities, and traders need to know the volatility of the underlying asset from now until the option expires. Existing models of volatility generally use either historical returns on assets or option-implied volatility to generate forecasts. As academic research suggests, the incorporation of option-implied volatility and trading volume into sophisticated econometrics models leads to outperformance over alternative forecast approaches.[1]

While portfolio trade scheduling tools and market participation models require upcoming raw volumes, other algorithms make use of predicted volume percentages. The volume-weighted average price (VWAP) algorithms send smaller pieces of the large order to the market in proportion with the intraday volume curve. By design, VWAP is a static strategy and will remain constant throughout the day.

However, it can be made dynamic and benefit from volume predictions in a given minute or hour as a percentage of the full day's trading volume. Clearly, volume percentage forecasts are regarded as a different type of prediction problem than raw volume forecasts.

An ensemble of predictive algorithms

What makes prediction of FX volumes so challenging? The difficulty of FX volume forecasting is explained by the unique characteristics of the market, which are over-the-counter, geographically dispersed, continuously operated 24 hours a day (except weekends), and by the variety of factors that affect exchange rates. All this creates distinct and currency-specific volume profiles which are frequently distorted by unexpected events or announcements. As a result, volume predictions need to be not only accurate, but also robust. This can be achieved through various modeling methods.

Deploying ensemble methods of forecasting uses multiple algorithms to achieve better predictive performance than could be obtained from any of the constituent forecasting algorithms alone (see Figure 1).

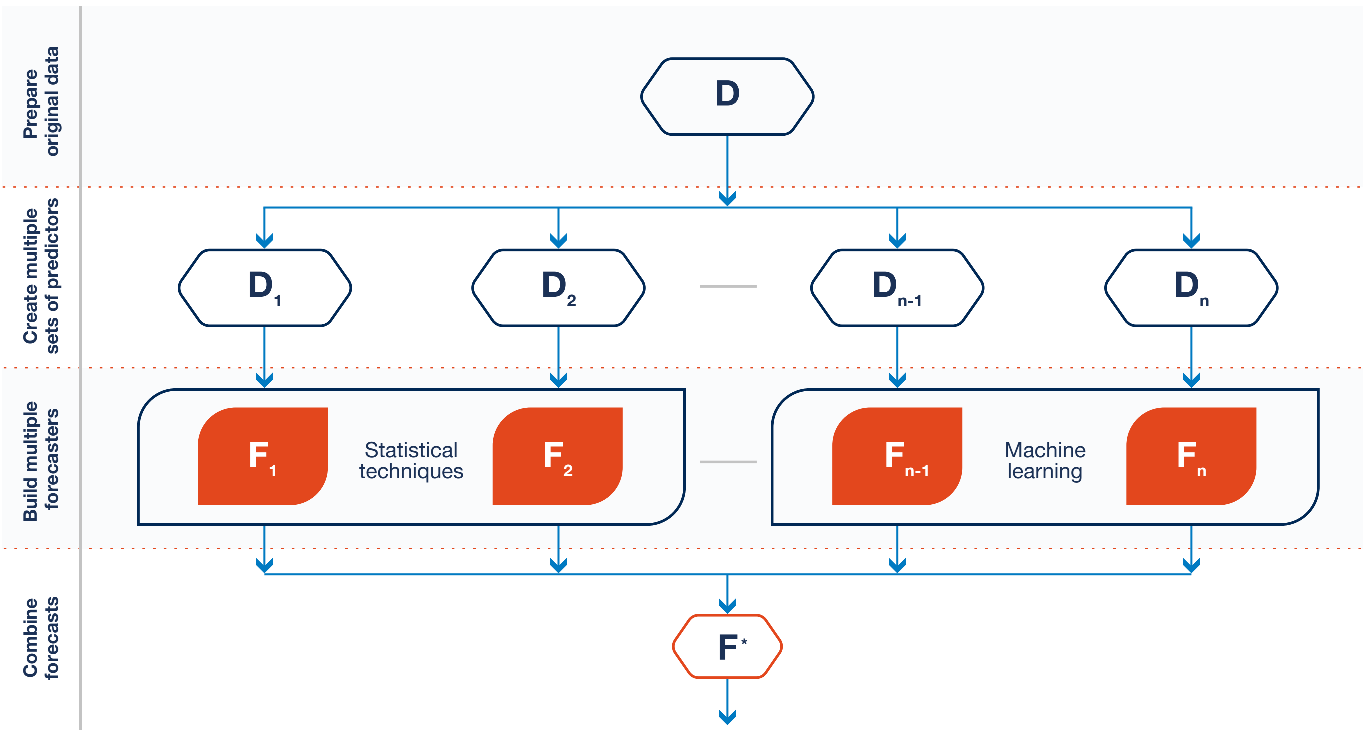
Auto-regressive models, a type of least-squares regressions, specify that the output variable depends linearly on its own previous values, some exogenous explanatory variables, and on a stochastic term. Least-squares regression fits a model by minimizing the sum of the squared differences between past actual values of the output variable and the model's own predictions of those values given the input variables at the time. Regularized least-squares auto-regressive models aim to overcome a well-known generalization problem that past actions may not reflect future behaviour, and penalize for overfitting to historical data.

Standard and regularized linear regressions are global models, where there is a single predictive formula holding over the entire data space. Regression model trees are an alternative modeling approach in which the entire data space is subdivided (partitioned) into smaller regions using a decision-tree induction algorithm. In the second stage, a simple regression model is then built for each cell of the partition. A combination of the predictions from different models along the same path through the tree is then used in prediction.

Support vector machines (SVMs) are a machine-learning technique grounded in statistical learning theory, which characterizes the properties of learning machines that enable them to generalize well to unseen data. They were first developed for classification problems but later generalized to regression-type problems. The fundamental model underlying an SVM is a separating hyperplane in a high or infinited-dimensional feature space mapped from the input data space by a nonlinear kernel function. While SVMs can be complex black-box functions, similar to neural networks, they have the ability to model highly nonlinear processes.

All four forecasts use the same historical and future data, but they are built on different methodologies. The final forecast is taken to be the average of the four model predictions, which is aimed at improving accuracy over the individual forecasts.[2]

Figure 1: Ensemble-based forecasting



[1] Bates, J. M. and Granger, C. W. J. 1969. The combination of forecasts. OR 20(4), 451–468.
[2] Le, V. and Zurbruegg, R. 2010. The role of trading volume in volatility forecasting. Journal of International Financial Markets, Institutions and Money 20(5), 533–555.

CLS and CLSMarketData

We are the world's leading provider of FX settlement services. Launching in 2002, it transformed FX with its innovative approach to multilateral netting and settlement.

Our specialists have worked to reduce systemic risk ever since, while creating operational efficiencies and significant cost savings for its clients. Today, we settle an average USD5.5 trillion of payment instructions every day on behalf of over 70 of the world's most important financial institutions, and more than 25,000 of their clients.

Today, we're leveraging our experience and substantial market intelligence to address the market's wider settlement, processing, and data needs. Our products make the trading process faster, easier, safer, and more cost effective – empowering our clients' success.

The ability to recognize and react to market changes is critical to FX market participants success. They rely on robust transaction data to gain insight into market trends. CLS processes an average of 500,000 FX trades per day, worth over USD1.55 trillion. Since its launch we have warehoused the details of every trade, to create the largest single source of FX executed trade data available to the market.

CLS data solutions support reporting requirements and empower growth – providing the insights its clients need to improve their trading strategies, make smarter business decisions, and enhance FX risk management.

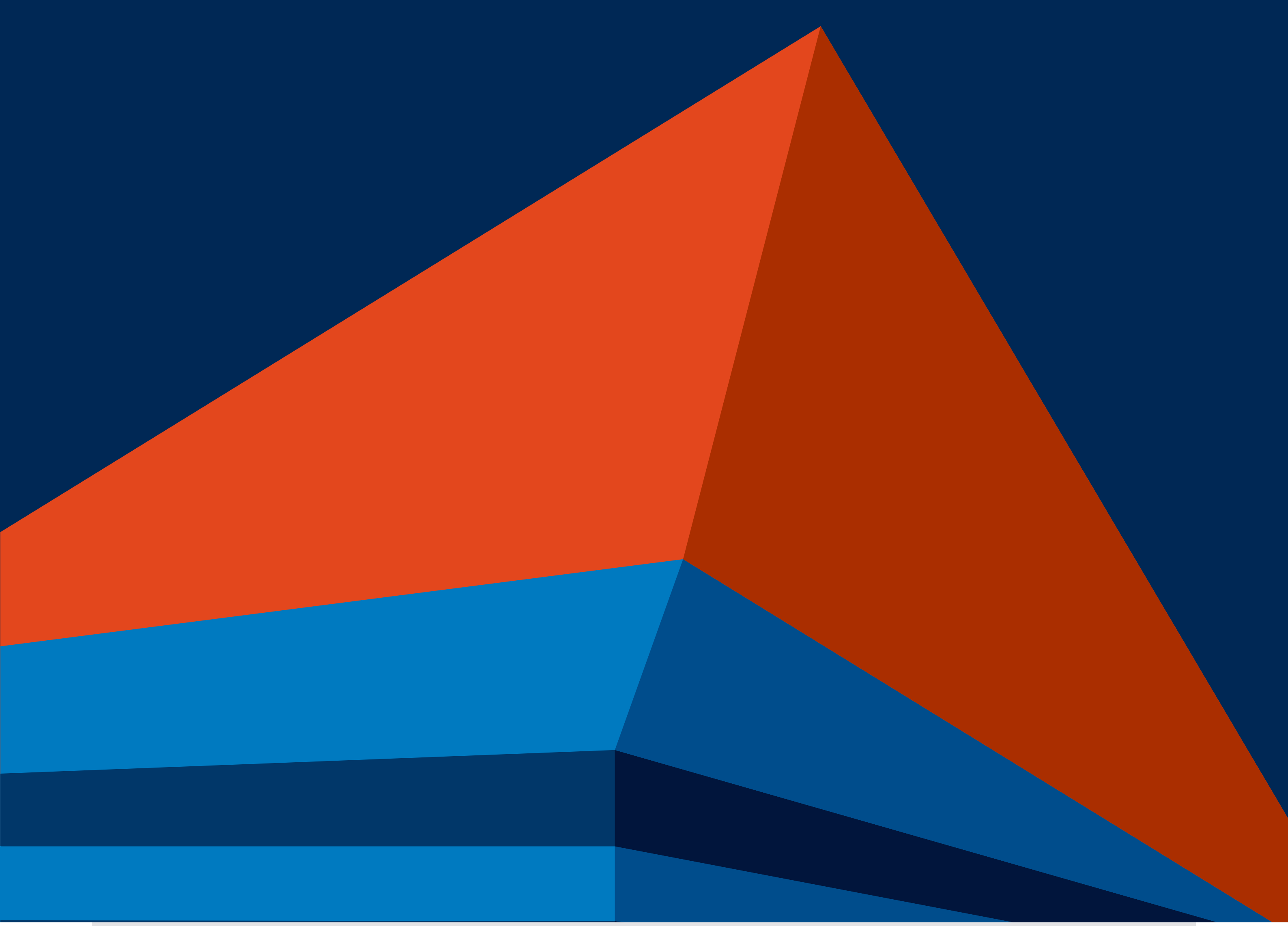
CLSMarketData is an award-winning, comprehensive suite of data products designed to provide quality insight and analytics, both in historical and intraday formats, for financial efficiency, visibility, and control. All CLS data is transaction based, platform agnostic, and dynamic.[3]

Since 2016, we have made our data available to subscribers in standardized and easy-to-use datasets directly from CLS or via an open data delivery platform distributor:

- Data is available for subscription in the form of separate reports, delivered on an intraday, end-of-day or daily bases. The data reports contain executed trade volume, in terms of both the number of trades and the total value in USD. The data is aggregated by trade instrument (spot, swap, and outright forward) and currency pair
- Data is adjusted to follow the reporting convention used by the Bank for International Settlements (BIS) and the semi-annual foreign exchange committee market reports.

Find out more at cls-group.com

[3] FX Week e-FX 2019 Awards, July 2019 "Best FX market data provider."



Spot volume forecast

Suppose we would like to forecast the volume of spot trades (in US dollars) over a 120-hour period, from 22:00 GMT on a Sunday to 21:00 GMT on the following Friday. This forecast may be useful for market risk managers and liquidity managers as it will identify upcoming periods of possible price volatility and hours of market illiquidity. In addition, the forecast may help central bankers responsible for monetary policies to detect abnormal market conditions historically.

We consider eight major currency pairs (AUD/USD, EUR/GBP, EUR/JPY, EUR/USD, GBP/USD, USD/CAD, USD/CHF, USD/JPY). For each currency pair, we measure the total value of spot trades that arrives at CLS within consecutive time intervals (e.g., one-hour intervals).

It is assumed that the measured series of traded spot volumes realizes some unknown discrete-time stochastic process. The accuracy of predictions depends on the choice of explanatory variables. For spot volume forecasting, the input variables include auto-regressive terms, seasonality factors (hourly, weekly and monthly), scheduled economic events (e.g., GDP and CPI announcements) and regional holidays.

"The accuracy of predictions depends on the choice of explanatory variables."

Figure 2 shows five predictions of GBP/USD spot volumes for the week beginning 11 June, 2018. Four of the volume profiles have been produced by alternative forecasting algorithms, and the fifth one is the average of the four.

The comparison of predicted volumes with the actual ones is shown in Figure 3. Clearly, the forecast is capable of capturing a daily seasonality of traded volumes and spikes at the announcement time of major economic indicators, such as UK average earnings and US retail sales.

However, there are also periods of increased and subdued volatility, such as was observed in the hours after the announcement of UK CPI. The algorithms were unable to capture volumes in these periods because they resulted from unexpected news or represented slower than usual recoveries from previous overreactions.

The performance of each model over a 26-week period (1 January to 30 June 2018) is given in Tables 1 and 2 for two common metrics: mean absolute percentage error (MAPE) and root mean square error (RMSE). Depending on the performance metric and the currency pair, some models are seen to outperform others. However, the average model consistently meets or outperforms the best model across both performance metrics and all currency pairs.

Figure 2: Alternative predictions of GBP/USD spot volumes

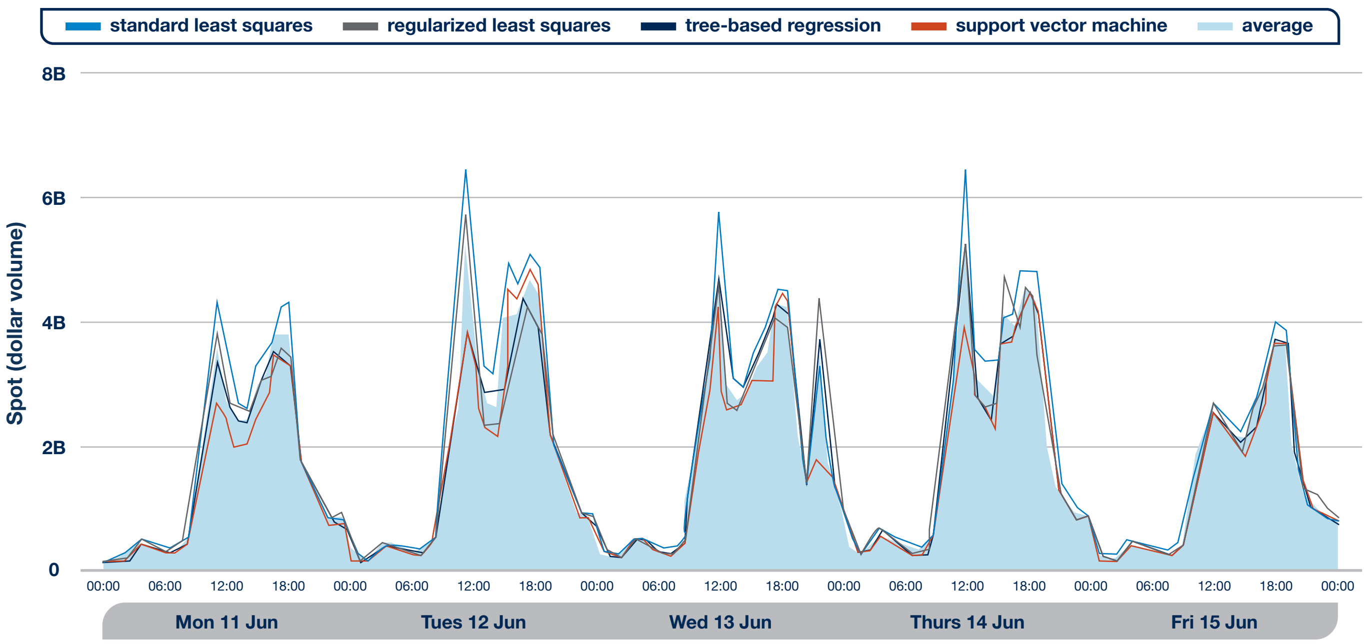


Figure 3: Predicted and actual GBP/USD spot volumes

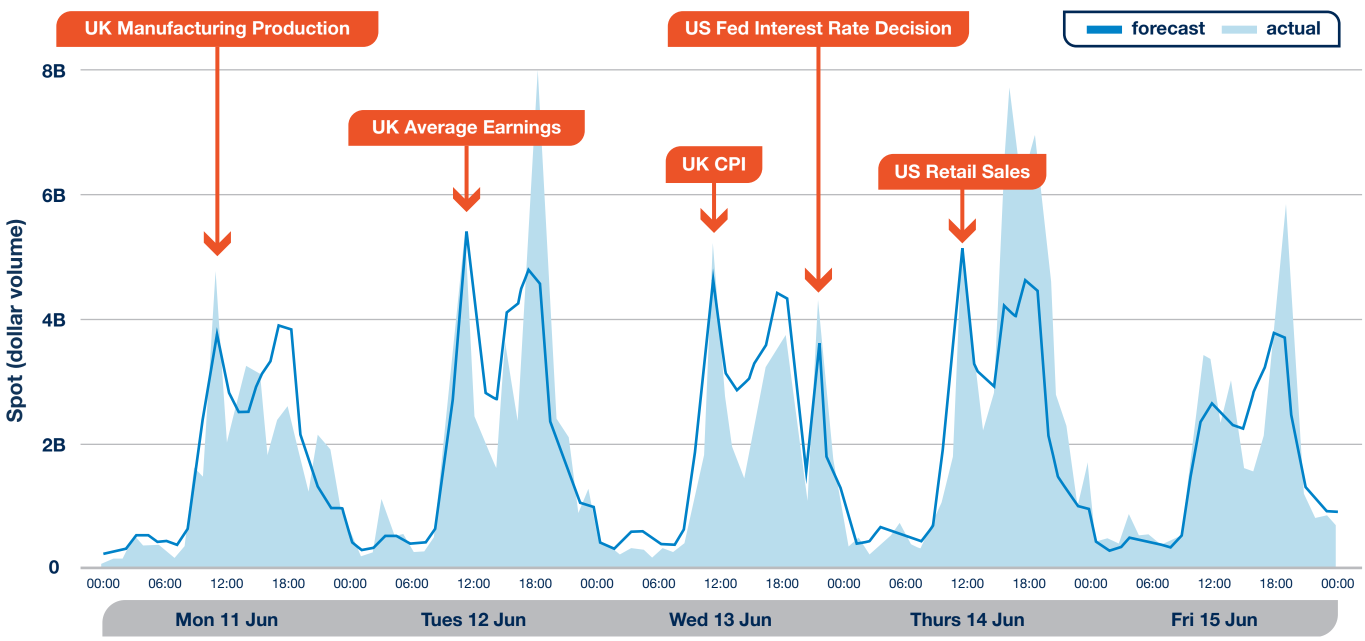


Table 1: Mean absolute percentage error (MAPE) of each model over the 26-week period from 1 January to 30 June 2018. MAPE = $E[|f_t - a_t| / a_t]$, where f_t and a_t denote predicted and actual volumes, respectively. The bold numbers highlight the best individual forecast (with lowest error) for each currency pair.

Forecast	AUD/ USD	EUR/ GBP	EUR/ JPY	EUR/ USD	GBP/ USD	USD/ CAD	USD/ CHF	USD/ JPY
Standard least squares	35.38	**61.27**	49.53	34.12	40.70	53.98	47.13	46.88
Regularized least squares	35.95	85.26	56.42	33.82	41.48	**42.37**	51.54	37.80
Tree-based regression	**34.00**	65.85	**46.99**	**31.36**	**35.53**	52.43	**43.24**	37.86
Support vector machine	34.58	91.19	58.71	32.40	38.82	45.49	49.99	**35.66**
Average	32.96	74.20	50.54	30.80	37.36	46.77	46.04	37.22

Table 2: Root mean square error (RMSE), as a percentage of actual volumes, of each model over the 26-week period from 1 January to 30 June 2018. RMSE (%) = $E_t[(f_t - a_t)2] / Et[at]$
The bold numbers highlight the best individual forecast (with lowest error) for each currency pair.

Forecast	AUD/ USD	EUR/ GBP	EUR/ JPY	EUR/ USD	GBP/ USD	USD/ CAD	USD/ CHF	USD/ JPY
Standard least squares	50.27	**72.11**	79.97	57.04	59.55	52.17	76.83	56.18
Regularized least squares	**47.71**	76.23	**76.55**	**52.28**	**59.24**	51.66	**72.91**	**51.38**
Tree-based regression	53.90	75.71	85.43	61.13	62.71	**51.52**	77.56	54.02
Support vector machine	49.99	79.21	77.85	54.61	61.82	52.35	74.98	53.34
Average	48.27	73.46	78.25	53.53	58.28	49.41	73.80	50.80

Deviations between predicted volumes and actual volumes should be reduced by using volume data at higher frequencies or by re-running the ensemble of forecasting algorithms intradaily. Alternatively, there is potential to improve performance by growing the ensemble to include more diverse algorithms or by simply combining the current forecasts in more sophisticated ways, beyond the simple average.

[1] Bates, J. M. and Granger, C. W. J. 1969. The combination of forecasts. OR 20(4), 451–468.
[2] Le, V. and Zurbruegg, R. 2010. The role of trading volume in volatility forecasting. Journal of International Financial Markets, Institutions and Money 20(5), 533–555.
[3] FX Week e-FX 2019 Awards, July 2019 "Best FX market data provider."

Conclusion

The need for data, particularly in the FX market, is greater than ever due to ongoing market structure and regulatory change, which have created pressure on bank balance sheets, liquidity fragmentation, and the need for best execution, necessitated by MiFID II. Datasets such as CLS's volume and forecast data, combined with a choice of volume-forecasting models should help market participants navigate this changing market environment.

Settlement Processing **Data**

For more information please email sales@cls-group.com

London
Hong Kong
Tokyo

New York
New Jersey

info@cls-group.com
cls-group.com