



COMILLAS
UNIVERSIDAD PONTIFICIA

ICAI

GRADO EN INGENIERÍA EN TECNOLOGÍAS INDUSTRIALES

TRABAJO FIN DE GRADO
REGRESSION MODELS AVERAGING UNDER
INTERVAL-VALUED DATABASES.
APPLICATIONS IN FINANCIAL MARKETS

Autor: Helena Isabel Álvarez García

Director: Carlos Maté Jiménez

Madrid

Julio de 2019

AUTHORIZATION FOR DIGITALIZATION, STORAGE AND DISSEMINATION IN THE NETWORK OF END-OF-DEGREE PROJECTS, MASTER PROJECTS, DISSERTATIONS OR BACHILLERATO REPORTS

1. Declaration of authorship and accreditation thereof.

The author Mr./Ms. Helena Isabel Álvarez García

HEREBY DECLARES that he/she owns the intellectual property rights regarding the piece of work: Regression models averaging under interval-valued databases. Applications in financial markets that this is an original piece of work, and that he/she holds the status of author, in the sense granted by the Intellectual Property Law.

2. Subject matter and purpose of this assignment.

With the aim of disseminating the aforementioned piece of work as widely as possible using the University's Institutional Repository the author hereby **GRANTS** Comillas Pontifical University, on a royalty-free and non-exclusive basis, for the maximum legal term and with universal scope, the digitization, archiving, reproduction, distribution and public communication rights, including the right to make it electronically available, as described in the Intellectual Property Law. Transformation rights are assigned solely for the purposes described in a) of the following section.

3. Transfer and access terms

Without prejudice to the ownership of the work, which remains with its author, the transfer of rights covered by this license enables:

- a) Transform it in order to adapt it to any technology suitable for sharing it online, as well as including metadata to register the piece of work and include "watermarks" or any other security or protection system.
- b) Reproduce it in any digital medium in order to be included on an electronic database, including the right to reproduce and store the work on servers for the purposes of guaranteeing its security, maintaining it and preserving its format.
- c) Communicate it, by default, by means of an institutional open archive, which has open and cost-free online access.
- d) Any other way of access (restricted, embargoed, closed) shall be explicitly requested and requires that good cause be demonstrated.
- e) Assign these pieces of work a Creative Commons license by default.
- f) Assign these pieces of work a HANDLE (*persistent* URL). by default.

4. Copyright.

The author, as the owner of a piece of work, has the right to:

- a) Have his/her name clearly identified by the University as the author
- b) Communicate and publish the work in the version assigned and in other subsequent versions using any medium.
- c) Request that the work be withdrawn from the repository for just cause.
- d) Receive reliable communication of any claims third parties may make in relation to the work and, in particular, any claims relating to its intellectual property rights.

5. Duties of the author.

The author agrees to:

- a) Guarantee that the commitment undertaken by means of this official document does not infringe any third party rights, regardless of whether they relate to industrial or intellectual property or any other type.

- b) Guarantee that the content of the work does not infringe any third party honor, privacy or image rights.
- c) Take responsibility for all claims and liability, including compensation for any damages, which may be brought against the University by third parties who believe that their rights and interests have been infringed by the assignment.
- d) Take responsibility in the event that the institutions are found guilty of a rights infringement regarding the work subject to assignment.

6. Institutional Repository purposes and functioning.

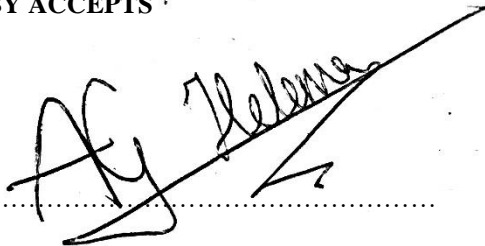
The work shall be made available to the users so that they may use it in a fair and respectful way with regards to the copyright, according to the allowances given in the relevant legislation, and for study or research purposes, or any other legal use. With this aim in mind, the University undertakes the following duties and reserves the following powers:

- a) The University shall inform the archive users of the permitted uses; however, it shall not guarantee or take any responsibility for any other subsequent ways the work may be used by users, which are non-compliant with the legislation in force. Any subsequent use, beyond private copying, shall require the source to be cited and authorship to be recognized, as well as the guarantee not to use it to gain commercial profit or carry out any derivative works.
- b) The University shall not review the content of the works, which shall at all times fall under the exclusive responsibility of the author and it shall not be obligated to take part in lawsuits on behalf of the author in the event of any infringement of intellectual property rights deriving from storing and archiving the works. The author hereby waives any claim against the University due to any way the users may use the works that is not in keeping with the legislation in force.
- c) The University shall adopt the necessary measures to safeguard the work in the future.
- d) The University reserves the right to withdraw the work, after notifying the author, in sufficiently justified cases, or in the event of third party claims.

Madrid, on 11 of July, 2019

HEREBY ACCEPTS

Signed.....



Reasons for requesting the restricted, closed or embargoed access to the work in the Institution's Repository

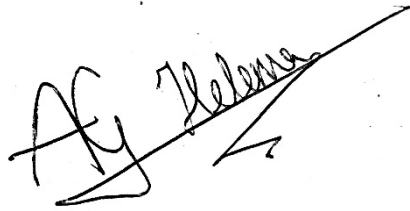
Declaro, bajo mi responsabilidad, que el Proyecto presentado con el título
**REGRESSION MODELS AVERAGING UNDER INTERVAL-VALUED
DATABASES. APPLICATIONS IN FINANCIAL MARKETS**

en la ETS de Ingeniería - ICAI de la Universidad Pontificia Comillas en el
curso académico 2018-19 es de mi autoría, original e inédito y
no ha sido presentado con anterioridad a otros efectos. El Proyecto no es
plagio de otro, ni total ni parcialmente y la información que ha sido
tomada

de otros documentos está debidamente referenciada.

Fdo.: Helena Isabel Álvarez García

Fecha: 03/07/2019



Autorizada la entrega del
proyecto

EL DIRECTOR DEL PROYECTO

Fdo.: CARLOS MATE JIMENEZ

Fecha: 10/07/2019



Firma: Carlos Mate



COMILLAS
UNIVERSIDAD PONTIFICIA

ICAI

GRADO EN INGENIERÍA EN TECNOLOGÍAS INDUSTRIALES

TRABAJO FIN DE GRADO
REGRESSION MODELS AVERAGING UNDER
INTERVAL-VALUED DATABASES.
APPLICATIONS IN FINANCIAL MARKETS

Autor: Helena Isabel Álvarez García

Director: Carlos Maté Jiménez

Madrid

Julio de 2019

COMBINACIÓN DE MODELOS DE REGRESIÓN PARA DATOS DE INTERVALO. APLICACIONES EN MERCADOS FINANCIEROS

Autor: Álvarez García, Helena Isabel.

Director: Maté Jiménez, Carlos.

Entidad Colaboradora: ICAI – Universidad Pontificia Comillas.

RESUMEN DEL PROYECTO

Introducción

La predicción de valores en mercados financieros es una actividad de notable interés hoy en día. La necesidad de anticipar los valores de las acciones, junto con la dependencia que pequeñas y grandes empresas tienen en la precisión de estas estimaciones, convierten los modelos de predicción en un área de suma importancia para el funcionamiento de los mercados.

El incremento de las necesidades de predicción en el área financiera requiere del desarrollo de métodos de alto rendimiento, sobre todo para aquellos datos que proporcionen la información necesaria para describir las tendencias económicas actuales. Los datos de intervalo están basados en intervalos con límite superior en inferior para cada medida. La información proporcionada por este tipo de datos es normalmente más precisa y completa que la que se puede extraer de los datos sencillos, lo que convierte a los datos de intervalo en una herramienta muy valiosa cuando la precisión es una prioridad en el análisis de datos (Maté, 2012).

La combinación de modelos también desempeña un rol clave en la resolución de problemas que requieren métodos de regresión. Dicha combinación reduce el riesgo de pérdida de información al rechazar ciertos modelos por otros, por lo que puede ser una manera de mejorar la precisión de las predicciones en el campo de los mercados financieros (Moral Benito, 2015). La aplicación en estos mercados consiste en la combinación de modelos que sean capaces de predecir series temporales de intervalo (ITS), que son datos de intervalo registrados como una secuencia ordenada a lo largo del tiempo.

Estado de la Técnica

Desde que el análisis de intervalos (AI) fue usado por primera vez en 1959, varios estudios y aplicaciones han sido desarrollados para este tipo de datos. En cuanto a la predicción de datos, hay una gran gama de modelos lineales adaptados a los datos de

intervalo, tales como la técnica Lasso, el método minMax o la aproximación parametrizada. El estudio de Maceda (2018) incluye y describe varios de estos métodos. Distintos modelos de regresión no lineales han sido estudiados por Lima Neto y De Carvalho (2017) a través de los experimentos de Monte Carlo, demostrando mayor precisión y exactitud que los métodos lineales en ciertas bases de datos.

Las librerías de funciones de MATLAB suponen el mayor reto de este proyecto, ya que, actualmente, no es posible encontrar una herramienta que represente, analice y compare métodos de regresión, lineales y no lineales, y que permita diferentes combinaciones entre ellos para estudiar la combinación de modelos.

Objetivo del proyecto

Este proyecto busca crear una interfaz en MATLAB que funcione como herramienta de análisis, evaluación y comparación de modelos de regresión. Esta interfaz está formada por un conjunto de funciones que calculan los parámetros necesarios para la valoración, presentan los gráficos adecuados que representan y comparan modelos, y que desarrollan los modelos de óptimo ajuste. Los resultados obtenidos permiten extraer conclusiones sobre la promediación de modelos y profundizar en trabajos previos sobre el análisis de datos de intervalo.

La exactitud de las predicciones financieras tiene un papel crítico en el comportamiento de los mercados de la economía global de hoy en día. La capacidad de predecir una caída en los valores de una acción determinada puede suponer la diferencia entre evitar o no una crisis económica. Por esta razón, la finalidad del proyecto incluye la búsqueda de las técnicas de mayor rendimiento en la combinación de modelos, con el objetivo de aplicarlas a los mercados financieros y mejorar el desempeño de las predicciones tanto como sea posible.

Metodología

La primera sección de esta investigación comprende un resumen de los actuales métodos de regresión para datos de intervalo. Incluye fuentes sobre modelos lineales y no lineales, su definición, características y su rendimiento en relación con el resto.

A continuación, se presenta un marco teórico sobre la combinación de modelos, compuesto por el análisis y descripción de distintas técnicas de combinación y su aplicación a los datos de intervalo. La introducción del Puzle de Combinación de la Predicción da lugar a la comparación entre la combinación por media simple y combinaciones más complejas que difieren en los pesos asociados a cada modelo. Este dilema existe desde la primera introducción de la combinación de modelos por Bates y Granger (1969).

Establecer los parámetros evaluadores de cada modelo es un paso clave en este análisis. Estos estadísticos son muy valiosos en el contraste de modelos y su rendimiento. Los

parámetros seleccionados evalúan un rango amplio de propiedades de cada modelo, como la distancia entre la predicción y el valor real del intervalo, o el porcentaje de los datos que están cubiertos por la aproximación. Los criterios de evaluación seleccionados son los siguientes:

1. U de Theil (IU).
2. Tasa de Cubrimiento (CR).
3. Tasa de Eficiencia (ER).
4. Distancia Media de Error (MDE).

Una vez el problema está definido y se selecciona la base de datos, la metodología se aplica de acuerdo con los siguientes pasos:

1. Análisis del desempeño de cada modelo mediante los cuatro evaluadores seleccionados. Los valores obtenidos permiten la clasificación de los distintos modelos en cuanto a su precisión y rendimiento, proporcionando información útil sobre las estimaciones realizadas por ese modelo. Esto se lleva a cabo mediante varias funciones de MATLAB, *UTheil_k*, *CER_k* y *MDE_k*, que calculan la U de Theil, los radios de cubrimiento y eficiencia, y la distancia media de error, respectivamente.
2. Elección de los modelos que van a ser combinados. La combinación de modelos busca potenciar las propiedades de aquellos modelos que presenten mejores predicciones para obtener un nuevo modelo con características mejoradas. *DrawmultipleITSvdef* permite la representación de las series temporales de intervalo de los modelos seleccionados y su comparativa con la serie temporal real, arrojando luz sobre los posibles aspectos de mejora.
3. Implementación de la combinación de modelos. La función *Combination_k* devuelve la predicción combinada a partir de dos modelos predictores, mediante el método de pesos iguales y el de pesos óptimos. La función utilizada para calcular los pesos óptimos en una combinación de dos modelos es *OptimalW*.
4. Evaluación de las mejoras obtenidas mediante la combinación. Los estadísticos evaluadores se calculan para el nuevo modelo, consiguiendo nuevos valores que son evaluados y dan lugar a las conclusiones de este proyecto.

Resultados

Cuatro estudios de caso se analizaron siguiendo la metodología expuesta con anterioridad. Las cuatro bases de datos evaluadas incluyen los valores del índice S&P500 durante 2017 (uno de los índices de mercado más importantes a nivel global), los valores bursátiles de General Electric durante 2017 (multinacional con sectores energético, financiero, tecnológico e industrial), los valores de stock de Zardoya desde el inicio de 2018 al primer trimestre de 2019 (la principal productora de ascensores y escaleras automáticas en el mundo) y los valores de Visa durante 2017 (marca de pago líder).

A continuación, se presentan los resultados obtenidos para cada caso, incluyendo tablas de valores y gráficos representativos.

Índice S&P500

Las predicciones hechas con seis modelos diferentes de regresión lineal se compararon con el método del camino aleatorio (RW). Este método realiza la predicción exclusivamente con el valor del último periodo, y sirve de referencia para cotejar el resto de métodos.

Los resultados prueban que los seis métodos elegidos (minMax, CM, CRM, BSRM, CIRM y M Method) son mejores predictores que el camino aleatorio. La Tabla 1 muestra los valores de la U de Theil para cada método. El camino aleatorio siempre tiene de valor la unidad, mientras que valores inferiores a 1 significan un mejor rendimiento que este método de referencia.

Tabla 1: Caso 1. U de Theil.

M1	M2	M3	M4	M5	RW	M6
0.7715	0.7655	0.7484	0.7639	0.7492	1.0000	0.6194

En el cálculo de los pesos óptimos para combinar modelos, llama la atención la similitud con la combinación por pesos iguales. Los pesos son prácticamente iguales en cada combinación, con valor aproximadamente igual a 0.5. Esto significa que la combinación óptima implica asignar la misma importancia a todos los modelos.

Los métodos 1 (minMax) y 6 (método M) se combinaron para crear un nuevo modelo. Esta combinación resulta en un aumento global del rendimiento de los modelos. El método M aporta mejoras en los valores de la U de Theil, la tasa de eficiencia y la distancia media al error, mientras que el método minMax mejora la tasa de cubrimiento conjunta. Los resultados se muestran en la Tabla 2.

Tabla 2: Caso 1. Combinación de modelos. Método 1 vs Método 6.

MÉTODO	IU	CR	ER	MDE
M1	0.7715	0.6693	0.6600	9255.2
M6	0.6194	0.6143	0.8261	5688.7
Pesos Iguales	0.6624	0.6537	0.7402	6471.6
Pesos Óptimos	0.6625	0.6537	0.7402	6471.9

Debido a las similitudes en los pesos iguales y los pesos óptimos, los valores de los estadísticos de contraste presentan diferencias apenas distinguibles.

General Electric

Este caso analiza datos con una tendencia bajista y ligeras fluctuaciones a lo largo del tiempo. El estudio de estos datos demuestra que el método de regresión que mejor se ajusta es una curva polinómica de cuarto grado. La Figura 1 muestra el intervalo medio de los datos de General Electric (valor medio de los valores inferiores y valor medio de los valores superiores) frente al valor medio de la predicción dada por el ajuste polinómico. Se incluyen también las tasas de cubrimiento y eficiencia, ambas superiores al ochenta por ciento.

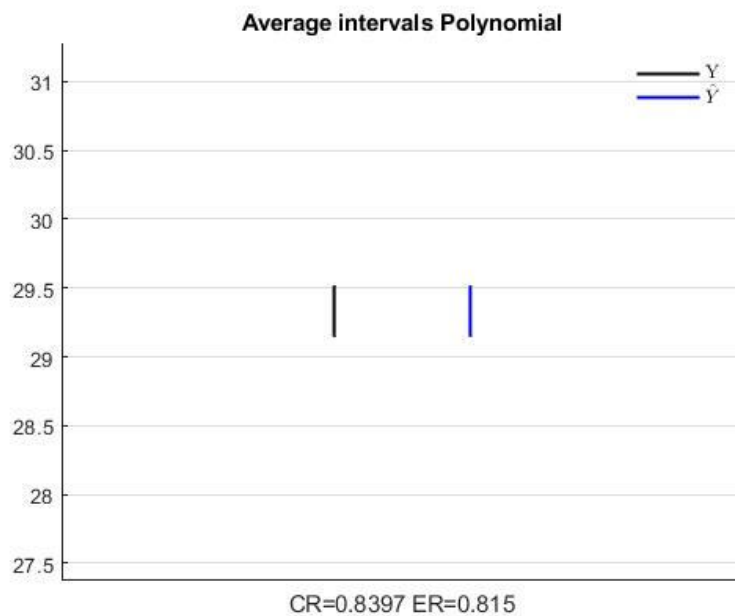


Figura 1: Caso 2. Intervalo medio real (Y) e intervalo medio de la predicción polinómica (\hat{Y}).

Este modelo basado en una aproximación polinómico se combina con el camino aleatorio mediante media simple y mediante pesos óptimos. Los resultados se presentan en la Tabla 3.

Tabla 3: Caso 2. Combinación de modelos: Camino Aleatorio vs Ajuste Polinómico.

METHOD	IU	CR	ER	MDE
RW	1.0000	0.1448	0.0929	4.7741
Polinomio	0.0942	0.8397	0.8150	0.0386
Pesos Iguales	0.5212	0.3029	0.2760	1.3579
Pesos Óptimos	0.0890	0.8507	0.8193	0.0340

Los valores obtenidos demuestran notables mejoras con respecto a las predicciones individuales para la combinación mediante pesos óptimos. Los resultados de la U de Theil y la distancia media de error son muy próximos a cero (indicando una distancia mínima

entre el valor real y la predicción) y los valores de las tasas de eficiencia y cubrimiento son superiores a los del modelo polinómico sin combinar.

Zardoya

Los datos de Zardoya se predicen utilizando métodos de tendencia, como la media, la mediana o los cuartiles. El comportamiento lateral de esta base de datos resulta en una falta de exactitud en las predicciones basadas en tendencia, con valores muy superiores a la unidad para la U de Theil. Los modelos evaluados son peores predictores que el camino aleatorio.

La media de Aumann y la mediana de Hausdorff se evalúan como métodos de predicción, comparándolos y combinándolos con el camino aleatorio. Los resultados demuestran que prueban que estos dos métodos son lo suficientemente similares en rendimiento y características para ser considerados equivalentes como predictores.

Las siguientes figuras ilustran las similitudes en la predicción de series temporales entre la media de Aumann (Figura 2) y la mediana de Hausdorff (Figura 3) para el caso Zardoya. La herramienta de MATLAB permite comparar de manera visual ambos métodos de predicción, sin limitarse a cálculos numéricos de estimadores.

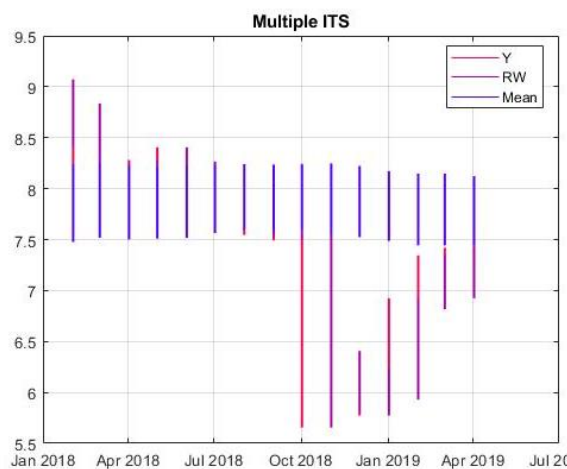


Figura 2: Caso 3. ITS de Zardoya (rojo), ITS del Camino Aleatorio (morado), ITS de la media de Aumann (azul).

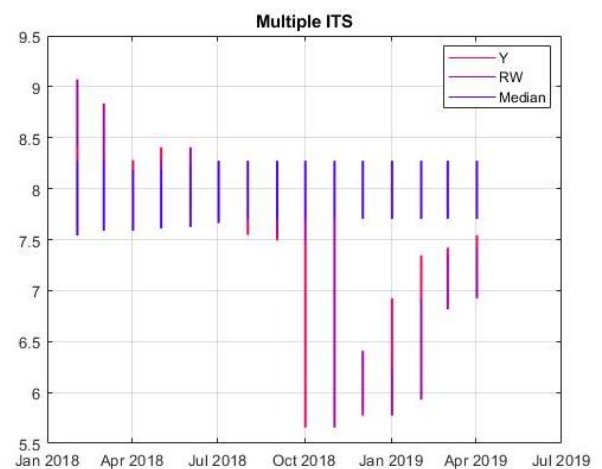


Figura 3: Caso 3. ITS de Zardoya (rojo), ITS del Camino Aleatorio (morado), ITS de la mediana de Hausdorff (azul).

Visa

El último estudio de caso predice datos con una marcada tendencia alcista. En este caso, las predicciones óptimas se obtienen a través de un ajuste exponencial. Este método de regresión no lineal proporciona un modelo con altos valores de radio de cobertura (CR)

y de eficiencia (ER), y valores bajos para la U de Theil y la distancia media de error. Los resultados de su combinación con el camino aleatorio se muestran en la Tabla 4.

Table 1: Case 4. Model Averaging: Random Walk vs Exponential Fit.

MÉTODO	IU	CR	ER	MDE
RW	1.0000	0.4902	0.5315	156.2200
Exponential	0.4016	0.8322	0.8225	25.2780
EW	0.6344	0.7018	0.7164	60.2200
OW	0.6861	0.6739	0.6932	70.4870

Los valores de los estadísticos para este método no mejoran respecto al ajuste polinómico. Sin embargo, este caso representa otra situación en la que la combinación de modelos enriquece las características de un determinado modelo, en este caso, el ajuste exponencial. La combinación se lleva a cabo con pesos iguales y pesos óptimos, dando lugar a una discrepancia respecto a los casos anteriores. En este caso, la combinación pesos óptimos no demuestra una clara superioridad, sino que la media simple proporciona mejores propiedades para todos los parámetros.

Conclusiones

Los resultados obtenidos enfatizan distintos aspectos sobre el rendimiento de los modelos de regresión lineales y no lineales para los datos de intervalo.

En cuanto al Puzle de Combinación de la Predicción, se ha evidenciado que la combinación con pesos iguales es el método más directo, sencillo y fiable mientras que los modelos que se combinan tengan características similares. No obstante, en una situación con un modelo que destaca sobre los demás en cuanto a nivel de rendimiento y precisión (por ejemplo, el ajuste polinómico en el caso de General Electric), la media simple deja de ser el método más fiable en la combinación de modelos. En estos casos, el cálculo de pesos óptimos basado en el desempeño de cada modelo refuerza las propiedades del modelo más preciso.

Los modelos predictivos basados en tendencias, como la media, la mediana o los cuartiles, han demostrado ser menos efectivos y exactos que otros métodos en conjuntos de datos con comportamientos laterales o fuertes tendencias crecientes. La ausencia de tendencia en los datos (laterales) dificulta la capacidad del modelo de predecir con exactitud los movimientos futuros de las acciones a partir de posibles tendencias pasadas. Por otro lado, tendencias que cambien muy rápidamente empeora la habilidad del modelo de adaptarse a los nuevos valores y hacer predicciones con precisión.

Se ha probado que los métodos de regresión no lineal, como las curvas polinómicas o exponenciales, presentan radios más altos de cubrimiento y de eficiencia. Esto se traduce en mayor eficiencia y mejores predicciones que los métodos lineales en datos con fuertes

tendencias alcistas o bajistas. El desarrollo de métodos no lineales podría ser el factor decisivo para obtener las mejores predicciones posibles para datos de intervalo.

Finalmente, en cuanto a las recomendaciones para futuros estudios, sería interesante crear una aplicación mediante el complemento App Designer de MATLAB, que incluyera todas las funciones programadas y desarrolladas en este proyecto. Esto permitiría obtener los resultados a través de una interfaz visual y sencilla, facilitando las comparaciones entre modelos y la combinación de los mismos.

Referencias

- Bates, J. M., & Granger, C. W. (1969). The Combination of Forecasts. *The OR Society*, 20(4), 451-468.
- Lima Neto, E. d., & de Carvalho, F. d. (2017). Nonlinear regression applied to interval-valued data. *Pattern Analysis and Applications*, 20(3), 809-824.
- Maceda Dal-Re, P. (2018). Basic analytic system for interval valued data. Madrid: Universidad Pontificia Comillas.
- Maté, C. (2012). El análisis de intervalos. Aplicaciones en la ingeniería. *Anales de mecánica y electricidad*, 89, 20-27.
- Moral Benito, E. (2015). Model averaging in economics: an overview. *Journal of Economic Surveys*, 29(1), 46-75.

REGRESSION MODELS AVERAGING UNDER INTERVAL-VALUED DATABASES. APPLICATIONS IN FINANCIAL MARKETS

Introduction

Forecasting financial markets is an essential economic interest nowadays. The necessity to predict future stock values and the reliance big and small companies have on the goodness of these predictions turn forecasting into a fundamental research area.

The more demanding needs of financial forecasters require the development of high-performance methods for a kind of data with the information necessary to describe the current economic problems. Interval-valued (IV) data is based on intervals, with lower and upper limits for each measure. The information given by interval data is usually more accurate and complete than the one that can be extracted from single data, so it is really valuable when precision is a priority in data analysis (Maté, 2012).

Model averaging also plays a key role in solving many different problems regarding regression methods, as the combination of predictive models reduces the risks of losing information when disregarding certain models. Its implementation on regression methods could be a way of improving forecasts' accuracy in the desired field (Moral Benito, 2015). The application in financial markets goes through the combination of models that predict Interval Time Series (ITS), which are IV data registered as an ordered sequence throughout time.

State of the Art

Since IV data was used for the first time, wide research has been done on methods and applications regarding this kind of data. Concerning data forecasting, there is an extensive variety of linear regression methods that have been studied for IV data, such as the Lasso technique, the minMax method or the parametrized approach, summarized and described by Maceda (2018). Non-linear regression methods were studied by Lima Neto and De Carvalho (2017) through the Monte Carlo experiments, demonstrating a greater accuracy and precision than linear regression methods in certain interval-valued datasets.

MATLAB's libraries imply the primary challenge, as it is not currently possible to find a tool that represents, analyzes and compares regression methods, linear and non-linear, allowing different combinations among them to study model averaging.

Project purpose

This project aims to build a MATLAB environment that works as a tool to analyze, evaluate and compare several linear and non-linear regression models. This environment comprises a set of functions that calculate the necessary parameters for the evaluation,

displays suitable graphs to compare the models and develop optimal fit models. The obtained results are used to draw conclusions on model averaging and to develop and deepen previous works on interval-valued data analysis.

Accurate financial forecasts have a critical role in the markets' behavior and the globalized economy we live in. This is why the main purpose includes finding the best performance techniques on model averaging in order to apply them to financial markets and improve the performance of the forecast as much as possible.

Methodology

The first section in this project comprises a review of the current regression methods that are suitable for interval-valued data. This bibliographic review includes sources on linear and non-linear models, their characteristics and definition, and their qualitative relative performance.

It is followed by a theoretical framework on model averaging. Different model averaging methods are analyzed and described for the subsequent application to IV data. Challenging the Forecast Combination Puzzle unleashes the comparison between simple averaging and more complex weight combinations, an existing dilemma since model averaging was raised for the first time by Bates and Granger (1969).

Establishing the model evaluators is another essential stage in this analysis. These statistic parameters are very valuable when contrasting models and comparing their performance. The selected evaluation parameters evaluate a range of properties of the model, such as the distance between the forecast and the real interval, or the percentage of interval covered by the prediction. The models are evaluated according to four criteria:

1. Theil's U.
2. Coverage Rate.
3. Efficiency Rate.
4. Mean Distance Error.

Once the problem is defined and the financial data is selected, the methodology follows these main steps:

1. Analyze each model's performance according to the four established evaluators. The obtained values allow the classification of the different models according to their accuracy and performance, providing useful information about the forecasts on the data. This is carried out using different MATLAB functions: *UTheil_k*, *CER_k* and *MDE_k*.
2. Choose the models that are being combined. Model averaging looks forward enhancing the properties of the better predictive models in order to obtain a new model with improved characteristics. *DrawmultipleITSvdef* allows the representation of the selected models' ITS and the comparison with the real ITS, shedding light to possible improvement features.

3. Carry out model averaging. The function *Combination_k* returns the combined forecast averaging two different model, for both equal weights and optimal weights. The function used to calculate the optimal weights in a two-model combination is called *OptimalW*.
4. Evaluate the improvements made by the averaging. The model evaluators are calculated again for the new model and results are evaluated. Conclusions are drawn on the different aspects observed through this evaluation.

Results

There are four different study cases that were analyzed and evaluated following the previously exposed methodology, obtaining diverse results. The four datasets evaluated include the S&P500 index values during 2017 (one of the most important market indexes in the world), stock values from General Electric in 2017 (an energetic, financial, technological and industrial multinational), Zardoya’s stock values during 2018 and the first term of 2019 (the main elevator manufacturer in the world) and stock values from Visa in 2017 (one of the leader payment brands).

Results obtained after evaluating each case follow in sequence.

S&P500 index

The data was evaluated using six different linear regression methods compared to the random walk. Results proved that all the six methods perform better than the random walk when forecasting the data. This is shown in Table 1, which presents the Theil’s U values for each model, all lower than 1.

Table 2: Case 1. Theil’s U.

M1	M2	M3	M4	M5	RW	M6
0.7715	0.7655	0.7484	0.7639	0.7492	1.0000	0.6194

When calculating the optimal weights, it stands out that the weights assigned to each model for an optimal combination are approximately equal in every case (0.5), this meaning that the best performer combination is the one that gives the same importance to every model.

Methods 1 and 6 where combined to create a new model. This resulting model was an increase in overall performance. Method 6 improves the Theil’s U, Efficiency Rate and Mean Distance Error values of the first method, while the first method model contributes increasing the Coverage Rate of the sixth model. These results are illustrated in Table 2.

The combination was carried out using simple averaging and optimal weights combination. Nevertheless, due to the values of the optimal weights when combining this case's models, the differences between both averaging methods are negligible.

Table 2: Case 1. Model averaging: Method 1 vs Method 6.

METHOD	IU	CR	ER	MDE
M1	0.7715	0.6693	0.6600	9255.2
M6	0.6194	0.6143	0.8261	5688.7
EW	0.6624	0.6537	0.7402	6471.6
OW	0.6625	0.6537	0.7402	6471.9

General Electric

This case comprises a downward tendency data with small fluctuations throughout time. It was demonstrated that the best possible forecast was obtained through a polynomial fit (non-linear regression method). Figure 1 depicts the average GE interval and the average interval from the polynomial regression model, including the Coverage Rate and the Efficiency Rate, which are both higher than eighty percent.

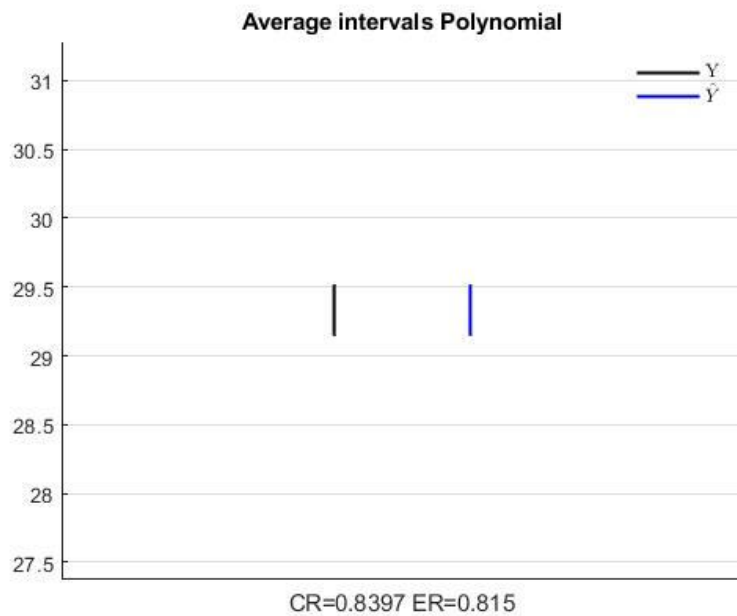


Figure 1: Case 2. Real average interval (Y) and polynomial forecasted average interval (\hat{Y}).

This polynomial forecast was combined with the random walk through simple and optimal averaging. Results are presented in Table 3.

Table 3: Case 2. Model averaging: Random Walk vs Polynomial.

METHOD	IU	CR	ER	MDE
RW	1.0000	0.1448	0.0929	4.7741
Polynomial	0.0942	0.8397	0.8150	0.0386
EW	0.5212	0.3029	0.2760	1.3579
OW	0.0890	0.8507	0.8193	0.0340

Obtained results showed remarkable improvements in the predictions when both models were combined using optimal weights. Their's U value and Mean Distance Error were close to zero (meaning the distance between the data and the forecast was minimal) and coverage and efficiency rates were even higher than the model's rates alone.

Zardoya

Zardoya's dataset was forecasted using tendency-based methods, such as the mean, the median or the quartiles. The data's lateral behavior resulted in lack of accuracy in the predictions based on tendency, with Theil's U values much greater than 1. The performance of the evaluated models was worse than the random walk.

Aumann's Mean and Hausdorff's Median were evaluated as forecasting methods, comparing and combining them with the random walk. Results proved that these two methods are similar enough on performance and characteristics to be considered equivalent when implementing data forecasting.

The following figures illustrate the similarities in ITS forecasting between the Aumann's Mean (Figure 2) and the Hausdorff's Median (Figure 3).

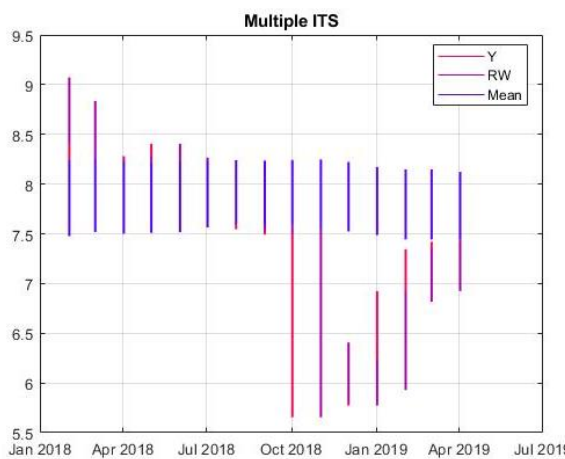


Figure 2: Case 3. Zardoya's ITS (red), Random Walk's ITS (purple), Aumann's Mean's ITS (blue).

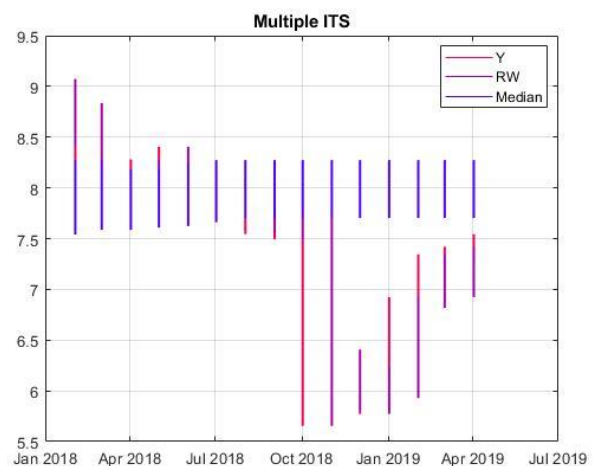


Figure 3: Case 3. Zardoya's ITS (red), Random Walk's ITS (purple), Hausdorff's Median's ITS (blue).

Visa

The final study case forecasts data with a steep upward tendency. In this case, the optimal predictions were obtained through an exponential approach. This nonlinear regression method provides a model with high CR and ER values, and low Theil's U and MDE values. Results are displayed in Table 4.

Table 3: Case 4. Model Averaging: Random Walk vs Exponential Fit.

METHOD	IU	CR	ER	MDE
RW	1.0000	0.4902	0.5315	156.2200
Exponential	0.4016	0.8322	0.8225	25.2780
EW	0.6344	0.7018	0.7164	60.2200
OW	0.6861	0.6739	0.6932	70.4870

Model evaluator's values are not as optimal as the polynomial approach. Nevertheless, this case shows another scenario where model averaging improves the characteristics of a certain model. Both optimal and equal weights were used for the combination, resulting in a discrepancy with the previous cases regarding the supremacy of optimal weights in model averaging: in this example, simple averaging provides better properties for the predictive model than the optimal weights combination.

Conclusions

Obtained results have highlighted various aspects about the performance of linear and non-linear forecasting methods for interval-valued data.

Regarding the Forecast Combination Puzzle, results have demonstrated that simple average is the most straightforward averaging method as long as the models that are being combined present similar characteristics. However, in a situation where a model outstands the rest (for example, the polynomial approach in the General Electric case), simple averaging stops being the most reliable method for model averaging. In these cases, the calculation of optimal weights based on each model's performance strengthen the properties of the most accurate model.

Predictive models based on tendency parameters, such as the average, the median or the quartiles, have proven to be less effective and accurate than other methods for datasets with lateral behaviors or steep upward trends. The absence of a tendency in the data hinders the model's ability to exactly predict future movements of the stock values and fast changing tendencies makes it hard for the model to adapt to the new values and make accurate predictions.

It was noticed that non-linear regression methods, like polynomial or exponential approaches, present higher coverage and efficiency rates. They have proven to be more accurate and to have better forecast abilities than linear methods on datasets with strong upward and downward tendencies. The development of non-linear methods might be the key to attain the best possible predictions for interval-valued data.

Finally, referring to potential future studies, it would be interesting to create an app using MATLAB's App Designer that includes all the functions programmed and developed in this project. This could

turn the outcome into a visual and user-friendly interface, facilitating the comparison between models and averaging methods.

References

- Bates, J. M., & Granger, C. W. (1969). The Combination of Forecasts. *The OR Society*, 20(4), 451-468.
- Lima Neto, E. d., & de Carvalho, F. d. (2017). Nonlinear regression applied to interval-valued data. *Pattern Analysis and Applications*, 20(3), 809-824.
- Maceda Dal-Re, P. (2018). Basic analytic system for interval valued data. Madrid: Universidad Pontificia Comillas.
- Maté, C. (2012). El análisis de intervalos. Aplicaciones en la ingeniería. *Anales de mecánica y electricidad*, 89, 20-27.
- Moral Benito, E. (2015). Model averaging in economics: an overview. *Journal of Economic Surveys*, 29(1), 46-75.

MEMORIA

Table of contents

1. Introduction.....	9
1.1. State of the art	10
1.2. Motivation.....	12
1.3. Objectives	12
1.4. Work Methodology	13
1.5. Resources.....	13
2. Interval Analysis	15
2.1. Interval data	15
2.2. IV data forecasting	16
2.3. Statistics measures for IV data	17
3. Regression methods for interval-valued data	21
3.1. Linear regression methods	21
3.2. Non-linear regression methods.....	26
4. Combining forecasts.....	29
4.1. Simple Average.....	29
4.2. Optimal weights	31
4.3. Bayesian Model Averaging	32
5. Model evaluation	35
5.1. Theil's U.....	35
5.2. Coverage Rate.....	36
5.3. Efficiency Rate.....	37
5.4. Mean Distance Error (MDE).....	38
6. Analysis in financial markets.....	41
6.1. Methodology	41
6.2. Case 1: S&P 500 index.....	48
6.3. Case 2: General Electric	53
6.4. Case 3: Zardoya	59
6.5. Case 4: Visa	66
7. Conclusions.....	71

References.....	75
Appendix I: MATLAB code	79

List of Figures

Figure 1: β Estimates from different candidate models. Source: Moral Benito (2015).....	10
Figure 2: Example of ITS forecasting. Source: Lin and González-Rivera (2016).....	11
Figure 3: ITS. Monthly electric power demand for Hour 1 in Spain, Year 2000. Source: García-Ascanio and Maté (2010).....	16
Figure 4: ITS example. Energy price in Spain. Year 2003-2004. Source: San Roque et al. (2007)	17
Figure 5: Example of equal weights combination for crisp data. Source: Prepared by the author.	30
Figure 6: Real interval (Y) and forecasted interval (\hat{Y}) with Coverage Rate and Efficiency Rate. Source: Prepared by the author.	38
Figure 7: UTheil_k.m Code.	41
Figure 8: CER_k.m Code	42
Figure 9: MDE_k.m Code.....	43
Figure 10: Example of Output from REPRESENT_k.m.....	43
Figure 11: REPRESENT_k.m Code.....	44
Figure 12: Example of Output from DrawmultipleITSvdef.m	45
Figure 13: OptimalW.m Code.....	45
Figure 14: Example of Output from MATLAB's code.....	46
Figure 15: PolynomialFit.m Code	47
Figure 16: ExponentialFit.m Code.....	47
Figure 17: Case 1. Average interval from every model. Source: Prepared by the author.....	49
Figure 18: Case 1. Multiple ITS: real, minMax, M. Source: Prepared by the author.....	51

Figure 19: Case 1. Median and Average Intervals. Source: Prepared by the author.....	52
Figure 20: Case 2. Average interval from every model. Source: Prepared by the author.....	54
Figure 21: Case 2. Polynomial fit. Source: Prepared by the author.....	56
Figure 22: Case 2. Average interval: polynomial forecast. Source: Prepared by the author.....	57
Figure 23: Case 3. Average interval from every model. Source: Prepared by the author.....	60
Figure 24: Case 3. Multiple ITS: real, Random Walk, Aumann's Mean. Source: Prepared by the author.	62
Figure 25. Case 3. Multiple ITS: real, Random Walk, Hausdorff's Median. Source: Prepared by the author.	63
Figure 26: Case 3. Average interval: combination of RW, Mean and Median. Source: Prepared by the author.	64
Figure 27: Case 4. Average interval from every model. Source: Prepared by the author.....	66
Figure 28: Case 4. Exponential fit. Source: Prepared by the author.	68
Figure 29: Case 4. Average interval: Exponential Fit. Source: Prepared by the author.....	69

List of Tables

Table 1: UNH Database	19
Table 2: Aumann's Mean and median intervals. UNH dataset.	20
Table 3: Case 1. Theil's U. Source: Prepared by the author.	48
Table 4: Case 1. Optimal weights. Source: Prepared by the author.....	50
Table 5: Case 1. MinMax vs M Method. Source: Prepared by the author. ...	50
Table 6: Case 2. Theil's U. Source: Prepared by the author.	53
Table 7: Case 2. MDE. Source: Prepared by the author.	54
Table 8: Case 2. Random Walk vs M3. Source: Prepared by the author.	55
Table 9: Case 2. Random Walk vs Polynomial Model. Source: Prepared by the author.....	57
Table 10: Case 3. Theil's U. Source: Prepared by the author.	59
Table 11: Case 3. Optimal weights. Source: Prepared by the author.....	60
Table 12: Case 3. Random Walk vs M2. Source: Prepared by the author. ...	61
Table 13: Case 3. Random Walk vs M3. Source: Prepared by the author. ...	62
Table 14: Case 3. Combination of three models: Random Walk, Aumann's Mean, Hausdorff's Median. Source: Prepared by the author.....	64
Table 15: Case 4. Theil' U. Source: Prepared by the author.	67
Table 16: Case 4. Random Walk vs Exponential Fit. Source: Prepared by the author.....	69

1. Introduction

Statistics applications are wider and more complex everyday thanks to the advances in data processing. Model averaging in particular is key in offering potential solutions to many problems regarding forecasting. The combination of predictive models can be the answer to challenges presented on areas with global reach, such as financial markets. The necessity to predict future stock values and the reliance big and small companies have on the goodness of these predictions makes the pursuit of solutions a fundamental economic interest.

Single or crisp data analysis and its application to these problems has been deeply developed over the years; nevertheless, its results are not always the simplest or easiest ones. Crisp data may lack on the information necessary to describe a certain value, providing incomplete forecasts and predictions.

The more demanding needs of financial forecasters and different day-to-day fields, such as risk evaluation in engineering, programming or robotics, require the development of methods for a different kind of data: interval-valued (IV) data. This type of data is based on intervals with lower and upper limits for each measure. The information given by interval data is usually more accurate and complete than that which can be extracted from single data, so it is highly valuable when precision is a priority in data analysis (Maté, 2012).

A remarkable use of intervals is the study of time-based data and predicting series. Time series can be found in data concerning weather evolution, markets movement or changes in energy prices. This leads to the development of Interval Time Series (ITS), IV data registered as an ordered sequence throughout time. Arroyo and Maté (2006) prove the versatility of ITS in the prediction problem and forecasting.

IV data can also provide a broader framework and better solutions to complex subjects regarding regression models. Forecasting financial markets, an essential activity for the economy, relies heavily on these regression models, so IV data and ITS can open new and interesting avenues here.

Moral Benito (2015) raised the issue of the existing uncertainty when choosing a regression model over the space of all possible models. The information lost when disregarding some of the regression models is not usually taken into account when calculating the uncertainty associated to the model. In an attempt to avoid this common mistake, model averaging will be considered in the design of a tool that assesses the different regression methods.

Moral Benito (2015) justified the uncertainty of regression models when estimating parameters depending on the chosen control variables. Figure 1 shows the β estimates from 65,536 regression models. These models were

obtained from different combinations of control variables of the same data and evaluated according to the Bayesian Information Criterion (BIC).

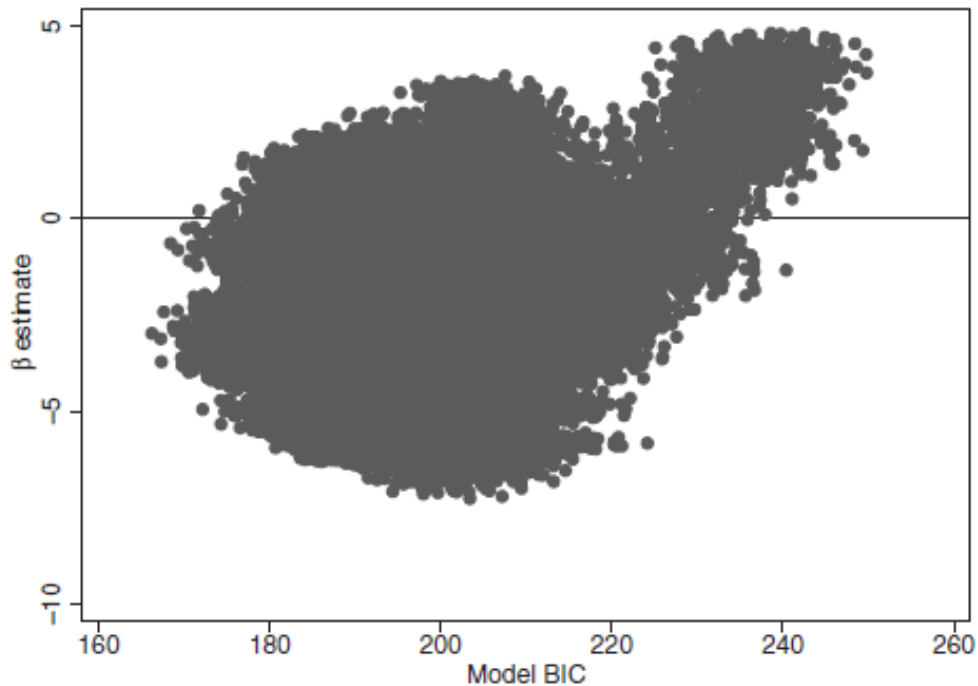


Figure 1: β Estimates from different candidate models. Source: Moral Benito (2015).

Figure 1 illustrates the variability of forecasting models depending on the selected variables. Selecting a model over another could lead to significant differences on results, as the amount of discarded information affects the predictive abilities of the model. Finding ways to combine the most advantageous properties of each model represents a potential improvement on regression methods, suitable to IV data models and forecasts.

This project will focus on linear and non-linear regression models in order to find new methods of working with both crisp and interval data, being able to obtain precise estimations and apply the results in the forecast of financial markets. The final expected output is a set of MATLAB functions that works as a tool to analyze and compare several models.

1.1. State of the art

Interval-valued data, despite of its many and useful possible applications, has not been explored as much as single data over time. According to Maté (2012), the first time someone used IV analysis as a tool was in 1959, when Ramon Moore applied IV to automatic error control. However, the research boost actually started in 1960, when methods and applications regarding interval-valued data came to the forefront.

One of the areas in which interval-valued data has had recent advances is in regression models. In relation to linear regression, different methods have been studied over the last years to determine their performance with IV data, such as the Lasso technique, the parametrized approach, the Constrained Center and Range Method (CCRM) and the MinMax method. Nevertheless, Maceda (2018) concluded that there is not a favored method yet among the existing ones.

Nonlinear regression and its method (NLM) for interval-valued data has been evaluated recently through the Monte Carlo experiments, demonstrating a greater accuracy and precision than linear regression methods in interval-valued dataset and different scenarios (Lima Neto & de Carvalho, 2017).

Extensive research has been done about Interval Time Series, mainly in the forecasting area. Lin and González-Rivera (2016) proposed an approach for data sets with interval data as the only relevant information. Assuming an underlying stochastic process for the IV time series, the authors developed a forecasting model for beef prices from 2010 to 2013 and evaluated the performance of this model. Figure 2 represents the actual and predicted beef prices (historical data provided by the United States Department of Agriculture).

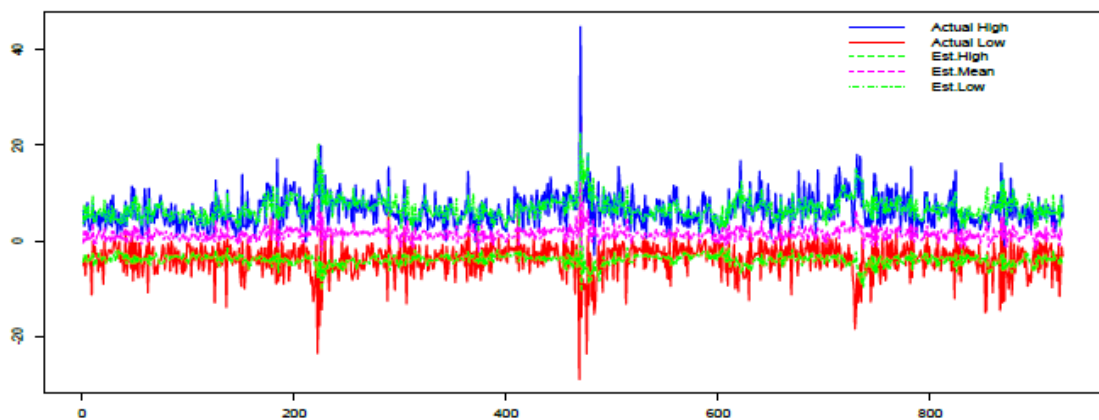


Figure 2: Example of ITS forecasting. Source: Lin and González-Rivera (2016).

Lin and González-Rivera (2016) exploited the extreme nature of the lower and upper bounds of the interval, proving that IV data approaches offer much more information than crisp ITS models.

Model averaging represents one of the most powerful techniques on regression models, but there is still a lot of work ahead the current State-of-the-Art. Moral Benito (2015) analyzed the Bayesian Model Averaging (BMA), the most developed method that emphasizes on economic applications and has a soaring relevance in software for model averaging, and the Frequentist

Model Averaging (FMA), a potential future tool when dealing with endogenous regressors.

The cutting-edge studies emphasize the usefulness of the BMA in situations of model uncertainty, due to its unambiguous, finite sample nature (Steel, 2019). The author highlights the applicability of model averaging in economics, with targets such as forecasting or scenarios modelling.

Nevertheless, the current challenge is related to MATLAB's libraries. So far, it is not possible to find a tool that represents, analyzes and compares all the explored regression methods, linear and non-linear, and allows different combinations of them to study model averaging.

1.2. Motivation

The best way of approaching the challenging problems that need to be faced every day in relation to data predictions is with solutions at the same level. Regression models have already plunged into crisp data, but there are still abundant possibilities regarding IV data.

As for forecasting, there is still uncertainty and lack of confidence on the results provided by the current methods. Analyzing more complex data such as IV data could provide the key factors for an accurate prediction.

Further exploration of regression model averaging in single and interval-valued data could aid the financial field great, particularly in forecasting financial markets. This is the reason why this project looks forward to analyzing existing approaches to regression models and creating a visual system that helps moving closer the target: accurate predictions for IV databases in financial markets.

With the final aim of creating a MATLAB environment that compares and evaluates different regression methods, this work expects to be a follow-up and expansion of previous works about interval-valued data and its applications.

1.3. Objectives

The main objectives of this project are listed below:

- Building a set of MATLAB functions. The goal is to create a system for exploring different methods of regression in an original and visual way.
- Compare the results provided by the current regression methods and those obtained by model averaging, in order to draw conclusions about the quality and precision of the different estimations.

- Develop and delve into previous works done on interval-valued data analysis. This means pushing forward the state-of-the-art on this field by exploring new possible applications of IV data analysis.
- Validate the regression models applications in particular problems in financial markets, using interval time databases.

1.4. Work Methodology

In the early stages, this project will be based on the study and processing of existing data about interval-valued data, linear and non-linear regression models, averaging among these models, and their applications to financial markets. Some of the sources will be found by research and some of them will be provided by the director, and they will include scientific articles, papers and prior thesis.

After the research is done, the work will focus on MATLAB programming, in order to create the necessary programs to explore, compare and average different regression model methods. At this stage, some functions will need to be created, as well as the tool and its interface.

1.5. Resources

The major resource used in this project will be MATLAB. The software from MathWorks will be the tool utilized to define the functions needed to create the different regression models and apply them to the existing data. The final goal is to produce a set of functions that will facilitate the averaging between models, both linear-linear averaging and linear-nonlinear.

The final results and conclusions will be supported with graphs and screenshots from MATLAB, aiming to open a new research path or a new perspective regarding model averaging and, specifically, model averaging for financial data.

2. Interval Analysis

Symbolic data is a kind of data with a more complex and complete structure than crisp data. It is not limited by the boundaries of a single value for each datum, providing wider possibilities to study existing problems in many different fields, such as engineering or financial markets. This type of data appeared for the first time in the second half of the twentieth century, in an attempt to fill the gaps left by single or crisp data.

Symbolic data analysis (SDA) has been considerably developed in recent years, as it gives a framework able to consider the variability of the data, significantly reducing errors (Norhomme-Fraiture & Brito, 2011). The authors distinguished several types of symbolic data, including categorical data (modal and multi-valued) and numerical data, which comprises interval-valued data, histogram variables and functions.

The subject of study in this project are interval-valued (IV) data, introduced by Moore (1966). The author described the interval as a number pair with a dual nature, representing an interval of real numbers $a \leq x \leq b$.

This section summarizes the applications of IV data in forecasting and the main descriptive statistic measures for interval-valued datasets.

2.1. Interval data

Interval data is the particular case of symbolic data with a lower and upper limit for each value. There are two options to represent an interval. The first one is with its lower and upper bounds: $[x] = [x^L; x^U]$, where $-\infty < x^L \leq x^U < \infty$. The other way to define an interval is by its center and radius (half of the range): $[x] = \langle x^C; x^R \rangle$, where $x^C = (x^L + x^U)/2$ and $x^R = (x^U - x^L)/2$.

The range comprised by interval data allows the consideration of wider ranges of information, reducing the information lost when assigning a single value to a certain figure (as in crisp data).

Interval Time Series (ITS) are a kind of databases described by Maté (2012) as a time-ordered sequence of IV data. ITS are denoted as $\{[x_t]\} = \{[x_t^L; x_t^U]\}$ for $t = 1, 2, \dots, n$. They are a really valuable tool when organizing interval values chronologically throughout time.

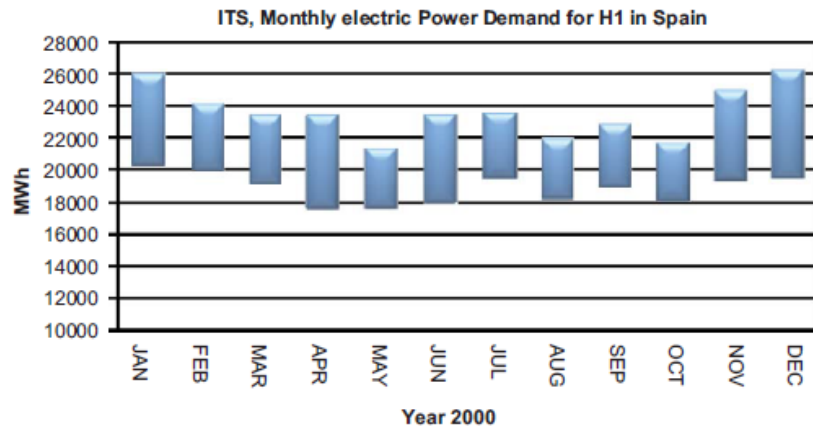
2.2. IV data forecasting

Like most mathematical areas, forecasting methods were originally developed for single valued data. These methods integrated regression models (linear and non-linear) or tendency-based methods, such as moving average models (García-Ascanio & Maté, 2010).

Regarding interval-valued databases, different techniques have also been studied in the last years. Since Moore (1966) introduced the importance of interval analysis as a tool to deal with relevant scientific problems, updates on this field have been continuously made.

A review of the main regression methods for IV data forecasting is included in Section 3. The evolution of these methods demonstrates the reliability they have on the previous ones and the importance of pushing the state-of-the-art if the best possible performer models are the goal.

Forecasting has been especially developed for ITS. García-Ascanio and Maté (2010) proposed an approach to an electric power demand forecasting problem (Figure 3) using two different solutions: the vector autoregressive (VAR) models and the multi-layer perceptron model (iMLP), both adapted to interval data.



Upper Bound	26091	24219	23513	23491	21373	23445	23609	22008	22842	21768	24956	26302
Lower Bound	20328	19920	19137	17679	17582	18000	19494	18198	18909	18141	19342	19523
	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC

Figure 3: ITS. Monthly electric power demand for Hour 1 in Spain, Year 2000. Source: García-Ascanio and Maté (2010)

The authors concluded that the use of symbolic data in the energy forecasting field could be a potential progress from crisp data, obtaining more complete and rigorous results.

San Roque et al. (2007) provided another example of an ITS in the energetic field. Figure 4 illustrates the evolution throughout time of the energy prices in the Spanish market.

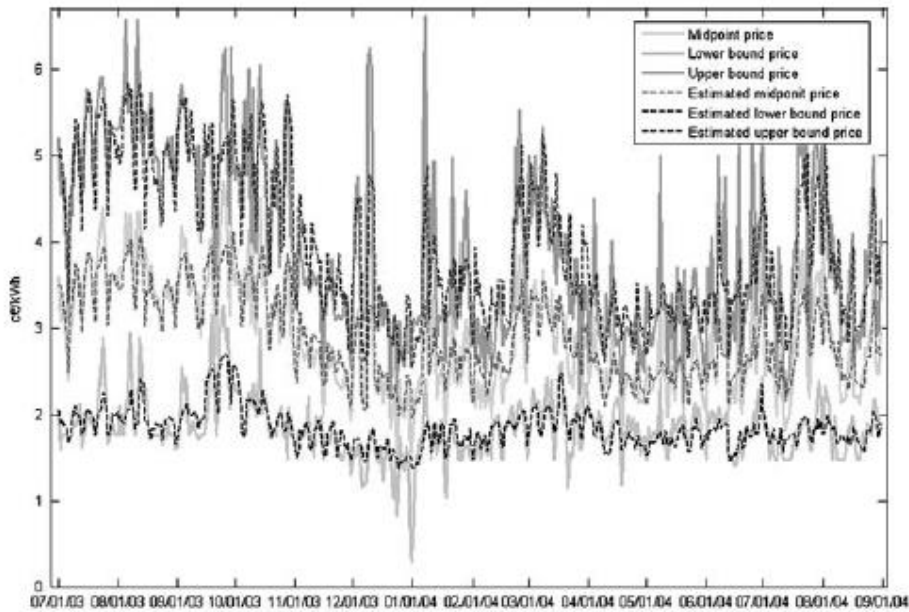


Figure 4: ITS example. Energy price in Spain. Year 2003-2004. Source: San Roque et al. (2007)

This time series was evaluated applying multi-layer perceptrons (iMLP) as predictive models for IV data. Results proved that this approximation was a valuable tool when minimizing the cost function defined by the discrepancies between the desired output intervals and the estimated ones.

Regarding financial applications, Arroyo, González-Rivera and Maté (2010) analyzed the effect of regression IV methods on S&P500 prices. The authors pointed out the quality of the simple smoothing techniques when applied to histogram time series.

Forecasting accuracy measures for IV data and ITS will be exposed in Section 5, including parameters like Theil's U, the Coverage Rate, the Efficiency Rate and the Mean Distance Error. These measures allow the evaluation of the quality of the forecasting models, providing criteria to discern accurate models from worse performer ones.

2.3. Statistics measures for IV data

Interval-valued databases are a source of new approaches to many kinds of problems, due to their descriptive nature on figures that need more than one value to be defined. For example, stock values over a period of time (that

contain maximum and minimum values) or weather forecasts (made of maximum and minimum temperatures).

However, this resource would be useless without proper analytic tools for this kind of data. The main statistics measures for IV datasets are presented hereafter, providing a theoretical framework for their further application.

(i) Aumann's Mean

This is a descriptive statistic measure for univariate interval-valued data. It is a central-tendency measure. Bertrand and Goupil (2000) defined the mean of an interval set as the mean of the centers, according to the formula that follows:

$$\bar{x}_1 = \frac{1}{2n} \sum_{i=1}^n (x_i^L + x_i^U) = \frac{1}{n} \sum_{i=1}^n x_i^C \quad (1)$$

Where x_i represents an observed random sample of size n from a random interval $[X]$.

This mean of the centers (\bar{x}^C) can be treated as the simple mean of an IV database, whereas considering only the interval centers has the risk of mistaking IV data for crisp data. The mean of the centers is more likely to be contemplated as the center of gravity from the database.

In order to include wider information about the dataset, the mean of the radii is defined as:

$$\bar{x}^R = \frac{1}{n} \sum_{i=1}^n x_i^R \quad (2)$$

According to Maté (2019), the Aumann's Mean is established as the mean value interval of the dataset:

$$[\bar{x}] = [\bar{x}^C - \bar{x}^R, \bar{x}^C + \bar{x}^R] \quad (3)$$

This central tendency statistic takes into account, not only the average of the interval, but also the range of the data. This will be the preferred measure to be applied in forecasts, as it is a more reliable representation of the information given by the interval.

(ii) 1-norm Median

The 1-norm median is one of the possible approaches for the median of a set of intervals. It is defined as the median of the interval-valued dataset, made

separately for the upper and the lower limits. Maté (2019) gathers its formulation as:

$$\begin{aligned} [Me([X])]_{1N} &= [Me\{[x_1], [x_2], \dots, [x_n]\}] \\ &= [Me\{x_1^L, \dots, x_n^L\}, Me\{x_1^U, \dots, x_n^U\}] \end{aligned} \quad (4)$$

Where x_i^L represent the lower limits of the interval and x_i^U stand for the upper limits.

(iii) Hausdorff's Median

This interval median type is defined as the median of the IV dataset in the center-radius formulation. According to Maté (2019), it can be stated as:

$$\begin{aligned} \langle Me([X]) \rangle_H &= [Me\{\langle x_1 \rangle, \dots, \langle x_n \rangle\}] \\ &= [Me\{x_1^C, \dots, x_n^C\} - Me\{x_1^R, \dots, x_n^R\}, \\ &\quad Me\{x_1^C, \dots, x_n^C\} + Me\{x_1^R, \dots, x_n^R\}] \end{aligned} \quad (5)$$

Where x_i^C represent the interval centers and x_i^R stand for the interval radii.

Hausdorff's median has as lower limit the difference between the centers' median and the radii's median, while the upper limit is obtained through adding the median of all centers and the median of all radii. 1-norm median and Hausdorff's median can be both equal and different, depending on the chosen interval.

Table 1 comprises a dataset of stock values from United Health Nation (an American healthcare company) for three months of 2017, with the maximum and minimum value for each month.

Table 1: UNH Database

	min	Max
April 2017	164.25	176.07
May 2017	166.65	178.89
June 2017	175.19	188.66

Aumann's mean, 1-norm median and Hausdorff's mean are calculated for this database and presented in Table 2. These statistics are calculated for the three-interval dataset and, in this case, both medians present the same values.

Table 2: Aumann's Mean and median intervals. UNH dataset.

	x^L	x^U
Aumann's Mean	168.70	181.21
1-norm Median	166.65	178.89
Hausdorff's Median	166.65	178.89

The exchangeability grade of these measures in interval forecasting will be addressed in following sections, comparing both statistic measures as predicting methods in relation to their performance and accuracy for different databases.

3. Regression methods for interval-valued data

This section comprises a bibliographic review of the existing regression methods for IV data. Linear and nonlinear models will be explained and analyzed according to their performance and current utility.

3.1. Linear regression methods

Linear regression attempts to be a linear approach to a model that relates a dependent variable with one or more independent or explanatory variables. It has an extensive development for crisp data, both for single and multiple variables, but linear regression methods for interval-valued data have not been deeply developed so far.

The most complex type of problems regarding regression methods are those where both dependent and independent variables are interval valued data. The equation of these problems should look like this: $\beta_0 + \beta_1[X_1] + \dots + \beta_n[X_n] = [Y]$, where the X are the independent variables, Y the dependent variable and β the parameters of the model.

The objective of linear regression methods is to minimize the sum of the squares of deviation, that is, the number obtained by adding the squares of the errors of all the equations in a model.

The main existing methods of interval linear regression (ILR) are explained in this section and follow in sequence.

(i) Centre Method (CM)

This is the oldest method, introduced in 2000 by Billard and Diday. It deals with IV data as if it was crisp data, by using the centers of the intervals like single data and then applying the model to the upper and lower limits.

The Sum of the squares of deviation is the square of the sum of the errors from the upper limit equation and the lower limit equation:

$$S_{CM} = \sum_{i=1}^n (\epsilon_{li} + \epsilon_{ui})^2 \quad (6)$$

Where ϵ_{li} and ϵ_{ui} come from:

$$\begin{aligned} \epsilon_{li} &= y_{li} - \beta_0 - \beta_1 x_{l1} - \dots - \beta_n x_{ln} \\ \epsilon_{ui} &= y_{ui} - \beta_0 - \beta_1 x_{u1} - \dots - \beta_n x_{un} \end{aligned} \quad (7)$$

The minimum value of S_{CM} is achieved with the values of the parameters ($\beta_0, \beta_1, \dots, \beta_n$) obtained with a linear regression of the centers of $[Y]$ and $[X_1, X_2, \dots, X_n]$. The resulting values of the parameters can predict \hat{Y}_{CM} :

$$\begin{aligned}\hat{y}_{li} &= \hat{\beta}_0 + \hat{\beta}_1 x_{l1} + \dots + \hat{\beta}_n x_{ln} \\ \hat{y}_{ui} &= \hat{\beta}_0 + \hat{\beta}_1 x_{u1} + \dots + \hat{\beta}_n x_{un}\end{aligned}\tag{8}$$

This method sets the basis for the ones that came after it. However, its lack of precision justifies its classification as rudimentary, as, by focusing on the centers, information such as the range of the interval is being missed.

(ii) MinMax Method

MinMax methods represent an improvement of the Center Method, introduced in 2002 by the same authors. The main difference with the previous work of the authors is the use of different parameters for the upper and the lower limits and handling them like different variables.

In this case:

$$\begin{aligned}\epsilon_{li} &= y_{li} - \beta_0^l - \beta_1^l x_{l1} - \dots - \beta_n^l x_{ln} \\ \epsilon_{ui} &= y_{ui} - \beta_0^u - \beta_1^u x_{u1} - \dots - \beta_n^u x_{un}\end{aligned}\tag{9}$$

The sum of the squares of deviation is the sum of the squares of the different errors in the upper and lower bounds:

$$S_{MinMax} = \sum_{i=1}^n (\epsilon_{li})^2 + \sum_{i=1}^n (\epsilon_{ui})^2\tag{10}$$

The MinMax method requires two separate linear regressions in order to obtain the values of the parameters and minimize S_{MinMax} , one for the lower values and another one for the upper values.

These parameters allow to predict \hat{Y}_{MinMax} :

$$\begin{aligned}\hat{y}_{li} &= \hat{\beta}_0^l + \hat{\beta}_1^l x_{l1} + \dots + \hat{\beta}_n^l x_{ln} \\ \hat{y}_{ui} &= \hat{\beta}_0^u + \hat{\beta}_1^u x_{u1} + \dots + \hat{\beta}_n^u x_{un}\end{aligned}\tag{11}$$

This model increases the precision in the estimation by considering more information of the intervals, such as the upper and the lower limits. Symbolic regression approach is preferred to classical techniques (Billard & Diday, 2002).

(iii) Center and Range Method (CRM)

Lima Neto and De Carvalho developed this method in 2008. It is also a development of the Center Method. The Center and Range Method looks forward to improving the model response to complex data by adding the ranges of the intervals to the existing centers. To simplify the mathematical expressions, IV data is represented by the center and radio of each interval (instead of using the lower and upper limits).

The sum to minimize in this model is the sum of the square of the errors of the centers and the square of the errors of the radiuses:

$$S_{CRM} = \sum_{i=1}^n ((\epsilon_{ci})^2 + (\epsilon_{ri})^2) \quad (12)$$

ϵ_{ci} and ϵ_{ri} can be expressed as:

$$\begin{aligned} \epsilon_{ci} &= y_{ci} - \beta_0^c - \beta_1^c x_{c1} - \dots - \beta_n^c x_{cn} \\ \epsilon_{ri} &= y_{ri} - \beta_0^r - \beta_1^r x_{r1} - \dots - \beta_n^r x_{rn} \end{aligned} \quad (13)$$

The values of the β parameters that minimize S_{CRM} are obtained by two a linear regression of the centers and an independent linear regression of the radiuses. \hat{Y}_{CRM} is predicted as follows:

$$\begin{aligned} \hat{y}_{ci} &= \hat{\beta}_0^c + \hat{\beta}_1^c x_{c1} + \dots + \hat{\beta}_n^c x_{cn} \\ \hat{y}_{ri} &= \hat{\beta}_0^r + \hat{\beta}_1^r x_{r1} + \dots + \hat{\beta}_n^r x_{rn} \end{aligned} \quad (14)$$

This model provides a scatter plot solution for the centers and another one for the radii. The additional information about the intervals (radii) compared to the Center Method arranges the Center and Range Method as more accurate and reliable than the first one.

(iv) Constrained Center and Range Method (CCRM)

CCRM is an improvement of the Center and Range Method made by Lima Neto and de Carvalho in 2010. The main difference is the addition of a non-negative restriction for the radii, forcing them to be positive.

The sum of the squares of deviation is the same as the CRM but including the constraint that makes the radiuses positive:

$$\begin{aligned} S_{CCRM} &= \sum_{i=1}^n ((\epsilon_{ci})^2 + (\epsilon_{ri})^2) \\ s. t. \beta_i^r &\geq 0, \quad i = 0, \dots, p. \end{aligned} \quad (15)$$

This method requires from two separate regression problems to find the optimal values of the parameters β : one for the centers and one for the radii. The centers are not affected by the new constraint, nevertheless, the estimate of β^r is solved as a constrained regression problem.

The prediction of \hat{Y}_{CCRM} can be expressed like this:

$$\begin{aligned}\hat{y}_{ci} &= \hat{\beta}_0^c + \hat{\beta}_1^c x_{c1} + \dots + \hat{\beta}_n^c x_{cn} \\ \hat{y}_{ri} &= \hat{\beta}_1^r X_r\end{aligned}\tag{16}$$

Compared to the CRM method, the CCRM provides the same solution for the centers' approximation, but a slightly different one for the radiuses. It can be considered a better prediction, as it avoids illogical results such as negative radii.

(v) Arithmetic-based simple linear regression (ABSLR)

This method, introduced by Sinova, Colubi and González-Rodríguez (2012), breaks with the previous trends by using an arithmetic approach for the regression and by being only for simple regression.

The approach to the sample Y is made by looking for a real value \hat{a} and an interval value \hat{B} and making $\hat{a}X + \hat{B}$ an estimation of the sample. The algorithm to find these values:

$$\frac{1}{n} \sum_{i=1}^n (d_{\theta}(Y, \hat{a}X + \hat{B}))\tag{17}$$

Once \hat{a} and \hat{B} are obtained, it is possible to predict \hat{Y}_{ABSLR} :

$$\hat{Y}_{ABSLR} = \hat{a}X + \hat{B}\tag{18}$$

The ABSLR method provides a prediction of the upper and lower limits of the interval, with the same slope as \hat{a}

(vi) Linear regression based on Lasso technique

Giordani brought in this method in 2011. The ultimate goal is to find a common set of regression coefficients for the center and the radius. These coefficients must be estimated in such a way that the coefficients for the midpoints are close to the ones for the radii as much as possible.

Being b_c and b_r the vector of the regression coefficients for the midpoint and the radius, respectively, and b_a the vector of the additive coefficients ($b_r = b_c + b_a$), the linear regression relationship can be expressed as:

$$\begin{aligned}
y_c &= b_c X_c + \epsilon_c \\
y_r &= b_r X_r + \epsilon_r = (b_c + b_a) X_r + \epsilon_r
\end{aligned}
\tag{19}$$

A constraint is also needed to ensure the positivity of the radii as well as the minimum value for b_a :

$$(b_c + b_a) X_r \geq 0 \tag{20}$$

The Lasso (Least Absolute Shrinkage and Selection Operator) technique is an estimation method that shrinks some regression coefficients while sets other ones to 0.

The estimation of \hat{Y}_{lasso} results in:

$$\begin{aligned}
\hat{y}_c &= b_c X_c \\
\hat{y}_r &= (b_c + b_a) X_r
\end{aligned}
\tag{21}$$

The linear regression method based on Lasso technique provides the same result as the CCRM for the radii (as they both use the same constraint for non-negative radiuses), but a different one for the centers.

(vii) Constrained center and range joint method (CCRJM)

The Constrained Center and Range Joint method, introduced in 2017 by Ha and Guo, is a development of the CCRM method: it also sets a coefficient for the center and the radius, but in a shared way, considering the relationship between these two elements.

Being the linear regression relationship the following:

$$\begin{aligned}
Y_c &= X\beta_c + \epsilon_c \\
Y_r &= X\beta_r + \epsilon_r
\end{aligned}
\tag{22}$$

The objective of the problem is to minimize the sum of the squares of the errors:

$$\begin{aligned}
\min \sum_{i=1}^n ((\epsilon_{ci})^2 + (\epsilon_{ri})^2) \\
s. t. X\beta^r \geq 0
\end{aligned}
\tag{23}$$

The prediction for \hat{Y}_{CCRJM} is given as follows:

$$\begin{aligned}
\hat{Y}_c &= X\beta^c \\
\hat{Y}_r &= X\beta^r
\end{aligned}
\tag{24}$$

This method gives a multiple regression solution that cannot be represented in a two-dimension plot, so the only way to test its effectiveness is numerically.

(viii) Parametrized Approach for Linear Regression of Interval Data

This is the most recent method, created by Souza et al in 2017. It goes a step forward from the MinMax method, jointing the lower and upper limits by using the lower limit values for the upper limit model and vice versa. As the CCRJM, reckons with the relationship among these two components of the interval.

The linear regression relationship would be:

$$\begin{aligned} Y_u &= X\beta_u + \epsilon_u \\ Y_l &= X\beta_l + \epsilon_l \end{aligned} \quad (25)$$

In this method, the sum of the squares of deviation needs to be minimized independently for the upper limit and the lower limit in order to obtain β^u and β^l :

$$\begin{aligned} S_l &= \sum_{i=1}^n (\epsilon_{li})^2 \\ S_u &= \sum_{i=1}^n (\epsilon_{ui})^2 \end{aligned} \quad (26)$$

After the coefficient matrix are operated, the resulting prediction for \hat{Y}_{param} can be expressed as:

$$\begin{aligned} \hat{Y}_l &= X\beta^l \\ \hat{Y}_u &= X\beta^u \end{aligned} \quad (27)$$

Like in CCRJM, this method's effectiveness can only be tested numerically, as the complexity of the results does not allow their representation in a 2D plot.

3.2. Non-linear regression methods

Linear regression models are not a valid method when the relation between the independent (X) and the dependent (Y) variables is a nonlinear function. These cases require non-linear regression methods, which contain nonlinear equations and usually need from iterative procedures to solve the model. These methods are useful when data presents a non-linear trend, such as quadratic or logarithm functions.

Non-linear regression methods are both more complex and more useful in certain problems, for example, when the function that relates the variables is not known. The advantages of these kind of models are now being proved with interval-valued data, as many real IV data problems include nonlinear relationships between the variables. Lima Neto and De Carvalho introduced in 2016 a nonlinear regression method (NLM) for IV variables, a method which uses two independent nonlinear regression models to fit the functions.

Considering Y the dependent variable and X_j ($j=1, \dots, n$) the independent predictor variables, the nonlinear regression relationship can be expressed like:

$$\begin{aligned} Y_i^c &= f_c(X_i^c, \theta^c) + \epsilon_i^c \\ Y_i^r &= f_r(X_i^r, \theta^r) + \epsilon_i^r \end{aligned} \quad (28)$$

Where ϵ represent the random errors, θ are vectors of unknown parameters, and f are non-linear functions.

The sum of the squares of deviations is the sum of the squares of the errors:

$$S_{NLM} = \sum_{i=1}^n (\epsilon_i^c)^2 + \sum_{i=1}^n (\epsilon_i^r)^2 \quad (29)$$

With the errors being:

$$\begin{aligned} \epsilon_i^c &= Y_i^c - f_c(X_i^c, \theta^c) \\ \epsilon_i^r &= Y_i^r - f_r(X_i^r, \theta^r) \end{aligned} \quad (30)$$

In the same way as linear regression models, the ultimate goal is to estimate the parameters that minimize the sum of the squares of deviations by using optimization methods.

Lima Neto and de Carvalho (2017) recognize and evaluate 4 optimization algorithms that will be explained in sequence.

1. The Broyden-Fletcher-Goldfarb-Shanno (BFGS) method uses an approximate Hessian matrix¹ for solving nonlinear optimization problems.
2. The Stochastic Gradient method and the Conjugate Gradient method are based on the gradient of an objective function. There are Sub-Gradient methods that back-up the cases when the gradient cannot be obtained.
3. The Conjugate Gradient approach differs in the absence of Hessian matrix evaluations, inversion and storage.

¹ Square matrix of second-order partial derivatives of a scalar-valued function.

4. Finally, the Simulated Annealing method only utilizes the objective function values (and not derivatives). It is a global search algorithm whose utility can be found in the production of logical starting points for the rest of the methods.

Summarizing, non-linear regression models aim to supplement the results of linear regression models in those cases where linear is not enough, either because of complex relations among the variable or because of lack of information.

4. Combining forecasts

The more accurate a regression method is, the higher quality found in its predictions. The study and development of new regression techniques, both linear and non-linear, looks forward providing new models that will be able to discard old forecasts in favor of the newer and better ones.

However, this constant search for new methods could lead to increasing errors, as rejecting a model implies rejecting useful independent information in two ways: one model bases its forecast on variables or information that the other one has not considered, and different forecasts make different assumptions about the relationships between the variables (Bates & Granger, 1969).

The combination of different forecasts is an appealing alternative to dismissing a certain method, offering the opportunity to bring together the properties of two or more models, increasing this way the quality of the forecast. By combining several forecasts, the estimates of the figures of interest are usually improved, and it is more likely to achieve unbiased results. Model or forecasts averaging bypasses the understatement of the real uncertainty of the models, as it emphasizes the value of the lost information when disregarding forecasts (Moral Benito, 2015).

There are numerous procedures when combining forecasts, being some simpler or more accessible than others. Assigning equal weights to every model is definitely a simple and effective mean, yet common sense urges us to allocate more weight to those models who seem to hold the lower errors.

According to Clemen and Winkler (1986), the simple average (forcing the weights assigned to the forecasts to be equal) is a successful approach to the dilemma of models' combination. The Bayesian Model (BMA), which contracts the weights that would be used a priori toward equality, is another method that avoid the rise of extreme weights, being also a quite robust system.

This section will analyze some of these techniques, providing examples and advantages of each of the averaging methods.

4.1. Simple Average

Simple averaging is the most elemental combining procedure, nonetheless, the combination of models using equal weights has proven to be one of the most effective. Under equal weights averaging, each model is given equal weight to obtain the combined forecast. It is particularly effective when the forecasts that are being combined present similar variance and correlations.

Being $\hat{Y}_1, \hat{Y}_2, \dots, \hat{Y}_k$ the k forecasts that are being combined, and Y the real variable being forecast, the combined forecast \hat{Y}_C can be estimated like:

$$\hat{Y}_C = \frac{\sum_{i=1}^k \hat{Y}_i}{k} \quad (31)$$

This method deals with the forecasts from different models as if they were exchangeable, which is not always the case. However, the arithmetic average often serves as a useful benchmark and is sometimes more valid than more complicated combining attempts (Clemen & Winkler, 1986).

Figure 3 illustrates an example of simple average combination with crisp data. The data belongs to Abbott Laboratories, an American drug company founded in 1888 and included in the S&P 500 index. The S&P 500 is one of the most important stock indexes in the US, often considered the most representative index of the current market situation.

The graph shows two different predictions of the centers of the intervals containing the maximum and minimum values of the stocks over a certain month, as well as the real data and the combined forecast (\hat{Y}_C) using equal weights. Time frame goes from December 2016 to February 2018, using data from the previous two years to make each prediction. Forecast 1 (\hat{Y}_1) is made with a random walk, while the second method (\hat{Y}_2) predicts using the mean of the last 24 values.

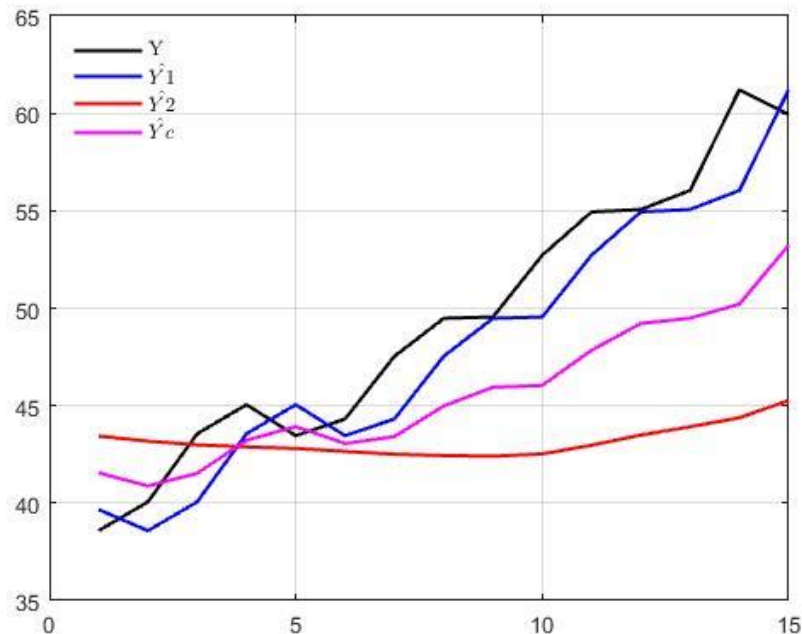


Figure 5: Example of equal weights combination for crisp data. Source: Prepared by the author.

The Forecast Combination Puzzle stands the theory on how simple combinations of forecast are repeatedly found to beat and outperform elaborated weighted combinations

According to Smith and Wallis (2009), the explanation for the fact that simple combinations outperform sophisticated combinations in empirical applications relies in the effect of finite-sample error when estimating the combining weights. A possible reason is the instability that arises when the weights are calculated, as this estimation is highly dispersed. The authors proved that optimal combining weights that are close to equality tend to be more accurate and fit better the resolution of the puzzle.

Exploring the possibility of beating the simple average as a forecast combining method has been a branch of research since the combining forecasts dilemma was raised by Bates and Granger in 1969. Some of the study cases in this project will also follow that path, trying to find a better weighted combination or, on the other side, attempting to provide evidence that supports the superiority of simple average versus complex models.

4.2. Optimal weights

Model averaging using optimal weights aims to assign different weights to different models based on each model's performance, aiming to minimize errors in the combined forecast.

Bates and Granger (1969) suggested a weight calculation based on the variance of the errors of a certain model (σ_i) as the performance of the forecast, and the correlation coefficient between the errors of two models (ρ).

In a two-model scenario, proportionate weight w given to the first model was defined as:

$$w = \frac{\sigma_2^2 - \rho * \sigma_1 * \sigma_2}{\sigma_1^2 + \sigma_2^2 - 2 * \rho * \sigma_1 * \sigma_2} \quad (32)$$

The combined forecast of the two models was calculated as:

$$\hat{Y}_C = w * \hat{Y}_1 + (1 - w) * \hat{Y}_2 \quad (33)$$

Methods that present the lower mean-square errors are typically the ones with weights that adapt quickly to new values and that barely vary about the optimum value. This is crucial when dealing with financial data, as sudden

changes in the tendency are likely to occur (either upward or downward) and the model needs to be prepared to face them.

A variant of this method will be proposed in the next chapter, using the Theil's U as the estimator for the model's performance.

Benefits of using optimal instead of equal weights when combining forecasts include creating a compelling method for cases with dissimilar forecasts. It takes into account the "quality" of each model, seeking for a combination of models with the minimum error.

4.3. Bayesian Model Averaging

BMA is another forecast combination method. These methods represent the alternative methodology to factor modelling when forecasting with large datasets. They have a big impact on different fields like economy, ecology or energy.

The Bayesian perspective assumes that there is a distribution of models where each model has a different uncertainty probability. It suggests a direct way of building model weights for forecast combination, based on the idea of different models contributing to the posterior odds of the chosen one (Kapetanios, Labhard, & Price, 2008).

The output of a Bayesian analysis provides a probability distribution for the quantified focus of interest for the analysis of a given set of models. BMA focuses on the probability of a certain model being the optimal model, providing a better average predictive ability than predictions made using just one model. It disregards the assumption of some methods that apply weights to every considered model: BMA not necessarily uses all the available models.

Bayesian Model Averaging was one of the approaches proposed by Nowotarski et al. (2014) in their comparison of methods in forecasts combination. Their study is applied to the energetic market, as their dataset comprises four price time series from three major power markets (European Energy Exchange, Nord Pool and the Pennsylvania-New Jersey-Maryland Interconnection).

In a dataset with M models, 2^M possible forecast combinations can be considered. In this evaluation of the energy prices, the resulting forecast combination is given by:

$$\hat{P}_t^c = \sum_{l=1}^{2^M} w_{lt} \mathbb{E}(P_t | m_l, \theta_l) \quad (34)$$

Where m_l stands for a particular model combination from the 2^M that are possible, θ_l represents the collection of parameters required for that combination and w_{lt} is given by the likelihood of a price (P_t) in a certain combination (m_l) and data (D), according to the following formula:

$$w_{lt} = \frac{L(P_t | m_l, D) \rho(m_l)}{\sum_{j=1}^{2^M} L(P_t | m_j, D) \rho(m_j)} \quad (35)$$

After examining the Weighted Mean Absolute Error (WMAE) obtained through different combination schemes, results showed that, even though the BMA was not the least accurate method, it generally underperforms, giving greater errors than simple averaging.

This conclusion reinforces the Forecast Combination Puzzle, proving the most elaborated averaging methods are not always a better option than simple average, that considers every model as equal when combining their properties.

5. Model evaluation

The evaluation of the accuracy and performance of regression models is an essential step when these methods are used for forecasting and prediction.

There are different statistic parameters to compare the performance of the different models. They are very valuable when contrasting two models between them or one model to a random walk. They can be used to determine if the adopted regression method provides a good approximation to the data, or if the information provided by the model is worse than the figures obtained arbitrarily.

This project is going to focus on four evaluation and comparison parameters: Theil's U (iU), Coverage Rate (CR), Efficiency Rate (ER) and Euclidean Distance (MDE). These set of parameters represent a wide analysis on a certain prediction, as they focus on different aspects of the model that is being evaluated. While Theil's U and the Mean Distance Error measure the distance between the real interval values and the forecast, CR and ER judge the models depending on the percentage of interval covered by the prediction.

5.1. Theil's U

Maia and De Carvalho (2011) proposed the utilization of the Theil's U as a comparison tool, a statistic originally used in finance to measure accuracy and quality of forecasts. Its formula for IV data is given by:

$$U_I = \sqrt{\frac{\sum_{t=2}^M (x_t^L - \hat{x}_t^L)^2 + \sum_{t=2}^M (x_t^U - \hat{x}_t^U)^2}{\sum_{t=2}^M (x_t^L - x_{t-1}^L)^2 + \sum_{t=2}^M (x_t^U - x_{t-1}^U)^2}} \quad (36)$$

Where x^L and x^U represent the lower and upper limits of the real interval, respectively, and \hat{x}^L and \hat{x}^U are the estimations of the interval limits in a certain model.

The optimal values of the U_I when evaluating a model are the ones lower than 1: this means the model's performance is better than the random walk (an approximation that considers the value of the last period as the predictor for the current one). A U_I equal to 1 reveals a model with the same performance as random walk, while U_I greater than 1 represents a model that performs worse than a random walk.

When combining two models, this project is going to focus on two approaches: equal weights or optimal weights.

In the case of only two models and equal weights combination, the combined predictor will be calculated as:

$$\hat{Y}_C = 0.5 * \hat{Y}_1 + 0.5 * \hat{Y}_2 \quad (37)$$

On the other hand, the optimal combination of two different models can be determined by defining the optimal weight of each model. Being \hat{Y}_1 the estimator for model 1 and \hat{Y}_2 the estimator for model 2, the optimal combined estimator \hat{Y}_C would be:

$$\hat{Y}_C = w * \hat{Y}_1 + (1 - w) * \hat{Y}_2 \quad (38)$$

Where

$$w = \frac{(U_{I,2})^2 - ACEU_{I,1}U_{I,2}}{(U_{I,1})^2 + (U_{I,2})^2 - 2 * ACEU_{I,1}U_{I,2}} \quad (39)$$

$$ACEU_{I,1}U_{I,2} = \frac{\sum_{t=1}^T [(x_t^L - \hat{x}_{t,M1}^L)(x_t^L - \hat{x}_{t,M2}^L)]}{2T} + \frac{\sum_{t=1}^T [(x_t^U - \hat{x}_{t,M1}^U)(x_t^U - \hat{x}_{t,M2}^U)]}{2T} \quad (40)$$

Even though it may seem that a certain model has the lowest U_i and therefore it is not worth it to combine it with worse-performer models, model averaging and forecast combination are best long-term option when looking for accuracy and reliable predictions. The Theil's U value is calculated based on data from a limited amount of time, so it cannot assure the model is the best performer for existing every time period. Consequently, the other model can provide a better approach in future forecasts.

5.2. Coverage Rate

The coverage rate is another parameter for testing the goodness of fit. As explained in Maté (2015), the CR for a certain model is determined by calculating the average of all the CR for every estimation. The Coverage Rate for a single prediction is computed with the following formula:

$$CR_t = \frac{UV_t - LV_t}{x_t^U - x_t^L} \quad (41)$$

Where LV stands for Lower Value and UV stands for Upper Value:

$$LV_t = \max(x_t^L, \hat{x}_t^L) \quad (42)$$

$$UV_t = \min(x_t^U, \hat{x}_t^U) \quad (43)$$

In a model with k predictions, the CR for the model is calculated as:

$$CR = \frac{\sum_{i=1}^k CR_i}{k} \quad (44)$$

A prediction is more accurate the closer its CR is to 1. A Coverage Rate of 1 indicates that the forecasting interval covers the whole real interval (all the real values are contained in the model's interval). However, the Coverage Rate has to go along with the Efficiency Rate: both parameters need to be close to 1 to name an estimation as great.

5.3. Efficiency Rate

Similar to the Coverage Rate, the Efficiency Rate for a prediction is calculated as follows:

$$ER_t = \frac{UV_t - LV_t}{\hat{x}_t^U - \hat{x}_t^L} \quad (45)$$

Being the average of all the ER the ER of a model:

$$ER = \frac{\sum_{i=1}^k ER_i}{k} \quad (46)$$

The optimal value regarding the ER is also 1, as this indicates that a forecast is completely included in the actual interval. If both CR and ER are equal to 1, the forecast is excellent, as the predicted interval perfectly matches the real interval.

Figure 4 exemplifies the CR and the ER of a regression model. The regression method is the MinMax method and the IV data comprises forecasts of the S&P500 index during 2017. The values of the Coverage Rate and the Efficiency Rate are very similar, this meaning the rate of real interval not predicted by the forecast and the rate of the forecast that predicts a different thing are quite balanced.

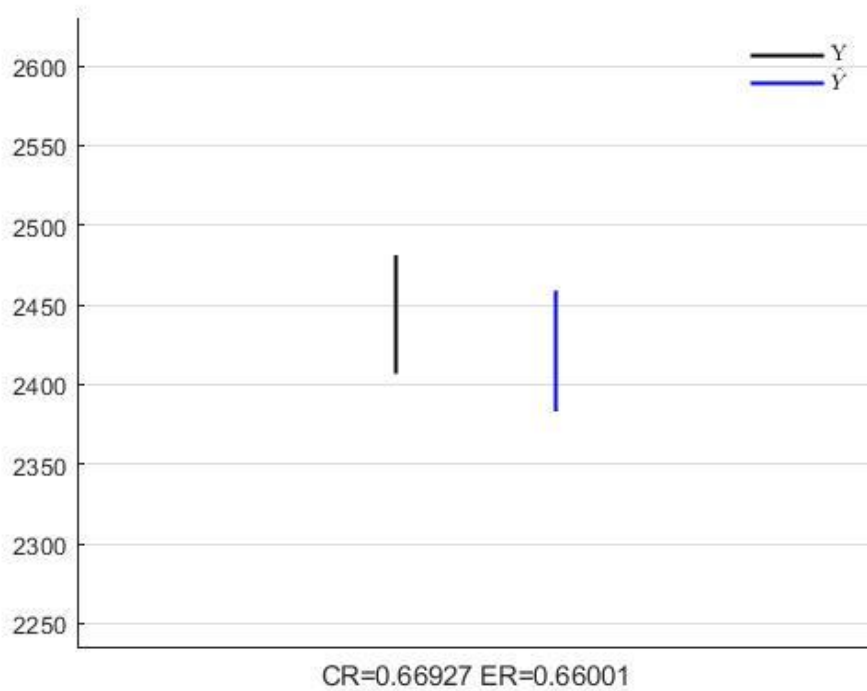


Figure 6: Real interval (Y) and forecasted interval (\hat{Y}) with Coverage Rate and Efficiency Rate. Source: Prepared by the author.

5.4. Mean Distance Error (MDE)

Mean Distance Error (MDE) method computes the distance between the forecast and the real data for every prediction available. Subsequently, these calculated distances are averaged to obtain the model's MDE value. This comprises another valuable way to evaluate the performance of a regression model.

Different distances can be used to reckon the MDE, such as Hausdorff distance, Euclidean distance or Ichino-Yaguchi distance.

This project calculates the MDE applying the Euclidean distance, following the definition in Arroyo, Espínola and Maté (2011):

$$MDE^2 = \frac{1}{2} \frac{\sum_{t=1}^T (d_E([x_t], [\hat{x}_t]))^2}{T} \quad (47)$$

Where the Euclidean distance is defined as:

$$d_E([x_t], [\hat{x}_t]) = \sqrt{(x_t^L - \hat{x}_t^L)^2 + (x_t^U - \hat{x}_t^U)^2} \quad (48)$$

The smaller the MDE and the Euclidean distance are, the better the forecast is, so the goal is minimizing the distance between records and predictions.

This estimator is affected by the magnitude of the data it is analyzing. For example, a value of 10 for the MDE is not the same for a model which is forecasting data ranging between 0.1 and 1, than for another model that deals with data between 100 and 1000. The higher the values of the data, the higher the admissible MDE is.

6. Analysis in financial markets

This section analyzes and predict various financial datasets with IV data and different trends (lateral, upward, downward). Different regression models (linear and non-linear) are applied to each data, trying to obtain the best possible prediction by combining models using model averaging.

6.1. Methodology

Given datasets and several regression models, the first step is analyzing each model's performance following the four criteria exposed in the previous section: Theil's U, Coverage Rate, Efficiency Rate and Mean Distance Error.

This is carried out using different MATLAB functions that are described hereunder. The code of the complete program is included in Appendix I.

- *UTheil_k*. Calculates the Theil's U for k models and gives back a single row array where each column is the U_i of a model. The inputs for this function are the real intervals (included in matrix *Y*) and the predictions from each method (in matrix *Yhat*). This parameter is calculated according the formula presented in section 4.1. This function allows to differentiate those models which perform better than the random walk ($U_i < 1$) and those which are worse performers than the reference method ($U_i > 1$).

```
function iutheil=IUTHEIL(Y,Yhat)
%Takes value of interval-valued data Y and its prediction Y^ according to
%a method used and calculates U-Theil value.

[r,~]=size(Y);

sum1=0;
sum2=0;
sum3=0;
sum4=0;

for i = 2:r
    sum1=sum1+(Y(i,1)-Yhat(i,1))^2;
    sum2=sum2+(Y(i,2)-Yhat(i,2))^2;
    sum3=sum3+(Y(i,1)-Y(i-1,1))^2;
    sum4=sum4+(Y(i,2)-Y(i-1,2))^2;
end

iutheil=sqrt((sum1+sum2)/(sum3+sum4));

end
```

Figure 7: *UTheil_k.m* Code.

- *CER_k*. Starting from matrix *Y* (the real intervals) and matrix *Yhat* (the forecasts from every model), this function returns the Coverage Rate and Efficiency Rate for every prediction, as well as the average CR and ER for each model. It also displays a graph contrasting the average real interval with the average interval of the forecasts, including the average values of the CR and the ER. This allows the comparison and visual evaluation of the predictions' accuracy.

```

function [CR,ER,CRAvg,ERAvg]=CER_k(Y,Yhat)

[r,~]=size(Y);
[s,t]=size(Yhat);
k=t/2;

CR=zeros(k,r);
ER=zeros(k,r);

for i = 1:k
    for j = 1:r
        LV=max(Y(j,1),Yhat(j,(2*i-1)));
        UV=min(Y(j,2),Yhat(j,2*i));

        if LV > UV
            CR(i,j)=0;
            ER(i,j)=0;
        else

            CR(i,j)=(UV-LV)/(Y(j,2)-Y(j,1));
            ER(i,j)=(UV-LV)/(Yhat(j,2*i)-Yhat(j,(2*i-1)));

        end
    end
end

CRAvg=mean(CR,2);
ERAvg=mean(ER,2);

end

```

Figure 8: *CER_k.m* Code

- *MDE_k*. Calculates the Mean Distance Error for all the models, returning a single column array where each row is the MDE of a model. This output is the average of the MDE from every prediction from the different methods.

```

function mde=MDE(Y,Yhat)

[r,~]=size(Y);
[s,t]=size(Yhat);

k=t/2;

sumDE=zeros(k,1);
mde=zeros(k,1);

for j=1:k
    for i=1:r
        sumDE(j)=sumDE(j)+sqrt((Y(i,1)-Yhat(i,(2*j-1)))^2+(Y(i,2)-Yhat(i,(2*j)))^2);
    end

    mde(j)=(sumDE(j))^2/(2*r);

end

end

```

Figure 9: MDE_k.m Code

Once the single models are evaluated, a global comparison is carried out.

- **REPRESENT_k**. Displays a figure containing the average interval for each model, depicting the similarities between real data and the forecasts. It is a way to graphically compare the models, detecting similarities and outliers. The graph consists of k models along X axis and the stock values along Y axis.

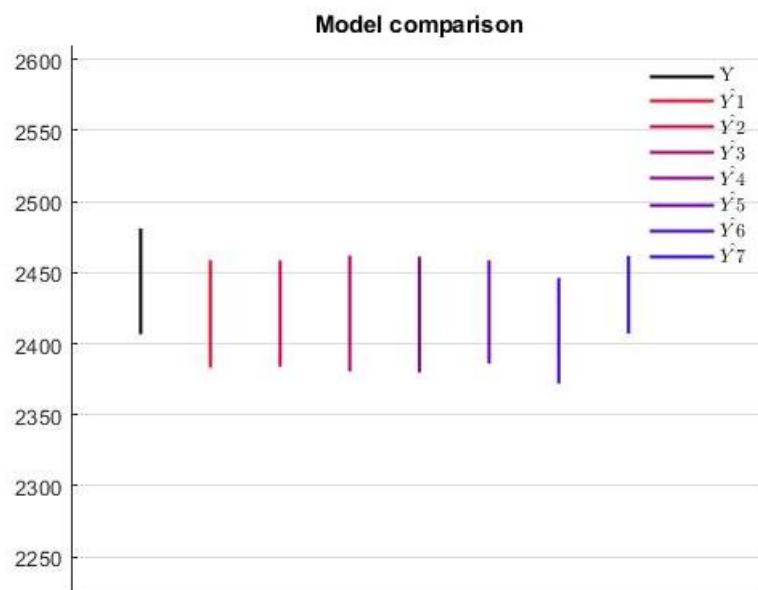


Figure 10: Example of Output from REPRESENT_k.m

```

function REPRESENT_k(Y,Yhat)

[r,~]=size(Y);
[~,p]=size(Yhat);

k=p/2;

Ylvavg=mean(Y(:,1));
Yuvavg=mean(Y(:,2));

for j=1:k

    Yhatlvavg(j)=mean(Yhat(:,(2*j-1)));
    Yhatuvavg(j)=mean(Yhat(:,2*j));

    % 7x1 array, each column is the average predictive
    %interval of the model

end

figure
axis([0,10,min(Yhatlvavg)-0.06*Yuvavg,max(Yhatuvavg)+0.06*Yuvavg]);
hold on;

%real
X1=linspace(1,1);
Y1=linspace(Ylvavg,Yuvavg);
plot(X1,Y1,'k','Linewidth',1.5);
hold on

for i=1:k

    plot((linspace(1+i,1+i)),(linspace(Yhatlvavg(i),Yhatuvavg(i))));
    plot('Linewidth',1.5,'Color',[(k-i+1)/k, 0, i/k]);
    hold on

end

```

Figure 11: REPRESENT_k.m Code.

- *DrawmultipleITSvdef.* Plots the interval time series from two selected models, comparing them with the real ITS. This is the most visual outcome to contrast the different regression models and their performance over the course of time.

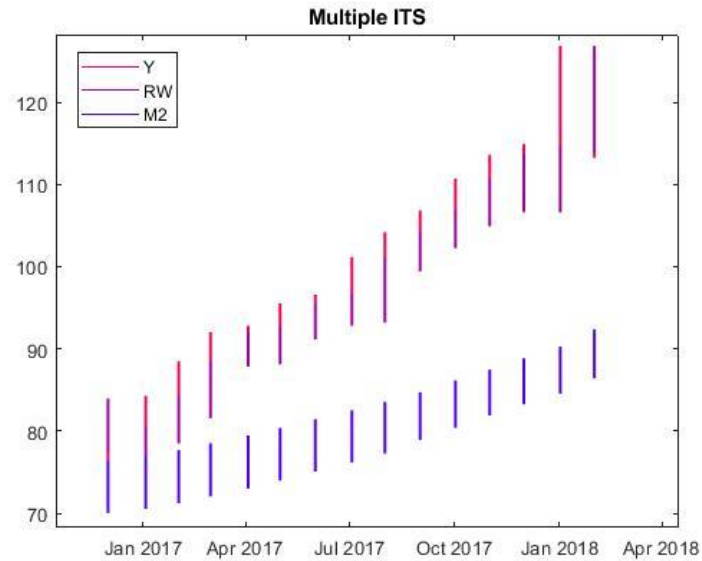


Figure 12: Example of Output from DrawmultipleITSvdef.m

- *OptimalW*. This function is used to calculate the optimal weights in a two-model combination. The output is a matrix w size $k \times k$, being k the number of models. Element ij represents the weight of model i when combined with model j , while element ji represents the weight of model j in the same combination.

```
function optimalw=OPTIMALW(Y,Yhat1,Yhat2,U1,U2)

[r,~]=size(Y);

sum1=0;
sum2=0;

for i = 2:r
    sum1=sum1+((Y(i,1)-Yhat1(i,1))*(Y(i,1)-Yhat2(i,1)));
    sum2=sum2+((Y(i,2)-Yhat1(i,2))*(Y(i,2)-Yhat2(i,2)));
end

ACE=(sum1+sum2)/(2*r);

optimalw=(U2^2-ACE)/(U1^2+U2^2-2*ACE);

end
```

Figure 13: OptimalW.m Code

- *Combination_k*. Returns the combined forecast averaging two different models, for both equal weights and optimal weights. The output is a matrix with two columns (lower and upper values) and as many rows as time periods are being forecasted. For optimal weights, it uses matrix w as the input for calculating the optimal combination.

Finally, the four estimators (U_i , CR, ER, MDE) are calculated for the combined model, displaying the results in a table. This last step allows the examination of model averaging, evaluating if this technique actually improves the properties of the forecasts and the way it does it.

	<u>U</u>	<u>CR</u>	<u>ER</u>	<u>MDE</u>
Method 1	0.77147	0.66927	0.66001	9255.2
Method 2	0.74836	0.72348	0.65409	8582.5
Equal weights	0.75722	0.69668	0.657	8856.1
Optimal Weights	0.75722	0.69668	0.657	8856.1

Figure 14: Example of Output from MATLAB's code.

There are two additional functions that define nonlinear regression methods. They are used in particular cases to obtain a better fit for certain datasets.

- *PolynomialFit*. Using matrix Y as the input, the function creates a polynomial fit for the selected database. It returns a graph illustrating the polynomial curve versus the real ITS. It also calculates the values of the four evaluators and gives the possibility of combining the new model with another existing one. The final output is a table containing the values of the parameters for the single model and its combination via equal weights and optimal weights.
- *ExponentialFit*. Using matrix Y as the input, the function creates an exponential fit for the IV data. The output is a graph representing the exponential curve versus the real ITS. The four evaluating parameters are calculated for the new model. It also allows the combination of the exponential model with another selected one. The final table includes the values of the parameters for the exponential model and its combination via equal weights and optimal weights.

```

function [Yhatp,U_p,CRAvg_p,ERavg_p,MDE_p]=PolynomialFit(Y)

[r,~]=size(Y);

x=linspace(1,r,r);
y_lv=transpose(Y(:,1));
y_uv=transpose(Y(:,2));

p_lv=polyfit(x,y_lv,4);
p_uv=polyfit(x,y_uv,4);

x1=linspace(1,r);
y1_lv=polyval(p_lv,x1);
y1_uv=polyval(p_uv,x1);
figure
plot(x,y_lv,'r')
hold on
plot(x,y_uv,'r')
hold on
plot(x1,y1_lv,'b')
hold on
plot(x1,y1_uv,'b')

```

Figure 15: PolynomialFit.m Code

```

function [Yhatexp,U_exp,CRAvg_exp,ERavg_exp,MDE_exp]=ExponentialFit(Y)

[r,~]=size(Y);

x=transpose(linspace(1,r,r));
y_lv=Y(:,1);
y_uv=Y(:,2);

exp_lv=fit(x,y_lv,'exp2');
exp_uv=fit(x,y_uv,'exp2');

x1=linspace(1,r);
figure
plot(exp_lv,x,y_lv,'b')
hold on
plot(exp_uv,x,y_uv,'b')
hold on

```

Figure 16: ExponentialFit.m Code

6.2. Case 1: S&P 500 index

The first model averaging compares six different regression models with the random walk:

1. MinMax Method (mM).
2. Center Method (CM).
3. Center and Range Method (CRM).
4. Bivariate Symbolic Regression Method (BSRM).
5. CIRM.
6. Random Walk (RW).
7. M Method (M).

The dataset comprises intervals representing the maximum and minimum S&P500 index every month during 2017 (12 rows) and its forecast with the different models. The comparison of the 7 models results in 21 different pairs, and 21 possible optimal estimators.

The index value in 2017 has an upward tendency. The S&P 500 index value in January is [2245.13; 2300.99], while in December it has augmented until [2605.52; 2694.97]. This is something that should be taken into account when drawing conclusions.

Results obtained throw light to different aspects.

In the first place, all 6 regression models used individually perform better than the Random Walk, as all their U_i range from 0.6194 to 0.7715 (smaller than 1). Table 1 presents the Theil's U obtained for each model. The best model according to this criterion is the M Method, a regression method specifically designed for IV data.

Table 3: Case 1. Theil's U . Source: Prepared by the author.

mM	CM	CRM	BSRM	CIRM	RW	M
0.7715	0.7655	0.7484	0.7639	0.7492	1.0000	0.6194

Figure 5 illustrates the average interval for each model. These intervals are computed by calculating the average of the lower and the upper values for each model. It is a visual representation of the Coverage Rate and the Efficiency Rate. Y represents the average real interval, while \hat{Y} represents the average interval for a certain model.

In this case, it stands out how all the seven models cover the low values of the real interval, however, any of the forecast models are able to predict the highest ones. The upward tendency of the model explains this: as regression

models are usually based on previous data, it is easy for them to predict values that have previously occurred. Notwithstanding, as the data keeps going up, the references used by the prediction models are no longer valuable.

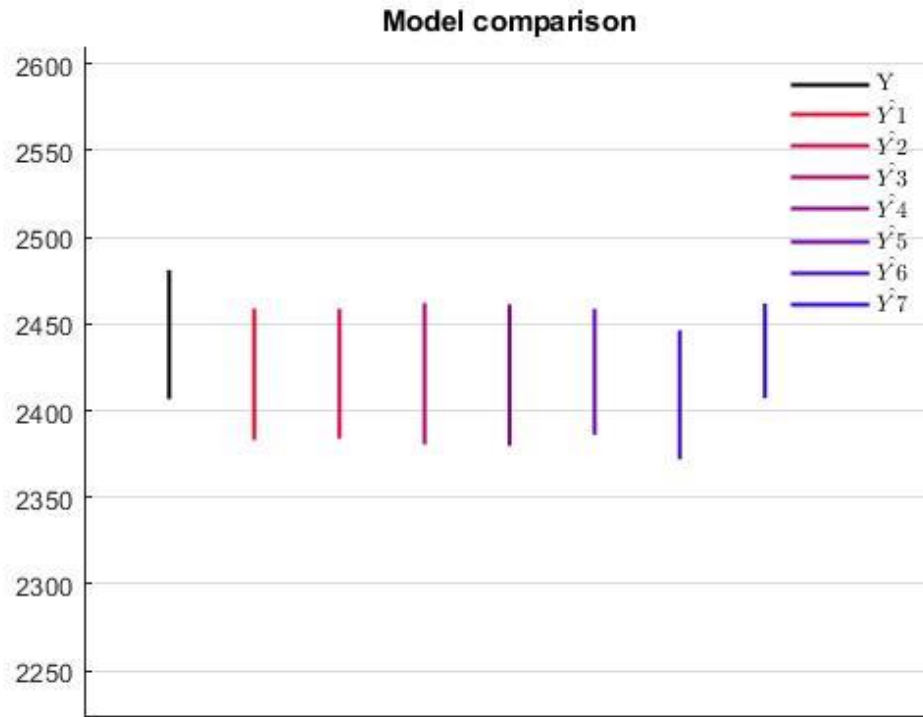


Figure 17: Case 1. Average interval from every model. Source: Prepared by the author.

Regarding optimal weights when combining methods, they are approximately 0.5 for all 21 combinations, as shown in Table 2. This means the best possible combination of two models is the one where both have the same importance. Given this, it can be stated that simple average beats weight averaging with Theil's U, at least for data with a noticeable upward tendency and prediction models based on regression.

This result is really pertinent if addressing the Forecast Combination Puzzle. For six different models that are good enough performers (better than the Random Walk), assigning more weight to one of them does not make a difference when combining forecasts in terms of improving the quality of the forecast. Notice that all the utilized models are regression-based models. These results could change when introducing methods based on tendency, such as the median or the average.

Table 4: Case 1. Optimal weights. Source: Prepared by the author

ω	M1	M2	M3	M4	M5	M6	M7
M1	0.5000	0.5000	0.5000	0.5000	0.5000	0.4999	0.5001
M2	0.5000	0.5000	0.5000	0.5000	0.5000	0.4999	0.5001
M3	0.5000	0.5000	0.5000	0.5000	0.5000	0.4999	0.5001
M4	0.5000	0.5000	0.5000	0.5000	0.5000	0.4999	0.5001
M5	0.5000	0.5000	0.5000	0.5000	0.5000	0.4999	0.5001
M6	0.5001	0.5001	0.5001	0.5001	0.5001	0.5000	0.5002
M7	0.4999	0.4999	0.4999	0.4999	0.4999	0.4998	0.5000

Random Walk (model 6), usually considered the benchmark for every other model, is not really accurate in this type of data, as its focus on the values from the last period hampers its ability to follow a rising tendency. This is why the chosen methods for a comparative analysis comprise the minMax method (model 1) and the M method (model 7).

Table 3 comprises the final table displayed in MATLAB comparing two chosen models, both with equal weights and optimal weights. It provides information about the four model evaluators, comparing the performance of the single models and the combination.

Table 5: Case 1. MinMax vs M Method. Source: Prepared by the author.

METHOD	IU	CR	ER	MDE
M1	0.7715	0.6693	0.6600	9255.2
M7	0.6194	0.6143	0.8261	5688.7
EW	0.6624	0.6537	0.7402	6471.6
OW	0.6625	0.6537	0.7402	6471.9

The difference between equal and optimal weights is negligible. This was the expected output due to the closeness of the optimal weights to equality.

There is a clear increase in overall performance when combining the two models. M method improves the Theil's U, Efficiency Rate and Mean Distance Error of the minMax method alone, while the minMax model contributes increasing the Coverage Rate of the M model.

A reason for such high MDE relies on the magnitude of the data values. The current stock values range between 2000 and 3000, so distances higher than 1000 are relatively not that high.

Figure 6 depicts the Interval Time Series for the data and its forecasts with models 1 and 7. The real intervals are plotted in red. The graph shows how, even though the models are capable of following the upward trend, the upper values of the last months are never reached by the forecasts.

It can also be seen that the lower values of the S&P500 index are always higher than the lower values provided by the forecasts. When there is a strong data tendency (in any direction), it is usual to find this kind of bias among regression models based on prior data.

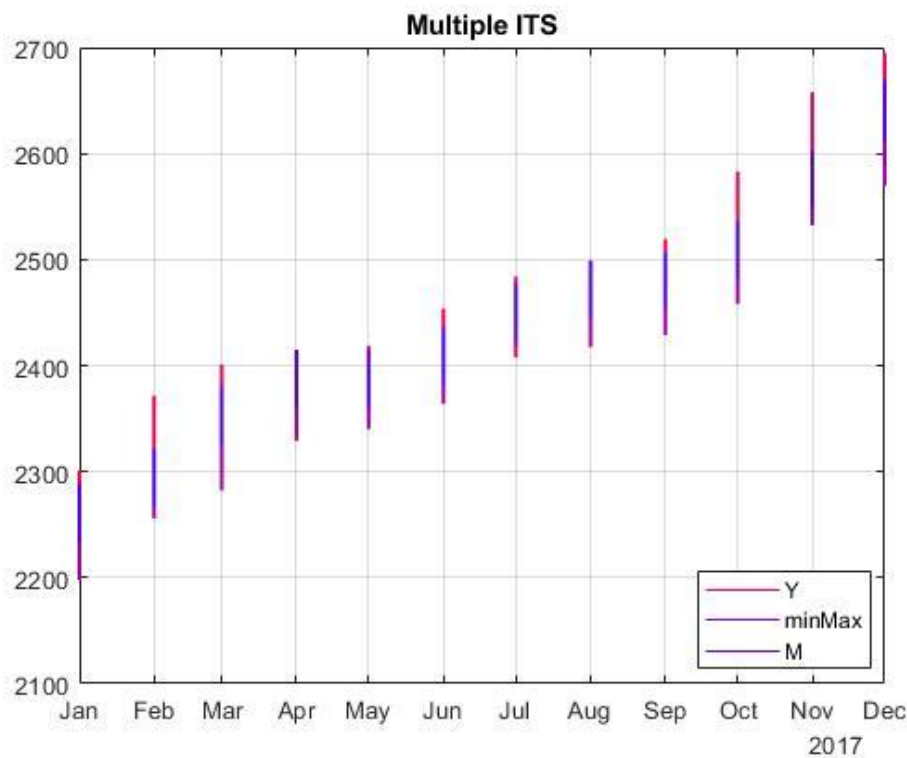


Figure 18: Case 1. Multiple ITS: real, minMax, M. Source: Prepared by the author.

Figure 7 offers a more detailed comparison of the models by representing the median and the average for each ITS. The third ITS (corresponding to M Method) is more similar to the data compared to the second one (minMax

Method). These similarities were detected through the data analysis and the calculation of the model evaluators.

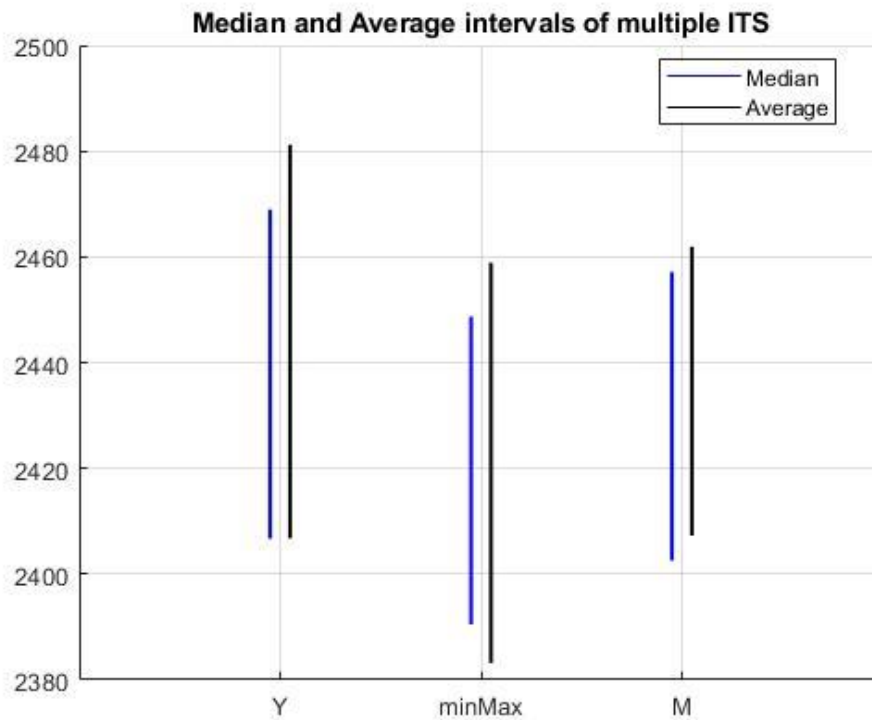


Figure 19: Case 1. Median and Average Intervals. Source: Prepared by the author.

The analysis of this first case can be concluded by highlighting the competence of simple averaging in a model bank with similar properties among the models. Without complex weights computations, model averaging that assigns equal weights to every model provides new models with improved capabilities and estimations for financial data.

This S&P500 case is an argument in favor of the Forecast Combination Puzzle: it is actually hard to beat the simple average approach when combining forecasts, even for optimal weights based on a reliable estimator such as the Theil's U.

Next cases will focus on non-linear models and their combination with linear models (or the Random Walk), and tendency-based models, such as the median or the average.

6.3. Case 2: General Electric

The second case study takes data from the stock values of General Electric. General Electric Company, also known as GE, is an American multinational based in Connecticut, whose fields of operation include energy, capital finance, technology infrastructure and consumer and industrial.

GE's stock values are included in the S&P 500 index. The data selected for the study includes the lowest and highest stock value of a month for the first seven months of 2017 (from January 2017 to July 2017).

The data presents a clear downward trend, starting with a value of [31.400; 31.835] in January and ending with [27.060; 27.590] in July. This company's declining tendency started in the middle of 2016 and keeps being the main direction of its value, even though it has settled in the last five months of 2019.

This data was evaluated using ten different models, including the Random Walk (M1). As shown in Table 4, the forecast depending on the average is the worst performing model, with a Theil's U much bigger than one (3.390). The rest of the models have a performance similar to Case 1: they all work better than the Random Walk without being excellent predictors.

Table 6: Case 2. Theil's U. Source: Prepared by the author.

M1	M2	M3	M4	M5	M6	M7	M8	M9	M10
1.000	3.390	0.771	0.738	0.722	0.843	0.727	0.867	0.734	0.722

Figure 8 shows how the second model is an obvious outlier compared to the rest. It has a zero coverage rate and a zero efficiency rate (the forecast does not match the interval at any point), and its U_i indicates it is better to use the Random Walk rather than this model. These are the signs of a bad forecast: graphically apart from the interval, high values of Theil's U and low values of CR and ER.

The rest of the models are better predictors, with reasonable distances to the real interval (MDE illustrated in Table 5). The closer the forecast is to the real interval (like in Model 3), the lower its MDE is. Model 2 shows a striking MDE value in comparison to the rest of the models.

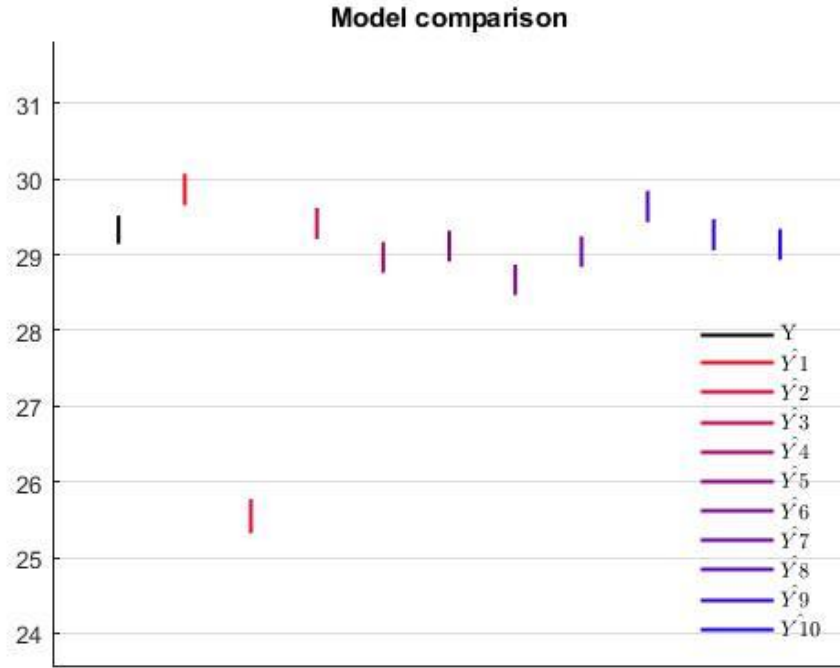


Figure 20: Case 2. Average interval from every model. Source: Prepared by the author.

It is noticeable how the magnitude of the Mean Distance Errors has considerably shrunk compared to Case 1. This does not mean that all the forecasts are much better and have less errors, but, as the magnitude of the real interval is smaller, the absolute value of the error decreases even though its relative value might be even higher.

Table 7: Case 2. MDE. Source: Prepared by the author.

M1	M2	M3	M4	M5	M6	M7	M8	M9	M10
4.77	100.35	3.65	4.41	3.97	5.90	4.08	3.90	3.79	3.94

The model chosen to compare against the RW is model 3, as it has the lowest MDE. In this case, optimal weights differ from the equality, so results are diverse depending on the type of averaging used (equal or optimal weights). Table 6 shows the comparison between these two models.

Table 8: Case 2. Random Walk vs M3. Source: Prepared by the author.

METHOD	IU	CR	ER	MDE
RW	1.0000	0.1448	0.0929	4.7741
M3	0.7707	0.1837	0.1838	3.6458
EW	0.8674	0.2227	0.1827	3.9008
OW	0.7344	0.0816	0.1114	3.7802

Theil's U is notably improved by the combination: the optimal combination attains a value of 0.7344, much lower than the U_i from both of the models. Regarding MDE, it is also improved through the averaging. However, CR and ER values are too low to consider any of these models a good approximation to this General Electric Data. Values that close to zero indicate that the forecasted interval barely overlaps the real one.

In order to improve the forecast of the IV data from General Electric during 2017, a non-linear regression model was introduced. Non-linear models are useful when the relationship between the dependent and independent variables is not a linear relationship. In this case, non-linear regression is required because the downward tendency does not follow a straight decreasing line: the stock value in March increases from [29.560; 29.810] to [29.820; 30.350], but goes down again in April to [29.745; 30.000].

Using the function *PolynomialFit*, a fourth-grade polynomial was created, with a much better fit than the rest of the models. Figure 9 shows the ITS of the data in red and the polynomial approach in blue.

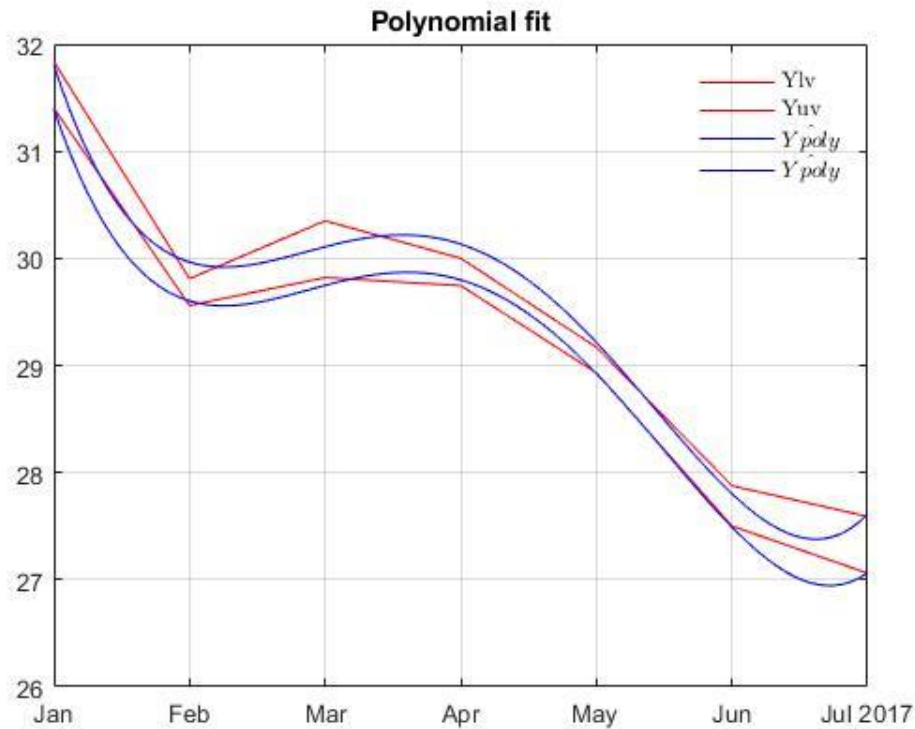


Figure 21: Case 2. Polynomial fit. Source: Prepared by the author.

This non-linear regression method is able to accurately follow the decreasing trend of the data, achieving a much better forecast than previous models. In order to make it suitable for IV data, the function *PolynomialFit* calculates the polynomial fit for both the upper and lower values of the intervals, similar to the minMax method in linear regression. Another possible approach (that could lead to new testing paths) consists in creating the polynomial from the interval centers and adjusting the radii with a linear regression.

According to Figure 10, that represents the average GE interval and the average interval from the polynomial regression model, the Coverage Rate and the Efficiency Rate are both higher than eighty percent, the greatest value achieved so far. Unlike the previously analyzed models, the polynomial approach successfully matches the real data from this case.

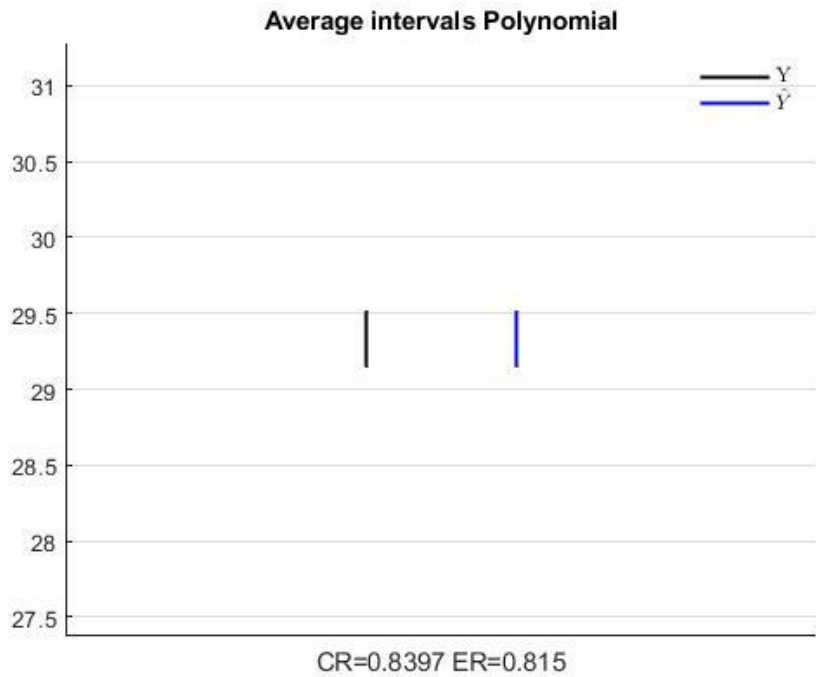


Figure 22: Case 2. Average interval: polynomial forecast. Source: Prepared by the author.

The following table shows the values of the estimators for the polynomial forecast, as well as its comparison with the Random Walk. Even though the polynomial approach was exceptional, these results prove that model averaging is able to upgrade even the greatest models.

Table 9: Case 2. Random Walk vs Polynomial Model. Source: Prepared by the author.

METHOD	IU	CR	ER	MDE
RW	1.0000	0.1448	0.0929	4.7741
Polynomial	0.0942	0.8397	0.8150	0.0386
EW	0.5212	0.3029	0.2760	1.3579
OW	0.0890	0.8507	0.8193	0.0340

It is noteworthy how, when calculating optimal weights for this case, the function returns a weight of approximately 1 for the polynomial model and approximately 0 for the random walk.

By combining the models with this optimal weight, Theil's U is reduced (obtaining a value extremely close to the optimum) while CR and ER are increased, covering a higher percentage of the real interval. MDE is almost insignificant: the distance error between the data and the forecast can be considered zero.

This analysis proves that, for a forecast model that clearly outstands the rest, equal weights are not the best option. It is true that simple averaging implies a relative enhancement from the Random Walk, but it is far from being the optimal solution when looking for the most accurate forecast.

Optimal weights, on the other side, improve the properties of the good forecast, achieving the desired close-to-perfection combined model. Different factors have probably influenced in beating the simple average.

In the first place, the performance gap between the polynomial model and the rest was significant. Models with similar properties are not compared anymore, so these differences among models can easily be omitted if simple averaging is used.

Secondly, this case presents a linear-nonlinear combination. The approach provided by each model addresses the forecast problem in a different way. Non-linear methods are more flexible and a good approach when data follows a non-linear tendency. This is a compelling reason to not equalizing both models but trying to find the optimal weight combination.

6.4. Case 3: Zardoya

The third dataset comprises the stock values of Zardoya Otis SA from February 2018 to April 2019 (15 months). This company was founded in New York in 1853 and constitutes the main elevator and escalator manufacturer in the world.

In the chosen period, the dataset has no clear trend: small rises are followed by short decreases and vice versa. The stock values of the company are therefore considered to have a lateral behavior. For this reason, it might be an interesting approach trying to predict this data with forecast models based on tendency.

The six models applied to the data are:

1. Random Walk (RW).
2. Aumann's Mean.
3. Hausdorff's Median.
4. Center-Radii Median.
5. Third Quartile.
6. First Quartile.

In general, the tendency-based models (Table) are not good performers when forecasting data without a tendency. The Random Walk (M1) behaves as the benchmark, with a Theil's U value equal to 1. All the tendency methods present a greater U_i value.

Table 10: Case 3. Theil's U. Source: Prepared by the author.

M1	M2	M3	M4	M5	M6
1.0000	1.7275	1.8791	1.8686	2.3760	1.6086

This is sustained by Figure 11, which illustrates the average interval for every model. Random Walk (\hat{Y}_1) is the closest forecast for Zardoya's stock values, as the rest of the models tend to overestimate the real values, leading to really low CR and ER values. The worst performer is Model 6 (first quartile) because not only it has the highest U_i value, but it is also located the furthest apart from the real interval (increasing the distance error).

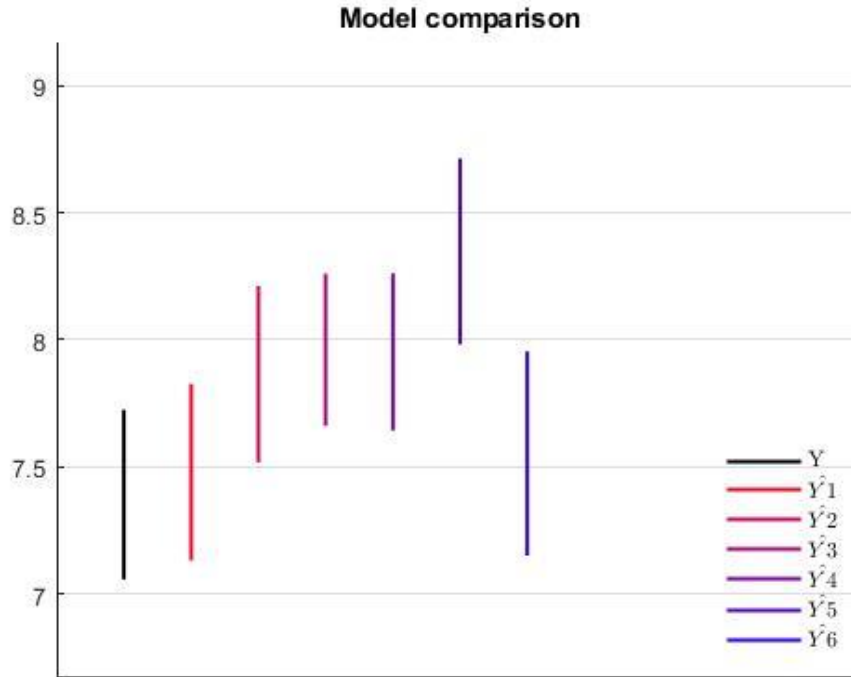


Figure 23: Case 3. Average interval from every model. Source: Prepared by the author.

Analyzing optimal weights among the models, results are not favorable for the simple average (Table 9). Weights assigned to each model vary from 0 to 1 for all the fifteen possible combinations between models.

Table 11: Case 3. Optimal weights. Source: Prepared by the author.

ω	M1	M2	M3	M4	M5	M6
M1	0.5000	0.7688	0.7975	0.7955	0.8678	0.7362
M2	0.2312	0.5000	0.5571	0.5534	0.7035	0.4526
M3	0.2025	0.4429	0.5000	0.4962	0.6540	0.3985
M4	0.2045	0.4466	0.5038	0.5000	0.6576	0.4021
M5	0.1322	0.2965	0.3460	0.3425	0.5000	0.2690
M6	0.2638	0.5474	0.6015	0.5979	0.7310	0.5000

There is a notable tendency to assigning more weight to the best performer model. In this case, Random Walk always gets weights higher than 0.5 when combined with the rest of the models (that have higher Theil's U value). This

is represented in the first row of Table 9, that shows the weight of the Random Walk when combined with the Aumann's Mean, the Hausdorff's Median, and so on.

Contrastingly, the model which has, on average, less weight assigned in the combination is model 5 (third quartile), the one with the least quality performance and worst predictive results.

In the light of these results, it would be interesting to compare the combination of the Random Walk with the mean (Model 2) and with the median (Model 3) separately, along with the combination of the three models using equal weights.

First of all, models 1 and 2 are combined through equal weights and optimal weights. The values for the model estimators are presented in Table 10. The differences between the equal and the optimal combination are remarkable: optimal weights combination achieves a model that performs better than the RW ($U_i < 1$), covers a greater percentage of the interval and has a lower error distance.

In this case, simple average has been clearly outperformed by the optimal weights combination. It is probably due to the fact that the models being compared (Random Walk and Aumann's Mean) are very dissimilar. While Random Walk only considers the interval value from the last period, the Mean is a tendency method that considers the data from the previous two years to make a forecast.

Table 12: Case 3. Random Walk vs M2. Source: Prepared by the author.

METHOD	IU	CR	ER	MDE
RW	1.0000	0.6435	0.6184	2.0371
M2	1.7275	0.4646	0.3649	6.6566
EW	1.1267	0.6475	0.5628	2.3476
OW	0.9721	0.6850	0.6371	1.6857

Results from combining the Random Walk with the Hausdorff's Median, included in Table 11, are almost identical: the combined model is only improved when using optimal weights.

Table 13: Case 3. Random Walk vs M3. Source: Prepared by the author.

METHOD	IU	CR	ER	MDE
RW	1.0000	0.6435	0.6184	2.0371
M3	1.8791	0.4026	0.3632	7.7958
EW	1.1795	0.5903	0.5459	2.6328
OW	0.9735	0.6654	0.6387	1.7193

Figure 12 illustrates the different behaviors of the two models: while the Random Walk follows the data tendency (a time period behind), Aumann's Mean stagnates around the same values. This implies the tendency-based method is not capable of adapting to the data fluctuations in the absence of a clear trend, so it keeps forecasting in the same limited range.

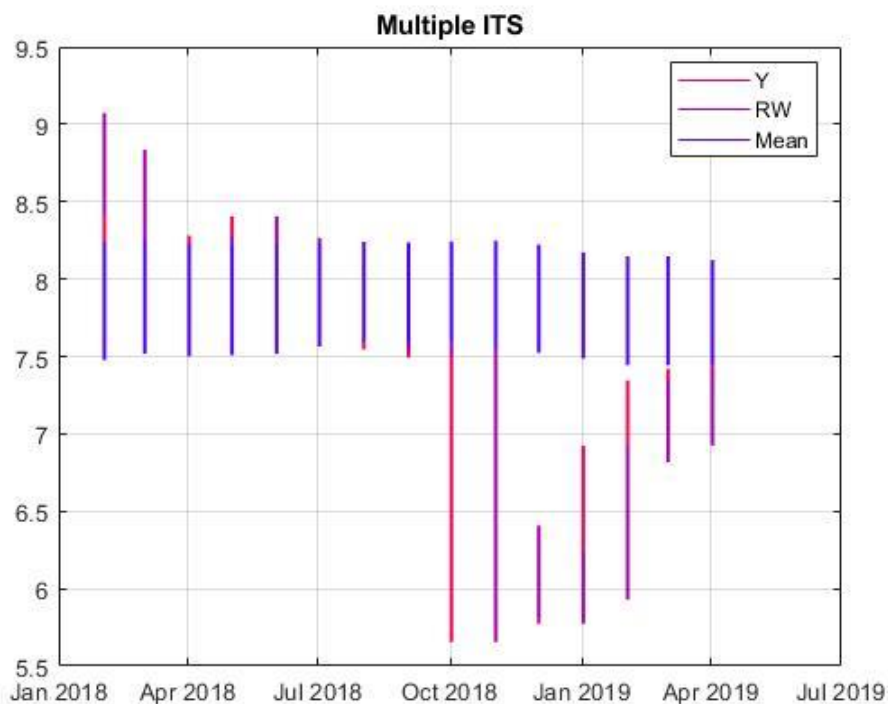


Figure 24: Case 3. Multiple ITS: real, Random Walk, Aumann's Mean. Source: Prepared by the author.

The behavior of the ITS forecasted by the Hausdorff's Median follows the same patterns. Represented in Figure 13, the third model analyzed for Zardoya's case (plotted in blue) is not able to properly forecast the interval time series established by the stock values.

While the average value of the real ITS ranges in the interval [7.058; 7.725], the average interval of the forecast stays at [7.662; 8.260]. Notice that the lower value of the forecast is practically the same as the upper value of the real interval. However, by combining the Median model with the Random Walk, an average interval of [7.240; 7.914] is obtained for the forecast. This is a clear example of how model averaging can help improving a forecast without completely giving up a model that initially looked like a bad performer.

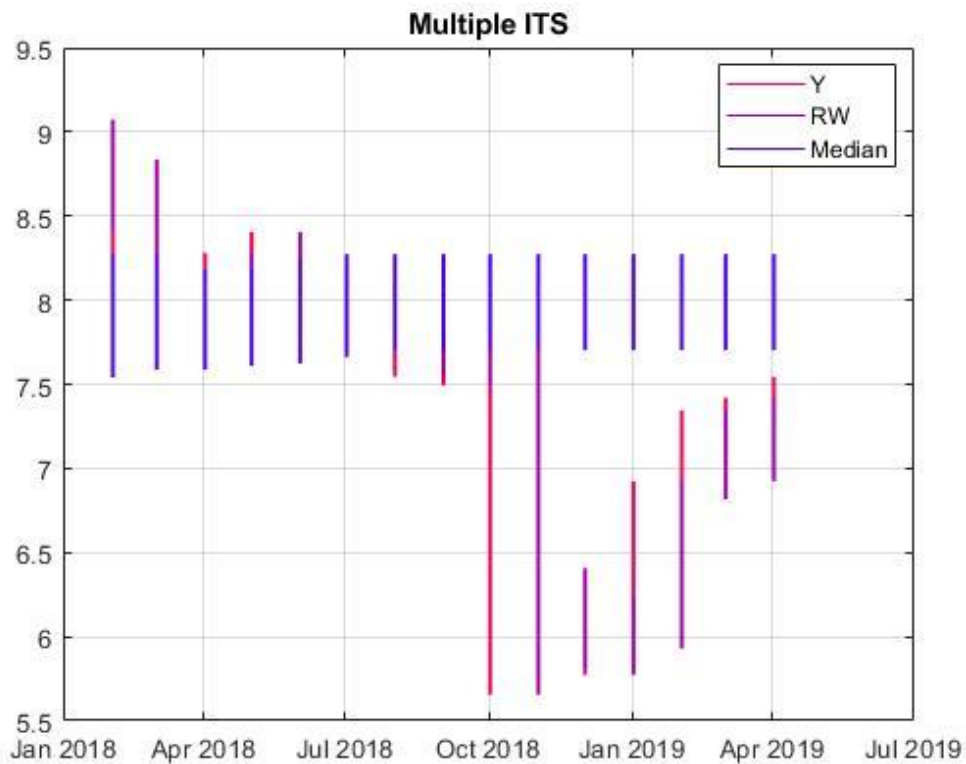


Figure 25. Case 3. Multiple ITS: real, Random Walk, Hausdorff's Median. Source: Prepared by the author.

The last test made for this data is a combination of the three models together: Random Walk, the Aumann's Mean and the Hausdorff's Median. This combination was implemented assigning equal weights to every method and the model evaluators were calculated for the new combination (Theil's U, Coverage Rate, Efficiency Rate, Mean Distance Error). Obtained results are presented in the following table.

Table 14: Case 3. Combination of three models: Random Walk, Aumann's Mean, Hausdorff's Median. Source: Prepared by the author.

IU	CR	ER	MDE
1.3376	0.5417	0.4666	3.6631

The outcomes of combining three methods instead of two using equal weights are not satisfactory. The performance of the combined model, according to the Theil's U criterion, is worse than the simple random walk, so it is pointless utilizing this new model to forecast interval data when looking for accuracy and legitimacy in the predictions. Coverage and efficiency rates are also reduced from the original RW forecast: the average predicted interval is now less similar to the real one. In addition, error distance is higher, so the four criteria agree in the bad quality of the new combination. This relationship between Zardoya's average interval and the combined model's average interval is plotted in Figure 14.

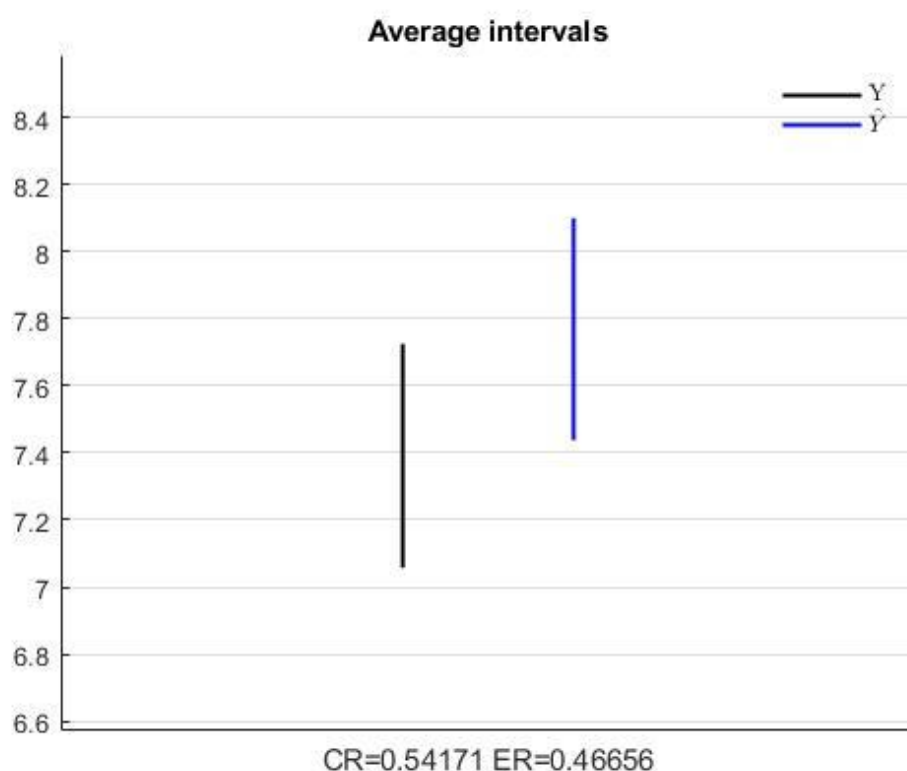


Figure 26: Case 3. Average interval: combination of RW, Mean and Median. Source: Prepared by the author.

This case proves that, even though tendency models are hardly useful with data without a trend, model averaging is a helpful tool to improve the forecasts and their properties. Optimal weights combination has demonstrated to be the

best averaging option, making a noticeable improvement in results and beating simple averaging.

In an scenario where simple average has been beaten, it has been verified that it does not matter how many models are being combined. Developing a formula to assign optimal weights to more than two models could be a new MATLAB function worthy of being developed.

6.5. Case 4: Visa

The fourth case study focuses on data with a steep upward tendency. The dataset comprises Visa's stock values along 15 months, from December 2016 to February 2018.

Visa International Service Association is a joint venture of 21000 financial institutions offering Visa products. Founded in 1958 and based in San Francisco, is one of the leader payment brands in the world. Although it plunged during the last four months of 2018, Visa Inc has maintained its crescent trend since 2014 until today.

The dataset was forecasted using the same tendency models as in case 3: Aumman's Mean, Hausdorff's Median, Center-Radii Median, and third and first quartiles. Visa's stock values present a defined trend, so it is interesting to analyze to what extent this makes a difference regarding the forecast performance of the model.

Figure 15 represents the average interval of the predictions versus the average interval of the real data. Along with the information on Table 13 about the Theil's U of the models, it can be stated that the selected models do not work at all with the selected dataset.

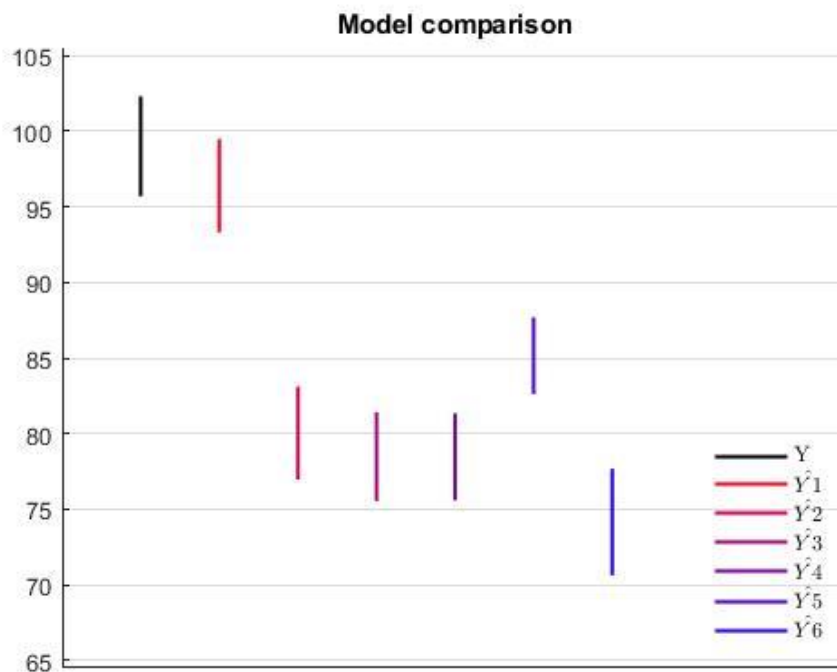


Figure 27: Case 4. Average interval from every model. Source: Prepared by the author.

Average forecasting intervals are about twenty percent lower than the average real interval, with zero coverage and efficiency (the real values are not

contained in the predictions, neither are the predictions in the real interval). As the tendency models make predictions given the previous events, they are not suitable for predicting a company with a severe growing trend, as the stock values grow too fast for the models to adapt.

Table 15: Case 4. Theil' U. Source: Prepared by the author.

M1	M2	M3	M4	M5	M6
1.0000	5.3323	5.9138	5.9192	4.0233	6.8728

The U_i values of the models indicate that the random walk is a better approximation to the data than any of the previously exposed methods. Therefore, it is necessary to find a different approach to forecast this increasing dataset.

Similarly to case 2, the chosen solution implies addressing a non-linear method. Pursuing the best fit for the increasing data, an exponential model was created with the function *ExponentialFit*.

This MATLAB function creates a regression model based on an exponential curve from the initial Visa data. As in the polynomial approach, the exponential curve is constructed simultaneously for the lower and upper values of the interval. Again, another option would be creating a fit for the interval centers and subsequently, adjusting the model for the interval radius. The function also returns the values of the four model evaluators calculated for this new exponential model.

The performance of the forecast is clearly improved compared to the previously analyzed tendency models, specially the Coverage Rate and Efficiency Rate. Figure 16 illustrates the data ITS in blue and the exponential curve in red. The exponential fit has, as expected, a much more quality performance ability, being capable of forecasting the upward tendency of the data.

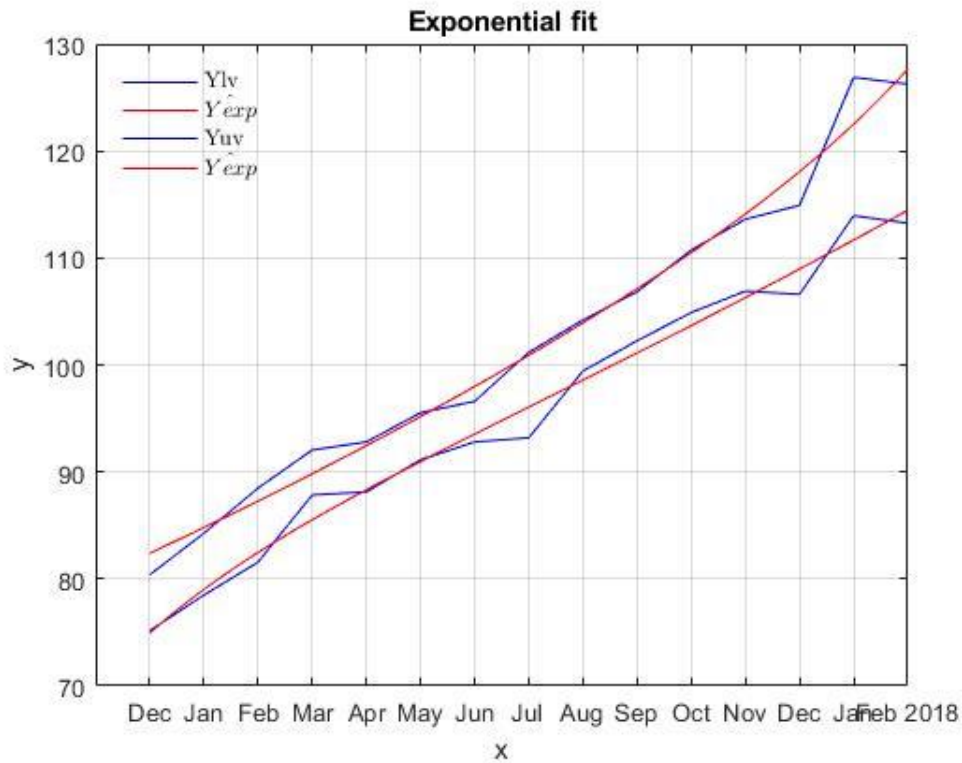


Figure 28: Case 4. Exponential fit. Source: Prepared by the author.

This is another case where non-linear regression methods are the solution to a problem with a more complex relationship between the variables that handicaps the forecast. Figure 17 depicts the average interval from the data and the average forecasted interval with the exponential method.

The matching is incredibly accurate: the prediction covers over eighty-three percent of the real data ($CR=0.8322$) and more than eighty-two percent of the forecast is contained in the real interval ($ER=0.8224$).

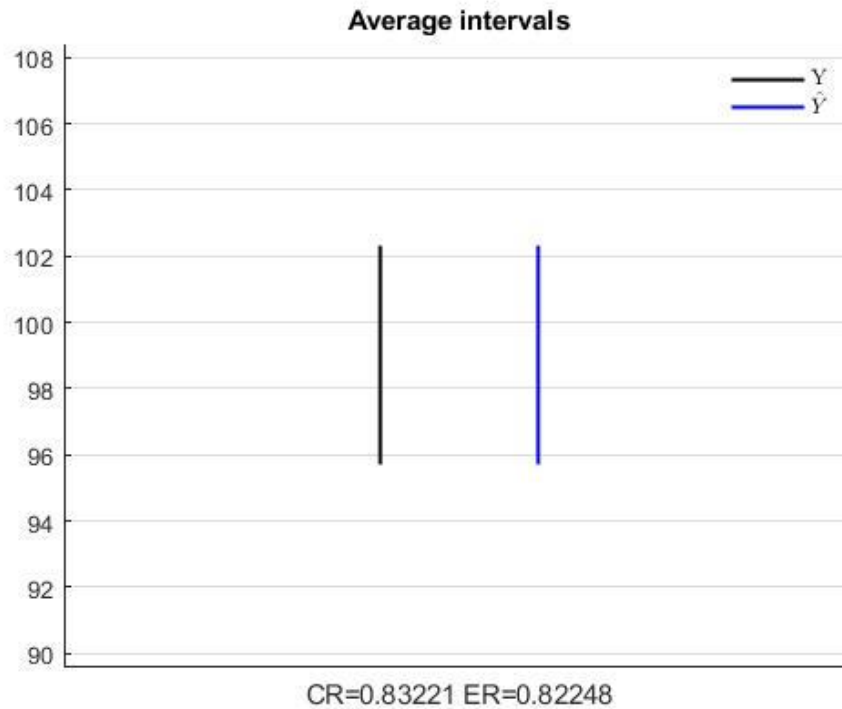


Figure 29: Case 4. Average interval: Exponential Fit. Source: Prepared by the author.

Following the model averaging tendency, this new model was combined with the random walk looking forward the potential improvement of the model forecast abilities. Table 14 shows the results of this combination for optimal and equal weights.

Table 16: Case 4. Random Walk vs Exponential Fit. Source: Prepared by the author.

METHOD	IU	CR	ER	MDE
RW	1.0000	0.4902	0.5315	156.2200
Exponential	0.4016	0.8322	0.8225	25.2780
EW	0.6344	0.7018	0.7164	60.2200
OW	0.6861	0.6739	0.6932	70.4870

By combining the Random Walk and the exponential curve, the properties obtained through the non-linear method wane. Either by using equal weights or optimal weights, the combined model increases its U_1 and its MDE, and decreases it CR and ER compared to the exponential fit.

There is a discrepancy with the previous cases regarding the supremacy of optimal weights in model averaging: in this example, simple averaging provides better properties for the predictive model than the optimal weights combination. The difference is not extremely significant, as the weight assigned to the first model over the second one was not too far from the equilibrium (0.5763). Anyhow, it cannot be ignored that equal weights, even they also worsen the exponential model, are not as harmful as optimal weights in this combination.

The optimal forecast for Visa's stock values was achieved through the non-linear exponential curve. A low U_1 value and high coverage and efficiency rates indicate that this is a better approach than the Random Walk. If the tendency of this company was continuing to be this way, the exponential model would be the safest bet for a good forecast.

7. Conclusions

This project has focused on model averaging applied to financial interval valued data, evaluating diverse models according to different pre-determined criteria.

Different results have been obtained and compared with the theoretical framework, shedding light on various aspects of the performance of linear and non-linear forecasting methods for IV data.

In the first place, the Forecast Combination Puzzle was discussed over different cases, to demonstrate if simple averaging is really that hard to beat. Results showed that simple average is the most straightforward averaging method (the simplest and most effective) as long as the models that are being combined present similar characteristics. The case of S&P500 index is very illustrative: for a set of linear regression methods with comparable performance, calculating optimal weights does not improve the model combination compared to assigning equal weights to each model.

Nevertheless, in a situation where a model outstands the rest (for example, the polynomial approach in the General Electric case), simple averaging stops being the most reliable method for model averaging. Models with a noticeable improvement in the forecast demand more weight in the combination in order to enhance their properties in the final combined model. In these cases, the calculation of optimal weights based on the Theil's U is a good approach, as it prioritizes the model with a lower U_i value, this being the model with the better performance according to this criterion.

Regarding possible future studies, it would be interesting to include new ways of defining optimal weights in a model combination. This study was limited by the Theil's U as the only approach to optimal weights estimation, while finding different formulas and methods to assign weights to models depending on their characteristics could expand the project and the state of the art of model averaging.

These potential formulas could be based on model evaluators, such as the coverage rate or the efficiency rate included in this project, or in model properties like the variance or the standard deviation. Comparing different ways of assigning weights could also call into question the accuracy of the current method and whether simple averaging is actually better than optimal weights or only better than this particular calculation of optimal weights.

Tendency-based forecasting methods and their utility in predicting certain kinds of data are another remarkable conclusion in this project. Predictive models based on tendency parameters, such as the average, the median or the quartiles, have proven to be less effective and accurate than other methods for datasets with lateral behaviors or steep upward trends.

Applying this kind of methods to data without a trend (for example, Zardoya values during 2018) has resulted in models unable to improve the Random Walk prediction. The absence of a tendency in the data hinders the model's ability to exactly predict future movements of the stock values.

On the other hand, datasets with rising values, such as Visa during 2017, entail a broad challenge for tendency-based methods due to their tendency to drift apart from previous values. The faster the values grow, the harder it is for the model to adapt to the new values and make accurate predictions.

A possible approach that could be made on further research would be studying the effect of methods based on tendency on data that present declining trends. The difference with the cases that have been studied relies on the direction of the data values: on a downward trend, predictions could be less likely to stagnate around certain values, as the reference figures that reach the highest values have already occurred.

Focusing on methods based on tendency, the research implemented in this project can conclude that the Aumman's mean and the Hausssdorff's median are similar enough on performance and characteristics to be considered equivalent methods. The analysis carried out with Zardoya's dataset resulted in differences smaller than five percent in the values of the model evaluators (Theil's U, CR, ER). The similarities on the forecasting ability of both methods could be a starting point for further exploration on how similar models could be substituted by each other on a combination of models depending on the available data and the desired parameters of the model.

Looking for more accurate forecasts and the increase of the coverage percentage of the interval, it was noticed that non-linear regression methods present higher coverage and efficiency rates. Studies like the one from Lima Neto and de Carvalho (2017) emphasize the importance of these methods when working with IV data. This project attempted to apply and evaluate models with non-linear relationships between the variables, such as polynomial or exponential fits. They have proven to be more accurate and to have better forecast abilities than linear methods on datasets with strong upward and downward tendencies.

Polynomial-based forecasting model demonstrated to be particularly accurate adjusting to data with a descending tendency and some fluctuations over time. Getting rid of the linear bindings when defining the model's equations allowed a better fitting to the data behavior, achieving a predictive model with an extraordinary performance.

Regarding exponential fit, it was a useful tool for interval-valued data with a tendency to increase. Compared to linear tendency-based regression methods, the predictive value was strongly improved.

Both polynomial and exponential fits have been applied to IV data using the upper and lower values of the intervals as the references for the regression model. However, substituting these parameters by the interval centers or the interval range may lead to new approaches that provide even better fits and are more suitable to create regression methods.

As a recommendation for future studies, non-linear methods seem to be the key to accurate and reliable forecasting in financial markets. The development of a tool for model averaging among several non-linear models might be the optimal solution to achieve the best possible predictions for interval-valued data.

Finally, all the MATLAB functions programmed to develop this project are the starting point to create an app using MATLAB's App Designer. This could turn the outcome into a visual and user-friendly interface, allowing to choose among various models and averaging options to forecast the desired data. It could also facilitate the comparison between models and averaging methods, displaying the pertinent graphs and figures.

References

- Arroyo, J., & Maté, C. (2006). Introducing Interval Time Series: Accuracy Measures. *COMPSTAT 2006, proceedings in computational statistics*, 1139-1146.
- Arroyo, J., Espínola, R., & Maté, C. (2011). Different approaches to forecast interval time series: a comparison in finance. *Computational Economics*, 37(2), 161-191.
- Arroyo, J., González-Rivera, G., & Maté, C. (2010). Forecasting with interval and histogram data. Some financial applications. *Handbook of empirical economics and finance*, 247-280.
- Bates, J. M., & Granger, C. W. (1969). The Combination of Forecasts. *The OR Society*, 20(4), 451-468.
- Bertrand, P., & Goupil, F. (2000). Descriptive statistics for symbolic data. In *Analysis of symbolic data* (pp. 106-124). Springer.
- Billard, L., & Diday, E. (2000). Regression analysis for interval-valued data. *Data Analysis, Classification and Related Methods* (pp. 369-374). Berlin, Heidelberg: Springer.
- Billard, L., & Diday, E. (2002). Symbolic regression analysis. *Classification, Clustering, and Data Analysis* (pp. 281-288). Berlin, Heidelberg: Springer.
- Clemen, R., & Winkler, R. (1986). Combining Economic Forecasts. *Journal of Business & Economic Statistics*, 4(1), 39-46.
- García-Ascanio, C., & Maté, C. (2010). Electric power demand forecasting using interval time series: A comparison between VAR and iMLP. *Energy Policy*, 38(2), 715-725.
- Giordani, P. (2011). Linear regression analysis for interval-valued data based on the Lasso technique. *Proceeding of 58th World Statistical Congress*, (pp. 5576-5581). Dublin.
- Hao, P., & Guo, J. (2017). Constrained center and range joint model for interval-valued symbolic data regression. *Computational Statistics & Data Analysis*, 116, 106-138.
- Kapetanios, G., Labhard, V., & Price, S. (2008). Forecasting using Bayesian and information-theoretic model averaging: an application to UK inflation. *Journal of Business & Economic Statistics*, 26(1), 33-41.
- Lima Neto, E. d., & de Carvalho, F. d. (2008). Centre and range method for fitting a linear regression model to symbolic interval data. *Computational Statistics & Data Analysis*, 52(3), 1500-1515.

- Lima Neto, E. d., & de Carvalho, F. d. (2010). Constrained linear regression models for symbolic interval-valued variables. *Computational Statistics & Data Analysis*, 54(2), 333-347.
- Lima Neto, E. d., & de Carvalho, F. d. (2017). Nonlinear regression applied to interval-valued data. *Pattern Analysis and Applications*, 20(3), 809-824.
- Lin, W., & González-Rivera, G. (2016). Interval-valued Time Series Models: Estimation based on order statistics exploring the Agriculture Marketing Service data. *Computational Statistics & Data Analysis*, 100, 694-711.
- Maceda Dal-Re, P. (2018, July). Basic analytic system for interval valued data. Madrid: Universidad Pontificia Comillas.
- Maia, A. L., & de Carvalho, F. A. (2011). Holt's exponential smoothing and neural network models for forecasting interval-valued time series. *International Journal of Forecasting*, 27, 740-759.
- Maté, C. (2012). El análisis de intervalos. Aplicaciones en la ingeniería. *Anales de mecánica y electricidad*, 89, 20-27.
- Maté, C. (2015). Combining interval time series forecasts. An agenda for future research. *1st International Symposium on Interval Data Modelling: Theory and Applications*.
- Maté, C. (2019). Descriptive statistics measures for interval-valued datasets. *Unpublished manuscript*. Madrid: Comillas Pontifical University.
- Moore, R. E. (1966). *Interval analysis*. Englewood Cliffs, NJ: Prentice-Hall.
- Moral Benito, E. (2015). Model averaging in economics: an overview. *Journal of Economic Surveys*, 29(1), 46-75.
- Norhomme-Fraiture, M., & Brito, P. (2011). Far beyond the classical data models: symbolic data analysis. *Statistical Analysis and Data Mining: the ASA Data Science Journal*, 4(2), 157-170.
- Nowotarski, J., Raviv, E., Trück, S., & Weron, R. (2014). An empirical comparison of alternative schemes for combining electricity spot price forecasts. *Energy Economics*, 46, 395-412.
- San Roque, A. M., Maté, C., Arroyo, J., & Sarabia, Á. (2007). iMLP: Applying Multi-Layer Perceptrons to Interval-Valued Data. *Neural Processing Letters*, 25(2), 157-169.
- Sinova, B., Colubi, A., & González-Rodríguez, G. (2012). Interval arithmetic-based simple linear regression between interval data: Discussion and sensitivity analysis on the choice of the metric. *Information Sciences*, 199, 109-124.
- Smith, J., & Wallis, K. (2009). A simple explanation of the forecast combination puzzle. *Oxford Bulletin of Economics and Statistics*, 71(3), 331-355.

- Souza, L. C., Souza, R. M., Amaral, G., & Silva Filho, T. (2017). A parametrized approach for linear regression of interval data. *Knowledge-Based Systems*, 131, 149-159.
- Steel, M. (2019). *Model averaging and its use in economics*. Department of Statistics, University of Warwick.

Appendix I: MATLAB code

```
clc; clear all; close all;

[data, dates] = xlsread('SP500_2017forecasts.xlsx'); %data
extraction

k=7; %number of models
c=nchoosek(k,2); %number of combinations

dates = dates(:,1);

Y = [data(:,1) data(:,2)];

Yhat1 = [data(:,3) data(:,4)];
Yhat2 = [data(:,5) data(:,6)];
Yhat3 = [data(:,7) data(:,8)];
Yhat4 = [data(:,9) data(:,10)];
Yhat5 = [data(:,11) data(:,12)];
Yhat6 = [data(:,13) data(:,14)];
Yhat7 = [data(:,15) data(:,16)];
%Yhat8 = [data(:,17) data(:,18)];
%Yhat9 = [data(:,19) data(:,20)];
%Yhat10 = [data(:,21) data(:,22)];

Yhat=[Yhat1 Yhat2 Yhat3 Yhat4 Yhat5 Yhat6 Yhat7];

disp(Y)
%disp(Yhat)
[r,~]=size(Y);

Yavg=mean(Y);

sum=zeros(k,4); %initialize

%% general
for j = 1:k

    for i=2:r
        sum(j,1)=sum(j,1)+(Y(i,1)-Yhat(i,(2*j-1)))^2;
        sum(j,2)=sum(j,2)+(Y(i,2)-Yhat(i,(2*j)))^2;
        sum(j,3)=sum(j,3)+(Y(i,1)-Y(i-1,1))^2;
        sum(j,4)=sum(j,4)+(Y(i,2)-Y(i-1,2))^2;
    end

    U(j)=sqrt((sum(j,1)+sum(j,2))/(sum(j,3)+sum(j,4)));

end

display(U)

[CR,ER,CRAvg,ERAvg]=CER_k(Y,Yhat);
```

```

        %Y is a rx2 matrix, Yhat is a rx14 matrix, CR and ER are kxr
matrix

MDEmatrix=MDE_k(Y,Yhat);
    % MDE is a kx1 matrix, each row is the MDE for a model

REPRESENT_k(Y,Yhat)

%% equal weights

%choose the models that are about to be compared
a=1; % first model
b=3; % second model

eqYc=Combination_k(Yhat(:,(2*a-1):(2*a)),Yhat(:,(2*b-
1):(2*b)),0.5,0.5)

Ueq=IUTHEIL(Y,eqYc)
[eqCR,eqER,eqCRavg,eqERavg]=CER(Y,eqYc)
MDEeq=MDE(Y,eqYc)

DrawmultipleITSvdef('Zardoya_2018forecasts.xlsx',1,[2 2+2*a
2+2*b],[3 3+2*a 3+2*b],1)

%% optimal weights

sumal=zeros(k,k);
suma2=zeros(k,k);

for m=1:k
    for n=1:k

        w(m,n)=OptimalW(Y,Yhat(:,(2*m-1):(2*m)),Yhat(:,(2*n-
1):(2*n)),U(m),U(n));

        end

    end

end

display(w)

optYc=Combination_k(Yhat(:,(2*a-1):(2*a)),Yhat(:,(2*b-
1):(2*b)),w(a,b),w(b,a))

Uopt=IUTHEIL(Y,optYc)
[optCR,optER,optCRavg,optERavg]=CER(Y,optYc)
MDEopt=MDE(Y,optYc)

%% display in table

varU=[U(a); U(b); Ueq; Uopt];
varCR=[CRavg(a); CRavg(b); eqCRavg; optCRavg];

```



```
varER=[ERavg(a); ERavg(b); eqERavg; optERavg];  
varMDE=[MDEmatrix(a); MDEmatrix(b); MDEeq; MDEopt];  
  
T=table(varU,varCR,varER,varMDE,'VariableNames',{'U','CR','ER','MDE'  
'},'RowNames',{'Method 1';'Method 2';'Equal weights';'Optimal  
Weights'});;  
  
display(T)
```