ΠΑΝΤΕΙΟΝ ΠΑΝΕΠΙΣΤΗΜΙΟ ΚΟΙΝΩΝΙΚΩΝ ΚΑΙ ΠΟΛΙΤΙΚΩΝ ΕΠΙΣΤΗΜΩΝ

PANTEION UNIVERSITY OF SOCIAL AND POLITICAL SCIENCES

ΣΧΟΛΗ ΔΗΜΟΣΙΑΣ ΔΙΟΙΚΗΣΗΣ

ΤΜΗΜΑ ΟΙΚΟΝΟΜΙΚΗΣ ΚΑΙ ΠΕΡΙΦΕΡΕΙΑΚΗΣ ΑΝΑΠΤΥΞΗΣ

Realized and Implied Volatility Decomposition

ΔΙΔΑΚΤΟΡΙΚΗ ΔΙΑΤΡΙΒΗ

Ελευθερία Κ. Καφουσάκη

Αθήνα, 2023

Τριμελής Επιτροπή

Σταύρος Ντεγιαννάκης, Καθηγητής Παντείου Πανεπιστημίου (Επιβλέπων)

Θεοδόσης Παλάσκας, Καθηγητής Παντείου Πανεπιστημίου

Γιώργος Φίλης, Επίκουρος Καθηγητής Πανεπιστημίου Πατρών

Στους λατρεμένους μου γονείς Κωνσταντίνο και Καλλιόπη

που γέμισαν τη ψυχή μου με ανυπέρβλητη στωικότητα,

αστείρευτο φως και αιώνια αρμονία

Ευχαριστίες

Κάθε διδακτορικό συνοδεύεται και από τις απαραίτητες ευχαριστίες που είναι μόνο μια πάρα πολύ μικρή μνεία σε ανθρώπους που με τον τρόπο τους συνέδραμαν στο να υλοποιηθεί και να περατωθεί επιτυχώς μια πορεία μάλλον μοναχική με πολλές προκλήσεις και μοιραία αδιέξοδα. Για αυτή λοιπόν την πορεία θα εκφράσω την αιώνια ευγνωμοσύνη μου προς το άτομο που πίστεψε σε εμένα, τον καθηγητή μου Σταύρο Ντεγιαννάκη. Υπήρξε πάντα δίπλα μου, συνοδοιπόρος με συμβουλές πατρικές όπως μόνο μπορούν να χαρακτηριστούν που αποτέλεσαν πυξίδα στις αδιέξοδες διαδρομές.

Συνεχίζοντας, το μεγάλο πρόσωπο στο οποίο χρωστάω τα πάντα, είναι η μητέρα μου που η παρουσία της και μόνο στις πολύ δύσκολες στιγμές που ακολούθησαν την έναρξη του διδακτορικού μου, με αποκορύφωμα την απώλεια του πατέρα μου, με γέμισε με την απαραίτητη δύναμη για να συνεχίσω ακάθεκτη χωρίς παύση ούτε για μια στιγμή προς τον υπέρτατο στόχο. Δεν θα μακρηγορήσω καθώς υπάρχουν και εκείνες οι δυνάμεις που αν και δεν είναι απτές και δεν μπορούν να περιγράφουν με λόγια αλλά μόνο με τις αισθήσεις, αλλά συνδράμουν και αυτές στην εσωτερική ώθηση προς τους ατελείωτους στόχους που κάθε ένας από εμάς θέτει.

Κλείνοντας ευχαριστώ την συμβουλευτική μου επιτροπή για τα εύστοχα και με αγάπη διατυπωμένα σχόλια που συνέδραμαν στην βελτίωσή μου και στην ολοκλήρωση αυτής της πορείας που αν και συνοδεύτηκε από αρκετές δυσκολίες, στο τέλος επήλθε η πλήρωση.

Table of Contents

List of tables

List of figures

Realized and Implied Volatility Decomposition

Abstract

This thesis consists an integrated work on analyzing, decomposing, modelling, and forecasting stock market's implied and realized volatility, two intrinsic elements for policy agencies, governmental agencies, portfolio managers, market makers, researchers, and academics. By exploiting six diverse disaggregation techniques, namely the empirical mode decomposition (EMD), the embedded empirical mode decomposition (EEMD), the singular spectrum analysis (SSA), the Hilbert vibration decomposition (HVD), the empirical wavelet transform (EWT), and the variational mode decomposition (VMD), tailored exactly on the peculiar nature these measures have, thesis isolates their key characteristics in order to result in optimum modelling and profitable forecasts for those actively engaging in the markets. Modelling and forecasting were conducted via autoregressive processes (AR), the heterogeneous autoregressive model (HAR), the holt winters framework (HW) and the long-short term model (LSTM) of the deep learning environment. Through entire process no look ahead bias was committed, something innovative in decomposition techniques literature and a trading exercise on futures contracts was incorporated in order to highlight the substantial divergence there exists among forecast evaluation criteria. Thus, thesis targets on investigating the profits that can be accumulated by infusing the forecasting and trading practices with the proposed model architecture. Important implications are drawn out of these results that justify the choice to

simultaneously test, for first time, six decomposition techniques under the same model specifications and under the actual way markets operate. The trading strategy we followed, generated significant economic profits for markets participants for both implied and realized volatility measures, as proxied by VIX volatility index and high frequency returns of S&P500 futures index, respectively, for at least two out of the six decomposition-based models, highlighting some critical facts of proposed methodological procedure.

This thesis, also, was expanded to include another decomposition technique, the continuous wavelet transform (CWT), paired with wavelet coherency, but this time for conducting analysis in the time-frequency domain. The choice was inevitably guided by the need to further exploit the advantages decomposition methods, originating from other scientific fields can offer to the analysis of financial time series. Hence, we included four extra variables, namely the S&P500 stock index, the OVX crude oil volatility index, the high frequency returns of WTI futures index as proxy for the realized volatility of crude oil and the WTI crude oil price in order to analyze them in pairs and spot the strength and the direction of their co-movements. Findings tell for a strong implied-realized volatility relation and of a leading realized volatility against implied at different frequency scales revealing cycling patterns coinciding with the frequency investors reconstruct their positions. Findings also inform of a leading VIX against OVX and the stock market, again following altering time cycles. The most astonishing though finding was the absence of any significant daily and weekly cyclical co-movement for all pairs.

***Keywords*:** decomposition techniques, implied and realized volatility forecasting, ensemble modelling, wavelet analysis

Περίληψη

Η διατριβή αυτή αποτελεί μια ολοκληρωμένη μελέτη πάνω στην ανάλυση, διάσπαση, μοντελοποίηση και πρόβλεψη της τεκμαρτής και της επιβεβαιωμένης μεταβλητότητας των χρηματαγορών, δυο βασικών στοιχείων για τους φορείς οικονομικής πολιτικής, τους κυβερνητικούς φορείς, τους ανθρώπους της αγοράς, τους ερευνητές και τους ακαδημαϊκούς. Εξερευνώντας έξι διαφορετικές τεχνικές διάσπασης χρονοσειρών, και συγκεκριμένα το empirical mode decomposition (EMD), το embedded empirical mode decomposition (EEMD), το singular spectrum analysis (SSA), το Hilbert vibration decomposition (HVD), το empirical wavelet transform (EWT), και το variational mode decomposition (VMD), που διαμορφώνονται πάνω στην ιδιάζουσα φύση που έχουν οι μεταβλητές αυτές, η έρευνα μας απομονώνει τα κύρια χαρακτηριστικά τους στοχεύοντας στην διαμόρφωση ιδανικών μοντέλων και κερδοφόρων προβλέψεων για τους συμμετέχοντες ενεργά στις χρηματαγορές. Ο μοντελισμός και οι προβλέψεις πραγματοποιήθηκαν μέσω των υποδειγμάτων των autoregressive processes (AR), του heterogeneous autoregressive model (HAR), του holt winters framework (HW) και του long, short term model (LSTM) προερχόμενου από το περιβάλλον της βαθιάς εκμάθησης. Καθόλη τη παραπάνω διαδικασία δεν διαπράχθηκε ουδεμία μεροληψία, μια καινοτομία στις τεχνικές διάσπασης χρονοσειρώνν ενώ παράλληλα χρησιμοποιήθηκε συγκεκριμένη επενδυτική στρατηγική σε συμβόλαια μελλοντικής εκπλήρωσης, προκειμένου να επισημάνουμε την σημαντική απόκλιση που παρατηρείται μεταξύ των κριτηρίων αξιολόγησης των προβλέψεων. Επομένως, η διατριβή αυτή

στοχεύει στο να εξερευνήσει τα οφέλη που μπορούν να αποκομιθούν από την προσθήκη της προτεινόμενης αρχιτεκτονικής μοντελισμού στις επενδυτικές πρακτικές και στις διαδικασίες πρόβλεψης. Σημαντικά συμπεράσματα αντλούνται από αυτά τα αποτελέσματα που δικαιολογούν την επιλογή μας να εξετάσουμε ταυτόχρονα, για πρώτη φορά, έξι τεχνικές διάσπασης χρονοσειρών, κάτω από τις ίδιους προσδιοριστικούς παράγοντες και κάτω από τις πραγματικές συνθήκες λειτουργίας των χρηματαγορών. Η επενδυτική στρατηγική που ακολουθήσαμε, επέφερε σημαντικές οικονομικές απολαβές για τους συμμετέχοντες στις χρηματαγορές και για τις δυο υπό εξέταση μεταβλητότητες, όπως προσεγγίζονται από τον δείκτη τεκμαρτής μεταβλητότητας, VIX και τον S&P500 δείκτη συμβολαίων μελλοντικής εκπλήρωσης από τον οποίο κατασκευάζεται η πραγματοποιημένη μεταβλητότητα, σε τουλάχιστον δυο από τις έξι τεχνικές διάσπασης, προβάλλοντας κάποια σημαντικά χαρακτηριστικά που φέρει η προτεινόμενη διαδικασία.

Επίσης, η διατριβή αυτή, επεκτείνεται για να εσωκλείσει ακόμα μια τεχνική διάσπασης, το continuous wavelet transform (CWT) που συνοδεύεται από την wavelet coherency διαδικασία, αλλά αυτή τη φορά για να πραγματοποιήσουμε ταυτόχρονη ανάλυση στο χρόνο και στη συχνότητα. Η επιλογή αυτή οδηγήθηκε από την ανάγκη για περαιτέρω διερεύνηση των δυνατοτήτων που μας προσφέρουν οι τεχνικές διάσπασης, που προέρχονται από άλλα επιστημονικά πεδία και μπορούν να συνεισφέρουν στην ανάλυση των χρηματοοικονομικών μεταβλητών. Έτσι συμπεριλάβαμε 4 επιπλέον μεταβλητές, συγκεκριμένα το δείκτη S&P500, το δείκτη OVX, που αποτελεί την τεκμαρτή μεταβλητότητα της τιμής του αργού πετρελαίου, τις υψηλής

συχνότητας αποδόσεις του δείκτη WTI futures, ως αντιπροσωπευτικού της πραγματοποιθείσης μεταβλητότητας της τιμής του αργού πετρελαίου και τον WTI δείκτη της τιμής του αργού πετρελαίου, με σκοπό την μελέτη τους σε ζεύγη και των εντοπισμό της έντασης και της κατεύθυνσης της συνδιακύμανσης τους. Τα ευρήματα μαρτυρούν μια δυνατή σχέση μεταξύ τεκμαρτής και πραγματοποιημένης μεταβλητότητας καθώς και το γεγονός ότι η πραγματοποιθείσα μεταβλητότητα ηγείται της τεκμαρτής σε διαφορετικές συχνότητες αποκαλύπτοντας διαφορετικές κυκλικότητες που γειτνιάζουν με την συχνότητα με την οποία οι επενδυτές αναπροσδιορίζουν τις επενδυτικές τους θέσεις. Τα ευρήματα επίσης αποκαλύπτουν ότι ο VIX ηγείται έναντι του OVX αλλά και του χρηματιστηριακού δείκτη, παρουσιάζοντας μεταβαλλόμενες κυκλικότητες. Ωστόσο αυτό που προκαλεί εντύπωση είναι η παντελή έλλειψη οποιασδήποτε σχέσης σε ημερήσια ή εβδομαδιαία βάση κάτι που εντοπίζεται σε όλα τα ζεύγη των μεταβλητών.

Λέξεις-κλειδιά: τεχνικές διάσπασης χρονοσειρών, πρόβλεψη τεκμαρτής και επιβεβαιωμένης μεταβλητότητας, ensemble modelling, ανάλυση κυμάτων

Chapter 1

Thesis introduction

And all goes around volatility, a substantial metric of risk and uncertainty, that is capable to govern the powerful transmission mechanisms of the financial system. Cult, by now, measures of dispersion, implied volatility and realized volatility, have long been consolidated, in academic research, financial policies, economic policy and trading. The reason lies in their distinctive characteristics that shape their nature, and the difficulty to tame them, since both metrics manipulate the invisible channels from which information strikes stock markets, derivatives markets, or energy markets by adding complexity. This thesis consists an integrated work on the exploitation of disaggregation techniques, that allows to decompose the two subjects in fully alignment with their underlying processes. By decomposing, the characteristic peculiarities are eased, and volatilities can be optimally analyzed, modelled, and forecasted. Afterall, the ultimate target is the forecast process. Forecasts are the ones that will guide those actively engaging in financial markets, especially derivatives markets, away from trading deficiencies, and incremental loss of capital and the ones that will guide them through alpha generation and portfolio diversification.

Therefore, here we engage in the study of stock market's implied and realized volatility. We incorporate six different decomposition frameworks, mainly originating from the signal processing field, that assist in decomposing the two metrics into their intrinsic modes. Then intrinsic modes generated out of each technique and for each volatility, are modelled separately and

forecasted through four models of mixed origins, an elementary one as is the autoregressive model (AR), a benchmark for volatility modelling, namely the heterogeneous autoregressive model (HAR), a classic one as the holt winters model (HW) and one emanating from the deep learning environment. The aggregated forecasts are then evaluated based on a statistical loss function and an economic loss function that employes a trading strategy tailored to market's actual operations. The aim is to show on the first place, how the diverse disaggregation techniques should be added and best exploited in the forecasting process to prevent the inclusion of look ahead bias in our system and second, to highlight the contradictive outcomes between the loss functions that can influence and severely impact investors decision-making projects. Thus, we manage to add something extra in a rather crowded field and contribute to both ensemble modelling and implied and realized volatility modelling and forecasting literature.

The second chapter of this thesis trucks the path that marks the evolution conducted in volatility modelling and ensemble modelling. There, we engage in a thorough description of the decomposition methods incorporated, namely the empirical mode decomposition, the ensemble empirical mode decomposition, the singular spectrum analysis, the Hilbert vibration decomposition, the empirical mode decomposition and the variational mode decomposition and provide an extensive description on how modelling, forecasting, evaluation and validation is being conducted in favor of implied and realized volatility, as proxied by VIX Cboe volatility index and high frequency S&P500 futures returns, respectively, that are explicitly presented in the chapters that follow. We specify the details of the out-of-sample

forecasting of multiple days ahead forecast horizon process, and describe the rules of the followed trading strategy, as part of the evaluation process on roll adjusted VIX futures, the tradable part of VIX index, and present their special features. Validation is processed via Model Confidence Test and Direction of Change.

The third chapter evolves around implied volatility for the U.S stock market, the VIX Cboe volatility index. It is the chapter where the first important implications are drawn out of our findings and profitable forecasts are being recorded. The cumulative returns of the implemented trading practise tell for a significant return on capital allowing for further exploitation, that is conducted in chapter four. In the fourth chapter we present the study on realized volatility, the high frequency S&P500 futures returns, under the same model architecture specifications. The study confirms the promising results recorded previously. The same models that worked in generating profitable forecasts for market participants under the economic loss function, for the case of VIX index, also worked for realized volatility, a much more persistent measure. The fact that not all decomposition techniques ended with promising results it is attributed mainly to the aspects that deal with the underlying theory of each scientific field. And that was the initial reason why we employed six decomposition frameworks altogether, under the exact same conditions, a practice not previously recorded. We wanted to test simultaneously their efficiency in the actual markets' environment. Certainly, findings demonstrated here, shed light to previously untouched issues. The fact that some models worked for both volatility measures, creates a pattern critical for future experimentations.

The fifth chapter diverges slightly from initial modelling process. Here, we included the continuous wavelet transform (CWT), a decomposition method that was not incorporated in the previous chapters as it is best applied for conducting a detailed analysis of a time series in the time-frequency domain, simultaneously, as the first step prior to wavelet coherence. Wavelet coherence is the one that trucks the co-movement of two time series in the time-frequency domain and estimates the power and the direction of this co-movement. Since we deal with two variables, two volatilities, so closely related and there is definitely a long history on studies trucking the influence of one-another and their spillovers in the markets, there could not be absent a method like this, from a thesis dedicated on these extremely influential measures. Hence, we created six pairs of variables, and in this final chapter, we extended our study to include the S&P500 stock index, the implied volatility of WTI crude oil, the OVX Cboe volatility index, the realized volatility of the crude oil constructed by high frequency returns of WTI futures index, and the WTI crude oil price. We did so, in order to allow for a thorough investigation of the interrelation of other implied and realized volatilities pairs from other markets, as the energy market. Our study would be restricted if not for this inclusion. In order to be able to quantify findings there must be a comparison with other identical measures that too dominate the global financial chessboard. Thus, study would be biased if we neglected a close fellow of VIX index, at least for this thesis, that incorporates VIX index as proxy for implied volatility. The pairs created was the one of VIX index-OVX index, VIX index-S&P500 index returns, VIX index-S&P500 realized volatility, OVX index-WTI realized volatility and OVX index-WTI crude oil price. Findings

are quite illuminating on stressing how this relation unfolds through the time-frequency course. Many cyclical behaviors aligned with the frequencies with which markets participants reconstruct their positions, and of specific direction, were spotted. Realized volatilities were found to lead implied ones and there were evident moments of not significant relation at all. There was also a striking discovery. No daily or weekly periodic relation was spotted for any of the proposed pairs.

Concluding, the final chapter, chapter six sets the last testimonies on this doctoral thesis, the studies conducted, the interesting and important implications and of course the continuity of research and experimentation as an inextricable element of every science.

Chapter 2

Divide and conquer

2.1The evolution in volatility modelling

From a pure statistical tool to a tool for measuring systemic risk, shaping macroprudential policy and pricing derivatives products, volatility has definitely covered a long way all these years. Volatility, captivates and shapes the work and moves of researchers, columnists, analysts, academics, governmental agencies but most importantly market participants, for whom ad hearing to volatility on a daily basis can lead to successfully implemented trading and diversifying strategies.

Literature has to provide an extensive collection of studies evolving around volatility, stochastic, conditional, implied, or realized (Alizadeh *et al*., 2001; Adrian and Rosenberg, 2006; Nomikos and Pouliasis, 2011; Degiannakis and Filis, 2017; Kampouroudis *et al.*, 2021), with implied and realized volatility to naturally capturing more intensively the interest of academics, investors, and policy makers.

The concept of realized volatility dates back to the studies of Andersen and Bollerslev (1998), Andersen *et al.* (2006; 2012), Hansen and Lunde (2006), Barndorff-Nielsen and Shephard (2002), Barndorff-Nielsen *et al.* (2008), as the availability of ultra-high frequency data completely altered the modeling landscape and allowed the construction of non-parametric realized volatility measures that could approximate the concept of integrated volatility. Measures such as Quadratic variation, bi-power variation, quantile-based realized variance, Min Realized variance, Med realized variance or positive

and negative realized semi-variance, were only few of the proposed measures. That in turn brought forward the need to introduce the best model that could capture the fundamental statistical properties, realized volatility had (Andersen *et al*., 2003; Andersen and Benzoni, 2008). Thus, modelling and forecasting realized volatility and at the same time assessing the usefulness of volatility forecasts, became an unconditional trend.

On the other hand, the use of implied volatility, conceived as the natural forecast of integrated volatility, primary involved in forecasting realized volatility as is reported in the early studies of Christensen and Prabhala (1998), Jorion (1995) or Canina and Figlewski (1993). Steadily gained ground due to the incremental information expected volatility encloses for the future path of a market, the pricing of its products, but also for macroeconomic conditions, energy commodities, speculative strategies, portfolio optimization etc. (Fernandes *et al.,* 2014; Degiannakis and Filis, 2022).

In the vast research conducted searching for the best framework able to model the two prevalent volatility measures and assess the usefulness of volatility forecasts, various parametric concepts have extensively been proposed. From the classical linear model (Christensen and Prabhala, 1998), to frameworks like ARFIMAX, ARCH-family as in Angelidis and Degiannakis (2008) for forecasting the realized and the conditional volatility of three European stock indices, to ARIMA, ARFIMA, VAR, as in Konstantinidi *et al*. (2008), where they forecast seven U.S. and European implied volatility indices along with asset allocation practices, or ARMA-type as in Kambouroudis *et al*. (2016) for implied volatility of three major U.S stock indices, who compare forecast performance with GARCH-type, and finally Heterogeneous

Autoregressive model, the HAR, of Corsi (2009), modelling has covered a long way. HAR the state-of-the-art framework for modeling and forecasting volatility, was initially proposed for realized volatility due to its alignment with markets' participants diverse time horizons and extended likewise by Andersen *et al.* (2007), Corsi and Renò (2012) and Degiannakis and Filis (2017) to account for multivariate settings. HAR was able to deal with long memory, fat tails and other stylized facts of realized volatility (Bollerslev *et al.*, 2003; 2009). Many are the studies to incorporate HAR-X frameworks for realized volatility as Busch *et al.* (2011) of stock, bond, exchange rate markets or Haugom *et al.* (2014) for the energy market. Lately, it has been established for modelling and forecasting implied volatility as well. Degiannakis and Filis (2022) forecast implied volatility of oil price through HAR-X model specifications and provide evidence of efficiency in producing meaningful forecasts and economic profits when incorporated in trading practices involving implied volatility derivatives. The same applies for Delis *et al*. (2023), who forecast crude oil implied volatility by incorporating a HAR-X framework in order to implement a straddle trading strategy and validate the profitable forecasts.

Quite recently in volatility modeling and more general in the financial analysis, modeling theory and practice, entered, with increasing rates, non-parametric decomposition methodologies paired either with parametric models or non-parametric techniques, resulting in hybrid models. These hybrid frameworks, are the outcome of the evolution contacted in computer science that allowed the construction of powerful algorithms, liberating the scientific and engineering work processes. One of the major contributors of

such algorithms by far is the signal processing field, followed by the machine learning, deep learning, and artificial intelligence environments (Prasad and Bakhshi 2022; Vrontos *et al.*, 2021). Signal decomposition methodologies along with ensemble frameworks, suddenly flourished providing a more integrated analysis for time series data. These techniques, efficiently tackle with the restrictions raised from classic econometric models and effectively cope with the special characteristics, economic and financial data processes have, the non-linearity and non-stationarity (Huang *et al.,* 1998; Civera and Surace, 2021). The decomposition frameworks, can reveal the peculiar features that time series display and are capable of separating trend, high and low frequency components, periodicities, noise etc. Then, the retrieved components can be optimally analyzed and modeled separately, with the help of parametric models, if the nature of the retrieved component allows such, or with more advanced. Adaptive mode decomposition, variational mode decomposition, semi-variational mode decomposition, wavelet decomposition, convolution neural networks, recursive neural networks, deep neural networks are only few of the available algorithms paired together in order to conduct analysis, modelling and forecasting of complex systems, resulting in optimal statistical and economic gains (Yu *et al.,* 2008; Degiannakis *et al*., 2018).

From the many decomposition techniques that exist, in this thesis, we adopted the six more widely incorporated by researchers and most fitted for the financial analysis we conduct, namely the empirical mode decomposition (EMD) (Huang *et al.* 1998), the embedded empirical mode decomposition (EEMD) (Wu and Huang, 2009), the singular spectrum analysis (SSA) (Golyandina et al., 2001), the Hilbert vibration decomposition (HVD)

(Feldman, 2006), the empirical wavelet decomposition (EWD) (Gilles, 2013) and the variational mode decomposition (VMD) (Dragomiretskiy and Zosso, 2014). We paired these techniques with three parametric and long consolidated in time series analysis models, the Autoregressive model (AR), the HAR framework, the Holt Winters (HW) and a model coming from the deep learning environment, the long short-term memory model (LSTM). The evidence reported from the exploitation of above architecture is rather encouraging since it comes with significant statistical gains and this was the initial reason to dive into the extensive experimentation that a study, like the disaggregation and forecasting of implied and realized volatility, requires. Degiannakis *et al*. (2018) in their work following the SSA decomposition method and the HW framework, report superiority of the multiple days ahead generated forecasts out of the disaggregation and modeling concept, when compared to the ARFIMA and HAR framework, a study that was further enhanced by the economic gains achieved from implementing two trading strategies. Hassani *et al.* (2010) use SSA and its multivariate extension (MSSA) to predict the price and the direction of change of U.K. exchange rates against random walk models, with MSSA, outperforming naïve approaches. Lahmiri (2016), adopting the VMD approach and the general regression neural networks models, (GRNN), forecast the 1 day ahead prices of U.S. economic and financial indices and report forecast gains. In the same line, Zhanke Li *et al.* (2020), apply VMD along with LSTM against EEMD-LSTM to forecast oil price and report improved forecast accuracy. The same is valid for Shu and Gao (2020), who use EMD, in order to decompose the Shanghai Stock exchange composite index and then forecast it through

convolutional neural networks. EMD and feed forward neural networks are also incorporated by Yu *et al.* (2008) for the crude oil price series of WTI and Brend and report statistical gains. Plakandaras *et al.* (2015), adopt the EEMD process to decompose the time series of 5 exchange rates and use support vector machines models to model and forecast the exchange rates and report in-sample and out-of-sample forecast superiority against benchmark models. Of course, studies presented here, are indicative and consist only a small fragment on the investigation conducted in financial markets through this modelling architecture, giving rise to new trends.

There is, however, a small flaw in the way the decomposition of a time series takes place, and in the way, the intrinsic components are being modeled and forecasted and then evaluated. In most studies the complete sample is used to perform the decomposition and then the sample is split into training and testing. This fact permits a look ahead bias to be evident in the entire process, concluding in sometimes false predictive gains. In this thesis we diverge from above process. We decompose implied and realized volatility in a daily basis through a rolling sample and at the same time model and forecast the individual components of each technique. That excludes the inclusion of any future information in the process. Thus, we do not solely seek for forecast accuracy in statistical grounds, because the advantage earned when decomposing entire series is diminished. Here the future path is an unknown parameter. What we seek is the economic benefits, the economic usefulness of the generated forecasts if any. This can be retrieved from the divide and conquer architecture only when entered and tested under the actual way financial markets' function and financial decisions are being drawn.

For that reason, we employ an objective-based evaluation criterion along the classic mean square error (MSE) and incorporate the Model Confidence Set (MCS) of Hansen *et al*. (2011) and the Direction of Change (DoC), in order to further validate the final outcome. The objective-based evaluation criterion is the cumulative returns earned by following a simply trading rule on long and short positions in roll adjusted futures contracts of implied volatility. We choose the VIX futures as the appropriate product since in our studies we incorporate the VIX volatility index, the central implied volatility index for the U.S. stock market and the realized volatility of the S&P500 futures index.

Objective-based evaluation criteria are not new in volatility forecasting literature. They date back to the work of Engle *et al*. (1993) and Noh *et al*. (1994), who incorporate options trading strategies in order to assess the economic usefulness of volatility forecasts. And since volatility's movements constitute the key for pricing options, profitable forecasts of implied or realized volatility would always be the prime desire of trading practices in many studies. Ahoniemi (2006) produce 1 step ahead forecasts of VIX through ARIMA-type and binary probit models and simulate a straddle option strategy. Three trading strategies on VIX Volatility Index and S&P500 Index, apply Vrontos *et al*. (2021) to evaluate the economic significance of the 1 step ahead forecasts generated out of their machine learning approaches. Degiannakis and Filis (2022) and Delis *et al*. (2023) also proceed with options trading strategies, especially straddles, only this time, model specifications and trades are fully aligned with options markets' actual conditions, under which, trades of long and short positions take place. So, the trading practice we adopt on futures instead of options is also aligned with these conditions.

Hence, this thesis contributes both to the literature strand that uses objective-based evaluation criteria to evaluate the generated forecasts and to the literature strand that incorporates ensemble methods for economic and financial time series modeling and forecasting in order to enhance the economic gains.

In the sections that follow, we proceed by presenting the decomposition methods and the forecasting frameworks, along the evaluation and validation criteria for the studies that will be presented at chapters 3 and 4, where we decompose, model and forecast VIX and realized volatility of S&P500 futures index, respectively. Thus section 2.2 presents the disaggregation techniques. Section 2.3 unfolds the model frameworks that are adopted for modelling and forecasting the two volatility measures. Finally, section 2.4 presents the evaluation and validation concepts.

2.2 The disaggregation techniques

All disaggregation methodologies that participate in our divide and conquer modelling architecture, were chosen very carefully, considering the peculiar statistical properties that realized and therefore implied volatility, exhibit. We also, tried to enhance our study by enclosing diverse techniques most originating from the signal processing field and applications to medical, earth, marine, seismic, machinery vibrations etc., with different theoretical backgrounds and test their performance under the same market conditions. All of the proposed frameworks deal efficiently with non-stationary, no-linear, and noise-distorted systems and all of them inevitably come with advantages

and disadvantages[1], but after all, no method is flawless no matter if most of them always evolve and new extended versions come forth. Of course, we stick to the original versions of these techniques.

2.2.1 The empirical mode decomposition

It was the year 1998 when the study of Huang *et al.* (1998) published, proposing a new 2-step tool to cope with the difficulty faced when trying to analyze noisy, non-stationary and non-linear time series and partially replace the Fourier spectral analysis that due to its restrictions would not fit for the analysis of such data. The first step was the EMD that could decompose any complex data through a spline algorithm into a finite and small number of intrinsic mode functions (imfs) that were complete and orthogonal (Huang et al., 2003) and would localize an event on time. The second step was the Hilbert transform of the imfs that would localize the event in the frequency domain and would allow for a more efficient and meaningful analysis. Since then, EMD, has been widely applied and paired with parametric and non-parametric models for decomposing and modeling wind-speed data, crude-oil price time series and other financial time series (Ali *et al*., 2023; Shu and Gao, 2020; Jin *et al*., 2022).

EMD is an adaptive and direct algorithmic method for decomposing a signal. The signal in order to be decomposed into a zero-mean amplitude modulated imfs and a residue, must comply with three assumptions (Huang et al., 1998), 1) have at least two extrema, one minimum and one maximum, 2)

[1] All techniques have pros and cons, but our intention is not to extent to every issue such methods face, either mathematical or technical, rather provide a basic understanding of the underlying theory. Those who desire a more thorough learning can always refer to the original papers of their initiators.

the characteristic time scale to be defined by the lapse of time between the extrema and 3) when there is a total absence of extrema and only the presence of inflection points, then differentiate one or more times in order to reveal the extrema. Furthermore, an imf must satisfy two conditions, 1) the number of extrema and the number of zero-crossings in the entire sample to be either equal or differ by one[2] and 2) at any point the mean value of the envelope defined by the local maxima and the envelope defined by local minima to be zero. Thus, in order to identify an imf, EMD follows an iterative process that is better described in five steps:

1. For any signal $\mathrm{X_t} = \{\mathrm{x_1}, \mathrm{x_2}, \dots, \mathrm{x_t}\}$, which in our case is the $\mathrm{VIX_t}$, we identify all extrema.
2. Then, we connect all local maxima with a cubic spline and create the upper envelope. We do the same for the local minima in order to create the lower envelope. Upper and lower envelopes should cover all data points, after all, the creation of the envelopes is the one that will help in separating components with close frequencies (Feldman, 2011).
3. We designate the mean of the upper and lower envelopes as $\mathrm{m_1} = (\mathrm{env_{max}} - \mathrm{env_{min}})/2$. We subtract m_1 from original signal and retrieve the first component $\mathrm{c_1} \rightarrow \mathrm{X_t} - \mathrm{m_1} = \mathrm{c_1}$

[2] Huang *et al*. (2003) argue that imfs retrieved based on this condition are orthogonal and not over sifted. Basically, one of the cons that is being addressed to EMD is the lack of a mathematic formula and the fact that although imfs are indeed orthogonal, orthogonality itself cannot be proved.

4. Does the component comply with imposed restrictions? If so, then it is an imf, if not we replace X_t with c_1 and iterate over i times until component equals an imf. That is called the sifting process.
5. In the case the resulted component satisfies the characteristics of an imf, then $imf_1 = c_i$ (with $i = 1, 2, \dots$). And we get the first imf. After having the first imf we subtract it from original series $X_t - imf_1 = r_1$. The r_1 now takes the place of signal and process starts over until all imfs are retrieved and only a residue is left, res_t, that is a monotonic function. At that point EMD terminates. The first imf is the one with the highest frequency, and progressively frequency of ongoing imfs lessens as gradually moves towards the final monotonic residue.

The summation of all imfs and residue returns original series. Among proposed and adopted decomposition techniques, EMD is the only complete method whose, the summation of components (imfs and residue) returns the exact same signal $\rightarrow X_t = \sum_{k=1}^{t} imf_k + res_t$.

There have been proposed various extensions to the above process regarding cubic spline end conditions or stopping criteria. There are three cubic spline end criteria. The first, as proposed in Huang *et al*. (1998), is to add characteristic waves at the ends, where cubic splines tend to have wide swings. The second, comes from Rilling *et al.* (2003), who propose to mirrorize the close to the end extrema instead of padding extra data. In this study we follow a distinct literature branch that adopts the third cubic spline end criterion, that is to place the end points of the slope of the cubic spline equal to 0 (Peel *et al*., 2007). According to Peel *et al.* (2007) this condition

returns fewer imfs and tends to be more efficient compared to other cubic spline rules proposed for EMD (Rilling *et al.*, 2003; Huang *et al.*, 1998; Flandrin *et al*., 2004). We evidenced the same findings[3] and ended with 5 and 6 components depending on the rolling sample, for both implied and realized volatility.

Regarding the stopping criterion of the sifting process, many different methods have occasionally appeared, the standard deviation of Huang *et al.* (1998), the threshold method of Rilling *et al.* (2003), the S-number of Huang *et al.* (1999) and the Energy Difference Tracking of Junsheng *et al.* (2006). In this thesis, we follow the original work of Huang *et al.* (1998), where a component is recognized as an imf, by the standard deviation function, that has the following classic form, with a value less than 0.2:

$$St.Dev = \sum_{t=0}^{T}[\frac{|(\left(c_{1(i-1)}(t)-c_{1i}(t)\right)|^2}{c_{1(i-1)}^2(t)} < 0.2$$

Overall, EMD constitutes an effective and fully data-driven decomposition method. There are definitely some disadvantages, the most proclaimed is the absence of a robust mathematical background, since it is an algorithm, but that never posed any issues in its functionality, because it is being widely applied in diverse scientific fields with quite a success for almost 25 years. Another disadvantage is the mode mixing, but this is something that is being tackled through the EEMD of Wu and Huang (2009), that is the second approach that we incorporate. Of course, there are occasionally other issues dealing with the sifting process etc., but the intention of this thesis is

[3] Classic cubic spline conditions returned 10 and 11 components, the Zero ends returned 5 and 6 components interchangeably between the 1600 iterative samples for implied volatility and 1021 for realized volatility, due to the rolling window technique.

not to provide an extensive analysis on every issue that the decomposition techniques may face, but to clarify some of them in order to realize that no process is flawless and there will always be some alterations among the proposed EMD algorithms that may alter the final outcome of different studies.

2.2.2 The embedded empirical mode decomposition

EMD process as we mentioned, although being an adaptive and fully data driven decomposition method, widely implemented, it has frequently been accused of mode mixing, an outcome of signal intermittency (Huang *et al*., 1999). This can raise annoying issues in the analysis conducted in the resulting components that may lack in any physical meaning through the second step that is the Hilbert transform. In order to cope with this issue Wu and Huang (2009) proposed the EEMD based on the study of Flandrin *et al*. (2004). This noise-assisted signal extraction method infuses signal with white noise of finite amplitude that leads to efficient frequency separation. Thus, the resulting imfs constitute the mean of an ensemble of repeated iterations (Yeh *et al*., 2010). The process of EEMD is in the same line as EMD and follows the same steps with the difference that:

1. We add a white noise series, $\mathrm{wn_t}$ to the original signal, $\mathrm{s_t}$: $\mathrm{X_t} = \mathrm{s_t} + \mathrm{wn_t}$.
2. We proceed with the decomposition as conducted in EMD and retrieve imfs.
3. We iterate[4] above steps many times adding different white noise series.

[4] The number of ensembles is dictated by the data itself. For this study, we went through 30 trials.

4. We calculate the ensemble means of the imfs returned from every iteration. The means returned are the final imfs of entire process.

According to Wu and Huang (2009) this repeated addition of different white noise series of finite amplitude not infinitesimal, in the original signal are cancelled out in the end result and the mode mixing problem is efficiently tackled without perturbating original signal. Moreover, the addition of finite amplitude white noise allows EMD to act as dyadic filter bank (Flandrin *et al.,* 2004), and accomplish meaningful imfs. Compared to EMD, the EEMD alternative is also extensively applied and a plentiful of studies end with promising results when EEMD is incorporated in hybrid models or EEMD extensions are used (Dong *et al.,* 2019; Sun *et al.,* 2018; Tang *et al*., 2018).

In this work EEMD inevitably is present, since both implied volatility, approximated by VIX Index and realized volatility, approximated by the 5 min returns of SP 500 futures index, have a persistent and volatile nature with long memory. We thought that the addition of EEMD in our decomposition techniques would allow to inspect whether EEMD compared to EMD gives more refined imfs with altered final economic recordings. In our case, both methods for both time series terminated at 5 and 6 components interchangeably, due to the alterations throughout the rolling scheme. The truth is that original EMD was efficient enough to decompose VIX, but for realized volatility only the EEMD method in the end is recorded. As for the reported economic gains, that are presented at the relative tables of sections 3 and 4. Based on the trading strategy we implemented, we should go on and state that profits, somehow accelerated for the EEMD case compared to EMD, although both frameworks returned remarkable cumulative returns for

most forecast horizons at least for the VIX, where both methods are present. Perhaps these infinitesimally small declines between the resulted components of the EMD and EEMD processes do seem to have a critical impact in the end. Practice can only tell.

2.2.3 The singular spectrum analysis

SSA, constitutes another non-parametric time series analysis and decomposition technique with various applications and many extensions (Golyandina and Zhigljavsky, 2013; Golyandina *et al.*, 2001; Hassani 2007) that approximately decomposes a time series into noise, periodicities, trend etc. Although SSA is demonstrated and primary used as a decomposition and filtering technique, it can also be used for forecasting purposes as in Degiannakis *et al.* (2018), Thomakos *et al.* (2002), Sulandari *et al*. (2020) to name a few, with remarkable results. In this study we are not going to use its forecasting ability, rather its efficiency to generate meaningful components. Afterall, the primary reason why one uses a decomposition technique is to separate the different harmonics, scales, frequencies, periodicities, trend, noise etc., that would reveal the inner nature of a time series and allow targeted analysis.

SSA is performed by following some distinctive steps of a robust mathematical path (Hassani and Thomakos, 2010). Hence, SSA has two stages, the decomposition and the reconstruction. The very first step in the SSA process is to construct the trajectory matrix and “transform” the time series into a multidimensional matrix:

Stage 1: Decomposition

1. Embedding. Let X_N be a time series of length N, $X_N = (x_1, x_2, \dots, x_N)$. We need two parameters, L, which is called embedding dimension and K, two integers that will constitute the number of rows and columns of the trajectory matrix respectively, with L being (1 < L < N)[5] and $K = N - L + 1$. We form a sequence of L and K lagged vectors out of the original sample. So, The X matrix of Eq. 1 constitutes a Hankel matrix and has equal anti-diagonal elements:

$$X = [X_1, X_2, \dots, X_K] = (x_{ij})_{i,j=1}^{L,K} = \begin{pmatrix} x_1 & x_2 & x_3 & \cdots & x_K \\ x_2 & x_3 & x_4 & \dots & x_{K+1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_L & x_{L+1} & x_{L+2} & \cdots & x_N \end{pmatrix}. \quad (2.2.1)$$

2. Decomposition. In the core of SSA lies the singular value decomposition (SVD). Having constructed the trajectory matrix, we perform SVD on X and decompose it into a sum of rank-one matrices:

$$X = X_1 + X_2 + \cdots + X_d, \quad (2.2.2)$$

where d denotes the rank of X and $X_i = \Sigma_i \sqrt{\lambda_i} U_i V_i^T$ $(i = 1, 2, \dots, d)$. SVD process returns the eigenvalues λ_i of XX^Tin decreasing order of magnitude, $(\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_L \geq 0)$, the left singular vectors of X, U_i and V_i the right singular vectors. The assortment of $(\sqrt{\lambda_i} U_i V_i)$ is called the ith eigentriple of SVD (Hassani *et al.* 2021).

Stage 2: Reconstruction

[5] According to Hassani et al. (2021) L plays an important role during the reconstruction phase and sets how well reconstructed time series approximates original one and how well efficient separation of components is dealt. Values of L varying in between (2 < L < N/2) have higher resolution.

1. Grouping. Grouping is the one that defines the way the individual eigentriples of Eq. 2, (1, 2, …, d), are going to be grouped with each other into m disjoint subsets $\mathrm{I}_1, \mathrm{I}_2, \dots, \mathrm{I}_\mathrm{m}$ called eigentriple grouping[6]:

$$\mathrm{X} = \mathrm{X}_{\mathrm{I}1} + \mathrm{X}_{\mathrm{I}2} + \cdots + \mathrm{X}_{\mathrm{Im}}. \qquad (2.2.3)$$

2. Diagonal averaging. Having grouped the eigentriples, the final step is to transform these X_{Ij} matrices into new series/components of length N, the length of original series. In order to achieve that we first transform these matrices into Hankel matrices Y, of $\mathrm{L} \times \mathrm{K}$ dimensions and elements $\mathrm{y}_{\mathrm{ij}}, 1 \le \mathrm{i} \le \mathrm{L}, 1 \le \mathrm{j} \le \mathrm{K}$. Then, via performing diagonal averaging on Y, we transform it into the desired series $(\mathrm{y}_1, \mathrm{y}_2, \dots, \mathrm{y}_\mathrm{N})$:

$$\widetilde{y_s} = \sum_{(l,k)\in A_s} y_{l,k} \,/|A_s|, \qquad (2.2.4)$$

where $\mathrm{A_s} = \{(\mathrm{l},\mathrm{k}): \mathrm{l}+\mathrm{k} = \mathrm{s}+1, 1 \le \mathrm{l} \le \mathrm{L}, 1 \le \mathrm{k} \le \mathrm{K}\}$ that is the averaging process of the antidiagonals and $|\mathrm{A_s}|$ the number of elements in the set of $\mathrm{A_s}$. Thus, from the original series of $(x_1, x_2, \dots, x_N)$, the entire process returns the reconstructed series of $\widetilde{\mathrm{X}}^{(\mathrm{k})} = (\tilde{\mathrm{x}}_1^{(\mathrm{k})}, \tilde{\mathrm{x}}_2^{(\mathrm{k})}, \dots, \tilde{\mathrm{x}}_\mathrm{N}^{(\mathrm{k})})$ that are the SSA m components and their summation approximates back original series:

$$x_n = \sum_{k=1}^{m} \tilde{x}_n^{(k)}, \qquad n = 1, 2, \dots, N. \qquad (2.2.5)$$

[6] For the case where $m = d$ and $I_j = \{j\}, j = 1, 2, \dots, d$, the corresponding group is called elementary.

Diagonal averaging completes the SSA decomposition process but still the step of grouping remains the most intrinsic one for efficient decomposition. That is why different frameworks have been proposed part of the separability issue faced in SSA and how different components are going to be grouped together. Golyandina *et al.* (2001) propose the weighted correlation concept along with the graphs of eigenvectors that can reveal components that seem to strongly interact-correlate. Hassani *et al.* (2021) against the classic weighted correlation propose hierarchical clustering methods and there is also the relative entropy concept that we adopt. Relative entropy, that in the probability and information theory is best known as the Kullback-Leibler divergence, is a measure that estimates the similarity between two probability density functions (pdfs), $P_a(x)$ and $P_b(x)$ and returns a value of divergence that is not the statistical distance between $P_A(x)$ and $P_B(x)$ (Kanishka et al., 2022). It is a non-negative value, but when it returns zero (0), then $P_a(x)$ = $P_b(x)$. Thus, the Kullback-Leibler divergence for two pdfs, $KL(A||B)$, is given by (Theodoridis, 2020):

$$KL(A||B) = \int_{-\infty}^{\infty} P_A(x) log \frac{P_A(x)}{P_B(x)} dx \tag{2.2.6}$$

Most commonly, relative entropy is used for optimization issues in the machine learning environments. Here, through its assistance, we end with 4 components for the VIX Index and 3 components for the realized volatility, constructed out of 5min returns of SP500 futures index. Now, regardless the preferred technique for the grouping step, effectiveness depends on the

reason why someone engages in SSA decomposition and whether final results comply with the research conducted.

2.2.4The Hilbert vibration decomposition

In the same spirit to EMD, that forms the first step for Hilbert Huang transform, Feldman (2006) introduced the HVD, as part of the evolution conducted in signal analysis. HVD is a decomposition method used especially for the analysis of mechanical vibrations, earthquake motions (Huang *et al.,* 2012), and for the analysis of Electroencephalograms (EEG), Electrocardiograms (ECG) or Seismocardiograms (SCG) (Shankar *et al.,* 2021; Singh *et al*., 2022) among others. HVD constitutes an iterative algorithm, where in each iteration the component with the highest energy is subtracted from original data relying on the synchronous demodulation process (Feldman and Braun, 2017). HVD as most methods, follows some rules and HVD must comply with three assumptions, 1) the underlying vibration to be the outcome of the superposition of quasi-harmonics functions, 2) the envelopes of each vibration component to differ, and 3) the total length of the vibration to include several longest periods of the corresponding slowest component (Feldman, 2006).

In the core of HVD lies the Hilbert transform (HT), a linear operator designated for the analysis of complex signals with varying amplitude and frequency through the course of time. For a vibration process**,** x(t), the HT takes the form of an integral transform considered a Cauchy Principal Value:

$$\mathrm{H}[\mathrm{x}(\mathrm{t})] = \tilde{\mathrm{x}}(\mathrm{t}) = \pi^{-1} \int_{-\infty}^{\infty} \frac{\mathrm{x}(\tau)}{\mathrm{t} - \tau} \mathrm{dt}, \qquad (2.2.7)$$

where $\tilde{x}(t)$is the HT of x(t)**.** Let us consider x(t) being a multicomponent signal, then it can be denoted as the sum of monocomponents of slow varying instantaneous amplitudes and frequencies:

$$x(t) = \sum_{l} a_l(t)\cos\left(\int \omega_l(t)dt\right), \tag{2.2.8}$$

where $a_l(t)$ represents the instantaneous amplitude (envelope) and $\omega_l(t)$ denotes the instantaneous frequency of the lth component. In order to estimate these parameters each time, in every successive iteration, we have to obtain the analytical signal, X(t), expressed by the summation of the real part along the imaginary part and by the exponential form (Feldman, 2011) both appearing in Eq. 7:

$$X(t) = x(t) + j\tilde{x}(t) = A(t)e^{j\varphi(t)}. \tag{2.2.9}$$

The instantaneous amplitude (envelope, magnitude) is then estimated as:

$$A(t) = |X(t)| = \sqrt{x^2(t) + \tilde{x}^2(t)} = e^{Re[\ln X(t)]}, \tag{2.2.10}$$

and its instantaneous phase as:

$$\varphi(t) = \arctan\left(\frac{\tilde{x}(t)}{x(t)}\right) = lm[\ln X(t)]. \tag{2.2.11}$$

From the first derivative of the instantaneous phase, we get the instantaneous angular frequency $\omega(t) = \dot{\varphi}(t)$. Now, the slow varying vibration component can be extracted by imposing a low pass filter on the

instantaneous envelope and frequency respectively that will help subtract the largest vibration component out of initial signal $x_{l-1}(t) = x(t) - x_l(t)$ and then treat $x_{l-1}(t)$ as the initial and repeat process until its termination. The cutoff frequency of the low pass filtering[7] is the one responsible for the frequency resolution of the HVD process and should be of a small value because on every successive iteration, after having subtracted the previous frequency, the frequency of the next to be extracted component dominates and so components with close frequencies must be able to be efficiently separated. The HVD process terminates when the standard deviation difference between two successive components becomes ≤ 0,01.

HVD has only few disadvantages. Braun and Feldman (2011), state that HVD is sensitive to additive noise, while Feldman (2008) report that HVD may not be efficient enough in short samples. In our study the sample was long enough to guarantee that decomposition process was optimal and terminated by providing 5 components for both implied and realized volatility. Overall, HVD constitutes an unpretentious process that is easily implemented and is considered to be computationally fast.

2.2.5 The empirical mode decomposition

Among the concepts that seem to dominate the signal processing field, lies the wavelet theory with a vast literature dedicated on wavelets' features and their use (Meyer, 1997; Daubechies, 1992). Inspired by the wavelet theory and the continuous and discrete wavelet transforms, Gilles (2013) proposed a new adaptive method for decomposing non-stationary and non-linear signals

[7] In our study we use the Elliptic low pass filter that fits perfect in order to efficiently manage the randomness that financial time series experience.

into amplitude modulated-frequency modulated (AM-FM) components, the EWT. Since then, EWT appeared in the analysis of power system signals, wind data, medical disease diagnosis, seismic data, machine fault diagnosis, image processing etc. (Liu *et al.,* 2016; Singh and Sunkaria, 2016; Beoula *et al,* 2017), but also, in combination with parametric and non-parametric frameworks for forecasting short-term wind speed, drought etc. (Hu and Wong, 2015; Shaari *et al.,* 2018).

EWT builds a family of wavelets that can easily be adapted to the signal to be decomposed. The very starting point and the most crucial one in the entire process, is how to segment the Fourier spectrum of the signal. Each segment denotes a mode, so there are equivalent segments corresponding to the inspected components. This very step seems to be and the most controversial as one has to predefine the modes the moment that any prior relevant information of the analyzed to be signal is absent. Thus, setting the boundaries of Fourier segments becomes demanding and various algorithms have been proposed to efficiently cope with this issue (Shi *et al*. 2017). The next step is to define the empirical scaling function and the empirical wavelets that will act as band pass filters on each predefined segment. Let us assume that we segment Fourier support [0, π] into N points. Let ω_n be the boundaries of the different segments with $\omega_0 = 0$ and $\omega_N = \pi$. Each segment lies in $[\omega_{n-1}, \omega_n]$. Around every ω_n a transition phase is defined, T_n, with width $2\tau_n$ and thus the empirical scaling function is defined by:

$$\widehat{\varphi}_n(\omega) = \begin{cases} 1 & \text{if } |\omega| \le \omega_n - \tau_n \\ \cos\left[\frac{\pi}{2}\beta\left(\frac{1}{2\tau_n}(|\omega| - \omega_n + \tau_n)\right)\right] & \text{if } \omega_n - \tau_n \le |\omega| \le \omega_n + \tau_n \\ 0 & \text{otherwise.} \end{cases} \quad (2.2.12)$$

And the empirical wavelets by:

$$\widehat{\psi}_n(\omega) = \begin{cases} 1 & \text{if } \omega_n + \tau_n \le |\omega| \le \omega_{n+1} - \tau_{n+1} \\ \cos\left[\frac{\pi}{2}\beta\left(\frac{1}{2\tau_{n+1}}(|\omega| - \omega_{n+1} + \tau_{n+1})\right)\right] & \text{if } \omega_{n+1} - \tau_{n+1} \le |\omega| \le \omega_{n+1} + \tau_{n+1} \\ \sin\left[\frac{\pi}{2}\beta\left(\frac{1}{2\tau_n}(|\omega| - \omega_n + \tau_n)\right)\right] & \text{if } \omega_n - \tau_n \le |\omega| \le \omega_n + \tau_n \\ 0 & \text{otherwise.} \end{cases} \quad (2.2.13)$$

Assuming $\beta(x)$ being an arbitrary function $C^k([0,1])$, where:

$$\beta(x) = \begin{cases} 0 & \text{if } x \le 0 \\ 1 & \text{if } x \ge 1 \end{cases} \text{and } \beta(x) + \beta(1-x) = 1 \quad \forall x \in [0,1]. \quad (2.2.14)$$

Setting τ_n proportional to ω_n, $\tau_n = \gamma\omega_n$, with $0 < \gamma < 1$, then above equations can be rewritten as:

$$\widehat{\varphi}_n(\omega) = \begin{cases} 1 & \text{if } |\omega| \le (1-\gamma)\omega_n \\ \cos\left[\frac{\pi}{2}\beta\left(\frac{1}{2\gamma\omega_n}(|\omega| - (1-\gamma)\omega_n)\right)\right] & \text{if } (1-\gamma)\omega_n \le |\omega| \le (1+\gamma)\omega_n \\ 0 & \text{otherwise.} \end{cases} \quad (2.2.15)$$

And

$$\widehat{\psi}_n(\omega) = \begin{cases} 1 & \text{if } (1+\gamma)\omega_n \le |\omega| \le (1-\gamma)\omega_{n+1} \\ \cos\left[\frac{\pi}{2}\beta\left(\frac{1}{2\gamma\omega_{n+1}}(|\omega| - (1-\gamma)\omega_{n+1})\right)\right] & \text{if } (1-\gamma)\omega_{n+1} \le |\omega| \le (1-\gamma)\omega_{n+1} \\ \sin\left[\frac{\pi}{2}\beta\left(\frac{1}{2\gamma\omega_n}(|\omega| - (1-\gamma)\omega_n)\right)\right] & \text{if } (1-\gamma)\omega_n \le |\omega| \le (1-\gamma)\omega_n \\ 0 & \text{otherwise.} \end{cases} \quad (2.2.16)$$

Now, in order to construct the EWT, denoted as $W_f^{\varepsilon}(n,t)$, the detail and approximation coefficients have to be estimated given by the inner products of empirical wavelets and scaling function respectively:

$$W_f^{\varepsilon}(n,t) = \langle f, \psi_n \rangle = \int f(\tau)\overline{\psi_n(\tau - t)}d\tau = (\hat{f}(\omega)\overline{\widehat{\psi}_n(\omega)})^V, \quad (2.2.17)$$

$$W_f^{\varepsilon}(0,t) = \langle f, \varphi_1 \rangle = \int f(\tau)\overline{\varphi_1(\tau - t)}d\tau = (\hat{f}(\omega)\overline{\widehat{\varphi}_1(\omega)})^V, \quad (2.2.18)$$

where $\widehat{\psi}_n(\omega)$ and $\widehat{\varphi}_1(\omega)$ are estimated above through Eq. and Eq. and the reconstruction:

$$\begin{aligned} f(t) &= W_f^{\varepsilon}(0,t) * \varphi_1(t) + \sum_{n=1}^{N} W_f^{\varepsilon}(n,t) * \psi_n(t) \\ &= (\widehat{W_f^{\varepsilon}}(n,\omega)\widehat{\psi}_n(\omega) + \sum_{n=1}^{N} \widehat{W_f^{\varepsilon}}(n,\omega) * \psi_n(\omega))^V. \end{aligned} \quad (2.2.19)$$

Finally, the empirical mode, f_k returned by the EWT is expressed as:

$$f_0(t) = W_f^{\varepsilon}(0,t) * \varphi_1(t), \quad (2.2.20)$$

$$f_k(t) = W_f^{\varepsilon}(k,t) * \psi_k(t). \quad (2.2.21)$$

The empirical modes returned out of EWT come in ascending frequency order. Process starts from separating the component with the lowest frequency and terminates to the one with the highest. Actually, the number of modes and the point that terminates process is user defined, after inspecting the Fourier spectrum. In our study, we defined 4 components for the VIX Index and 6 components for the realized volatility time series.

2.2.6The variational mode decomposition

The final decomposition method applied in our study is attributed to Dragomiretskiy and Zosso (2014), who proposed a non-recursive variational model that could address limitations rising from corresponding decomposition models such as EMD, HVD or EWT etc., the VMD. The use of this method has accelerated and many are the studies to incorporate VMD along neural networks, reconstruction techniques or optimization techniques in order to forecast oil price, short-term power load, financial data etc. (Lahmiri, 2016; Li *et al*., 2020; Ping Yu *et al*., 2021).

Dragomiretskiy and Zosso (2014), utilize three concepts, the Wiener filtering, the HT and the frequency mixing and heterodyne demodulation in order to synthesize the VMD process that decomposes a real signal f(x) into distinct modes. Each mode is required to be compact around a central frequency, ω_k determined through the process, so, they determine the bandwidth of the modes by following three steps, 1) they compute analytic signal of each mode through HT and obtain a unilateral frequency spectrum, 2) they shift mode's frequency spectrum to baseband, by mixing it with an exponential tuned to the respective estimated center frequency, and 3) they estimate the bandwidth through the H1 Gaussian smoothness of the demodulated signal. So, they end up in a constrained variational problem:

$$\min_{u_k,\,\omega_k,}\left\{\sum_k\left\|\partial_t\left[\left(\delta(t)+\frac{j}{\pi t}\right)*u_k(t)\right]e^{-j\omega_k t}\right\|_2^2\right\}\ \text{s.t.}\ \sum_k u_k=f(x), \qquad (6)$$

where u_k denotes the modes, ω_k denotes their central frequencies, * denotes convolution and the summation of all u_k approximates original signal. The solution to above problem is given through Lagrangian multipliers, λ and a quadratic penalty term with the assistance of the alternate direction method of multipliers (ADMM) (Gabay and Mercier, 1976; Eckstein and Bertsekas, 1992) an iterative sub-optimization method, so the Lagrangian takes the form:

$$\mathcal{L}(\mathrm{u_k}, \omega_\mathrm{k}, \lambda) = \mathrm{a} \sum_\mathrm{k} \left\| \partial_\mathrm{t} \left[\left(\delta(\mathrm{t}) + \frac{\mathrm{j}}{\pi \mathrm{t}} \right) * \mathrm{u_k}(\mathrm{t}) \right] \mathrm{e}^{-\mathrm{j}\omega_\mathrm{k}\mathrm{t}} \right\|_2^2 + \left\| \mathrm{f}(\mathrm{t}) - \sum_\mathrm{k} \mathrm{u_k}(\mathrm{t}) \right\|_2^2 + \langle \lambda(\mathrm{t}), \mathrm{f}(\mathrm{t}) - \sum_\mathrm{k} \mathrm{u_k}(\mathrm{t}) \rangle \,. \tag{7}$$

The process returns the predefined from the user modes. The Lagrangian multiplier when used along with the penalty term ensures that the summation of all returned components will reconstruct actual signal not perfect but approximately close. The inclusion of only the penalty term will slightly vary final outcome but will deal more sufficiently with noisy data. VMD is not without flaws, concerning the extraction of non-periodic components, the discontinuity of signals and the decomposition of wideband spectra (Civera and Surace, 2021) but in the end no technique is perfect. Regarding our study, after experimentation on the number of components that best approximates original series, we adopted the version that results in 5 modes for both inspected indices.

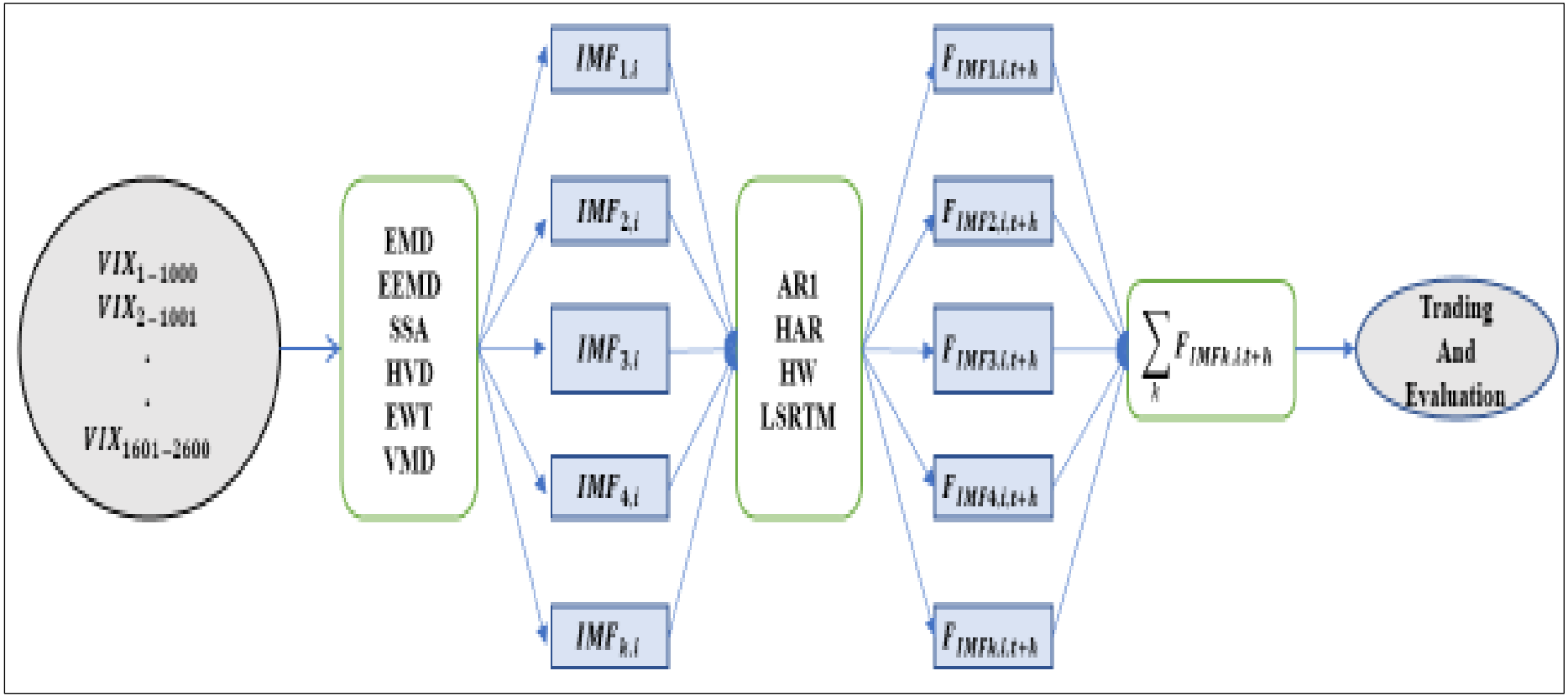
2.3Modelling and forecasting frameworks

The second step in the main target of the thesis, is to fit optimal models to the individual components each of the predefined decomposition methods granted for the two measures. The fitting is conducted according to the

distinctive nature of the diverse components after excessive experimentation, so the following models were not randomly selected. Forecasts for each of the separate components, are generated based on the model specification applied to each component and are processed by incorporating a rolling window approach with a fixed window of 1000 observations. The initial data sample of VIX index consists of $T = 2600$ trading days, while for the realized measure consists of $T = 2021$ trading days. Thus, we use the first $\tilde{T} = 1000$ observations for decomposition and for in-sample estimation, that is from 20th of August 2012 up to 12th of August 2016 for both time series. Then, we use the rest $\check{T} = 1600$ observations up to the 30th of November 2022, as for the components of VIX index and the rest $\check{T} = 1021$ observations up to 31st of August 2020 as for the components of the realized measure, for generating out-of-sample iterated forecasts for 1, 5, 10 and 22 days ahead horizon. The sample of the realized measure is dictated from data availability. Each day on the out-of-sample period, we decompose VIX and realized volatility, then model components and forecast for a completely unknown future path that alters each consecutive day. The forecasts produced for the individual components of each decomposition method, are then summed together in order to form the final aggregated forecasts for each trading day, for each technique and for each time series. Thus, study refrains entirely from using any future information and committing the often phenomenon of data leakage, best known as look ahead bias, during the decomposition process and inevitably during model estimation and forecast generation. Forecasts, are produced based on data that belong to the information set at time t.

All four models proposed in this section, as best fitting for our divide and conquer structure, are also applied to original series in order to investigate whether this tactic indeed adds in forecast accuracy not from a statistical perspective, but from the objective-based economic point of view that we apply. We end up with a total of 32 models for the case of VIX and 12 models for realized. When disaggregating VIX for the EMD approach we get 4 models, for EEMD we get 5 models, for SSA we end up with 6 models, for HVD with 5 models, for the EWT case, we have 4 models and finally for VMD we get 4 models. When disaggregating realized volatility, for the EEMD framework we have 3 models, for the SSA method, we get 2 models, for the HVD we have again 2 models and also two models we get for each of EWT and VMD techniques, respectively. The tables with the complete selection of the models, appear at the chapter of empirical findings, that is chapter 3.3 for VIX and 4.4 for the realized metric. The models of decomposition methods are reported based to each method's abbreviations. The 4 models of VIX index, as a unity that are included for comparison reasons, are conceived as benchmark models and the same is valid for the case of realized volatility. For components that constitute a white noise process, of course no modelling takes place. Fig. 2.1 illustrates superficially how the entire process is being conducted.

Figure 2.1 Decomposition, modelling and forecasting process



Note: The 1600 rolling samples of VIX with a fix window of 1000 observations are decomposed, modelled, and forecasted in order to get the aggregate forecasts of VIX for the forecast horizons of 1, 5, 10 and 22 days ahead. Then evaluation is conducted based on statistical and economic criteria. From the above, k stands for the number of components, i = {EMD, EEMD, SSA, HVD, EWT, VMD} and h is the forecast horizon, with h = {1,5,10,22} days ahead. For brevity, symbolism refer to the VIX index decomposition process, but the same applies for realized volatility as well.

2.3.1 The autoregressive model

The first model incorporated is the Autoregressive model (AR). The inclusion of an intrinsic econometric model along the more advanced ones in an ensemble modeling, helps to model and forecast components that bare a first or higher order autoregressive pattern. For brevity, we present the model's specification for the case of the first order, for modelling the logarithm of VIX components, for each of the decomposition techniques[8], that is:

$$\log\left(\mathrm{VIX}_{\mathrm{i,j,t}}\right) = \mathrm{w}_0^{(\mathrm{t})}\left(1-\varphi_1^{(\mathrm{t})}\right) + \mathrm{w}_1^{(\mathrm{t})}\log\left(\mathrm{VIX}_{\mathrm{i,j,t-1}}\right) + \varepsilon_{\mathrm{t}}, \tag{2.3.1}$$

where, $\mathrm{w}_0^{(\mathrm{t})}$ and $\mathrm{w}_1^{(\mathrm{t})}$ represent the rolling estimated coefficients, i is the number of the component, $\mathrm{j} = \{\mathrm{EMD}, \mathrm{EEMD}, \mathrm{SSA}, \mathrm{HVD}, \mathrm{EWT}, \mathrm{VMD}\}$ and ε_t are the residuals that thought to be normally distributed, $\varepsilon_{\mathrm{t}} \sim \mathrm{N}(0, \sigma_\varepsilon^2)$. When modeling and forecasting is conducted on realized volatility's components, then $\mathrm{VIX}_{\mathrm{i,j,t-1}}$ is substituted from $\mathrm{RV}_{\mathrm{i,j,t-1}}$ to adequately denote the components of the realized measure. Forecasting for the t+h horizon, with h = {1,5,10,22}is conducted through:

$$\mathrm{VIX}_{\mathrm{i,j,t+h|t}} = \exp\left(\mathrm{w}_0^{(\mathrm{t})}\left(1-\varphi_1^{(\mathrm{t})}\right) + \mathrm{w}_1^{(\mathrm{t})}\log\left(\mathrm{VIX}_{\mathrm{i,j,t+h-1}}\right)\right. . \tag{2.3.2}$$

Now, concerning the disaggregation of VIX and for the case of the EMD method and for the models of EMD-HW-AR, EMD-LSTM-AR and EMD-LSTM-AR-HAR, components 2, 3 and 4, components 2, 3 and 4 and

[8] Here equation is specified for VIX components. When modelling and forecasting the VIX as a unity, then specification becomes $\log(\mathrm{VIX}_{\mathrm{t}})$. The same applies for the rest of the benchmark models, but for brevity we refrain from complete presentation in the relevant sections and the same is valid when realized volatility is modelled and denoted as $\log(\mathrm{RV}_{\mathrm{t}})$.

component 2, respectively, are modelled and forecasted via AR processes. For the EEMD method and for models EEMD-HW-AR, EEMD-HW-AR-HAR, EEMD-LSTM-AR and EEMD-LSTM-AR-HAR, components 2, 3 and 4, component 2, components 2, 3 and 4 and component 2, respectively, were modelled and forecasted via AR processes. For the case of SSA and for the models of SSA-LSTM-AR and SSA-HW-AR, components 2 and 3, for both, are modelled and forecasted via AR processes. For the HVD method and for the models of HVD-HW-AR, HVD-HAR-AR, HVD-AR1-AR and HVD-LSTM-AR, components 3, 4 and 5, components 2, 3, 4 and 5, components 1, 2, 3, 4 and 5 and components 2, 3, 4 and 5, respectively, are modelled and forecasted through AR processes. To continue, for the case of EWT and for models EWT-HW-AR and EWT-LSTM-AR, components 3 and 4, for both, are modelled and forecasted based on AR processes. Finally, for VMD, and for the models VMD-HW-AR and VMD-LSTM-AR, components 3, 4 and 5, for both are modelled and forecasted via AR processes.

As for the disaggregation of realized volatility that returned fewer models for all proposed decomposition frameworks, for EEMD and EEMD-HW-AR and EEMD-LSTM-AR, components 2, 3 and 4 for both models, were modeled and forecasted via AR processes. For the HVD and HVD-AR all components were modelled and forecasted with AR processes. For the EWT and EWT-HW-AR, components 2, 3, 4, 5, 6, 7 were modelled and forecasted through AR processes.

2.3.2 The heterogeneous autoregressive model

Heterogeneous autoregressive model, HAR, the state-of-the-art framework for volatility modelling of Corsi (2009), is an additive linear combination of

indicators of volatility components at different time horizons, fully aligned with markets' fractal structure. HAR has proven to be one of the best performing models for generating forecasts, especially when these forecasts entail asset allocation aspects (Degiannakis and Filis, 2022). The simple HAR for the logarithm of VIX components, for each of the decomposition techniques, takes the form:

$$\log\left(\mathrm{VIX}_{i,j,t}\right) = w_0 + w_1 \log\left(\mathrm{VIX}_{i,j,t-1}\right) + w_2\left(5^{-1}\sum_{k=1}^{5}\log\left(\mathrm{VIX}_{i,j,t-k}\right)\right) + w_3\left(22^{-1}\sum_{k=1}^{22}\log\left(\mathrm{VIX}_{i,j,t-k}\right)\right) + \varepsilon_t, \tag{2.3.3}$$

where, w_0, w_1, w_2 and w_3 denote the rolling estimated coefficients and ε_t a normally distributed process, $\varepsilon_t \sim N(0, \sigma_\varepsilon^2)$. Again here, to avoid text complexity, we should mention that for the case of the realized volatility, the same model specifications are valid and $\mathrm{VIX}_{i,j,t-k}$ is replaced by $\mathrm{RV}_{i,j,t-k}$. Since we have used the log form for model estimation, forecasts for the h-days-ahead horizon, are given by:

$$\begin{aligned}\mathrm{VIX}_{i,j,t+h|t} = \exp\Big(&\widehat{w}_0 + \widehat{w}_1 \log\left(\mathrm{VIX}_{i,j,t+h-1|t}\right) \\ &+ \widehat{w}_2\left(s^{-1}\sum_{k=1}^{s-1}\log(\mathrm{VIX}_{i,j,t-k+h|t}) + (5-h)^{-1}\sum_{k=s}^{5}\log\left(\mathrm{VIX}_{i,j,t-k+h}\right)\right) \\ &+ \widehat{w}_3\left(s^{-1}\sum_{k=1}^{s-1}\log(\mathrm{VIX}_{i,j,t-k+h|t}) + (22-h)^{-1}\sum_{k=s}^{22}\log\left(\mathrm{VIX}_{i,j,t-k+h}\right)\right) + {}^{1}/_{2}\,\widehat{\sigma}_\varepsilon^2\Big).\end{aligned} \tag{2.3.4}$$

Through the HAR framework, when VIX Index is investigated, the following components are modelled and forecasted, 1) from the EMD and the model EMD-LSTM-AR-HAR, components 3 and 4, 2) from the EEMD and models EEMD-HW-AR-HAR and EEMD-LSTM-AR-HAR, components 3 and 4, for both, 3) from SSA and models SSA-LSTM-HAR and SSA-HW-HAR, components 2 and 3, for both, and 4) from HVD and model HVD-HAR-AR, component 1. When realized volatility is the object of interest, only the SSA concept incorporates the HAR framework and more specifically the model SSA-HAR where the first components is modelled and forecasted via HAR.

2.3.3The holt winters model

Apart from HAR and AR models, we further employ the Holt Winters (HW) framework as some components present features that best apply to HW specification. HW is a simple univariate procedure for producing forecasts based on the past and current values of a time series that utilizes triple exponential smoothing and allows to deal with both seasonal variation and trend (Winters, 1960; Chatfield, 1978). HW unfolds in two versions[9], an additive and a multiplicative where the seasonal effects are thought to be of constant size or proportional to the local mean respectively (Chatfield, 1978). There are three smoothing constants α, β, γ. The model specifications for the components of VIX, for the two distinctives versions have the following form[10]:

Additive

[9] Additive and multiplicative versions apply depending on the data and their values. For running the HW model we made use of the statsmodels package of the python programming language.

[10] Again $\log(VIX_{i,j,t})$ can be rewritten as $\log(VIX_t)$ to account for VIX as a unity.

$$\widehat{M}_t = \hat{a}_t\left(\log\left(VIX_{i,j,t}\right) - \widehat{T}_{t-c}\right) + (1-\hat{a})\left(\widehat{M}_{t-1} + \hat{F}_{t-1}\right),$$

$$\hat{F}_t = \hat{\beta}\left(\widehat{M}_t - \widehat{M}_{t-1}\right) + (1-\hat{\beta})\hat{F}_{t-1} \tag{2.3.5}$$

$$\widehat{T}_t = \hat{\gamma}_t\left(\log\left(VIX_{i,j,t}\right) - \widehat{M}_{t-1} - \hat{F}_{t-1}\right) + (1+\hat{\gamma})\widehat{T}_{t-c}.$$

Multiplicative

$$\widehat{M}_t = \hat{a}\left(\frac{\log(VIX_{i,j,t})}{\widehat{T}_{t-c}}\right) + (1-\hat{a})\left(\widehat{M}_{t-1} + \hat{F}_{t-1}\right),$$

$$\hat{F}_t = \hat{\beta}\left(\widehat{M}_t - \widehat{M}_{t-1}\right) + (1-\hat{\beta})\hat{F}_{t-1} \tag{2.3.6}$$

$$\widehat{T}_t = \hat{\gamma}\left(\frac{\log\left(VIX_{i,j,t}\right)}{\widehat{M}_t}\right) + (1-\hat{\gamma})\widehat{T}_{t-c},$$

where, $\widehat{M}_t$, $\hat{F}_t$, $\widehat{T}_t$ stand for the estimate of the de-seasonalized mean for time t, the estimated seasonal factor at time t and the estimated trend term at time t, respectively. The c term that appears as subindex in above system of equations denotes the number of observations included in a seasonal cycle (Chatfield, 1978). Again, as for all the models' specifications, $VIX_{i,j,t}$ is replaced by $RV_{i,j,t}$, to denote the methods and the components retrieved from each, for the case of realized volatility. Now, forecasts from the HW procedure can be generated for any horizon. Hence, forecasts of h days ahead of components in logarithmic form, for the additive and multiplicative versions respectively, are given by the formulas:

Additive

$$VIX_{i,j,t+h|t} = \exp\left(\widehat{M}_t + h\hat{F}_t + \widehat{T}_{t+h-c}\right). \tag{2.3.7}$$

Multiplicative

$$VIX_{i,j,t+h|t} = \exp\left(\widehat{M}_t + h\hat{F}_t\right) * \widehat{T}_{t+h-c}. \tag{2.3.8}$$

Through the HW method, the following components, coming from the 6 decomposition techniques, are modelled and forecasted, 1) from EMD and models EMD-HW and EMD-HW-AR, components 2, 3, 4, 5 and 6[11] and components 5 and 6[12], respectively, 2) from EEMD[13] and models EEMD-HW, EEMD-HW-AR and EEMD-HW-AR-HAR, components 2, 3, 4, 5 and 6, components 5 and 6 and components 5 and 6, respectively, 3) from SSA and models SSA-LSTM-HW, SSA-HW, SSA-HW-HAR and SSA-HW-AR, components 2 and 3, components 1, 2 and 3, component1 and component 1, respectively, 4) from HVD and models HVD-HW and HVD-HW-AR, components 1, 2, 3, 4 and 5 and components 1 and 2, respectively, 5) from EWT and models EWT-HW, EWT-HW-AR and EWT-LSTM-HW, component 1, 2, 3 and 4, component 1 and components 3 and 4, respectively, finally 6) from VMD and models VMD-HW, VMD-HW-AR and VMD-LSTM-HW, components 1, 2, 3, 4 and 5, component1 and components 3, 4 and 5 respectively.

When the time series of realized volatility is under investigation, then, 1) for the EEMD method, from the EEMD-HW model, all components were modelled and forecasted under HW framework, except for the first component that was naturally found to be best characterized by a white noise process, while from the EEMD-HW-AR model, components 5 and 6 were modelled and forecasted through HW, 2) for the SSA approach and SSA-HW model, all components were modelled and forecasted via HW, except again for the first

[11] When a rolling sample returns 5 components, then components 2, 3, 4 and 5 are modelled and forecasted through the HW framework.
[12] When a rolling sample returns 5 components, then components 4 and 5 are modelled and forecasted via HW.
[13] For EEMD as is the case for EMD, when a rolling sample returns 5 components, then components 2, 3, 4 and 5 are processed through HW for the case of EEMD-HW, components 4 and 5 for EEMD-HW-AR and EEMD-HW-AR-HAR, respectively.

one, 3) for the HVD and HVD-HW model applies the same as in the SSA-HW and EEMD-HW case, 4) for the EWT concept and EWT-HW and EWT-HW-AR, components 1, 2, 3, 4, 5, 6, and 7 and components 1 and 2 respectively, were modelled and forecasted through HW framework, 5) finally, for the case of VMD and models VMD-HW and VMD-HW-AR, components 1, 2, 3 and 4 and, components 1 and 2 were modelled and forecasted under HW.

2.3.4 The long-short term model

In the field of deep learning algorithms, there stands out the long short-term memory model (LSTM) (Hochreiter and Schmidhuber, 1997), that is the fourth framework we incorporate for modelling and forecasting the VIX components. Part of recursive neural networks (RNNs), LSTM came as a solution in a major drawback of RNNs, their inability to produce accurate predictions attributed to the explosive gradient descent when trying to learn the long-term dependencies of a series (Bengio *et al*., 1994). LSTM ever since has proven to be a powerful enough tool for time series forecast scenarios and many are the studies to include it (Michańków *et al*., 2022; Shu and Gao, 2020; Liu *et al*., 2022). Its main advantage lies in the addition of extra memory cells, extra layers, consisting of three gates, the input gate, ig_t, the output gate, og_t and the forget get, fg_t. Memory cells with the help of activation functions[14], such as sigmoid and hyperbolic tangent function, act as filters concerning the information that will eventually reach the next cell state. The mathematical expression of how information flows inside each memory cell and the role each gates plays is presented in the following equations:

[14] For setting the model and its parameters (layers, optimizers, epochs, etc.), we use the keras deep learning API in python programming language (keras.io) that is built on top of TensorFlow, a machine learning platform.

$$\mathrm{ig_t} = \mathrm{sigm}(\mathrm{weight_{ig}} \cdot [\mathrm{h_{t-1}}, \mathrm{x_t}] + \mathrm{bias_{ig}})$$

$$\mathrm{fg_t} = \mathrm{sigm}(\mathrm{weight_{fg}} \cdot [\mathrm{h_{t-1}}, \mathrm{x_t}] + \mathrm{bias_{fg}}) \quad (2.3.9)$$

$$\tilde{\mathrm{m}}_\mathrm{t} = \mathrm{hyptan}(\mathrm{weight_m} \cdot [\mathrm{h_{t-1}}, \mathrm{x_t}] + \mathrm{bias_m}),$$

where, $\tilde{\mathrm{m}}$ is the value of the memory cell in order to update the cell, $\mathrm{x_t}$ is the input data at time t, $\mathrm{h_{t-1}}$ is the output of the previous layer that enters as input to present state, the weight with subindices of the different gates, denotes the weights set every time by forward and back propagation (Hecht-Nielsen, 1992), that try to minimize the error condition[15] through successive efforts and finally, the bias term stands for the error term. The sigm and hyptan symbols, appearing in Eq. 2.3.11 and 2.3.12 denote the sigmoid and the hyperbolic tangent functions[16] respectively. The new value of the cell becomes:

$$\mathrm{c_t} = \mathrm{fg_t} * \mathrm{c_{t-1}} + \mathrm{ig_t} * \tilde{\mathrm{m}}_\mathrm{t}, \quad (2.3.10)$$

where * denotes the convolution and c_{t-1} denotes the value of the previous cell state. Now, the value of the output gate, $\mathrm{og_t}$, is given by:

$$\mathrm{og_t} = \mathrm{sigm}(\mathrm{weight_{og}}[\mathrm{h_{t-1}}, \mathrm{x_t}] + \mathrm{bias_{og}}, \quad (2.3.11)$$

Finally, the above steps result in the final layer's filtered output that is of the form:

$$\mathrm{h_t} = \mathrm{og_t} * \mathrm{hyptan}(\mathrm{c_t}). \quad (2.3.12)$$

The above procedure provides the steps occurring in each memory cell and the way they control information, progressively over and over again among layers until the final estimated value that minimizes the error term is

[15] LSTM utilizes different types of cost functions. A frequently used loss function is the "mean squared error" and is the one we also incorporate for constructing our model.

[16] For instance, the hyptan is a function that receives values between -1 and 1 and is denoted by $hyptan = \frac{e^{2x}-1}{e^{2x}+1}$, where x is a linear combination of the variables.

eventually reached. Perhaps it sounds a little complicated, but that is how RNNs work and LSTM being the improved version of RNNs, holds their core architecture. Now, LSTM contrariwise to RNNs thanks to its lookback period window and its ability to learn both the short-term and the long-term data characteristics, can generate multiple days ahead forecasts. Having produced the 1 day ahead forecast, incorporates forecasted value, updates data set and moves on to the next and so on, until all values for the specified horizon are gathered:

$$\mathrm{h_{t+h}} = \mathrm{og_{t+h}} * \mathrm{hyptan}(\mathrm{c_{t+h}}). \tag{2.3.13}$$

In this study the inclusion of LSTM is used for modelling and forecasting components that their nature excludes more elementary frameworks. Hence, through LSTM for the case of VIX index the following components were modelled and forecasted, 1) from EMD and models EMD-LSTM-AR and EMD-LSTM-AR-HAR, components 5 and 6, for both[17], 2) from EEMD and models EEMD-LSTM-AR and EEMD-LSTM-AR-HAR components 5 and 6 for both case, 3) from the SSA and models SSA-LSTM-AR, SSA-LSTM-HAR and SSA-LSTM-HW, component 1 for all three models, 4) from HVD and model HVD-LSTM-AR, component 1, 5) from EWT and models EWT-LSTM-HW and EWT-LSTM-AR, components1 and 2 for both and 6) from VMD and models VMD-LSTM-AR and VMD-LSTM-HW, component 1 and 2, for both. For the case of realized volatility only components 5 and 6 of the EEMD-LSTM-AR were modelled and forecasted through LSTM for horizon of 1 and 5 days. The nature of realized volatility resulted in not so optimal

[17] Here applies the same we mentioned at HW's section about components 5 and 6 of EMD and EEMD processes.

outcome so we refrained from including extra models of components modelled and forecasted via LSTM.

2.4 Criteria for forecast evaluation and model selection

2.4.1 Forecast evaluation criteria

One of the main scopes of this dissertation is to point out whether the proposed modelling architecture, can actually result in optimal forecast performance. We specify optimality through the production of meaningful and profitable forecasts for market participants, who engage dynamically in derivatives markets. Hence, in order to quantify forecast performance, we employ an objective-based forecast evaluation criterion, an economic criterion, along the classic statistical criterion of mean squared error, MSE, which inevitably appears in most studies that base comparison of their findings on its value. So, we develop an objective-based evaluation criterion shaped on the tradable part of VIX, the VIX futures. We should mention here, that regarding the fact that we also forecast realized volatility as of SP500 futures Index returns, a standard partner of VIX in studies of multivariate settings, the trading strategy will be also conducted on VIX futures, that are strongly linked to VIX Index and the equity market, due to the diverse properties they display for those enclosing these derivatives in their portfolios. But prior to unfolding our investing strategies, let us specify MSE that is estimated through:

$$\mathrm{MSE} = \widetilde{\mathrm{T}}^{-1} \sum (\mathrm{VIX}_{\mathrm{i,t+h|t}} - \mathrm{VIX}_{\mathrm{t+h}})^2. \qquad (2.3.14)$$

where, $\mathrm{VIX}_{\mathrm{i,t+h|t}}$ and $\mathrm{VIX}_{\mathrm{t+h}}$ are the forecasted and the actual values of VIX volatility index, respectively. $\widetilde{\mathrm{T}}$ is the out-of-sample forecast period and the

subindex i denotes the i=1, 2, …, 32 different models for the VIX case and i=1, 2, …, 11 for the realized volatility and it goes without saying that when we forecast realized volatility, $\mathrm{VIX}_{\mathrm{i,t+h|t}}$, becomes $\mathrm{RV}_{\mathrm{i,t+h|t}}$ and $\mathrm{VIX}_{\mathrm{t+h}}$ becomes $\mathrm{RV}_{\mathrm{t+h}}$ and, the same applies to the rest equations that follow, that for brevity are expressed only on VIX. Now, according to the objective-based criterion, we follow a simple, yet powerful trading rule best described through the following two conditions:

1. If $\mathrm{VIX}_{\mathrm{i,t+h|t}} > \mathrm{VIX}_{\mathrm{t}}$ then go long on VIX futures.
2. If $\mathrm{VIX}_{\mathrm{i,t+h|t}} < \mathrm{VIX}_{\mathrm{t}}$ then go short on VIX futures.

Depending on the forecasted values generated either from VIX or the aggregated forecasted components[18], every day for the next 1 to 22 days, for the entire out-of-sample period, we hold long or short positions on VIX futures. Short or long positions are translated into selling or purchasing VIX futures respectively. Thus, depending on the outcome of the trading rule, we calculate the cumulative returns for each model combination for the 1600 trading days[19]. The cumulative returns (CR) are estimated based to:

$$\mathrm{CR}_{\mathrm{i,h}} = \sum_{\mathrm{t=1}}^{\widetilde{\mathrm{T}}} \left(\mathrm{I}_{\mathrm{i,t}} \times \frac{\left(\mathrm{VIX}_{\mathrm{i,t+h|t}} - \mathrm{VIX}_{\mathrm{t}}\right)}{\mathrm{VIX}_{\mathrm{t}}} \right) \text{ and } \mathrm{I}_{\mathrm{i,t}} = \begin{cases} 1 & \text{if} \quad \mathrm{VIX}_{\mathrm{i,t+h|t}} > \mathrm{VIX}_{\mathrm{t}} \\ -1 & \text{if} \quad \mathrm{VIX}_{\mathrm{i,t+h|t}} \leq \mathrm{VIX}_{\mathrm{t}}. \end{cases} \tag{2.3.15}$$

It is important here to highlight a fact. We did not naively go for the VIX futures series, nor did calculations that would result in misleading profits or losses, as VIX futures are very powerful instruments in the hands of investors. We considered the way VIX futures market operate and considered the roll of

[18] Or realized volatility and its aggregated forecasted components of the different techniques.
[19] Or the 1021 trading days for the realized volatility index.

VIX futures. The third Tuesday of each month, the expiration of the present month contract takes place. The Wednesday that follows, the front month's contract is set in action. There is when the roll cost or the roll yield occurs and can be a source of losses or gains, due to the difference in the price levels between the two contracts with the spot price of VIX. When the price of the contract that expires is lower than the next month's contract (contango), then rolling comes with the cost of paying more in order to swap in the next month's contract. When the price of the contract that expires is higher than the next month's contract (backwardation), then rolling incurs gains due to the roll yield. Here, in this trading exercise, we actually take into consideration the roll of VIX futures no matter if that entails gains or losses and report the cumulative returns an investor incurs for the entire period under investigation. So, we how much this is gone to cost us in total, if not to benefit us.

Table 2.1 presents the price of the active contract and the price of next month's contract at the expiration date of the randomly chosen third week of the October of 2022. The price of October's contract is 30.86, while the one for November is 30.15. The following day the price of October's contract appears in the official records, while November is the active one. So, no matter what the final outcome shall be, for an investor that chooses as a settlement day the final trading day for the asset, she/he should close her/his position on October's contract and go for the November's on Tuesday.

Table 1.1 VIX futures contracts prices

18/10/2022	**Tuesday:** Expiration date of October's contract	**October's contract:** 30.86 **November's contract:** 30.15
19/10/2022	**Wednesday:** First day of November's contract	**October's contract:** 31.77 (**expired**) **November's contract:** 30.40 (**active**)

Note: Ttable reports the prices of VIX futures contracts at the expiration and the following day of the expiration. In most studies October's price is reported for both days, but especially for Wednesday where October's have ceased to exist is more appropriate November's to be officially reported since is the one traded. The values reported here, were retrieved from vixcentral.com.

2.4.2The model confidence set

A study that faces a multiple comparisons problem and has to deal with 32 models for the implied volatility and 11 models of realized and diverse forecast horizons, would be somehow diminished, if it did not employ a model selection procedure such the one proposed by Hansen *et al*. (2011), the model confidence set (MCS) test. MCS test, has turned out to be a valuable tool with several advances over other identical tests[20]. MCS test identifies the set of the best models, in terms of the evaluation criterion applied in a study and treats entire set of models equally and does not test against the value of a benchmark model. Here we have the statistical loss function of MSE and the economic criterion. MCS test will provide the p-values that will tell of the set of the best performing models or the ones that inevitably will be excluded, depending of course on the prespecified significance level. Thus, in the tables of findings presented in sections 3.3 and 4.4, we also report the p-values of the MCS test that help in justifying the density of the outcome.

[20] Here we refer to the superior predictive ability test (SPA) of Hansen (2006) that again is used in identical practices but it proceeds with its estimations for model selection, based on the values a benchmark model returned for a statistical or economic loss function.

2.4.3 Direction-of-change

Since in our study, we use two contradictive evaluation measures, with the one being an economic criterion directly involving profits and losses, we are somehow forced to include an extra forecast evaluation technique, the Direction-of-Change (DoC). DoC reports a proportion, that is the times that a forecast correctly predicted the direction of the implied volatility index. DoC becomes a substantial feature, especially for trading exercises such the one we engage in, due to our intention to test the ability of a forecast to generate economic profit. Economic profits from long or short positions accelerate only when correctly predicting market's direction. Hence, in order to validate the significance of our findings, in sections 3.3 and 4.4, we also report the values of DoC for the entire set of models. And being a little more precise with the conception of DoC, the proportion value, PV_i, for each model, i=1, 2…,32, and i=1, 2, …, 11 for the cases of VIX and S&P500 realized volatility respectively, is the outcome of the following process, where P_i, is a dummy variable that reports the times the forecast of the model i, correctly predicts the volatility movement. In Eq. 2.3.16, we present the test equation for the components of VIX, when realized volatility is studied, then $\mathrm{VIX}_{\mathrm{i,t+h|t}}$ is replaced by $\mathrm{RV}_{\mathrm{i,t+h|t}}$:

$$\mathrm{P_i} = \begin{cases} 1 \text{ if } \mathrm{VIX_{i,t+h|t}} > \mathrm{VIX_t} \\ 1 \text{ if } \mathrm{VIX_{i,t+h|t}} < \mathrm{VIX_t} \\ 0 \text{ otherwise,} \end{cases} \tag{2.3.16}$$

and

$$\mathrm{PV_i} = \frac{\sum_{i=1}^{\tilde{T}} \mathrm{P_i}}{\tilde{T}}. \tag{2.3.17}$$

Chapter 3

Disaggregating implied volatility

3.1 Introduction

Since the mid 00's when VIX futures and options were introduced as the tradable part of the VIX Cboe volatility index, the leading volatility index for the U.S. stock market, trading volumes on these derivatives have reached tremendous heights, especially during turbulent, for the global economy, moments. The reason lies in the special features these products display, especially futures and the strong ties with VIX index and more generally the equity market (Szado, 2020; Fernandes *et al.,* 2014). These characteristics make them quite appealing especially for investors and risk managers, who use them daily in order to hedge uncertainty and diversify portfolios, transforming trading strategies based on short and long volatility exposure in options and futures contracts into top financial risk management tools. Of course, the use of derivatives encloses high risk of capital loss and demand careful manipulation. Thus, engaging in derivatives markets' practices, requires the ability to accurately forecast the underlying implied volatility index, which is the key input for pricing them (Degiannakis *et al*., 2018).

Forecasting implied volatility, was and will continue to be a demanding and of immense importance task, not only for those directly involving in markets but also for stakeholders, researchers, and policy makers, who desire to navigate through past and present trends, but in the same moment desire to shape future regimes.

But what actually is implied volatility? Implied volatility is the market's current price for the expected volatility of an underlying asset. It is the volatility implied by the actual prices of options or futures trading on that underlying asset. Its calculation dates back to the cult models of Black and Scholes (1973) and Hull and White (1987) and was embellished through years to receive its later form. Here we incorporate VIX Cboe volatility index, the leading indicator of the broad U.S. stock market that is widely applied and is considered to be an unbiased and efficient proxy of implied volatility (Bollerslev *et al.* 2009; Bollerslev *et al.* 2011; Bekaert and Hoerova, 2011). VIX index is calculated accurately by using the midpoint of real time S&P500 (SPX) index options bid/ask quotes[21], providing an instantaneous measure of the S&P500 stock's index expected variation for the following 30 days from each tick. So, VIX forms an expectation[22] of markets' volatility and is the key factor for pricing VIX derivatives. Here, we exploit the disaggregation of VIX Index and then concentrate in the modeling and forecasting of the individual components, in order to invest in VIX futures.

But someone will argue why decompose and forecast VIX in order to invest in VIX futures. Well, VIX constitutes a major informative tool for implementing trading strategies with options and futures and has different structure from other volatility measures. Perhaps it is more sensitive to market movements, but by modelling VIX we avoid the frictions imposed by the constitution of VIX futures time series. Futures are characterized by jumps

[21] Detailed description on the methodology applied for VIX calculation is provided through Cboe's site. The interested reader could follow the provided link:
https://cdn.cboe.com/resources/vix/VIX_Methodology.pdf

[22] VIX is a widely recognized index and the way it is calculated has been widely covered and analyzed in numerous studies and for practical reasons it is omitted from the present study. For the not so informed user, one can always resort to https://cdn.cboe.com/resources/vix/VIX_Methodology.pdf where there is extensive analysis on the methodology is being applied for index's construction.

and discontinuities, since every month the tracked product alters, so it is not a proper instrument for modeling.

These facts make VIX more appropriate, especially for the proposed framework. On the other hand, VIX futures have unique return drivers (Moran and Dash, 2007) and unique properties (Szado, 2018). Their returns are highly correlated to VIX, but negatively correlated with equites not only necessarily the S&P500 equites, but also equites from major stock indices. Furthermore, what makes them extremely alluring is the expiration and the roll, that we already touched upon in the previous chapter, as VIX futures most of the times are in contango and only few in backwardation resulting in the roll cost and yield respectively. Overall, they constitute a powerful tool with accelerating trading volumes and whether they are used for hedging, capitalization, or portfolio diversification, are better fitted for volatility allocation practices. So, modeling occurs on VIX and trading on VIX futures. Afterall, there is a distinctive literature highlighting the importance of generating accurate in a profitable manner forecasts of implied volatility, for an efficient portfolio manipulation and diversification based on trading implied volatility derivatives (Degiannakis and Filis, 2022, Delis *et al.,* 2023)

In this thesis, we diverge from the classic modelling specifications. We import decomposition techniques prior to modelling and forecasting, in order to observe whether this attitude results in components of higher resolution that could be modelled and forecasted more optimally. The proposed decomposition frameworks are famed for revealing the distinctive characteristics of a time series that are only observed under different frequency scales (Huang *et al*., 1998; Gilles, 2013). Thus, being able to

efficiently separate the short, middle, or long term dependencies of volatile and persistent time series, we could probably enhance the generated final outcome and add in the literature strand that incorporates decomposition techniques and ensemble modelling in order to produce more accurate forecasts.

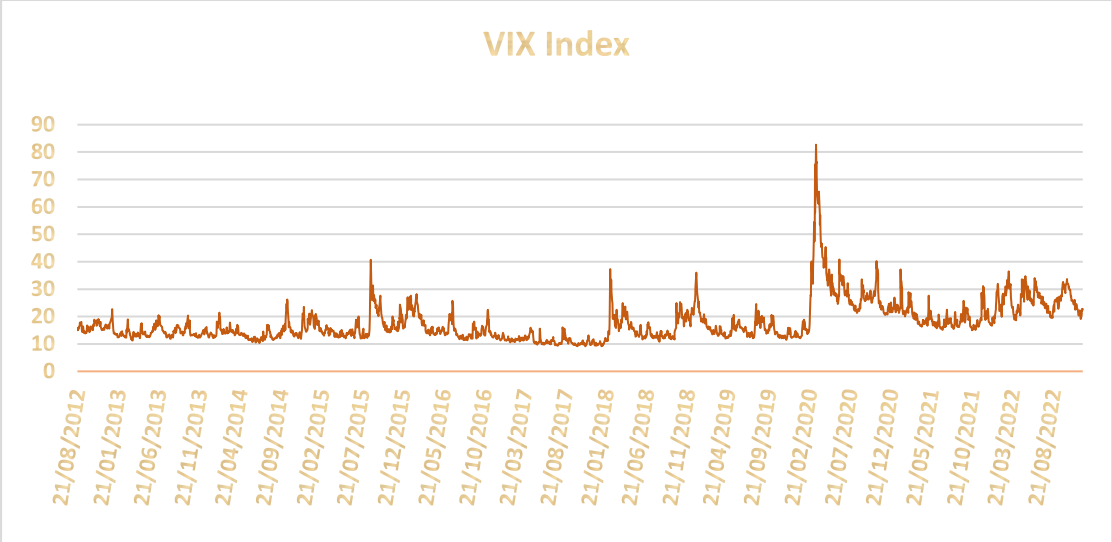
The findings of this study reward our intentions and verify that the proposed modelling architecture worths becoming part of the trading practice, since it produces profitable, for the market participant, forecasts. The 37 and 44 times return on capital based on trading roll adjusted VIX futures for the out-of-sample period, generated by the EEMD-based and EMD-based models, for most of our forecast horizons, is a very promising finding. But we should clarify something. We emphasized the word profitable, refraining from the word accurate. That is because based on the statistical evaluation criterion of the MSE, all models performed equally well, with no extra forecast gains, since most of the benchmark models were included on the MCS. But based on the objective-based evaluation criterion, that we presented in section 2.4.1, the difference was more than profound. Thus, it all depends on the reason why one engages in forecasting volatility and the practical use of these forecasts. Forecasts of financial instruments, owe to be aligned with the special conditions under which the respective markets operate, and conditions ask for capital profits and risk aversion, and so does the forecast validation.

The rest of the chapter is organized in the following way. Section 3.2 presents our data and their special features. Section 3.3 analyzes our findings, while section 3.4 concludes the study.

3.2 Data description

Daily data for both VIX and VIX futures were retrieved from CBOE[23]. Our samples consist of 2600 trading days spanning from 21st of August 2012 up to 30th of November 2022. Fig. 3.1 is the graphical visualization of VIX for the period under investigation.

Figure 3.1 Visualization of VIX



Note: VIX index for the entire period under investigation for conducting decomposition, modelling and forecasting. Throughout the entire period, there were definitely moments characterized by excess variability.

Table 3.1 reports the descriptive statistics of VIX, VIX futures, the logarithm of VIX and the correlation of VIX and VIX futures that reveals their strong relation (Daigler and Rossi, 2006; Szado 2009; Alexander *et al*., 2016), which is vital for the exercise we engage in. By inspecting Table 3.1, we can obviously confirm the non-normally distribution, the positive skewness and the leptokurtic structure mainly attributed to extreme volatility movements (Degiannakis and Filis, 2022). Also, inspect another recognizable feature extensively highlighted, the mean reverting property (Szado, 2020). Another recognizable fact is the more preferable features of the logarithm of VIX. The

[23] Data for VIX futures concerning the values of the first month contract at the day and the following day when settlement and roll takes place were retrieved from www.vixcentral.com.

logarithmic form is more appropriate for modeling and this is the reason why in our study we incorporate the logarithmic form of VIX, a much more preferable option and that is the reason why in the model description section, section 2.3, the frameworks' specification adopted the logarithmic form, not only because of the improved features, but also for the fact that some decomposition methods operate optimally under the transformed version.

Table 3.2 Descriptive Statistics of VIX, the logarithm of VIX and VIX futures

	VIX	Log. VIX	VIX futures
Mean	19.27	2.88	19.91
Median	17.28	2.84	18.17
Max	82.69	4.41	72.63
Min	9.14	2.21	9.88
St. Dev	8.52	0.38	7.44
Coef. Of Variation	0.44	0.13	0.37
Skewness	2.21	0.52	1.71
Kurtosis	9.17	0.10	5.94
J-Bera	6889.52	73.29	3135.72
ADF	-4.438*	-	-3.943*
Corr. with VIX	-	-	0.95

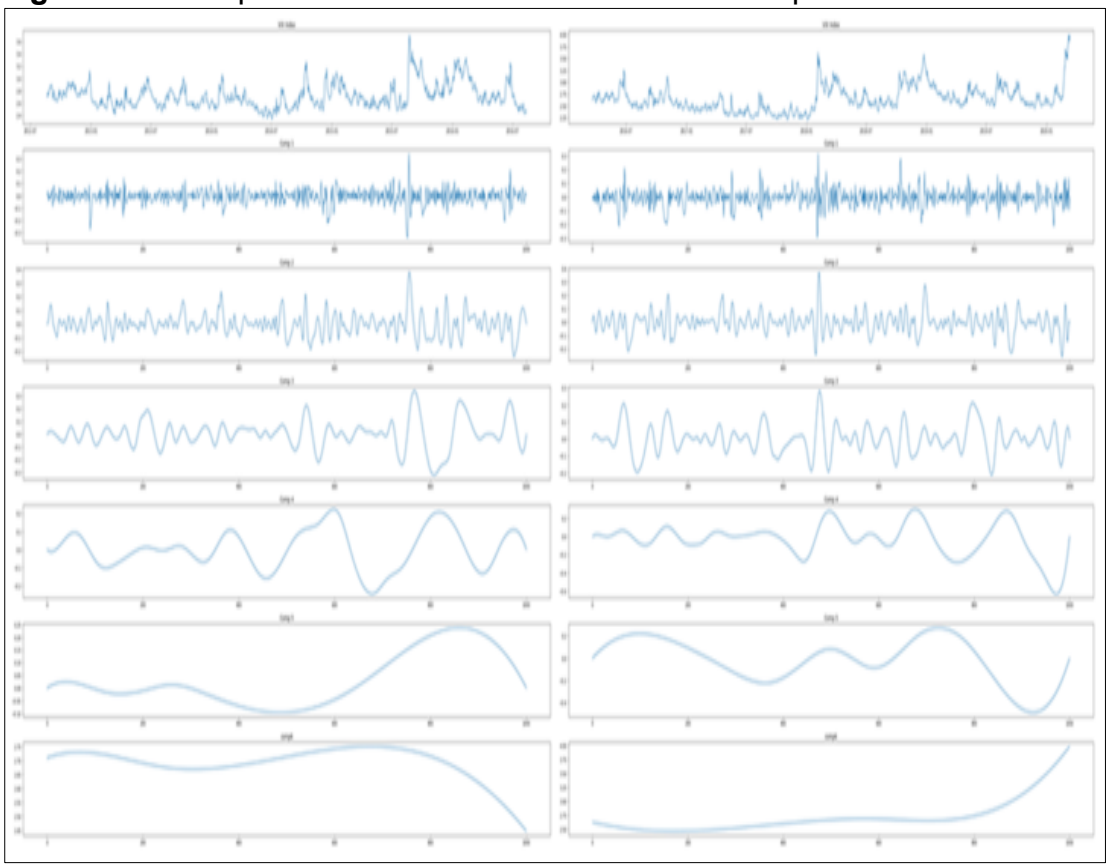
Note: Descriptive statistics for both VIX and VIX futures agree with the stylized facts of volatility. The correlation between VIX and VIX futures for the period under investigation also confirms their strong connection, crucial for trading exercises. The * denotes significance at the 1% level for the ADF, so both series do not exceed unit root issues

3.3 Empirical findings

Our analysis starts from Fig. 3.2 to 3.5, which depict the components of two randomly chosen rolling samples of VIX for the EMD, the EEMD, the HVD and the EWT techniques, respectively. We do so in order to start by revealing, how the nature and the trend of processes, completely alters through

successive rolling samples[24]. This fact raises the challenging bar of the entire process, but is the one that also validates the economic significance of our findings.

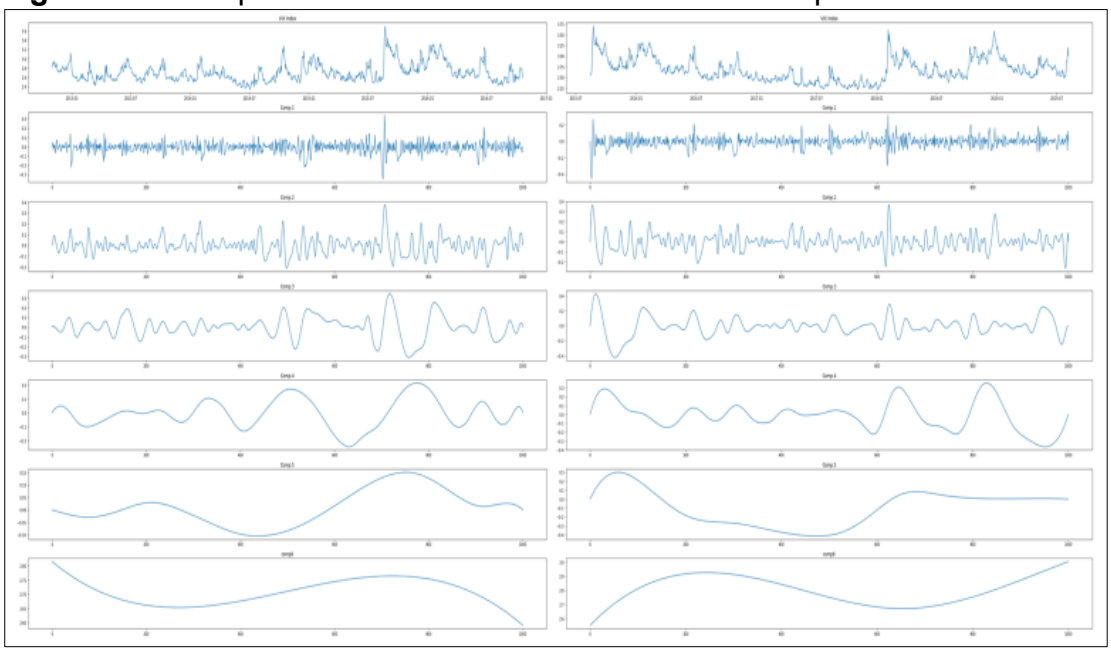
Figure 3.2. Components of EMD from two different samples



Note: At the top of each figure the VIX process of the chosen sample is depicted in order to present how radically components alter through successive samples and how unknown the future process becomes, something challenging for the forecasting frameworks that follow decomposition. The sample at the left is the 1st sample, while on the right appears the 1100th sample.

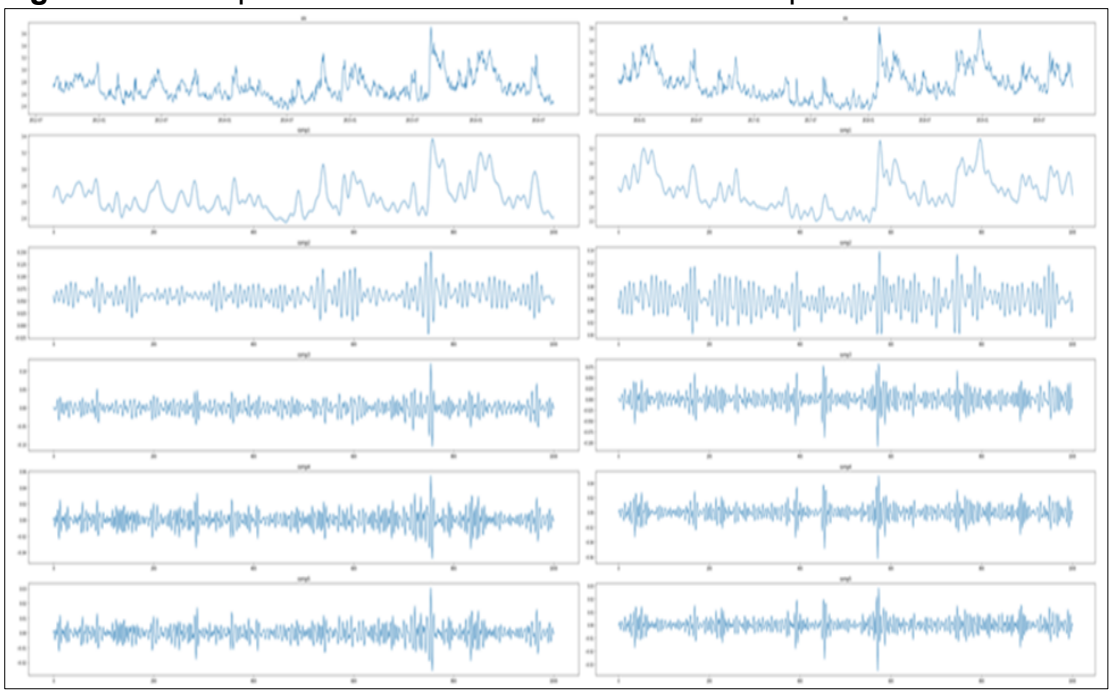
[24] The number of components for techniques that are not user specified alters between successive samplings. In some cases, there were 5 while in others the components were six. This of course adds a bit in computational difficulty but in that way, we recreate the exact way information arrives and that adds in the final outcome's validity.

Figure 3.3. Components of EEMD from two different samples



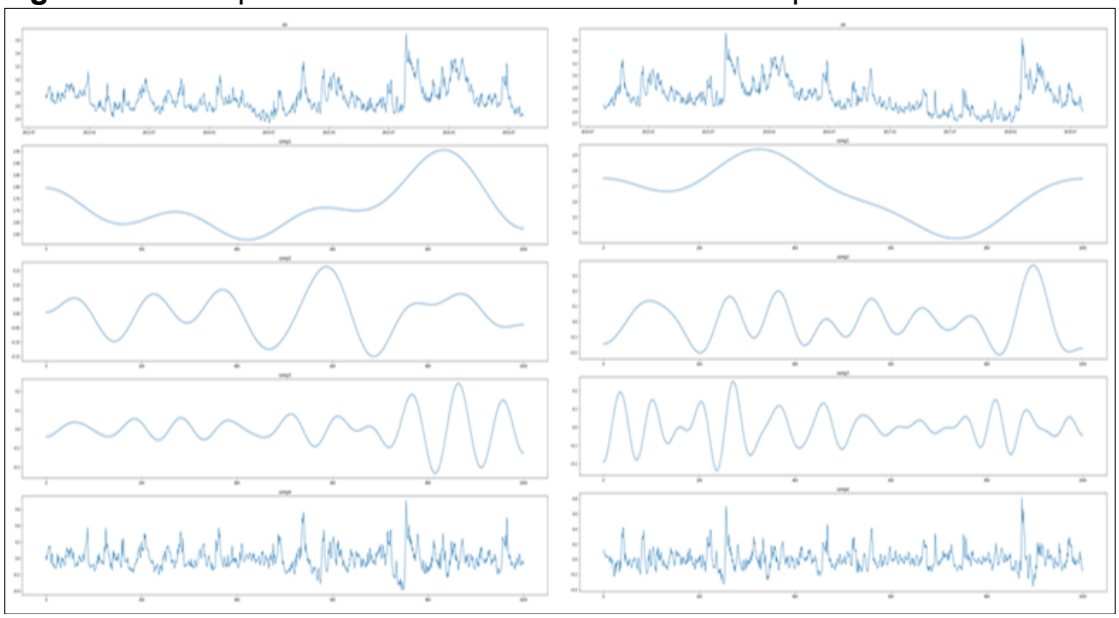
Note: The components of EEMD infused with finite white noise processes at the early steps of process, resemble those of EMD. We cannot tell if the one is better than the other in tackling with noisy data but overall, EEMD-based models generated the highest cumulative returns compared to competing models. At the top of each, appears the VIX process. On the left is the 30th sample, while on the right the 900nd sample is depicted.

Figure 3.4. Components of HVD from two different samples



Note: HVD process terminates when the difference of the standard deviation between two iterations becomes limited and that plays significant role during the approximate reconstruction of original time series. At the top is the VIX process. On the left appears the 1st sample and, on the right, the 800th sample.

Figure 3.5. Components of EWT from two different samples



Note: EWT is the method that the user must specify the segments of the Fourier spectrum. The correct choice of boundaries will return components that, when reconstructed give an optimum approximation of initial data. Components follow the wavelet theory and decomposition begins from the lowest energy component towards the highest. At the top is the VIX process. On the left, the 1st sample is presented and, on the right, the 500th.

One of the scopes of this study is to show that objective-based evaluation criteria are more informative compared to statistical loss functions. Table 3.2 reports the values of the MSE and the p-values of its respective MCS test. In the MCS of the best performing models, along the proposed frameworks also, are included the benchmark ones. It is clear that especially for horizons of 5 to 22 trading days ahead, we have infinitesimally small differences between the benchmark models and the ensemble ones, in order to claim we gain in predictive accuracy. Of course, the higher p-values of MCS are reported for the more sophisticated frameworks, but this is an indication that from a statistical perception, simple models can compete equally well with the more advanced ones.

Table 3.2 MSE values of forecasting models

Model	FORECASTING HORIZON							
	1 day		5 days		10 days		22 days	
	MSE	MCS	MSE	MCS	MSE	MCS	MSE	MCS
AR1	4.342	0.000	**16.960**	**0.815***	**30.91**	**0.422***	**53.12**	**0.836***
HAR	4.317	0.000	16.775	**0.845***	**30.37**	**0.422***	**51.28**	**0.859***
HW	4.257	0.000	111.609	0.000	1042.20	0.000	2553.59	0.000
LSTM	14.039	0.000	**15.978**	**0.901***	**29.523**	**0.773***	**59.150**	**0.715***
EMD-HW	**3.144**	**1.000***	32.803	0.000	185.492	0.000	297.844	0.000
EMD-HW-AR	4.145	0.000	**18.618**	**0.607***	45.175	0.001	119.973	0.000
EMD-LSTM-AR	6.703	0.000	**16.197**	**0.887***	**33.393**	**0.156***	**71.294**	**0.461***
EMD-LSTM-HW-AR	5.146	0.000	20.675	0.004	50.346	0.000	96.308	0.004
EEMD-HW	**3.145**	**0.652***	33.892	0.000	94.862	0.000	197.165	0.000
EEMD-HW-AR	4.129	0.000	16.791	0.845	47.860	0.000	139.849	0.000
EEMD-HW-AR-HAR	3.468	0.000	28.476	0.000	88.862	0.000	150.104	0.000
EEMD-LSTM-AR	7.094	0.000	20.948	0.004	38.448	0.075	**63.780**	**0.617***
EEMD-LSTM-AR-HAR	8.145	0.000	27.667	0.000	81.486	0.000	**82.054**	**0.371***
SSA-LSTM-AR	4.121	0.000	**15.671**	**0.901***	**28.007**	**0.773***	**49.665**	**0.901***
SSA-LSTM-HAR	4.106	0.000	**16.683**	**0.843***	**30.676**	**0.422***	**53.296**	**0.836***
SSA-LSTM-HW	4.139	0.000	**16.045**	**0.891***	**29.098**	**0.773***	**65.104**	**0.598***
SSA-HW	4.047	0.000	**16.413**	**0.867***	**30.547**	**0.422***	**66.653**	**0.598***
SSA-HW-HAR	4.063	0.000	**16.229**	**0.891***	**29.025**	**0.773***	**52.227**	**0.859***
SSA-HW-AR	3.971	0.000	**15.455**	**0.901***	**27.722**	**1.000***	**49.555**	**1.000***
HVD-HW	6.781	0.000	53.256	0.000	159.783	0.000	220.452	0.000
HVD-HW-AR	4.925	0.000	32.464	0.000	44.121	0.003	190.566	0.000
HVD-HAR-AR	5.658	0.000	**18.946**	**0.580***	39.863	0.060	**82.434**	**0.371***
HVD-AR1-AR	4.676	0.000	**13.916**	**1.000***	**29.792**	**0.773***	**60.441**	**0.715***
HVD-LSTM-AR	11.126	0.000	26.867	0.000	62.155	0.000	170.677	0.000
EWT-HW	6.802	0.000	20.778	0.004	37.577	0.089	**76.215**	**0.461***
EWT-HW-AR	6.445	0.000	25.071	0.000	44.505	0.003	89.045	0.063
EWT-LSTM-HW	6.583	0.000	20.636	0.004	37.778	0.089	**81.528**	**0.371***
EWT-LSTM-AR	4.728	0.000	25.813	0.000	46.461	0.001	89.040	0.063
VMD-HW	9.553	0.000	19.876	0.045	**32.956**	**0.231***	**66.874**	**0.598***
VMD-HW-AR	9.934	0.000	22.630	0.004	39.041	0.060	**78.114**	**0.461***
VMD-LSTM-AR	11.155	0.000	24.055	0.000	42.385	0.003	89.704	0.063
VMD-LSTM-HW	10.659	0.000	21.282	0.004	**34.560**	**0.156***	**69.392**	**0.483***

Note: Table reports the MSE for the 1 day and for the 5, 10 and 22 days ahead. With * and bold writing we denote the models that belong to the MCS. The choice of the models for the MCS test, depends on the threshold significant level we choose, as well other parameters involving MCS test framework. In this study we use the 0,15 level. We inspect, that based on MSE in the MCS of the best performing models do belong the AR1, the HAR and the LSTM for the 5, 10 and 22 days ahead horizons. Models' specifications:

EMD-HW, EMD-HW-AR, EMD-LSTM-AR, EMD-LSTM-AR-HAR: The EMD-based models with different modelling combinations. HW denotes that all components are modelled via the HW framework. HW-AR denotes that, two out of the six components are modelled via HW and the rest through AR processes. LSTM-AR denotes that, two components are modelled via LSTM and the rest through AR processes. LSTM-AR-HAR denotes that, 1 component is modelled via LSTM and the rest through AR and HAR processes.

EEMD-HW, EEMD-HW-AR, EEMD-HW-AR-HAR, EEMD-LSTM-AR, EEMD-LSTM-AR-HAR: Constitute the EEMD-based models. HW denotes that all components are modelled via the HW framework. HW-AR denotes that, two components out of the six, are modelled via HW and the rest through AR processes. HW-AR-HAR denotes that, two components are modelled via HW and the rest through AR and HAR processes. LSTM-AR denotes that, two components are modelled via LSTM and the rest through AR processes. LSTM-AR-HAR denotes that, one component is modelled via LSTM and the rest through AR and HAR processes.

SSA-LSTM-AR, SSA-LSTM-HAR, SSA-LSTM-HW, SSA-HW, SSA-HW-HAR, SSA-HW-AR: The SSA-based models. LSTM-AR denotes that, one component is modelled via LSTM and the rest through AR processes. LSTM-HAR denotes that, one component is modelled via LSTM and the rest two through the HAR framework. LSTM-HW denotes that, one component is modelled through LSTM and the rest with the HW framework. HW denotes that all components are modelled via the HW framework. HW-HAR denotes that, one component is modelled via HW and the rest through the HAR model. HW-AR denotes that, one component is modelled via HW and the rest through AR processes.

HVD-HW, HVD-HW-AR, HVD-HAR-AR, HVD-AR1-AR, HVD-LSTM-AR: The HVD-based models. HW denotes that all components are modelled via the HW framework. HW-AR denotes that, two out of the five components are modelled via HW and the rest through AR processes. HAR-AR denotes that, one component is modelled via HAR and the rest through AR processes. AR1-AR denotes that, 1 component is modelled via AR1 and the rest through AR processes. LSTM-AR denotes that, one component is modelled via LSTM and the rest through AR processes.

EWT-HW, EWT-HW-AR, EWT-LSTM-HW, EWT-LSTM-AR: The EWT-based models. HW denotes that all components are modelled via the HW framework. HW-AR denotes that, one component out of the four is modelled via HW and the rest through AR processes. LSTM-HW denotes that, one component is modelled through LSTM and the rest with the HW framework. LSTM-AR denotes that, one component is modelled via LSTM and the rest through AR processes.

VMD-HW, VMD-HW-AR, VMD-LSTM-AR, VMD-LSTM-HW: The VMD-based models. HW denotes that all components are modelled via the HW framework. HW-AR denotes that, one component out of the four is modelled via HW and the rest through AR processes. LSTM-AR denotes that, one component is modelled via LSTM and the rest through AR processes. LSTM-HW denotes that, one component is modelled through LSTM and the rest with the HW framework

From an economic point of view, the statistically imposed equality among forecast models, debases. Table 3.3 reports the cumulative returns of the models (not in % points) and their respective MCS test p-values, based on the objective-based economic criterion we propose and highlights how accuracy should be quantified. The reported cumulative returns at least for three of the proposed ensemble techniques, by far exceed those generated by the AR1, the HAR, the HW and the LSTM, something confirmed by the reported MCS test p-values. Most of the EEMD-based and EMD-based models outperformed for horizons of 5 to 22 trading days ahead, followed by SSA-based models for the 1 and 5 trading days ahead horizon and HVD-HAR-AR, HVD-AR1-AR and HVD-LSTM-AR for the 5 days ahead. EEMD-HW-AR-HAR generated the highest cumulative returns that reach the value of 44,17 for the 5 days and the 41,53 for the 10 days ahead horizon, that is 44 and 41 times the initial invested capital, respectively. EMD-based models follow, with remarkable cumulative returns ranging between 12,44 to 37,00 for horizons of 5, 10 and 22 days ahead[25].

[25] We do not place much attention on the 1 day ahead forecast horizon for the cases of the EEMD and EMD-based models due to the way the two techniques are implemented. At sections 3.1 and 3.2 we specified that we used the zero-end cubic spline condition, so literally the last observation in every rolling sample of the components, is excluded as no decomposition takes place.

Table 3.3. Cumulative returns of forecasting models

Model	FORECAST HORIZON							
	1 day		5 days		10 days		22 days	
	Cum. Returns	MCS	Cum. Returns	MCS	Cum. Returns	MCS	Cum. Returns	MCS
AR1	2.691	0.005	3.915	0.090	3.969	0.000	4.319	0.006
HAR	3.249	0.005	4.692	0.099	5.446	0.000	5.558	0.021
HW	0.247	0.000	3.505	0.085	-2.462	0.000	3.511	0.000
LSTM	3.238	0.005	0.188	0.000	2.140	0.000	-2.372	0.000
EMD-HW	**16.938**	**0.223***	**37.183**	**0.615***	**36.616**	**0.145***	**37.028**	**1.000***
EMD-HW-AR	8.993	0.053	**24.747**	**0.507***	20.935	0.032	**26.839**	**0.217***
EMD-LSTM-AR	-10.921	0.000	**12.442**	**0.488***	13.342	0.013	**18.634**	**0.144***
EMD-LSTM-HW-AR	4.029	0.011	**21.707**	**0.507***	**27.202**	**0.145***	**28.065**	**0.217***
EEMD-HW	**34.400**	**1.000***	**35.990**	**0.615***	**38.043**	**0.145***	**36.264**	**0.463***
EEMD-HW-AR	**27.528**	**0.937***	**24.686**	**0.507***	**32.704**	**0.145***	**31.281**	**0.427***
EEMD-HW-AR-HAR	**30.171**	**0.950***	**44.417**	**1.000***	**41.530**	**1.000***	**34.360**	**0.463***
EEMD-LSTM-AR	1.461	0.001	**9.252**	**0.470***	11.527	0.013	**19.608**	**0.144***
EEMD-LSTM-AR-HAR	0.647	0.000	**38.731**	**0.615***	**37.077**	**0.145***	**22.734**	**0.186***
SSA-LSTM-AR	**14.247**	**0.223***	**5.241**	**0.162***	8.594	0.004	6.225	0.060
SSA-LSTM-HAR	**12.409**	**0.223***	3.295	0.087	6.725	0.004	5.076	0.021
SSA-LSTM-HW	**10.486**	**0.223***	1.097	0.036	6.265	0.004	5.787	0.021
SSA-HW	8.219	0.053	**7.077**	**0.162***	6.191	0.004	4.174	0.015
SSA-HW-HAR	9.149	0.053	**6.940**	**0.150***	9.488	0.004	9.105	0.060
SSA-HW-AR	**10.143**	**0.233***	**9.660**	**0.470***	8.830	0.004	12.233	0.098
HVD-HW	1.316	0.001	0.160	0.000	-0.591	0.000	0.602	0.000
HVD-HW-AR	4.099	0.011	1.856	0.036	1.104	0.000	1.579	0.000
HVD-HAR-AR	4.789	0.011	**5.603**	**0.162***	4.652	0.000	3.069	0.000
HVD-AR1-AR	3.221	0.005	**9.881**	**0.470***	3.854	0.000	1.103	0.000
HVD-LSTM-AR	7.603	0.053	**22.198**	**0.507***	8.623	0.004	-	
EWT-HW	2.115	0.005	1.448	0.036	0.812	0.000	0.110	0.000
EWT-HW-AR	5.134	0.011	0.761	0.000	3.040	0.000	-0.217	0.000
EWT-LSTM-HW	3.034	0.005	1.247	0.036	2.081	0.000	-0.792	0.000
EWT-LSTM-AR	3.124	0.005	3.724	0.090	2.847	0.000	1.377	0.000
VMD-HW	1.432	0.001	0.544	0.000	-0.160	0.000	-0.115	0.000
VMD-HW-AR	2.349	0.005	0.296	0.000	1.396	0.000	1.473	0.000
VMD-LSTM-AR	2.641	0.005	2.025	0.085	2.636	0.000	2.415	0.000
VMD-LSTM-HW	3.437	0.005	1.099	0.036	0.487	0.000	-0.722	0.000

Note: Table reports the cumulative returns of trading VIX futures (are **not** expressed in % points) for 1, 5, 10 and 22 days ahead forecast horizon along with the p-values of the MCS test. With * and bold writing we denote the models that belong to the MCS. The choice of the models for the MCS test, depends on the threshold significant level we choose, as well other parameters involving MCS test framework. In this study we use the 0,15 level. According to the objective-based evaluation criterion, none of the AR1, HAR, HW and LSTM belong to the set of the best-performing models.

EMD-HW, EMD-HW-AR, EMD-LSTM-AR, EMD-LSTM-AR-HAR: The EMD-based models with different modelling combinations. HW denotes that all components are modelled via the HW framework. HW-AR denotes that, two out of the six components are modelled via HW and the rest through AR processes. LSTM-AR denotes that, two components are modelled via LSTM and the rest through AR processes. LSTM-AR-HAR denotes that, 1 component is modelled via LSTM and the rest through AR and HAR processes.

EEMD-HW, EEMD-HW-AR, EEMD-HW-AR-HAR, EEMD-LSTM-AR, EEMD-LSTM-AR-HAR: Constitute the EEMD-based models. HW denotes that all components are modelled via the HW framework. HW-AR denotes that, two components out of the six, are modelled via HW and the rest through AR processes. HW-AR-HAR denotes that, two components are modelled via HW and the rest through ar and HAR processes. LSTM-AR denotes that, two components are modelled via LSTM and the rest through AR processes. LSTM-AR-HAR denotes that, one component is modelled via LSTM and the rest through AR and HAR processes.

SSA-LSTM-AR, SSA-LSTM-HAR, SSA-LSTM-HW, SSA-HW, SSA-HW-HAR, SSA-HW-AR: The SSA-based models. LSTM-AR denotes that, one component is modelled via LSTM and the rest through AR processes. LSTM-HAR denotes that, one component is modelled via LSTM and the rest two through the HAR framework. LSTM-HW denotes that, one component is modelled through LSTM and the rest with the HW framework. HW denotes that all components are modelled via the HW framework. HW-HAR denotes that, one component is modelled via HW and the rest through the HAR model. HW-AR denotes that, one component is modelled via HW and the rest through AR processes.

HVD-HW, HVD-HW-AR, HVD-HAR-AR, HVD-AR1-AR, HVD-LSTM-AR: The HVD-based models. HW denotes that all components are modelled via the HW framework. HW-AR denotes that, two out of the five components are modelled via HW and the rest through AR processes. HAR-AR denotes that, one component is modelled via HAR and the rest through AR processes. AR1-AR denotes that, 1 component is modelled via AR1 and the rest through AR processes. LSTM-AR denotes that, one component is modelled via LSTM and the rest through AR processes.

EWT-HW, EWT-HW-AR, EWT-LSTM-HW, EWT-LSTM-AR: The EWT-based models. HW denotes that all components are modelled via the HW framework. HW-AR denotes that, one component out of the four is modelled via HW and the rest through AR processes. LSTM-HW denotes that, one component is modelled through LSTM and the rest with the HW framework. LSTM-AR denotes that, one component is modelled via LSTM and the rest through AR processes.

VMD-HW, VMD-HW-AR, VMD-LSTM-AR, VMD-LSTM-HW: The VMD-based models. HW denotes that all components are modelled via the HW framework. HW-AR denotes that, one component out of the four is modelled via HW and the rest through AR processes. LSTM-AR denotes that, one component is modelled via LSTM and the rest through AR processes. LSTM-HW denotes that, one component is modelled through LSTM and the rest with the HW framework.

Tables 3.2 and 3.3 for the same models, demonstrate two contradictive outcomes, an outcome of no predictive gain and an outcome of accelerated economic significance, hence, we raise an important issue. Loss functions such as mean square error, mean absolute error, root mean squared error etc., measure the distance of forecasts from the actual price level and report how much they diverge from it. But what happens when the values of these loss functions blur the truth of the directional accuracy, especially when economic and financial data are involved? Put it simply, it is very possible to inevitably direct towards incremental losses, especially in cases when an investor or risk manager follows a model-based trading strategy based on an implied volatility index and its tradable products, in order to gain portfolio optimization or neutrality. Table 3.4 reports the rates of the DoC evaluation test and validates the recordings of table 3.3. The EEMD-HW-AR-HAR model that reported the highest value in cumulative returns, according to DoC, correctly predicted the direction of VIX index for the 82% of times in the out-of-sample period of the 1600 trading days for the 5 days ahead horizon and about 78% of the times for the 10 days ahead horizon. Adequate values, also, are reported for the models that performed a little lower but correctly predicted the directional movement of the implied volatility index in a range spanning from 55% to 77%.

Table 3.4 Direction of Change rates of forecasting models

Model	1 day	5 days	10 days	22 days
	DoC	DoC	DoC	DoC
AR1	0.474	0.491	0.488	0.509
HAR	0.486	0.502	0.510	0.512
HW	0.450	0.479	0.405	0.489
LSTM	0.485	0.446	0.476	0.387
EMD-HW	**0.643***	**0.768***	**0.755***	**0.766***
EMD-HW-AR	0.596	0.663	0.650	0.680
EMD-LSTM-AR	0.397	0.553	0.564	0.602
EMD-LSTM-HW-AR	0.493	0.642	0.688	**0.710***
EEMD-HW	**0.749***	**0.754***	**0.769***	**0.756***
EEMD-HW-AR	**0.704***	**0.670***	**0.712***	**0.704***
EEMD-HW-AR-HAR	**0.723***	**0.821***	**0.778***	**0.731***
EEMD-LSTM-AR	0.445	0.534	0.538	0.608
EEMD-LSTM-AR-HAR	0.427	**0.766***	**0.722***	0.667
SSA-LSTM-AR	**0.608***	0.505	0.520	0.523
SSA-LSTM-HAR	0.581	0.488	0.512	0.511
SSA-LSTM-HW	0.575	0.469	0.508	0.517
SSA-HW	0.544	0.536	0.501	0.508
SSA-HW-HAR	0.551	0.513	0.528	0.530
SSA-HW-AR	0.572	0.541	0.520	0.541
HVD-HW	0.441	0.455	0.411	0.456
HVD-HW-AR	0.497	0.474	0.451	0.466
HVD-HAR-AR	0.499	0.518	0.487	0.479
HVD-AR1-AR	0.485	0.542	0.479	0.461
HVD-LSTM-AR	0.511	0.651	0.516	-
EWT-HW	0.469	0.473	0.430	0.440
EWT-HW-AR	0.489	0.450	0.465	0.411
EWT-LSTM-HW	0.482	0.469	0.453	0.402
EWT-LSTM-AR	0.482	0.479	0.459	0.453
VMD-HW	0.459	0.444	0.423	0.404
VMD-HW-AR	0.463	0.441	0.442	0.458
VMD-LSTM-AR	0.468	0.467	0.455	0.470
VMD-LSTM-HW	0.491	0.455	0.411	0.399

Note: The DoC is the extra forecast evaluation test we adopted for our study, since the MCS test of cumulative returns, highlighted the overperformance of some ensemble frameworks. DoC rates do confirm the superiority and the ability of some techniques to effectively predict VIX's correct directional movement. With bold writing and * we specify the models with rates over 70%.

EMD-HW, EMD-HW-AR, EMD-LSTM-AR, EMD-LSTM-AR-HAR: The EMD-based models with different modelling combinations. HW denotes that all components are modelled via the HW framework. HW-AR denotes that, two out of the six components are modelled via HW and the rest through AR

processes. LSTM-AR denotes that, two components are modelled via LSTM and the rest through AR processes. LSTM-AR-HAR denotes that, 1 component is modelled via LSTM and the rest through AR and HAR processes.

EEMD-HW, EEMD-HW-AR, EEMD-HW-AR-HAR, EEMD-LSTM-AR, EEMD-LSTM-AR-HAR: Constitute the EEMD-based models. HW denotes that all components are modelled via the HW framework. HW-AR denotes that, two components out of the six, are modelled via HW and the rest through AR processes. HW-AR-HAR denotes that, two components are modelled via HW and the rest through ar and HAR processes. LSTM-AR denotes that, two components are modelled via LSTM and the rest through AR processes. LSTM-AR-HAR denotes that, one component is modelled via LSTM and the rest through AR and HAR processes.

SSA-LSTM-AR, SSA-LSTM-HAR, SSA-LSTM-HW, SSA-HW, SSA-HW-HAR, SSA-HW-AR: The SSA-based models. LSTM-AR denotes that, one component is modelled via LSTM and the rest through AR processes. LSTM-HAR denotes that, one component is modelled via LSTM and the rest two through the HAR framework. LSTM-HW denotes that, one component is modelled through LSTM and the rest with the HW framework. HW denotes that all components are modelled via the HW framework. HW-HAR denotes that, one component is modelled via HW and the rest through the HAR model. HW-AR denotes that, one component is modelled via HW and the rest through AR processes.

HVD-HW, HVD-HW-AR, HVD-HAR-AR, HVD-AR1-AR, HVD-LSTM-AR: The HVD-based models. HW denotes that all components are modelled via the HW framework. HW-AR denotes that, two out of the five components are modelled via HW and the rest through AR processes. HAR-AR denotes that, one component is modelled via HAR and the rest through AR processes. AR1-AR denotes that, 1 component is modelled via AR1 and the rest through AR processes. LSTM-AR denotes that, one component is modelled via LSTM and the rest through AR processes.

EWT-HW, EWT-HW-AR, EWT-LSTM-HW, EWT-LSTM-AR: The EWT-based models. HW denotes that all components are modelled via the HW framework. HW-AR denotes that, one component out of the four is modelled via HW and the rest through AR processes. LSTM-HW denotes that, one component is modelled through LSTM and the rest with the HW framework. LSTM-AR denotes that, one component is modelled via LSTM and the rest through AR processes.

VMD-HW, VMD-HW-AR, VMD-LSTM-AR, VMD-LSTM-HW: The VMD-based models. HW denotes that all components are modelled via the HW framework. HW-AR denotes that, one component out of the four is modelled via HW and the rest through AR processes. LSTM-AR denotes that, one component is modelled via LSTM and the rest through AR processes. LSTM-HW denotes that, one component is modelled through LSTM and the rest with the HW framework.

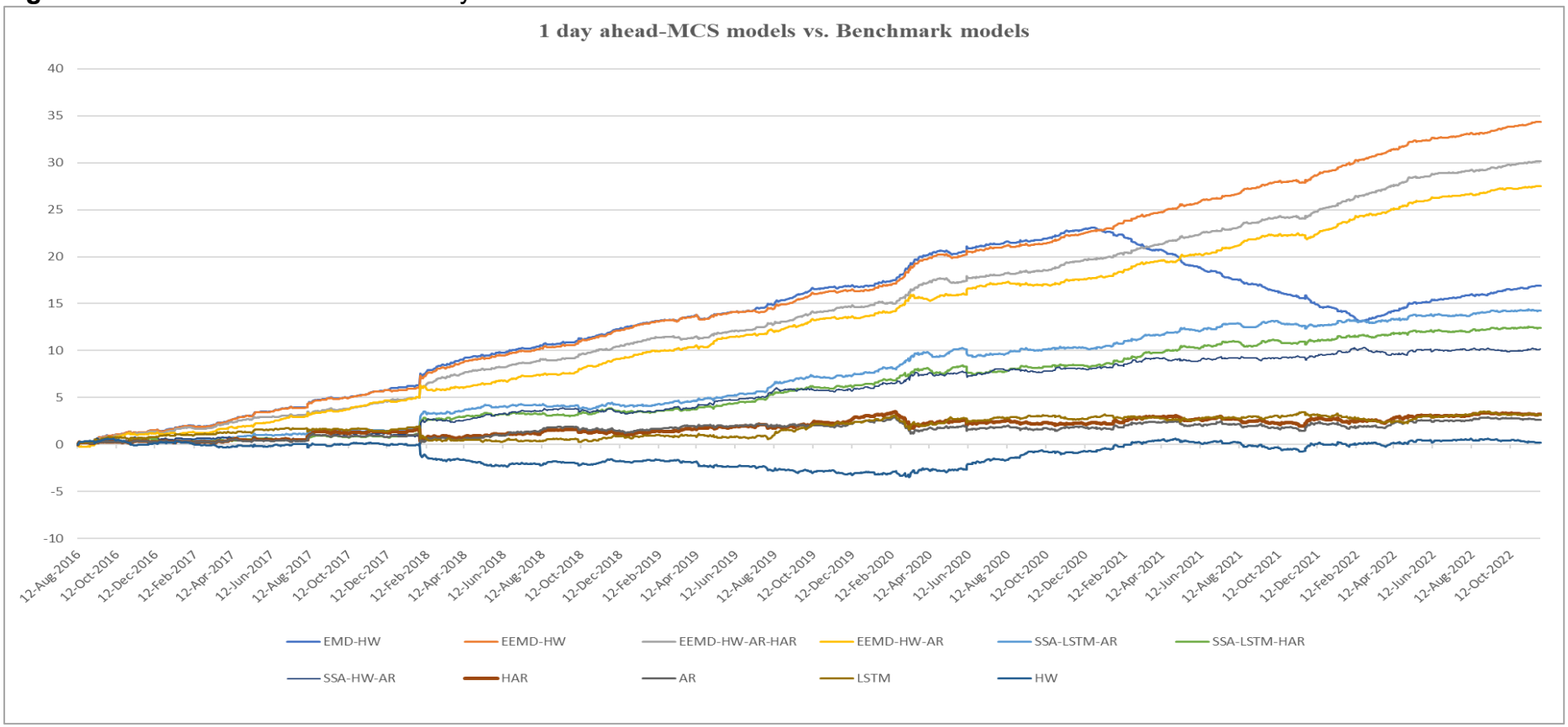
More importantly, though, the main scope of our study is to illustrate whether, the divide and conquer strategy, has proved to be rewarding. It has indeed proved to be rewarding and we actually do provide the required evidence. From the above findings, we can support that ensemble modelling accelerates cumulative returns and even boosts the performance of the classic and benchmark models, when used in the various modelling combinations of the separate modes. Perhaps not all ensemble models generated remarkable excess returns, but the ones that did, were high enough.

So, we deal with 3 aspects that may shed light to recorded findings. First off, from the proposed decomposition techniques, the EMD framework is the only one to be characterized complete (Huang *et al*., 1998). The EEMD follows in resolution as was the one with the lowest MSE[26] that of 0,02. SSA's rate was 0,42 and HVD's 0,49. The MSE for the EWT was 1,2 while for the VMD, by using 5 components, it became 2,4. Perhaps these small declines between methods, actually shape optimality of final forecasts and perhaps that is to blame for EWT-based and VMD-based models, resulting in minor or negative cumulative returns for most forecast horizons. Figures 3.6 to 3.9, visualize the cumulative returns of the MCS models, along the benchmark models for the 1, 5, 10 and 22 trading days ahead forecast horizon, while Fig. 3a to 3d at the end of this chapter, depict a mixture of most ensemble models for the same horizons, where the performance of EWT-based and VMD-based models, is also illustrated.

[26] Here MSE is the same statistical loss function used in our data only here it measures the divergence each decomposing method had from actual values of original input signal. That is how close the aggregated values of the components, approximate the values of the rolling samples of VIX index.

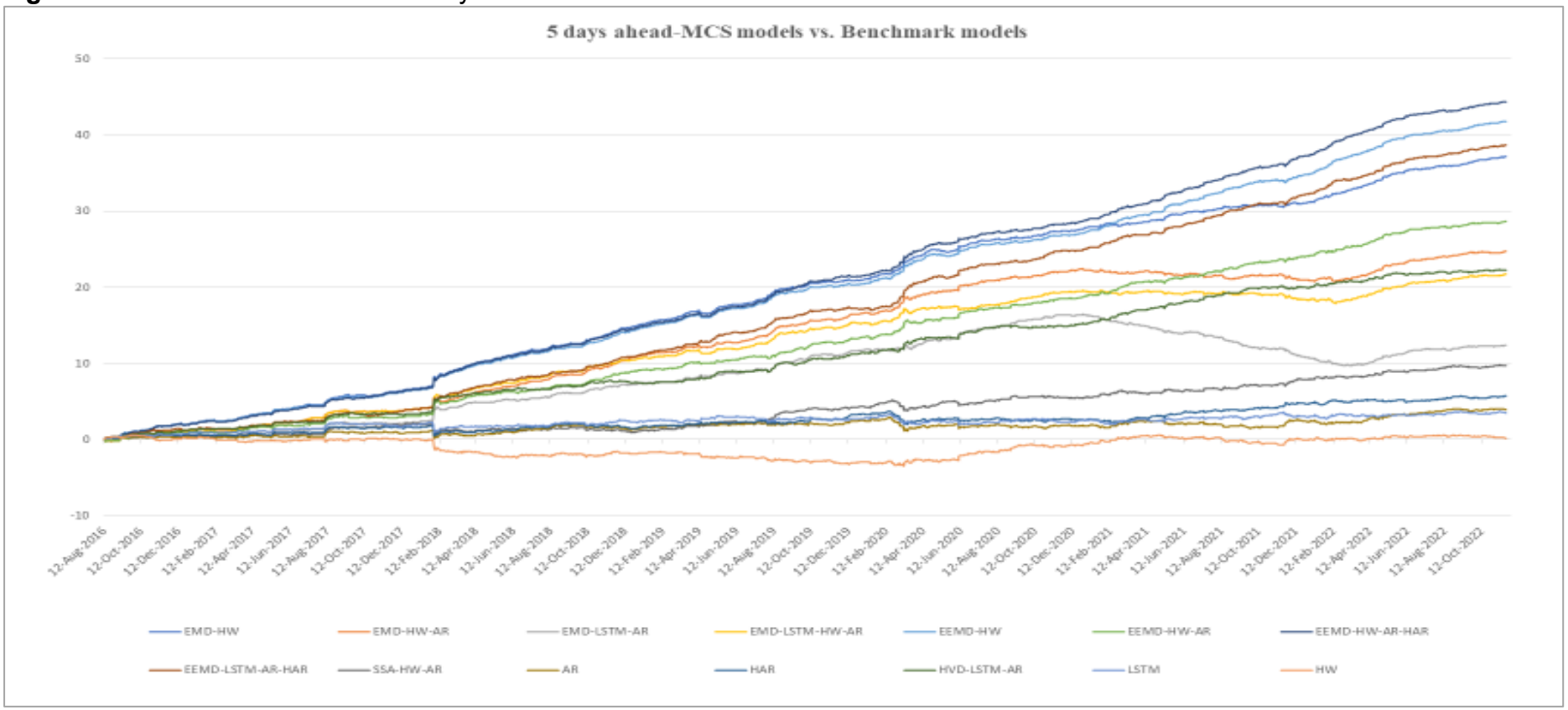
Second, an aspect that may play a very critical role, is the correct modelling combination of the components of the six decomposition techniques. Let's consider Fig. 3b and concentrate on EMD-HW and EMD-LSTM-HW-AR models for the 5 days ahead horizon. In the case where all components of the EMD process are modelled via HW framework, cumulative returns skyrocket at the rate of 37.18. When the same components are modelled through a combination of the LSTM model, the HW and the HAR, cumulative returns decrease to 21,70. The same findings are valid for EEMD-HW and EEMD-LSTM-AR-HAR for the 22 days ahead horizon, that from 36,36 lessens to 22,73.

Figure 3.6. Cumulative returns of 1 day ahead horizon



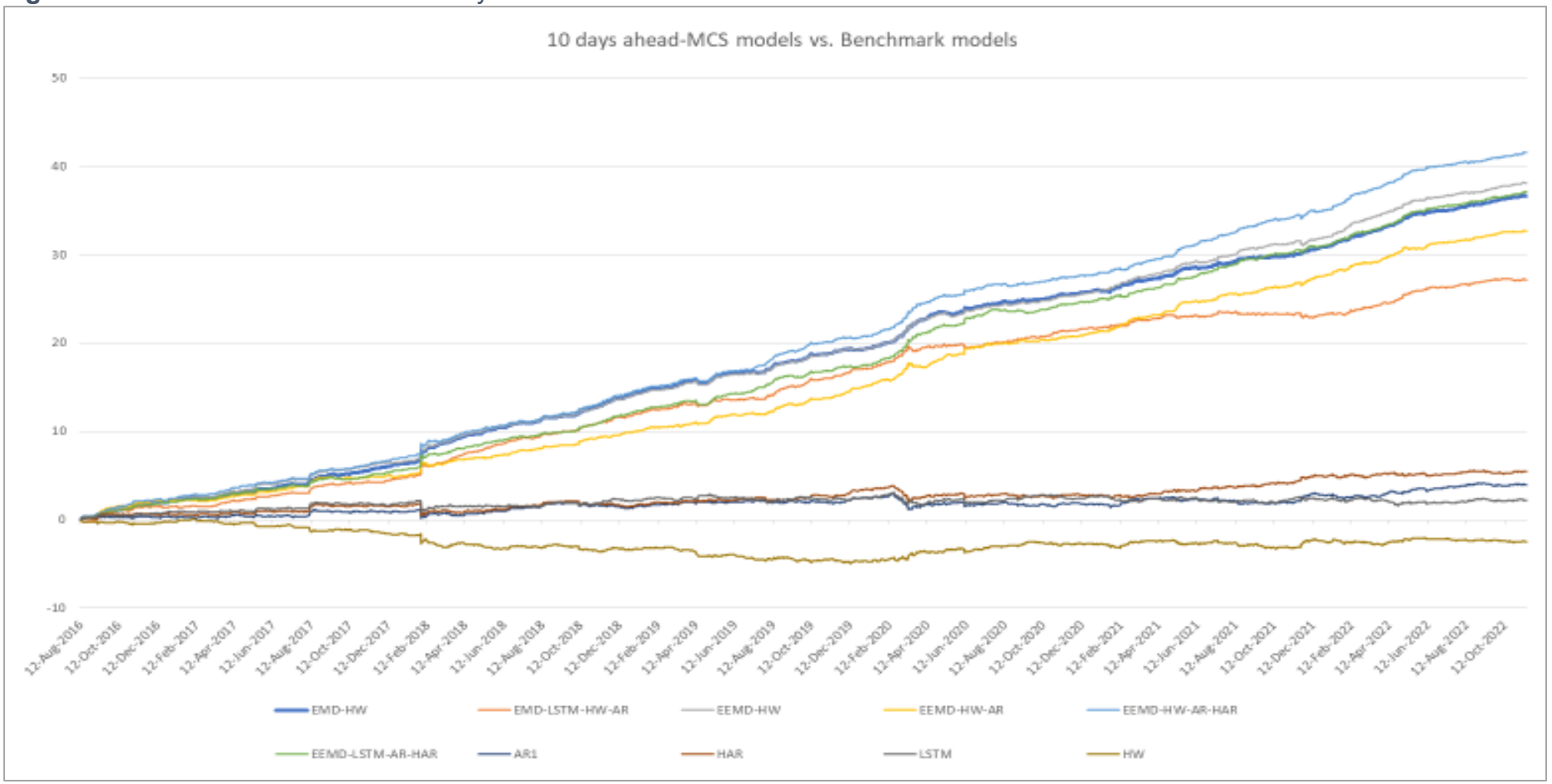
Note: Cumulative returns of the best performing models, according to MCS, along with the ones of the 4 benchmark models, for the 1 day ahead forecast horizon. All 4 benchmark models have mediocre or negative returns, compared to EMD-based, EEMD-based and SSA-based models, which moved with increasing rates during the period under investigation. The effect on cumulative returns based on the model combination is noticeable, but more noticeable is in the Fig. 3a at the end of chapter. For instance the model EMD-HW although profitable in the end definitely records losses during the COVID-19 pandemic period, where there is an obvious down trend movement not spotted in other models. The values reported in the y-axis are not in % points.

Figure 3.7. Cumulative returns of 5 days ahead forecast horizon



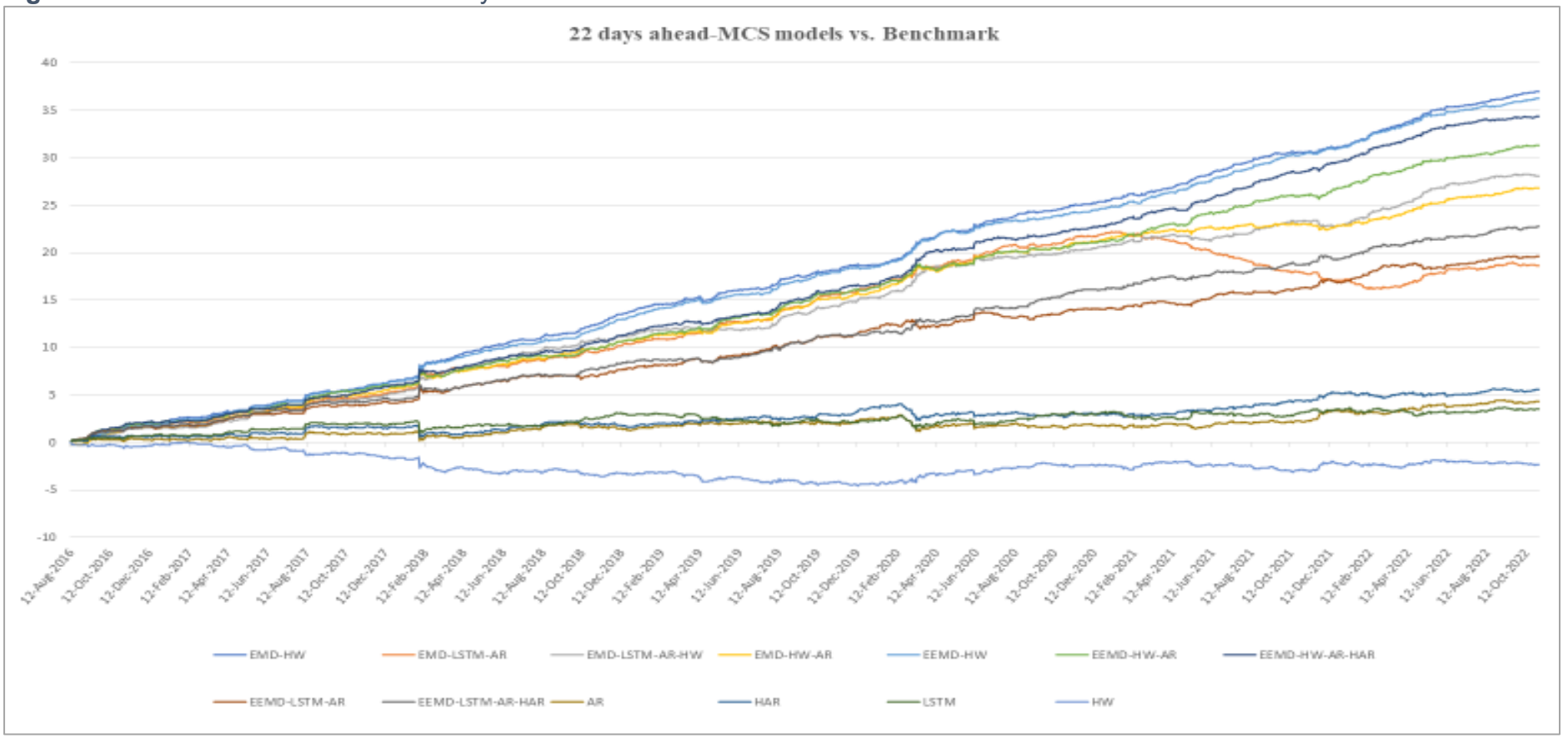
Note: For five days ahead forecast horizon, the cumulative returns earned by the trading strategy we implemented, do seem to diverge substantially between the ensemble models and the benchmark models. Of course, there were a few more models included in the MCS, but their cumulative returns were closely enough with others and their presentation would be rather messy, so we refrained from their inclusion. The values reported in the y-axis are not in % points.

Figure 3.8. Cumulative returns for 10 days ahead forecast horizon



Note: For the forecast horizon of ten days ahead fewer models were included in the MCS, but the ones that did, come with significant gains compared to benchmark models. The values reported in the y-axis are not in % points.

Figure 3.9. Cumulative returns for 22 days ahead forecast horizon



Note: All EMD-based and EEMD-based models were the only ones to be included in the MCS for the forecast horizon of the 22 days ahead. Generally, most of the EMD-based and EEMD-based models were amongst the best performing models for most horizons with significant increased cumulative returns, while benchmark models performed mediocre with HW being the one with systematically negative returns for entire period under investigation. On the other hand, EMD-HW and EEMD-HW, whose all components were modelled and forecasted via the HW framework, were the best performing. The values reported in the y-axis are not in % points.

Third, there is the roll. No matter if for most of the period under investigation VIX futures were in contango, a fact not that promising for investors, the performance of EMD-based, EEMD-based and SSA-based models, was accompanied with gains sufficiently high for all horizons.

Hence, we can claim 3 findings: 1) VIX futures indeed have unique return drivers (Moran and Dash, 2007; Szado, 2018), when the actual market conditions are considered, 2) modelling combinations of occurring components, do play a critical role and 3) most importantly, the proposed modelling architecture, paired with objective-based evaluation criteria, worth becoming part of model-based trading practices adopted by interested parties, since it is multi beneficial.

3.4 Conclusion

The aim of this study and of this thesis widely, is to contribute to the literature strand that uses ensemble methods to forecast not only implied volatility indices, but other financial time series data as well, by illustrating their predictive superiority. However, financial time series are not plain data, they are core instruments of financial markets and therefore of the global financial system. They hold vital information for traders, risk and portfolio managers, policy agencies, so research must be targeted to their utility, especially when forecast is involved. Forecasts must serve a purpose, a realistic economic purpose and their accuracy, should be evaluated according to this purpose. Thus, our study serves this purpose and certifies their superiority and usefulness, under economic standards. We produce useful for the market participants forecasts that are involved in model-based trading practices, by pairing six decomposition techniques, the EMD, the EEMD, the SSA, the

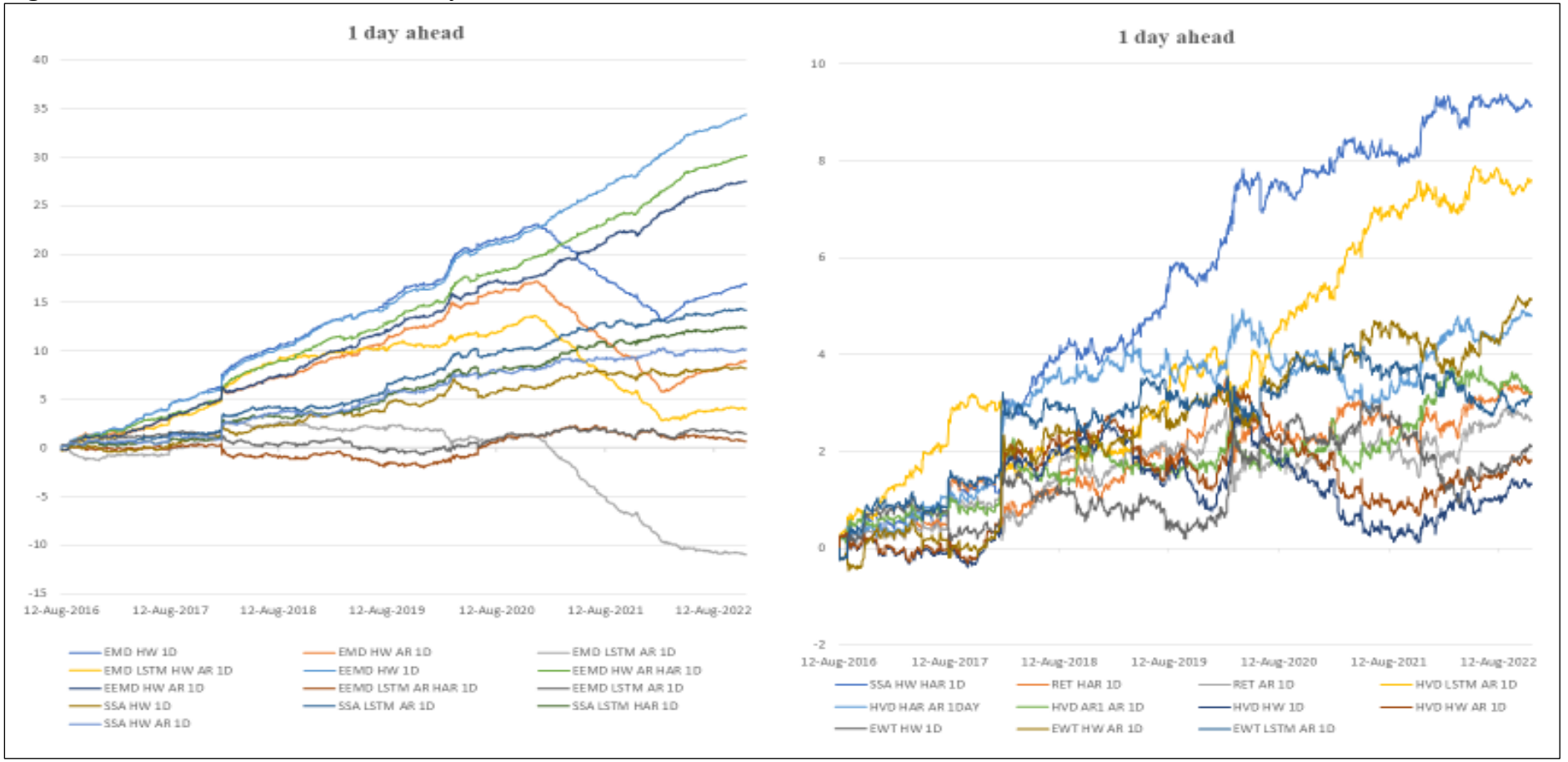
HVD, the EWT and the VMD, with parametric and non-parametric modelling frameworks, namely the AR, the HAR, the HW and the LSTM. On a daily basis, we decompose and forecast VIX Cboe volatility index for 1, 5, 10 and 22 trading days ahead horizon, resulting in a total of 32 models. Then, we certify the usefulness of forecasts by engaging in trading VIX futures, a dynamic tool in investors' arsenal, used massively for hedging, diversification and even capitalization aims.

The findings of our study are quite illuminating. We manage to provide significant evidence on the advancement of ensemble methods and of their different modelling combinations, at least for three out of the six decomposition methods, but only through the objective-based evaluation criterion we applied. The objective-based economic criterion was the cumulative returns earned from trading roll-adjusted VIX futures, by taking long or short positions each day, depending on the forecasted value of VIX. The EMD-based, the EEMD-based and the SSA-based models were amongst the best performing models, with EEMD-HW-AR-HAR by far exceeding expectations, returning 44 times the invested capital for the 5 days ahead horizon and 41 times for the 10 days ahead. The fact that not all decomposition techniques performed equally well, only intensifies the different theoretical and empirical backgrounds that could take a toll when modelling financial data and reveal some shortcomings.

Thus, we point out that ensemble methods do deserve becoming part of forecasting and trading practices when objective-based evaluation criteria are applied. We also stress that the modelling combinations of the produced modes out of the decomposition techniques, do play a critical role in the

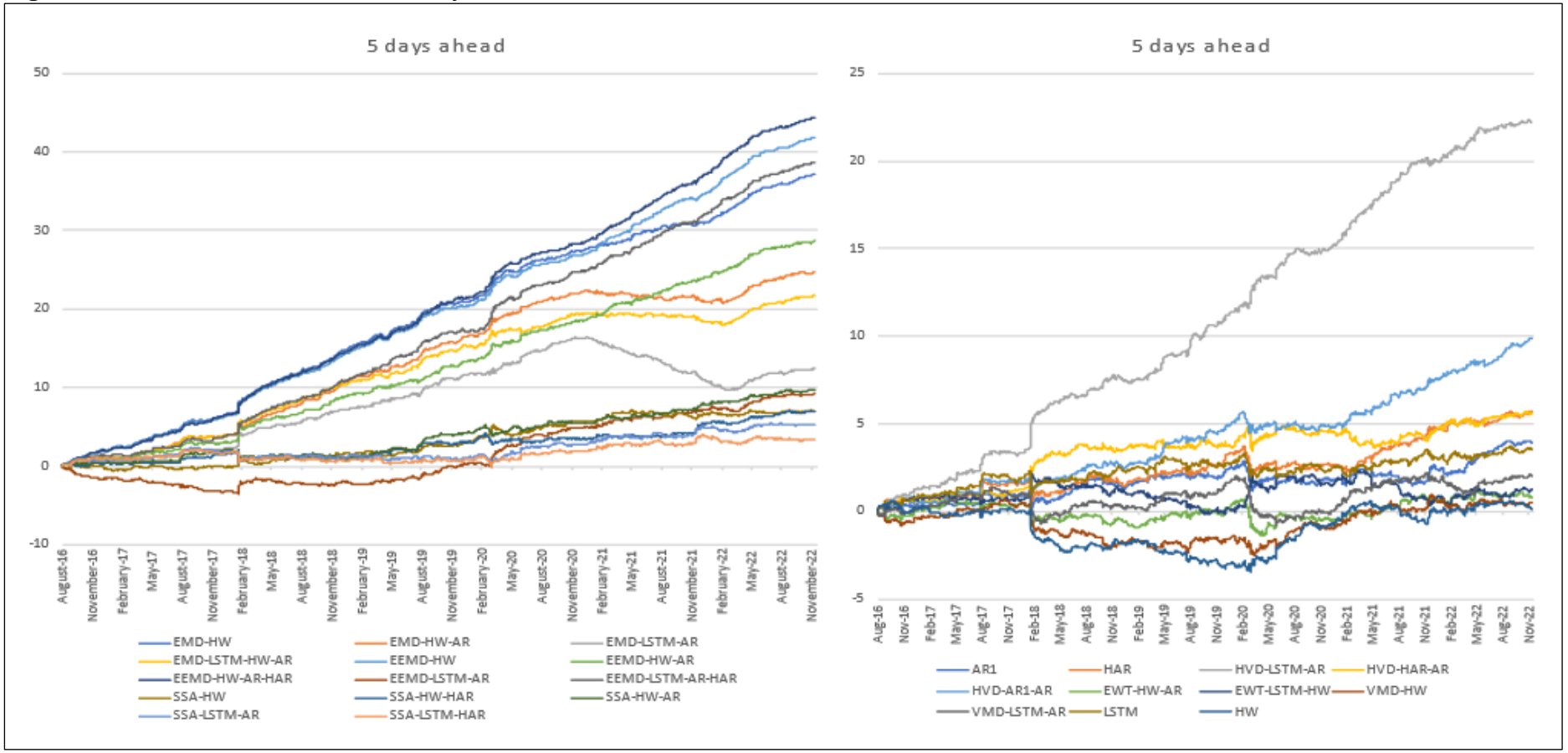
generated cumulative returns, but we definitely highlight the use of VIX futures and their unique return drivers that allowed accelerating cumulative returns accelerate in even higher levels, no matter if the roll of VIX futures was included. Of course, future research can shed new light on above findings by incorporating different decomposition techniques, different forecasting frameworks or different objective-based evaluation criteria. Afterall, markets' offer endless opportunities for experimentation.

Figure 3a. Cumulative returns of 1 day ahead



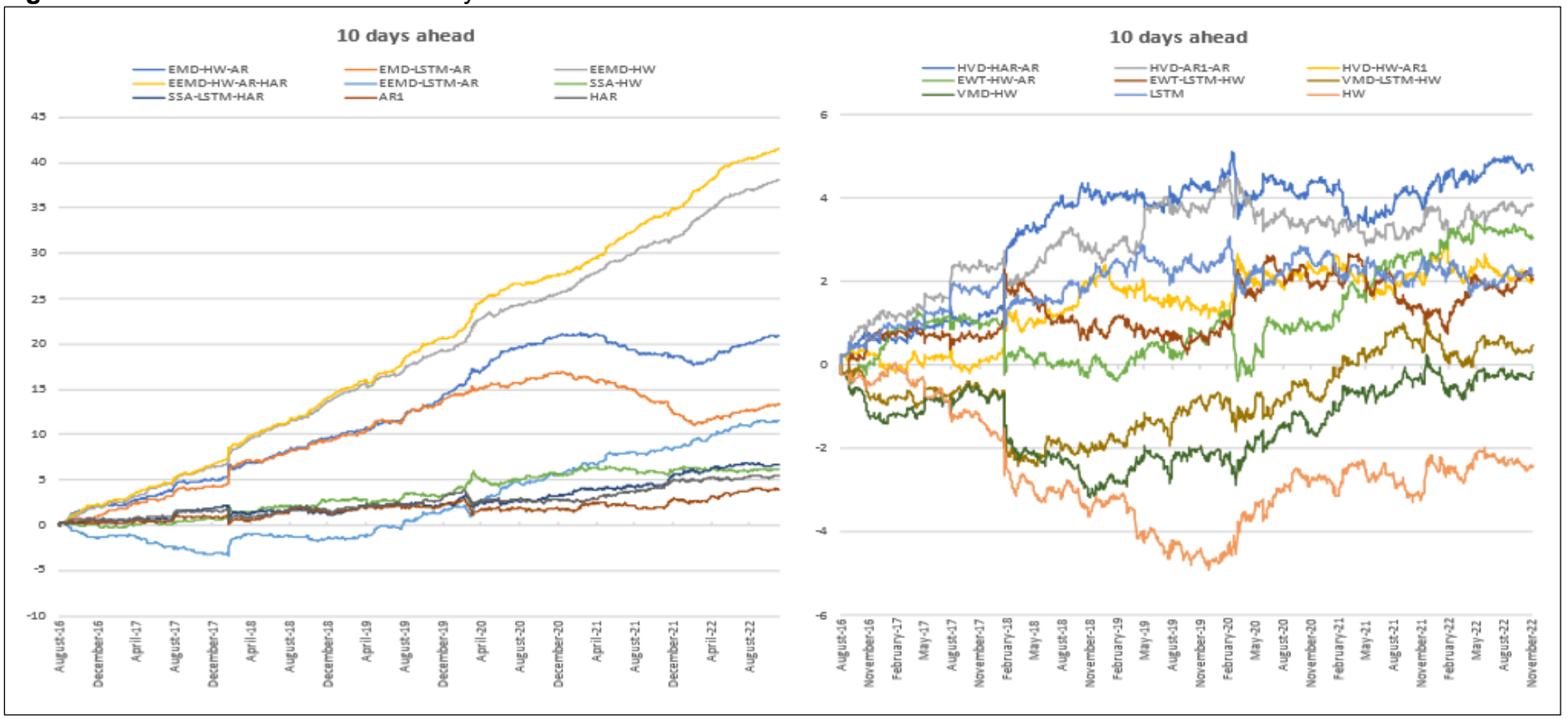
Note: Figure presents in two groups the cumulative returns of some of the various forecasting models for the 1 day ahead forecast horizon. Here we can see clearly how modelling the separate components of the six diverse techniques, with alternating combinations of AR, HAR, HW and LSTM frameworks, results in completely different outcomes, even for the same technique and observe how some models although profitable in the end, recorded a declining trend during the 2020 period where COVID-19 brought global imbalances. The values reported in the y-axis are not in % points.

Figure 3b. Cumulative returns of 5 days ahead



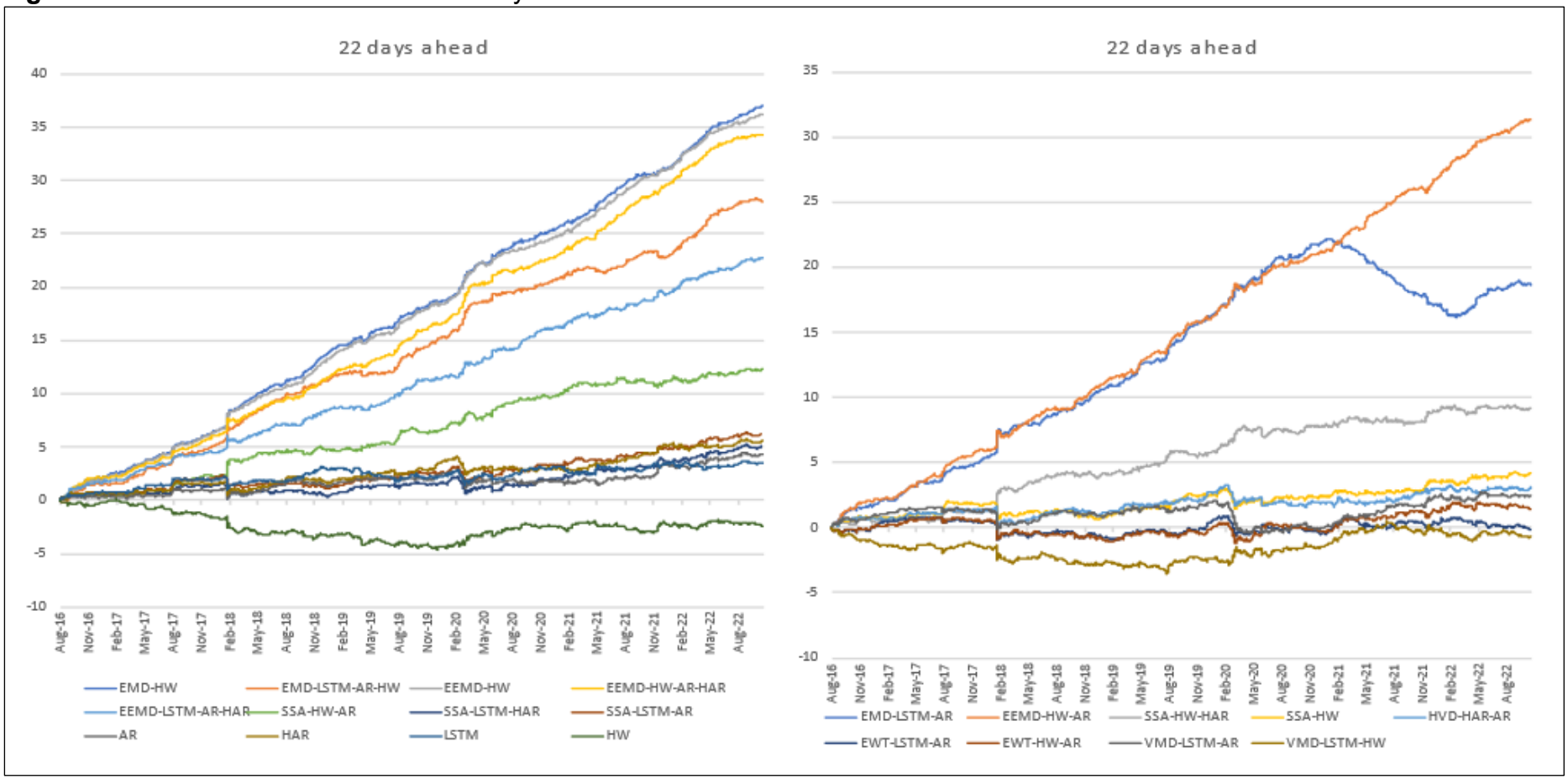
Note: The cumulative returns of some of the proposed models, for the 5 days ahead forecast horizon, appear in two groups. We definitely observe how outcome alters among different combinations and perhaps that is the purpose of experimentation, to find the optimal modelling combinations of components, belonging to the same decomposition technique. The optimal combination is the one that will boost returns. The values reported in the y-axis are not in % points.

Figure 3c. Cumulative returns of 10 days ahead



Note: Different model combinations, result in different cumulative returns. Here the model combinations with higher cumulative returns are more restricted. For the 10 days ahead horizon as for 5 days ahead, EEMD-HW-AR-HAR was the best performing model. Due to space and presentations limitations, we include only few of the 32 models in total. The values reported in the y-axis are not in % points.

Figure 3d. Cumulative returns of the 22 days ahead



Note: Even for the 22 days ahead horizon the group of the best performing models consists of the EEMD-based and EMD-based models. The models that belong to each of the two groups presented here, are randomly chosen between the 32 models. It was impossible to include all proposed models in above figures. The values reported in the y-axis are not in % poin

Chapter 4

Disaggregating realized volatility

4.1 Introduction

Studying realized volatility definitely is not a plain sailing task. We have already touched upon this topic in the introductory chapters, but the time has come to present the study that was conducted on behalf of realized volatility itself. Let us start by giving the exact dimensions with which, research activity has dressed realized volatility.

Realized volatility of asset returns price, to put it simply, is a non-parametric measure of the recorded variation of that asset (Andersen and Benzoni, 2008). It is a measure of dispersion and has gathered immense attention during the past two decades, due to the evolution conducted in econometrics and in volatility estimation and forecasting. The availability of high and ultra-high frequency data changed radically this scientific field. Iconic studies came forth to validate and enhance the construction of model-free realized volatility measures that could replicate the true stochastic process of volatility.

The studies of Andersen and Bollerslev (1998), Andersen *et al.* (2001a; 2001b; 2006; 2012), Barndorff-Nielsen and Shephard (2001; 2002; 2007), Hansen and Lunde, (2006) or Barndoff-Nielsen *et al.* (2008) to name a few, rose a revolutionary flag and reintroduced the concept of realized volatility to new standards. Quadratic variation, bi-power variation, quantile-based realized variance, Min Realized variance or Med realized variance, were some of these measures. These measures could efficiently approximate the

integrated volatility concept and deal with issues such as the systemic bias and the microstructure noise, that were mainly attributed to the way markets operate (Huang and Tauchen, 2006; Hansen and Lunde, 2005). Furthermore, the availability of realized measures in different sampling frequencies, made it possible to choose the best one that could result in unbiased forecasts (Andersen and Benzoni, 2008; Forsberg and Ghysels, 2007; Ait-Sahalia and Mancini, 2006). Hence, realized volatility became a powerful and more informative tool in the hands of all interesting parties, able to compete equally well with implied and even more, modelled along implied. As a result, there was and still is an accelerating research activity that builds upon it, as estimating, interpreting, and forecasting realized volatility remains an ultimate target (Degiannakis and Filis, 2017; Kambouroudis *et al.*, 2021; Degiannakis, 2018).

But this is not the only reason for one to engage in realized volatility analysis and forecasting. There is another reason, that intensifies and makes research on this topic so alluring, a constant reason that is always present and waits for the moment a crisis outburst to intensify, uncertainty that encloses risk. Realized volatility compared to implied is directly linked with the underlying assets as is their actual variability recorded in an almost continuous way. It is a measure of risk for the future returns of the representative asset (Angelidis and Degiannakis, 2008). And conditions in the microenvironment of the markets, have radically changed allowing for higher risk and complexity. Daily trading volumes of derivatives have escalated, new even more complex financial innovations have been launched, new services and operational techniques have been diffused, all resulting in markets'

dynamics to dramatically change, adding extra variability. So, portfolio management and diversification ask for accurate risk prediction and efficient manipulation.

Thus, the aim of this study is to incorporate the disaggregation techniques in order to decompose realized volatility and then model and forecast the individual components, as we did with the VIX index. It is also a chance to prove whether the encouraging findings accomplished through VIX can be quantified here as well. That would prove that the proposed divide and conquer model architecture, worths indeed to became part of the modelling and trading practices. Moreover, let us not forget that modelling realized volatility is a rather challenging and imperative task. That is mainly because realized volatility has an even more persistent nature compared to implied, since it is composed by high frequency data, so it encloses all the abrupt tick movements of an asset during its trading hours. So, being able to generate meaningful and profitable forecasts would justify our choice. Here, we use the 5min returns of S&P500 futures index as a proxy for stock market's realized volatility (Degiannakis and Filis, 2020; Liu *et al.,* 2015) and employ the same model specifications as we did in the previous chapter.

Findings justify our choice. The 16 and 18 times return on invested capital through the EEMD-based and HVD-based models is in close proximity with the findings recorded for the VIX index as well. So, this study adds in both realized volatility forecasting literature and in ensemble modeling.

The remaining body of this study is organized in the following manner. Section 4.2 expands a little the integrated volatility concept. Section 4.3

provides a scarce description of the chosen realized measure. Section 4.4 reports findings, while 4.5 concludes our practice.

4.2 The integrated volatility concept

Prior to disaggregating and modeling realized volatility, it is important to superficially touch upon the way this measure is constructed and the reason that led to its construction. In most studies, the most common way by which realized volatility is addressed is that “realized volatility provides a consistent non-parametric measure of integrated volatility” (Andersen and Benzoni, 2008). We have already referred to few of the various realized measures that were proposed as proxies of the integrated volatility. Thus, we have to dive all little more to what actually all these measures do. They try to replicate as closely as possible an underlying process, the true process, the continuous-integrated process, that every financial element undergoes, and volatility is the variation of that latent, unforeseen process. For the case of implied volatility, the approximation of that process was already active due to the classical pricing models of Black and Scholes (1973) or Hull and White (1987) that were used for its calculation. For realized that approximation and thorough description, dates back only two and so decades, when availability of high frequency data, gave rise to critical studies (Andersen and Bollerslev, 1998; Barndoff-Nielsen *et al*. 2008; Hansen and Lunde, 2006b). So, what is the notion behind model-free realized volatility? At least we are obliged to give a formal representation of that concept, through equation (4.2.1), denoting the stochastic differential equation of an asset’s price P(t) logarithm returns process:

$$dlog(P_t) = \mu_t d_t + \sigma_t dW_t, \tag{4.2.1}$$

where $dlog(P_t) = (logP_{tj} - logP_{tj-1}$, μ(t) is the drift process and σ(t) an instantaneous volatility process, independent of dW_t, that denotes the standard Brownian motion. For this diffusion process the integrated variance over a specific time interval [t-1, t] is:

$$\sigma^2_{[t-1,t]} = \int_{t-1}^{t} \sigma_t^2 dt. \tag{4.2.2}$$

Considering realized volatility is actually the square root of realized variance, that is the square root of the sum of squared intra-day returns, ($RVar_t^d = \sum_{j=1}^{t}(logP_{tj} - logP_{tj-1})^2$), integrated volatility is then approximated by realized volatility, given by:

$$RVol_t^d = \sqrt{\sum_{j=1}^{t}(logP_{tj} - logP_{tj-1})^2}, \tag{4.2.3}$$

where j=1, 2, …, t, stand for the subintervals when prices of an asset are recorded. With indicator d, we indicate that we calculate 1-trading day realized volatility. So, as these subintervals increase, realized volatility converges in probability to the integrated volatility:

$$\lim_{t\to\infty}(RVol_{[t-1,t]}) = \sigma_{[t-1,t]}. \tag{4.2.4}$$

Based on above justifications and on equation (4.2.3), we are allowed to construct the daily realized volatility, that in this thesis is the daily realized volatility of S&P500 futures index prices, recorded at 5-min intraday subintervals for each trading day. And because, through our modeling architecture we use the annualized form of realized volatility, equation (4.2.5), returns the annualized metric:

$$AnRVol_t = \sqrt{252} * RVol_t^d. \quad (4.2.5)$$

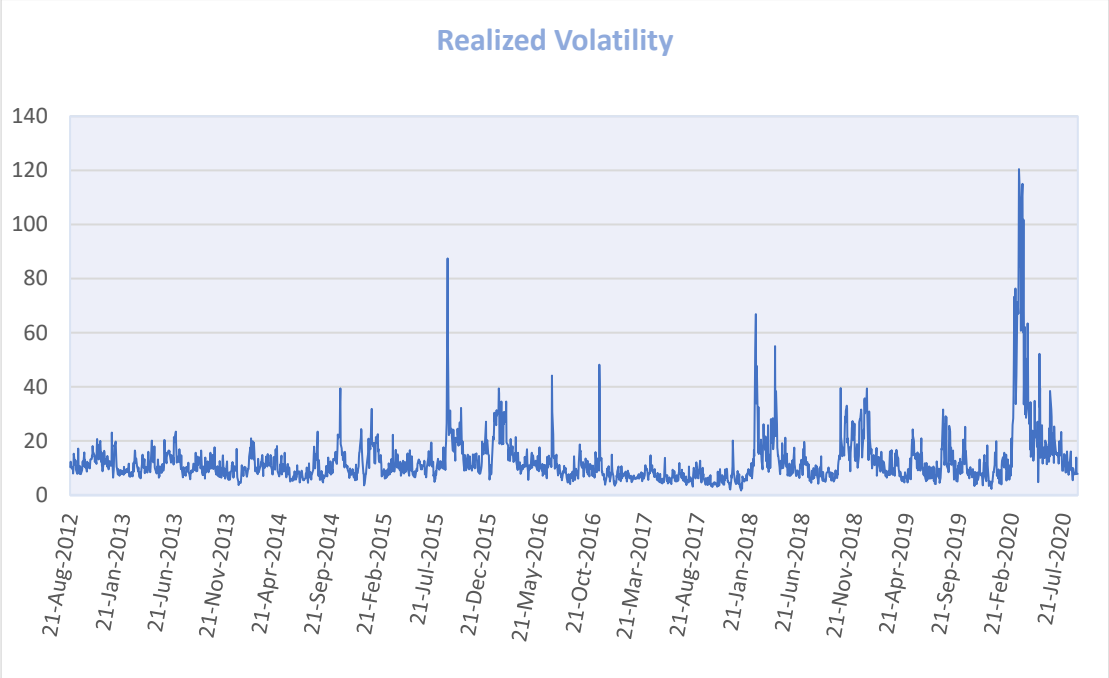
4.3 Data description

For our study we use high frequency (5-min) returns of the S&P500 futures index, for constructing the daily realized volatility. High frequency data were obtained through TickData. The sampling period, based on data availability spans from 21st of August 2012 up to 31st of August 2020. Data are cleaned in a way to reflect the actual trading days within a year, meaning weekends, bank days, 4th of July, Christmas, Easter Monday, Thanksgiving, and the day after, no-trading hours, etc. are excluded. Overall, a total of 2021 daily observations form our sample.

The choice of the 5-min intra-day returns of S&P500 futures index as a proxy for replicating the daily realized measure is dictated by the strong evidence provided by the studies of Andersen *et al.* (2005), Degiannakis and Filis (2020), Liu *et al.* (2015), Bollerslev *et al.* (2011) to name a few, where they experienced sampling frequencies of returns ranging from 5-min and 15-min to 30 min, to effectively eliminating microstructure noise, and other

recorded systemic frictions. Fig. 4.1 gives an optical representation of the constructed realized measure for the period under investigation.

Figure 4.1 The daily constructed realized volatility



Note: the realized volatility constructed out of 5 min returns of the S&P500 futures index for the period under investigation. Realized volatility definitely is more persistent compared to VIX.

Since, we proceed in line with the process implemented on VIX index, it is wise to present the table with the descriptive statistics of realized volatility and the VIX futures that were incorporated for the trading strategy that was thoroughly reported at chapter 2. We will only state that, for the recordings of Table 4.1, the sample of VIX futures extends from 21st of August 2012 up to 31st of August. For the validation of the out-of-sample period and for estimating the cumulative returns of the trading strategy, the sample extending from 12th of August 2016 up to 31st of August 2020 is only used.

From the inspection of table 4.1, we do confirm the same findings that apply also for VIX index. The realized measure is right-skewed, non-normally distributed, leptokurtic, more persistent, and volatile compared to implied.

Moreover, evident is also the mean reverting property of realized (Andersen and Benzoni, 2008).

Table 4.4. Descriptive Statistics of Realized Volatility, the logarithm of Realized volatility and VIX futures

	RV	Log. RV	VIX futures
Mean	12.66	3.76	16.84
Median	10.19	2.32	15.35
Max	120.48	4.79	72.63
Min	1.74	0.53	9.88
St. Dev	9.91	0.52	5.89
Coef. Of Variation	0.78	0.22	0.35
Skewness	4.82	0.77	3.71
Kurtosis	34.58	1.83	20.46
J-Bera	108559.9	484.77	39850.36
ADF	-6.878*	-	-4.241*
Corr. with RV	-	-	0.76

Note: Descriptive statistics for both realized volatility (RV) and VIX futures agree with the stylized facts of volatility. The correlation between RV and VIX futures for the period under investigation also confirms a strong connection, not that strong as VIX but is critical for the trading exercise. The * denotes significance at the 1% level for the ADF, so both series do not exceed unit root issues.

The correlation with VIX futures is another added asset, although not as strong as in the case of VIX index (0,95), which is natural to occur, but strong enough to allow experimentation with the trading strategy we follow as a means of objective-based forecast evaluation. Again, we come across with the model-friendly statistical properties of the logarithmic transformation of the realized volatility, that is the form with which realized volatility is being optimally decomposed, modeled and forecasted.

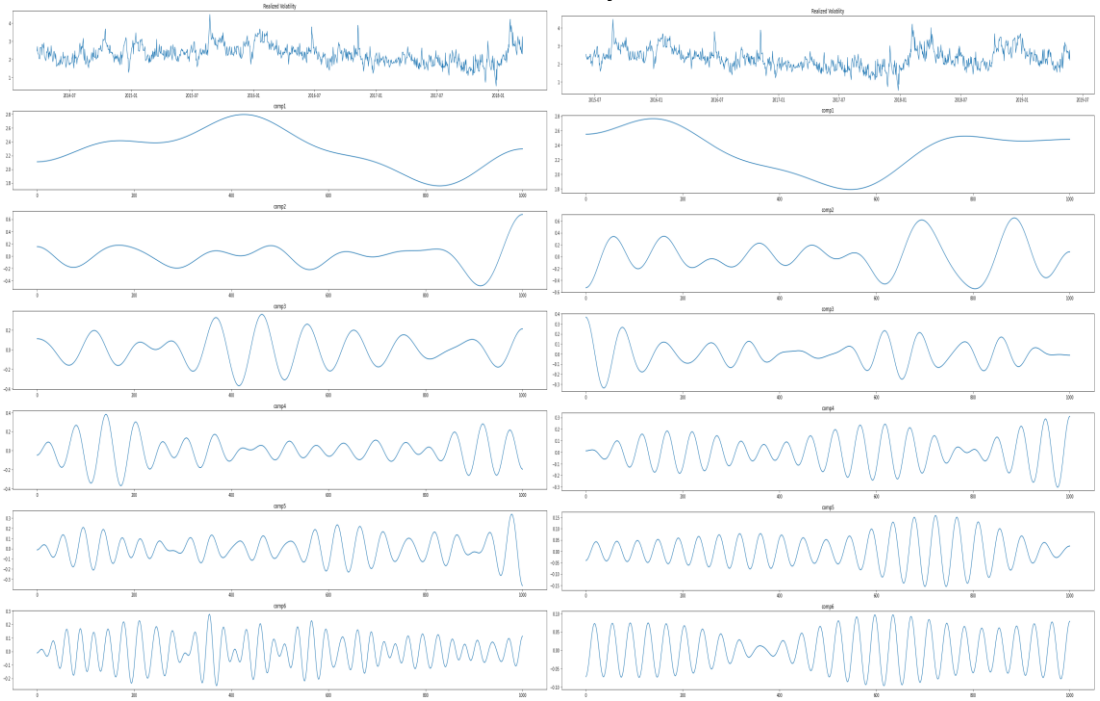
4.4 Empirical findings

This study is in close proximity with the study reported in chapter 3, for the implied volatility, captured by the VIX Index. What we want to achieve is the possible resemblance between these two metrics in the final outcomes and validate whether the profitable outcome of the cumulative returns earned by holding long or short positions in VIX futures according to forecasted values,

mainly attributed to our modeling architecture, could be replicated to other volatility measures, as well. It is wise, an analysis of disaggregation methods to start from visualizing the components of each technique. We work with rolling samples of a fix rolling window of 1000 observations, hence, in each iteration 1000 observations are used for the daily decomposition, modeling and forecasting of the components of the logarithm of realized volatility and the logarithm of realized for the benchmark models. The rest 1021 observations are kept for the out-of-sample validation.

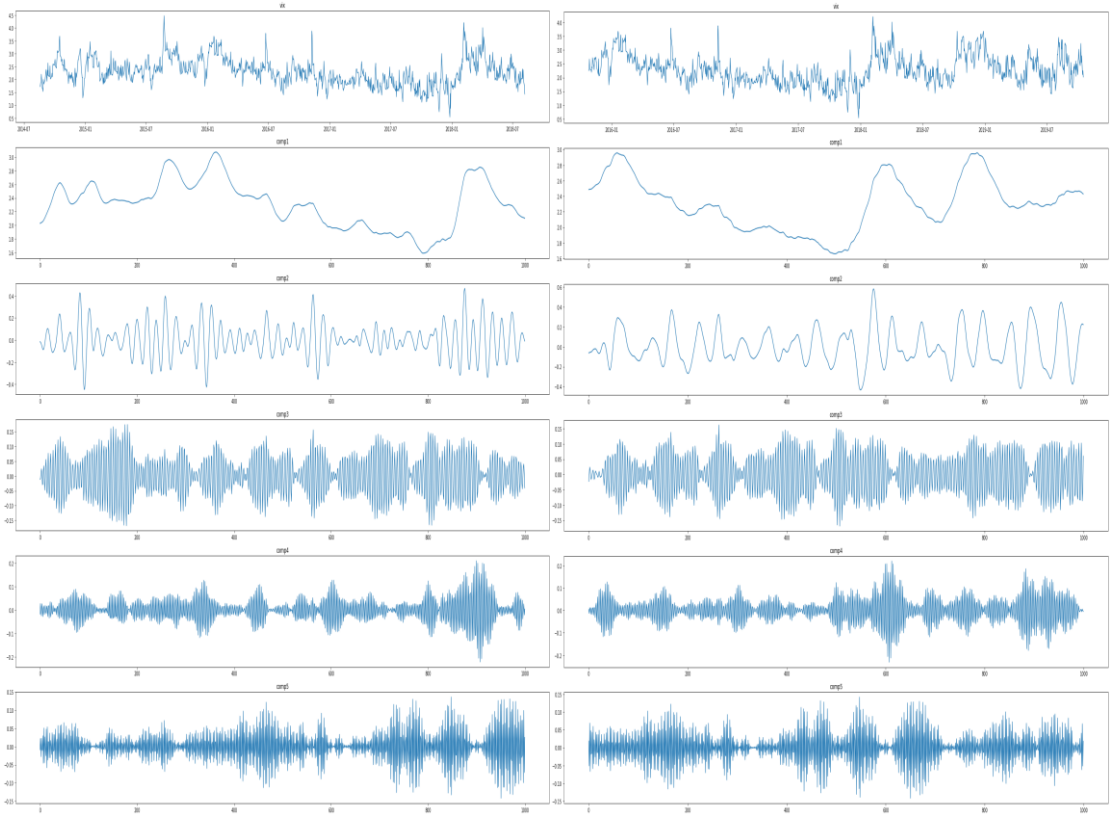
Fig. 4.2 to 4.6, depict the components of two randomly chosen rolling samples of realized volatility for the EWT, the VMD, the SSA, the HVD and the EEMD techniques, respectively. It is an important first step as it the one that reveals the hidden processes original series display and guides the choice of the most appropriate model, able to capture non-linearities, trend, long memory processes etc.

Figure 4.2 Components of realized volatility via EWT



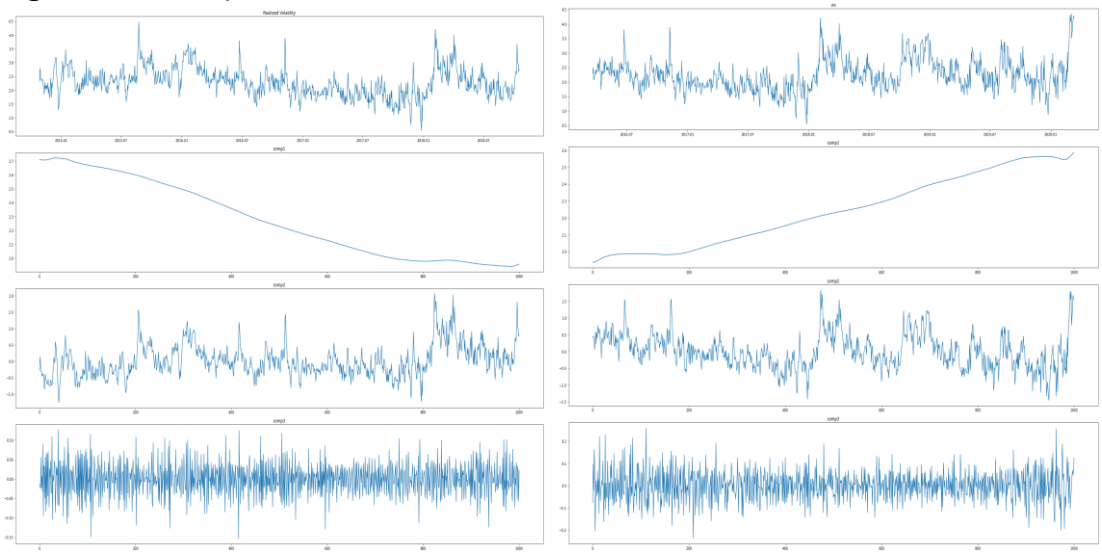
Note: Realized volatility was extremely persistent so we had to go for 6 components for the method of EWT by observing the Fourier spectrum and by estimating the MSE of the reconstruction process. On the left side the 10th sample appears, while on the right the 420th sample is presented.

Figure 4.3 Components of realized volatility via VMD



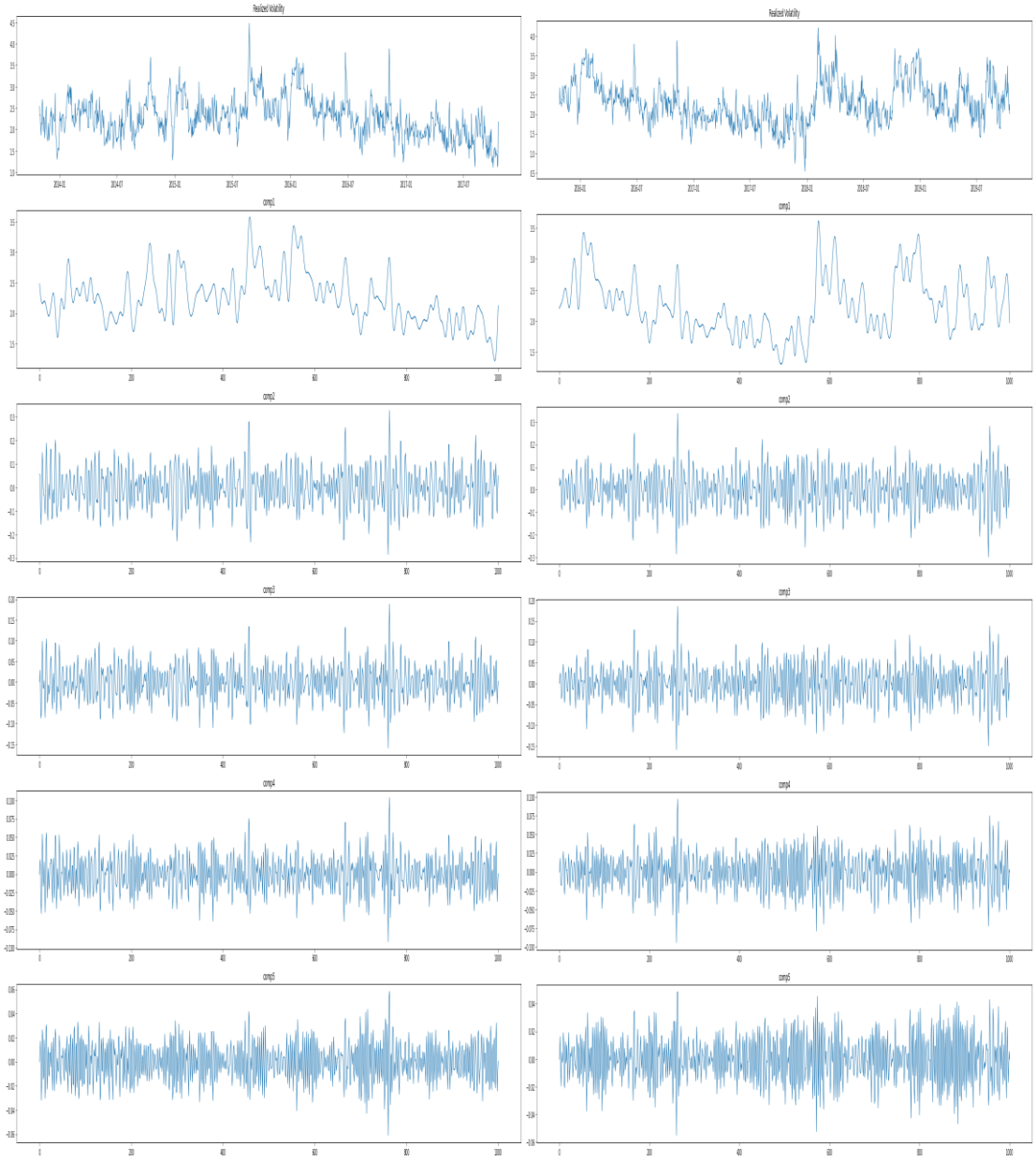
Note: VMD was efficient enough to decompose realized volatility, but components have not generated forecasts that were prosperous enough during the implemented trading strategy. On the left appears the 40th decomposed sample, while on the right the 400th.

Figure 4.4 Components of realized via SSA



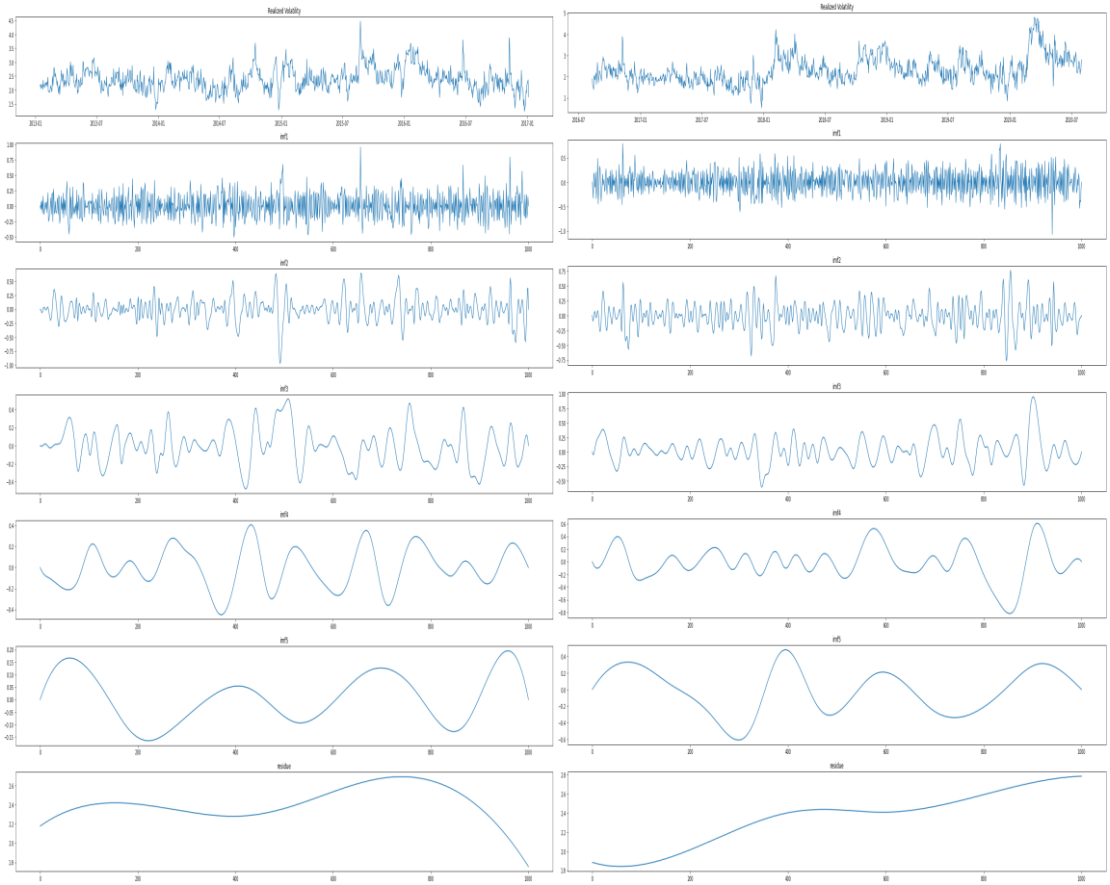
Note: Here we actually gathered ended with three components for the SSA method as for the relative entropy implementation on the best way for grouping. The first component is the trend, but the second component also escape some dynamics of the realized process as its very persistent variable and imposed moments of excess difficulty in clearing the mixed modes. On the there appears the 20th decomposed sample and on the right appears the 550th sample.

Figure 4.5 Components of realized volatility via HVD



Note: HVD due to its low pass filter lying to its core was the most efficient in decomposing the persistent realized volatility and at the same time was the one to generate forecasts that led to the highest cumulative returns for all horizons compared to the rest included models. On the left the 50th decomposed sample is presented, while on the right the 470th.

Figure 4.6 Components of realized volatility via EEMD



Note: EEMD used instead of EMD due to the nature of realized volatility and the fact that EEMD is more suitable for extremely noisy data. EEMD along HVD generated the highest cumulative returns. On the left the first sample is presented, while on the right 800th decompose sample is included.

The second step is the modeling and forecasting and the evaluation of generated forecasts through an objective-based criterion against the classic loss function of the MSE. In this study we are called to answer the same question raised in chapter 3. We are called to specify whether objective-based evaluation criteria are best fitted for studies of economic interest, or for dynamic portfolio construction decisions, and when someone desires a deeper understanding on the way financial markets operate and interact, compared to the statistical ones. We managed to show that for the study conducted on VIX Index, now it is time to test this for the realized volatility.

The table that follows, Table 4.2, reports the values of the MSE and the p-values of its respective MCS test. We have 11 models originating from the ensemble modeling and 3 benchmark models. The MCS, the values of which are formed based on the MSE value, records that benchmark models for all forecast horizons are part of the MCS. Although the proposed frameworks do seem to have returned fair MSE values, definitely HW and HAR are among the winners of the best performing models. Statistically, the advanced model architecture does not provide any gain in predictive accuracy. Thus, statistically speaking there is no reason for a market participant that follows a model-based trading strategy to engage in a modeling framework that adds in complexity but does not add in efficiency, when she/he can always resort to the simpler alternatives.

Table 4.2 MSE and MCS of benchmark and ensemble models

Model	FORECASTING HORIZON							
	1 day		5 days		10 days		22 days	
	MSE	MCS	MSE	MCS	MSE	MCS	MSE	MCS
AR1	63.46	0.027	135.44	0.050	164.61	0.088	**176.71**	**0.210***
HAR	54.06	0.068	**107.58**	**0.172***	141.77	**0.787***	**166.42**	**1.000***
HW	**41.79**	**0.194***	**76.07**	**1.000***	116.44	**0.876***	222.76	0.051
EEMD-HW	**13.94**	**1.000***	137.29	0.050	1717.74	0.001	2627.63	0.006
EEMD-HW-AR	**32.78**	**0.194***	148.87	0.045	1194.43	0.003	1768.11	0,010
EEMD-LSTM-AR	78.83	0.002	**96.36**	**0.600***				
SSA-HW	**40.57**	**0.194***	**78.18**	**0.959***	113.11	**1.000***	214.63	0.051
SSA-HAR	**44.87**	**0.194***	**93.95**	**0.600***	133.37	**0.787***	227.35	0.051
HVD-HW	**22.07**	**0.600***	146.42	0.045	253.13	0.040	318.48	0.038
HVD-AR	**27.49**	**0.194***	**100.23**	**0.172***	164.06	0.088	**193.98**	**0.210***
EWT-HW	55.27	0.068	**101.38**	**0.172***	197.88	0.055	343.95	0.038
EWT-HW-AR	52.76	0.068	**103.34**	**0.172***	160.55	0.088	343.29	0.038
VMD-HW	65.11	0.027	**82.22**	**0.859***	120.23	**0.787***	339.04	0.038
VMD-HW-AR	65.74	0.027	**87.52**	**0.859***	121.51	**0.787***	214.64	0.051

Note: Above benchmark and ensemble models present the MSE of the generated forecasts of the decomposed realized volatility produced out of 5min returns of the SP500 index. The p-value for the MCS of the best performing models is set on 0.15. Interchangeably AR1, HAR and HW are in the set of the best performing models.

EEMD-HW, EEMD-HW-AR, EEMD-LSTM-AR: Constitute the EEMD-based models. HW denotes that all components are modelled via the HW framework. HW-AR denotes that, two components out of the six, are modelled via HW and the rest through AR processes. LSTM-AR denotes that, two components are modelled via LSTM and the rest through AR processes.

SSA-HW, SSA-HAR: The SSA-based models. HW denotes that all components are modelled via the HW framework. HAR denotes that, all components were modelled via the HAR framework.

HVD-HW, HVD-AR: The HVD-based models. HW denotes that all components are modelled via the HW framework. AR denotes that, all components were modelled through AR processes.

EWT-HW, EWT-HW-AR: The EWT-based models. HW denotes that all components are modelled via the HW framework. HW-AR denotes that, one component out of the four is modelled via HW and the rest through AR processes.

VMD-HW, VMD-HW-AR: The VMD-based models. HW denotes that all components are modelled via the HW framework. HW-AR denotes that, one component out of the four is modelled via HW and the rest through AR processes.

But here, again we experience the same paradox that was recorded in the study of chapter 3. According to the objective-based economic criterion that we use, the statistical equality among forecast models, is diminished. Table 4.3 reports the cumulative returns of the models (not in % points) and the respective MCS test p-values. Table 4.3 definitely shows 4 winners for all forecast horizons and these winners by far are not the benchmark models. The two EEMD-based models, the EEMD-HW and the EEMD-HW-AR and the HVD-based, the HVD-HW and the HVD-AR outperformed all other frameworks, economically speaking, followed by SSA-HW for the 1 trading day ahead horizon, the EEMD-LSTM-AR for the 5 days ahead horizon and EWT-HW for the 22 trading days ahead horizon. The reported cumulative returns for the four more prevalent models by far exceed those generated by the rest. The reported MCS test p-values also speak of their superiority. The four models, the two HVD-based and the two EEMD-based, generated cumulative returns for all forecast horizons spanning from 7,91 up to 18,21, that is 7,91 to 18,21 times the initial invested capital. Of course, cumulative returns for the study conducted on VIX index, were higher compared to realized, but realized has a more persistent nature and it was satisfying that some of our ensemble methods generated remarkable cumulative returns compared to the rest frameworks. In this way we provide evidence for two conditions. The first one is that for realized and implied volatility, there is a common framework that generated best results for both and that is the EEMD-based architecture and more precisely the EEMD-HW and EEMD-HW-AR. The second is the fact that indeed objective-based evaluation criteria quantify forecast accuracy on the economic point of view, something that matches with the financial markets' environment and the way decisions are drawn.

Table 4.3 Cumulative returns and MCS values of benchmark and ensemble models

Model	FORECASTING HORIZON							
	1 day		5 days		10 days		22 days	
	Cum. Ret	MCS	Cum. Ret	MCS	Cum. Ret	MCS	Cum. Ret	MCS
AR1	2.77	0.120	2.73	0.102	2.61	0.059	2.58	0.106
HAR	0.12	0.120	0.36	0.016	1.16	0.003	2.12	0.106
HW	-0.54	0.000	-0.48	0.000	-0.56	0.000	-0.33	0.000
EEMD-HW	**14.67**	**0.600***	**13.88**	**0.849***	**17.22**	**1.000***	**13.71**	**0.786***
EEMD-HW-AR	**7.91**	**0.548***	**13.06**	**0.849***	**14.86**	**0.929***	**13.91**	**0,786***
EEMD-LSTM-AR	-11.19	0.000	**11.59**	**0.709***				
SSA-HW	**3.95**	**0.411***	1.76	0.016	4.16	0.140	4.46	0.106
SSA-HAR	2.31	0.120	**3.72**	**0.172***	3.73	0.127	2.80	0.106
HVD-HW	**18.21**	**1.000***	**18.04**	**1.000***	**17.10**	**0.957***	**16.14**	**1.000***
HVD-AR	**15.76**	**0.600***	**14.98**	**0.958***	**12.41**	**0.929***	**3.07**	**0.186***
EWT-HW	1.26	0.120	2.93	0.102	0.92	0.003	**4.79**	**0.190***
EWT-HW-AR	0.77	0.120	3.21	0.102	-0.51	0.000	2.75	0.106
VMD-HW	-1.61	0.000	1.25	0.016	1.08	0.005	-3.84	0.000
VMD-HW-AR	-1.52	0.000	-0.37	0.000	0.24	0.001	-0.87	0.000

Note: Table 4.3 clearly shows the difference in performance when objective-based evaluation criteria are present. The cumulative returns reported along with the p-values of the MCS test limit the number of the best performing models to the EEMD-based and HVD-based models for all forecast horizons. The p-value for the MCS of the best performing models is set on 0.15.

EEMD-HW, EEMD-HW-AR, EEMD-LSTM-AR: Constitute the EEMD-based models. HW denotes that all components are modelled via the HW framework. HW-AR denotes that, two components out of the six, are modelled via HW and the rest through AR processes. LSTM-AR denotes that, two components are modelled via LSTM and the rest through AR processes.

SSA-HW, SSA-HAR: The SSA-based models. HW denotes that all components are modelled via the HW framework. HAR denotes that, all components were modelled via the HAR framework.

HVD-HW, HVD-AR: The HVD-based models. HW denotes that all components are modelled via the HW framework. AR denotes that, all components were modelled through AR processes.

EWT-HW, EWT-HW-AR: The EWT-based models. HW denotes that all components are modelled via the HW framework. HW-AR denotes that, one component out of the four is modelled via HW and the rest through AR processes.

VMD-HW, VMD-HW-AR: The VMD-based models. HW denotes that all components are modelled via the HW framework. HW-AR denotes that, one component out of the four is modelled via HW and the rest through AR processes.

In order to escape from being graphical, we end in the exact same conclusion recorded for the study of chapter 3, and that is the MSE values, and the cumulative returns, demonstrate two contradictive outcomes. Tables 4.2 and 4.3 for the same models, speak of an outcome of no predictive gain and an outcome of accelerated economic significance. Thus, again statistical loss functions fail to predict the direction of market movements resulting in less profits or incremental losses for the market participant or the portfolio manager that would base his investing decisions on the recorded values of a loss function. Table 4.4 reports the rates of the DoC evaluation test and validates the recordings of table 4.3. The two HVD-based and the two EEMD-based models, according to DoC, correctly predicted the direction of realized volatility in a rate approximately ranging from 60% to 66% of times in the out-of-sample period of the 1021 trading days. That range for the VIX index was a little bit higher but cumulative returns were also higher. So, we end that DoC supports findings.

Table 4.4 Direction of Change of benchmark and ensemble models

Model	FORECASTING HORIZON			
	1 day	5 days	10 days	22 days
	DoC	DoC	DoC	DoC
AR1	0.465	0.476	0.481	0.480
HAR	0.482	0.464	0.470	0.476
HW	0.472	0.471	0.474	0.472
EEMD-HW	**0.610***	**0.595***	**0.643***	**0.601***
EEMD-HW-AR	**0.590***	**0.596***	**0.610***	**0.612***
EEMD-LSTM-AR				
SSA-HW	0.510	0.488	0.499	0.497
SSA-HAR	0.477	0.493	0.492	0.478
HVD-HW	**0.661***	**0.658***	**0.632***	**0.646***
HVD-HAR	**0.632***	**0.591***	**0.605***	0.493
EWT-HW	0.485	0.488	0.486	0.510
EWT-HW-AR	0.478	0.491	0.479	0.493
VMD-HW	0.468	0.483	0.478	0.452
VMD-HW-AR	0.469	0.469	0.477	0.453

Note: Direction of Change informs that EEMD-HW, EEMD-HW-AR, HVD-HW and HVD-HAR that were among the best performing models for most horizons for the trading strategy we implemented had forecasted correctly the directional movement of realized volatility for rates ranging between approximately 58% to 66% of our sample of 1021 trading days.

EEMD-HW, EEMD-HW-AR, EEMD-LSTM-AR: Constitute the EEMD-based models. HW denotes that all components are modelled via the HW framework. HW-AR denotes that, two components out of the six, are modelled via HW and the rest through AR processes. LSTM-AR denotes that, two components are modelled via LSTM and the rest through AR processes.

SSA-HW, SSA-HAR: The SSA-based models. HW denotes that all components are modelled via the HW framework. HAR denotes that, all components were modelled via the HAR framework.

HVD-HW, HVD-AR: The HVD-based models. HW denotes that all components are modelled via the HW framework. AR denotes that, all components were modelled through AR processes.

EWT-HW, EWT-HW-AR: The EWT-based models. HW denotes that all components are modelled via the HW framework. HW-AR denotes that, one component out of the four is modelled via HW and the rest through AR processes.

VMD-HW, VMD-HW-AR: The VMD-based models. HW denotes that all components are modelled via the HW framework. HW-AR denotes that, one component out of the four is modelled via HW and the rest through AR processes.

Gathering all together, once more we experience strong evidence on the fact that divide and conquer strategies worth becoming part of the trading practice, although they may come with a little more complexity than classical econometric modeling. Ensemble architecture indeed accelerates cumulative returns and boosts the performance of the benchmark models, when incorporated in the disaggregation techniques and in the modelling combinations of the separate modes. Of course, not all disaggregation-based techniques were efficient but that deals with another issue that we will further discuss shortly after.

As we mentioned throughout this study, that was conducted as part of this PhD thesis, we moved precisely as we did in the study conducted for the implied volatility index. This thesis main purpose was to disaggregate both implied and realized volatilities, two measures with distinctive characteristics that intensively draw research activity, but also are in the heart of financial markets, as major players of financial operations. We wanted to investigate, whether the inclusion of a divide and conquer approach would be more lucrative, for investors, risk managers, etc., than classic methods, when a trading strategy, would be applied. The disaggregation was processed with the help of 6 methods for the case of VIX and 5 for the case of realized, since it was a bit noisier and EMD was excluded. The modelling was processed with three classic models and one from the deep learning environment. The way components are formed deals with the underlaying theory of each method. The addition of each separate component approximates very closely original series, but only for the EMD we have completeness. The choice of the number of components for the techniques that were user-defined, was made

solely depending on the achieved resolution and how well the summation would approximate back the input variable. Thus, it is important to note that final outcome may depend on these issues. Perhaps, the difference among techniques is the one that will guide the final generated profits or losses and perhaps not all techniques are suitable for financial market practices and operations.

Apart from above jurisdiction that applies to both cases, there is once again the fact, that the modelling combinations of individual components for shaping the final aggregated forecast for each method, does have an impact on recorded findings. Figures 4.7 to 4.14, at the end of chapter, visualize the cumulative returns of the MCS models, along the benchmark models and all models together for the 1, 5, 10 and 22 trading days ahead forecast horizon. We do observe how cumulative returns alter between EEMD-HW and EEMD-HW-AR, for all forecast horizons, where modeling combinations of individual components changes. For instance, for the 1 day ahead horizon cumulative returns for the EEMD-HW models are 14,67, while for the EEMD-HW-AR where some components modelled via AR process, cumulative returns diminish to 7,91, and the same applies for the 10 days ahead horizon and for the HVD-based models. So, the correct modelling combination does play a critical role.

And finally, our trading partner, VIX futures and their roll, that played a significant role in trading both realized and implied volatility. We went through experimentation and outcome favored us even though the inclusion of roll and the fact that contracts at that day were mostly in contango, indicating a roll cast for futures markets participants.

4.5 Conclusion

The main scope of this thesis is to contribute to the literature strand that uses ensemble methods for forecasting financial time series data and also contribute to the literature strand that applies objective-based evaluation criteria in order to validate their forecasts. Findings that were confirmed by two successive studies on implied and realized volatility indices more than support this effort. The basic rule is to ad hear to data itself and find ways to explore them based on their nature. We have stressed that before, but it is important to bear in mind the peculiarities that financial data come with and what is their role inside the financial system. An index as VIX or as the constructed realized measure, out of the 5min returns of the S&P 500 futures index, are powerful informative tools used extensively for speculation, hedging, diversification, etc., hence, they are dynamically present in most investing decisions. So, forecasting studies should evolve under the way markets operate. By doing so, findings are better aligned with reality and actually contribute to the relative literature. Here, once more we produce useful for the market participants forecasts that are involved in model-based trading practices.

For forecasting realized volatility, we paired HVD, EEMD, SSA, EWT AND VMD with four parametric and non-parametric frameworks, the AR, the HAR, the HW and the LSTM. Every day for a total of 1021 trading days, we decomposed the realized metric and forecasted for horizons spanning from 1 to 22 days ahead, resulting in 11 models. The evaluation of forecasts was conducted by trading VIX futures contracts, a top financial risk management tool. Findings are aligned with those reported for VIX index, where EEMD-HW

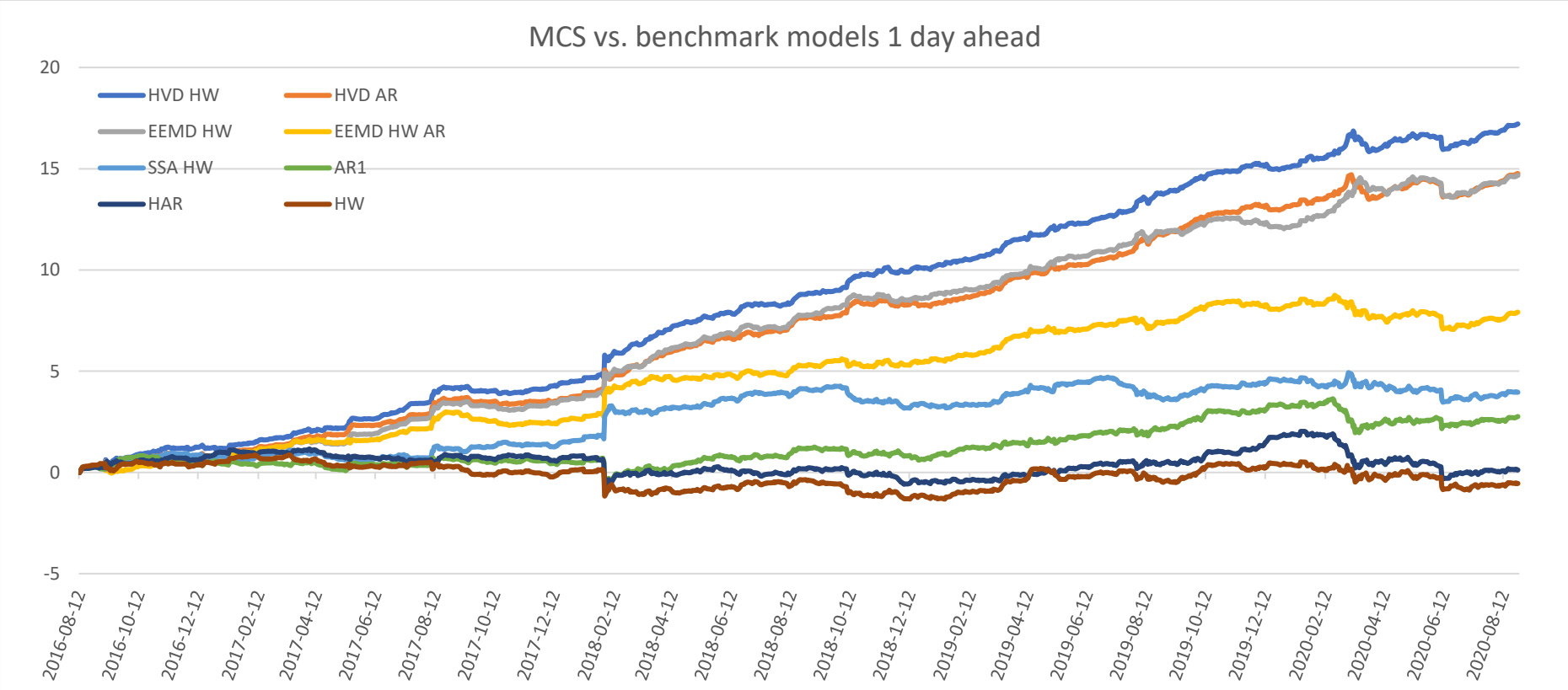
and EEMD-HW-AR were found to be amongst the best performing models for both indices and correctly forecasted the market trend for the 60% to 75% of the times, based to the out-of-sample period, while two out of the 5 decompositions-based approaches concluded in significant profitable economic gains based on the objective-based evaluation criterion.

Study provides significant evidence on the application of ensemble methods, since four models, the two HVD-based and the two EEMD-based, generated cumulative returns for all forecast horizons spanning from 7,91 up to 18,21, evidence justified only by the objective-based evaluation criterion and not through the loss function of MSE. The objective-based economic criterion was the cumulative returns earned from trading roll-adjusted VIX futures, by taking long or short positions each day, depending on the forecasted value of realized volatility. Of course, not all decomposition techniques outperformed but this is perhaps solely attributed to the different theoretical and empirical backgrounds that can signify that not all techniques are suitable for financial data.

Concluding, we do support the fact that ensemble methods do deserve becoming part of forecasting and trading practices, when objective-based evaluation criteria are applied and should recapitulate the three findings that were recorded for both studies 1) indeed VIX futures have unique return drivers (Moran and Dash, 2007; Szado, 2018), when the actual market conditions are considered and roll is included, 2) modelling combinations of occurring components, do play a critical role and 3) most importantly, the proposed modelling architecture, paired with objective-based evaluation criteria, worth becoming part of model-based trading practices adopted by

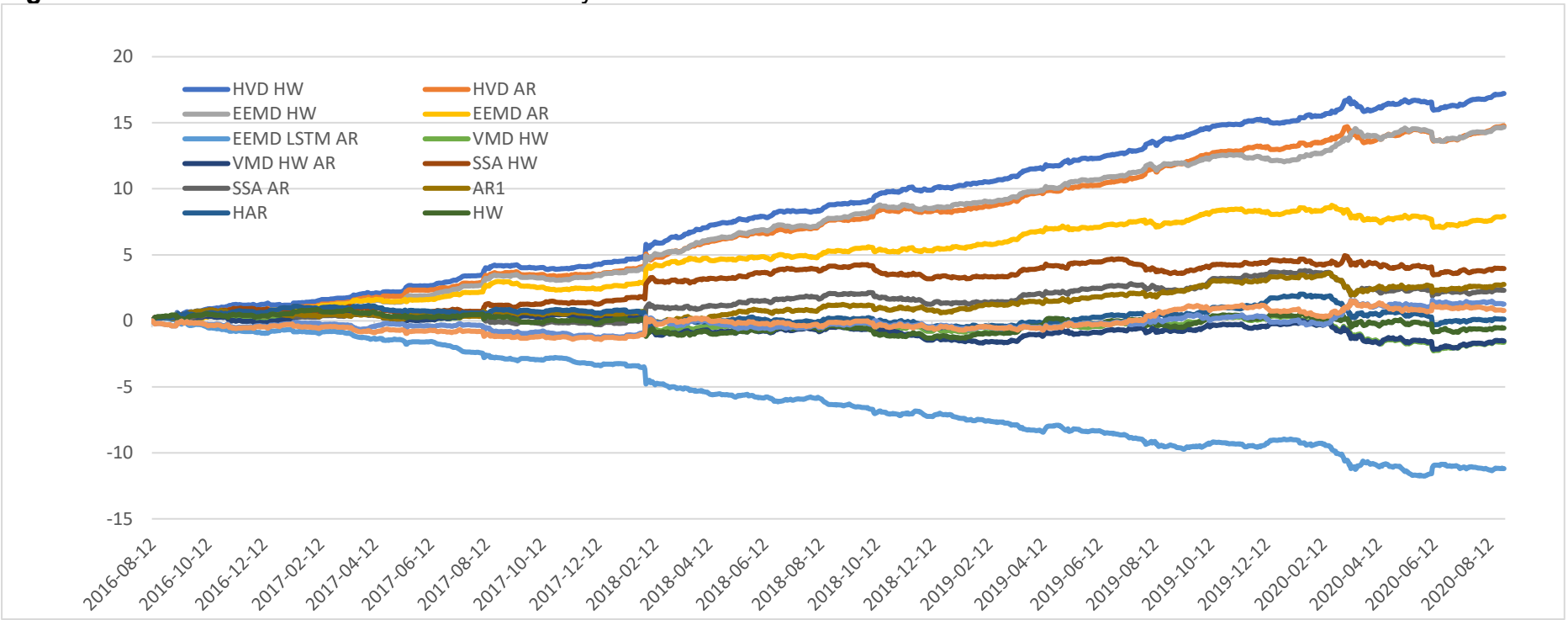
interested parties, since it is multi beneficial and both studies resulted more or less in the encouraging findings.

Figure 4.7 Cumulative returns of MCS models of 1 day ahead



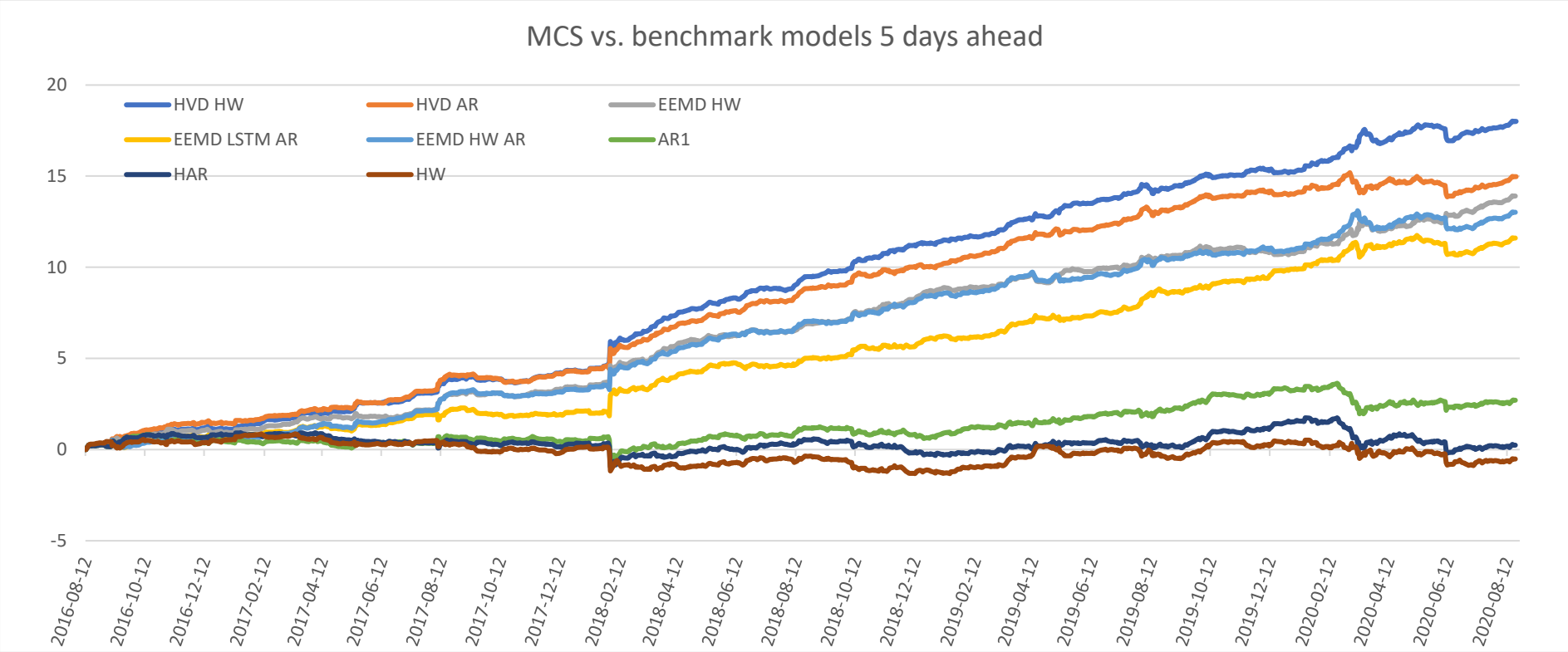
Note: Cumulative returns of the best performing models, according to MCS, along with the ones of the 3 benchmark models, for the 1 day ahead forecast horizon. All 3 benchmark models have mediocre or negative returns, compared to EEMD-based and HVD-based models, which moved with increasing rates during the period under investigation. The values reported in the y-axis are not in % points.

Figure 4.8 Cumulative returns all models of 1 days ahead



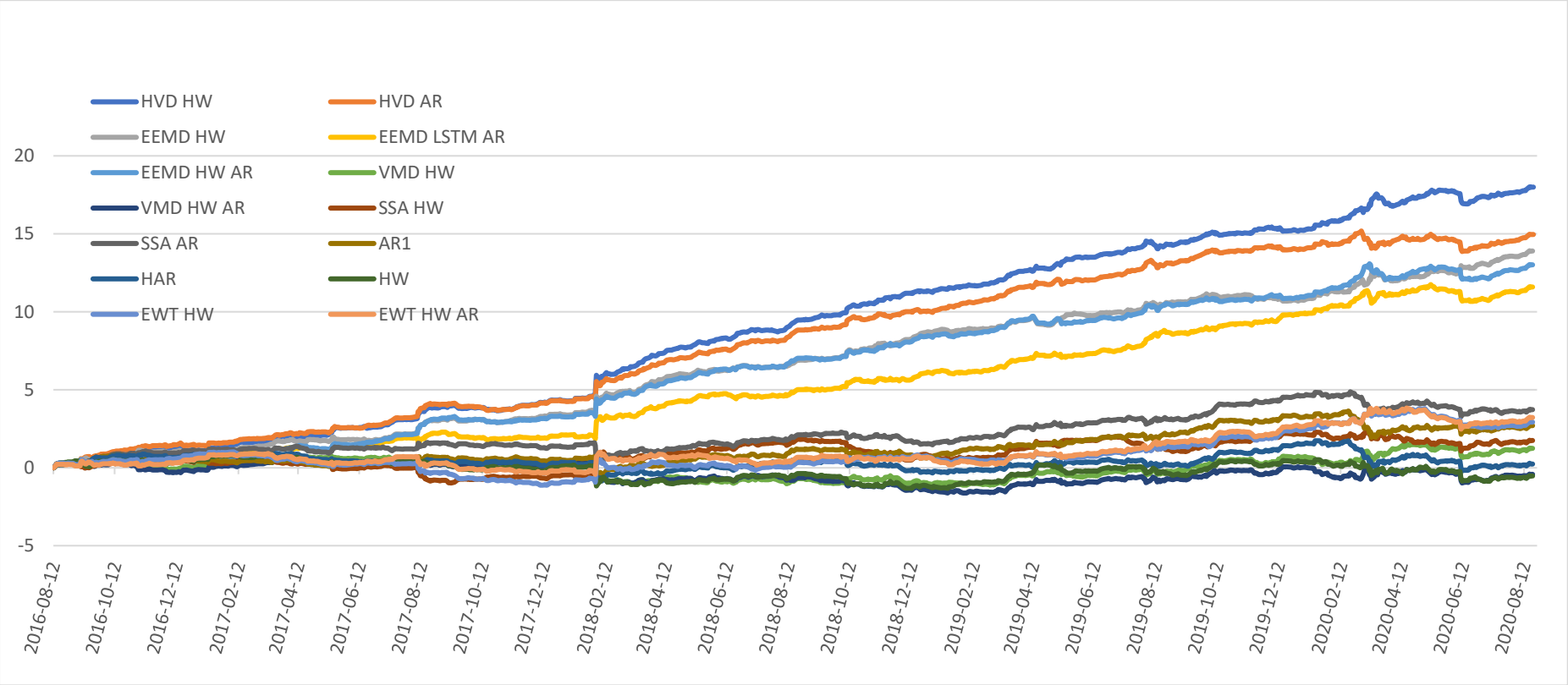
Note: Cumulative returns of all proposed models, for the 1 day ahead forecast horizon. there are significant divergencies between the different theoretic underground models, with most recording mediocre or negative returns for entire period. Here the EWT-based are not recorded. The values reported in the y-axis are not in % points.

Figure 4.9 Cumulative returns of MCS models of 5 days ahead



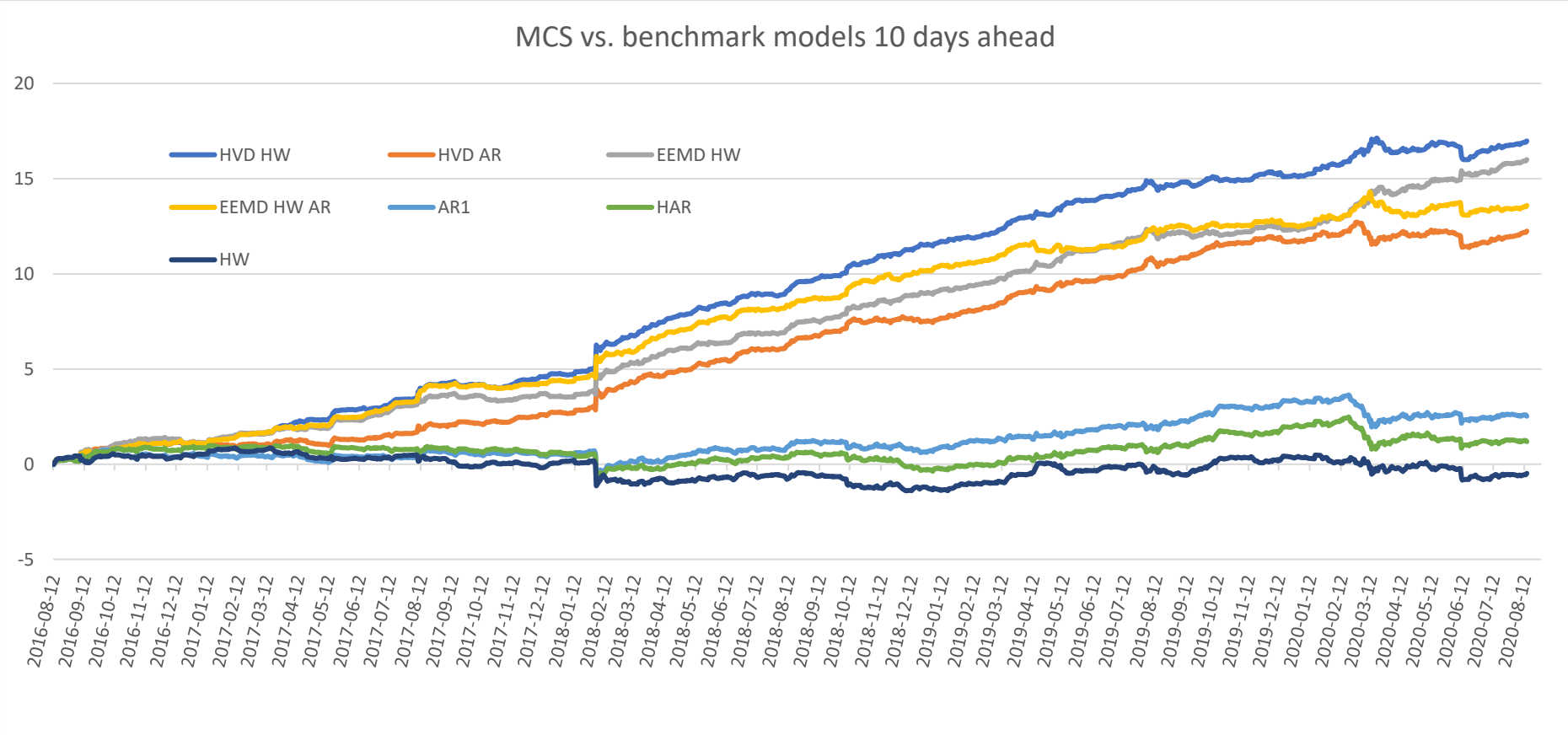
Note: Cumulative returns of the best performing models, according to MCS, along with the ones of the 3 benchmark models, for 5 days ahead forecast horizon. All 3 benchmark models have mediocre or negative returns, compared to EEMD-based and HVD-based models, which moved with increasing rates during the period under investigation. In this study conducted for realized volatility, we don't have the various model combinations we had when modelling VIX and we don not have EMD-based models, so we do not observe the bizarre downward trends spotted in the study of VIXThe values reported in the y-axis are not in % points.

Figure 4.10 Cumulative returns all models of 5 days ahead



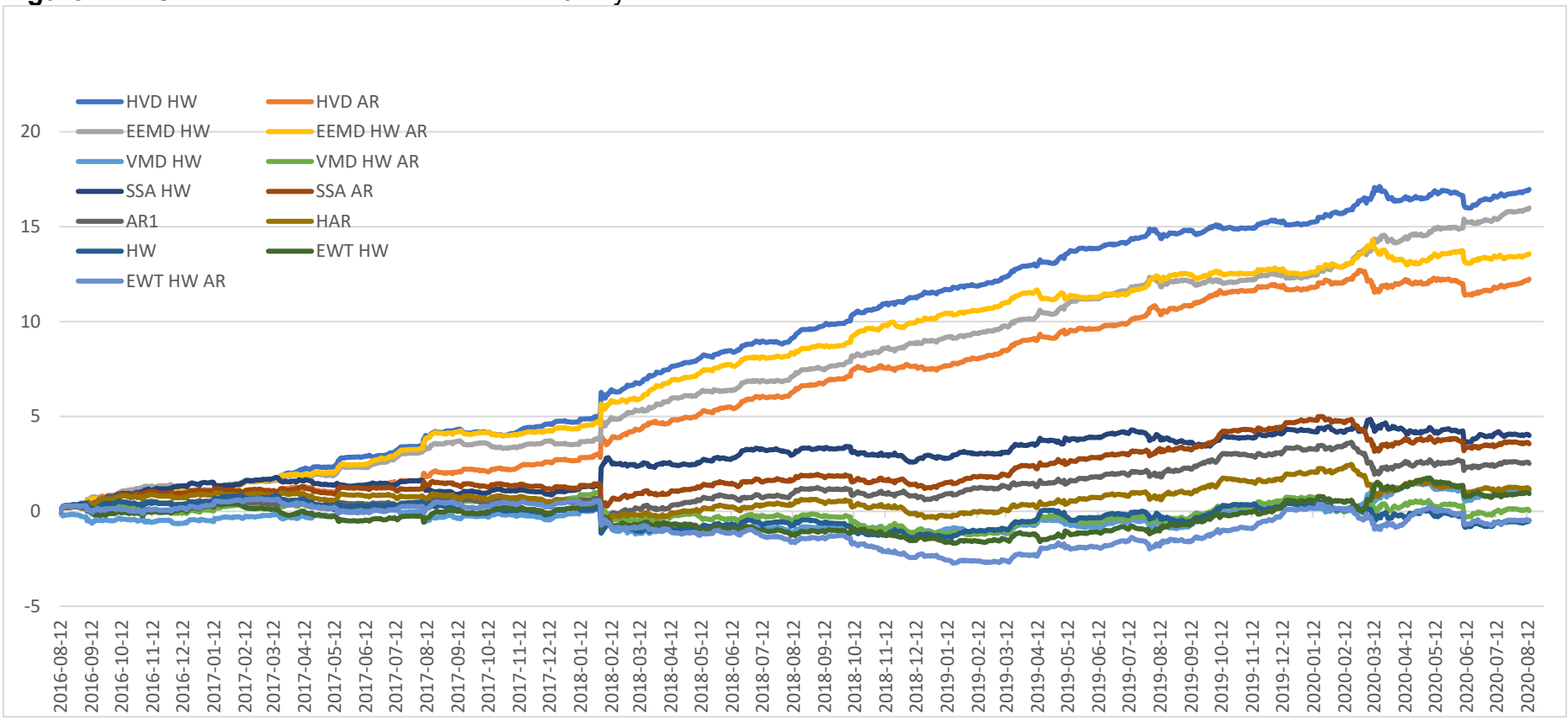
Note: Cumulative returns of all proposed models, for 5 days ahead forecast horizon. Again, here we observe that only the four models that were included in the MCS are the only with constantly increasing significant cumulative returns for entire period under investigation. The values reported in the y-axis are not in % points.

Figure 4.11 Cumulative returns of MCS models of 10 days ahead



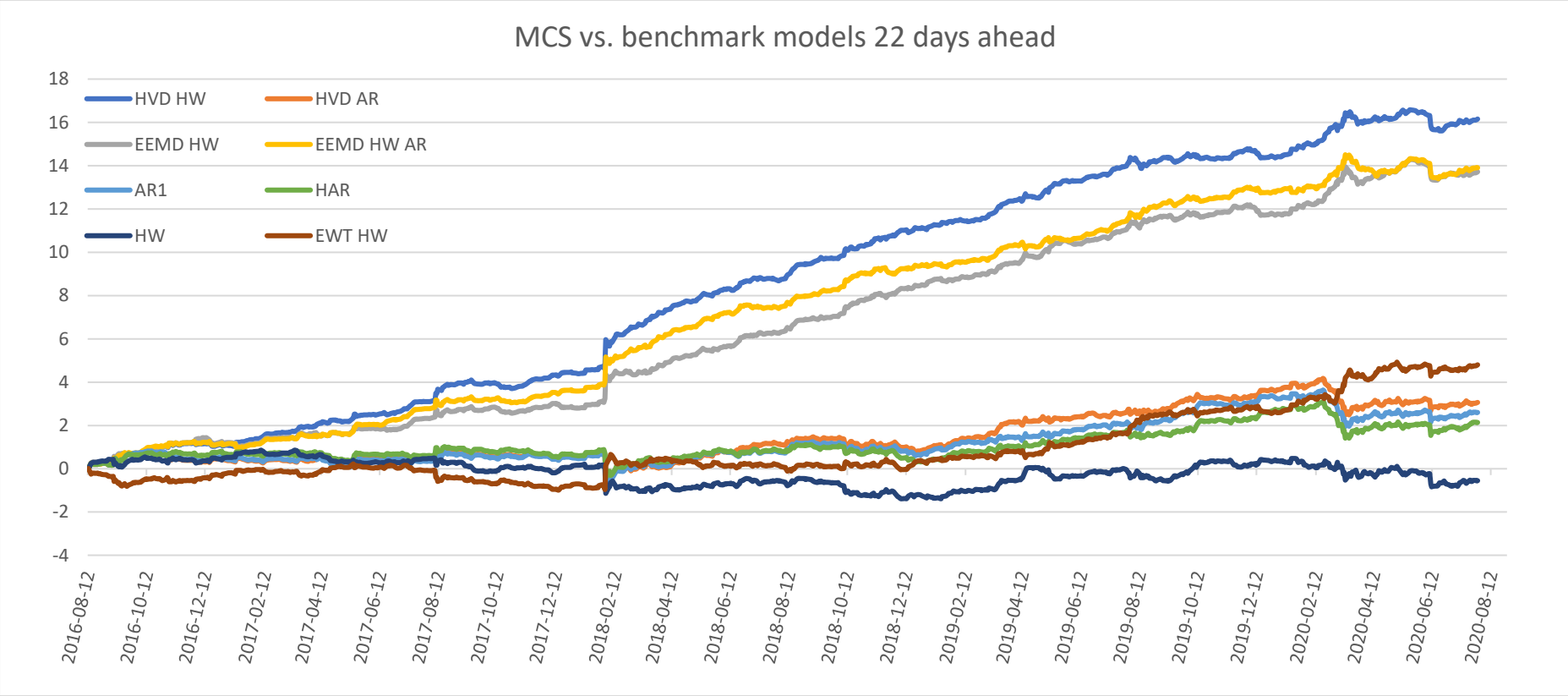
Note: Cumulative returns of the best performing models, according to MCS, along with the ones of the 3 benchmark models, for the 10 days ahead forecast horizon. Two of the benchmark models have mediocre cumulative returns, while HW has negative returns for most of the period, compared to EEMD-based and HVD-based models, which moved with increasing rates. The values reported in the y-axis are not in % points.

Figure 4.12 Cumulative returns all models of 10 days ahead



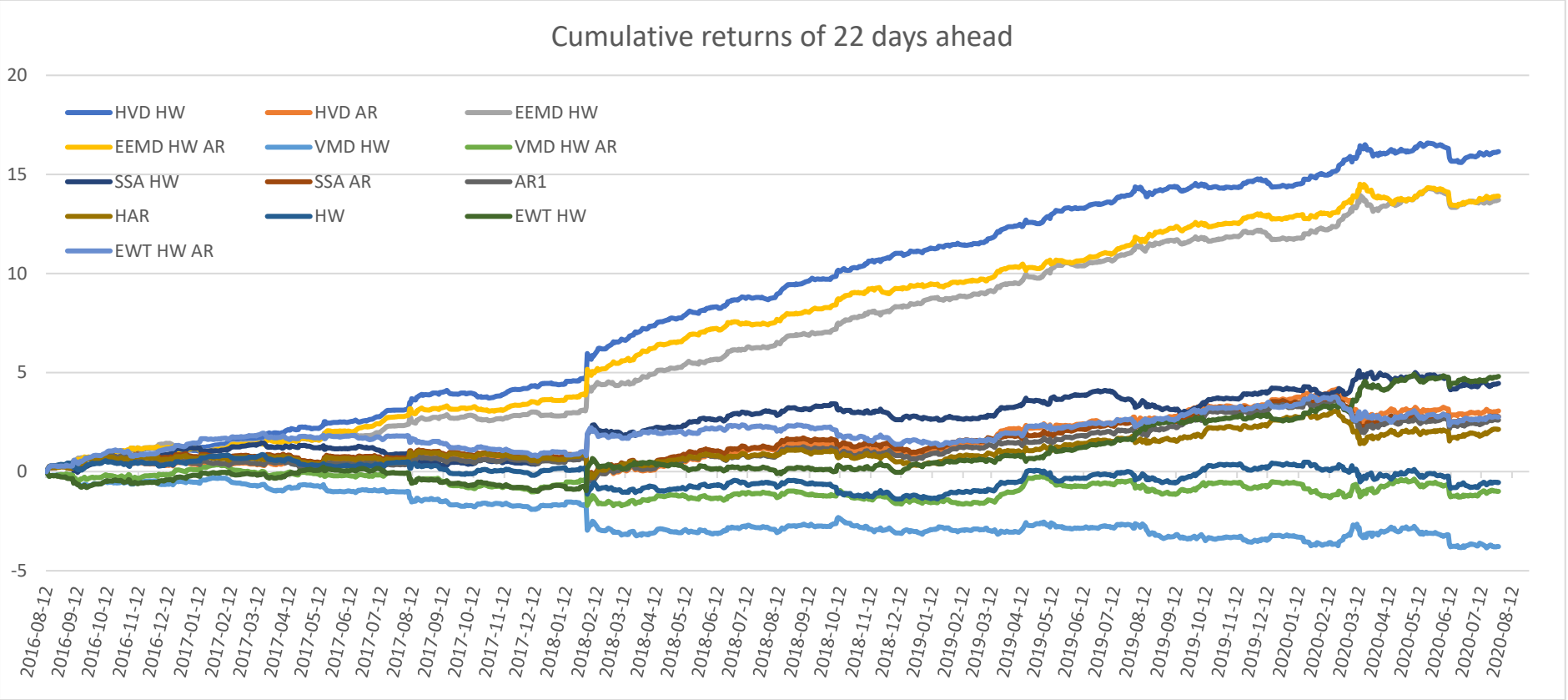
Note: Cumulative returns of all proposed models, for the 10 days ahead forecast horizon. Again, only the four models that were included in the MCS are the only with constantly increasing significant cumulative returns for entire period under investigation without abrupt increasing or decreasing trends. The rest models although some record steadily positive cumulative returns, did not present the accelerating pattern the MCS models, present. The values reported in the y-axis are not in % points.

Figure 4.13 Cumulative returns MCS vs benchmark models of 22 days ahead



Note: Cumulative returns of the best performing models, according to MCS, along with the ones of the 3 benchmark models, for the 22 days ahead forecast horizon. Again, the same pattern is illustrated that highlights the distance between the 3 out of the 4 MCS models and the benchmark models. the EWT-HW that was included in the MCS for the 22 days ahead horizon has only mediocre cumulative returns at least for the last 6 months of the period under investigation and significantly diverges from the constant upward trend recorded for the 3 MCS models. In the 22 days forecast horizon HVD-AR ill performed. The values reported in the y-axis are not in % points.

Figure 4.14 Cumulative returns of all models of 22 days ahead



Note: Cumulative returns of all proposed models, for the 22 days ahead forecast horizon. Only the 3 models out of the 4 that were included in the MCS are the ones with constantly increasing significant upward trend cumulative returns for entire period under. The rest models although some record steadily positive cumulative returns, did not present the accelerating pattern the MCS models, present. The values reported in the y-axis are not in % points.

Chapter 5

Multiresolution analysis

5.1 Introduction

In a thesis oriented in decomposition techniques, as the first step for a more integrated analysis on modeling and forecasting financial time series, the desire to further expand practice and unearth the extra possibilities the signal processing field has to offer, evolves almost naturally. In chapter 2 we introduced the decomposition tools, that were best applicable to financial data and were able to return components that could efficiently be forecasted. During this process we had to exclude some techniques that seemed not to suit in our modelling architecture and stick to the ones generated meaningful outcomes. One of these techniques was the discrete wavelet transform, DWT, that we instead employed the EWT framework. Through the DWT we came across the continuous wavelet transform, CWT, that in most research activities is paired along with the wavelet coherence concept, WC, for a detailed time-frequency domain analysis. It was then that we realized of the extent and practical applications, mathematical tools originating from the wavelet theory, can have in diverse scientific fields.

CWT and WC, appear in many studies targeted in the analysis, decomposition and interrelation of non-stationary and non-linear systems, signals, images, seismic, geological, meteorological, engineering, medical data, etc. (Rodriguez-Murillo and Filella, 2020; Pino *et al.*, 2016; Labat, 2008; Chevalier *et al.*, 2014). However, the last fifteen and so years the appearance in studies that incorporate above frameworks in order to examine business

cycle synchronization, intermittency or contagion between countries, financial markets etc., steadily gained ground towards classic modelling perspectives. The main reason was the resolution of the misconception between qualitative and quantitative findings, when the inclusion of statistical significance tests added robustness to the wavelet analysis concepts (Torrence and Combo, 1998). One of these studies is of Crowley and Mayes (2008), who examine in pairs the possibility of growth synchronization of Germany, France and Italy, by employing CWT and WC. They spot the periodicity of longer and shorter growth cycles and present how closely or not respective economies co-move. Business cycle synchronization for the euro area, through wavelet analysis, also study Aguiar-Conraria and Soares (2011) and demonstrate how synchronization, among neighbouring countries and more distinct ones, varies. Their findings are important for the implementation of the common policies initiatives. Rua and Nunes (2009) explore the co-movement of the German, Japan, United Kingdom, and United States stock markets returns on different frequencies, by incorporating wavelet analysis in an effort to highlight the profits of international portfolio diversification. Risk management, study Gençay *et al*. (2005) for Germany, United States and United Kingdom, at different time scales via multiscale analysis and observe how systematic risk of market portfolio alters at different scales, something essential for asset allocation practices. Vacha and Baruník (2012) incorporate wavelet coherence to investigate co-movements among energy commodities. Their aim is to study how this time-varying dependence impacts portfolio's risk exposure. Their findings suggest that portfolio managers should account for the strong co-movements of crude oil, heating oil and gasoline at various

investing horizons, for a successful portfolio diversification. From the perspective of portfolio diversification and economic policy implementation Yang *et al.* (2016) discriminate between contagion and interdependence observed among foreign exchange markets at different time scales with the help of WC. They do report the existence of contagion between euro/yen and euro/British pound during the global financial crisis and the European dept crisis and strong interdependence between the pound and the euro at all periods. The same discrimination, follow Alqaralleh and Canepa (2021) and proceed in a multiscale analysis for the stock markets returns of six global markets. They distinguish between contagion occurring in higher frequencies and interdependence at lower frequencies. Stock markets' returns co-movement also study Aloui and Hkiri (2014) for the case of the Gulf Cooperation Council countries through wavelet coherence and end up with interesting findings.

Obviously, we the number of studies embracing the power of wavelets for a time-frequency analysis of economic variables, and financial markets, steadily escalates. The fact that we are allowed to investigate not only their interrelations and co-movements, but also the lead-lag effect, makes these frameworks even more appealing, especially when forecast is present. There is a strand in volatility forecasting literature that examines the interrelation of implied and realized volatility, in multivariate modelling settings, by employing implied volatility in order to forecast realized and vice versa for stock markets, derivatives markets or energy markets (Christensen and Prabhala, 1998; Busch *et al.,* 2011; Haugom *et al.,* 2014). The ability to possible spot the exact period, these two metrics seemed to co-move and to spot the direction this

relationship had, will possible shed light to how strong this relation is, when it occurred and why some studies enhance their models' forecast accuracy, with the inclusion of the one or the other measure and why others fail to give evidence (Jorion, 1995; Degiannakis and Filis, 2020; Canina and Figlewski, 1993; Delis *et al*., 2023). Moreover, let us not forget of the volatility risk premium that incurs by subtracting realized volatility out of implied. Risk premium constitutes a key element for traders, especially those engaging in derivatives markets, that by far are the most active. Risk premium directly involves in shaping the appropriate trading/hedging strategies depending on the preferred time horizon and the accepted risk an investor is willing to undertake (Bollerslev and Todorov, 2011; Andersen *et al*., 2009; Bekaert and Hoerova, 2014) but also has an important impact on the macroeconomy variants as well (Barndorff-Nielsen *et al*., 2010; Bekaert *et al.,* 2013).

Hence, discovering the extent and the power of the interrelation of these two vital elements of the financial system will always come forth as imperative as ever. For this reason, here, we incorporate two implied volatility indices, namely the VIX volatility index and the OVX volatility index, two realized volatility indices, the one constructed out of 5min returns of S&P500 futures index and the one constructed out of 5min returns of the WTI futures index, the S&P500 stock market index and finally that WTI crude oil price index. With the VIX index and the realized volatility of the S&P500 futures index, we are already familiar, from chapters 3 and 4. The OVX volatility index is a new inclusion. OVX is the 30-day volatility index for the WTI crude oil price. For the OVX index, the suited realized volatility is the one constructed from a respective index, the WTI futures index. As for the other two indices,

the S&P500 index will be paired with VIX and WTI crude oil price index, will be paired with OVX, as representatives of the stock and energy market, respectively. So, we create 6 pairs, two implied-realized volatility pairs, two implied volatility-market indices pairs and a combination of two implied volatility indices. We choose the stock and the energy market, as both are key elements of the international financial chessboard, both boast a wide and advanced network of derivatives products and there exists an extensive literature dedicated to investigating volatility's transmission mechanisms (Degiannakis and Filis, 2017; Chatziantoniou *et al.,* 2020).

The rest of the chapter is unfolded in the following manner. Section 5.2 provides the appropriate mathematical equations that form the theory of CWT and extent to the WC between two time series. Section 5.3 gives a proper description of the implied and realized measures incorporated, while 5.4 examines the findings of the implemented techniques. Finally, section 5.5 concludes the present study.

5.2 Wavelet analysis

5.2.1 The continuous wavelet transform

A primitive CWT dates back to 1940's and the work of Ricker (1940) when he introduced the term wavelet. Then CWT was officially established by Grossman and Morlet (1984), who introduced the term continuous to stress out, the ability of a wavelet to be transformed in any time scale. Of course, CWT became more compact through the numerous studies that followed and demonstrated (familiarized) its use and its features even further and intensified the ability to analyse a signal/time series in both time and frequency domain (Mallat, 1989; Daubechies, 1992; Meyer, 1997). WC on the

other hand, appeared in the study of Torrence and Webster (1997) and soon became the preferred method for examining possible coherence between two time series, by "correlating" their CWTs.

In order to understand the mechanisms behind CWT we could start from the very own nature of a wavelet that is defined as a zero-mean, of short duration, oscillating function, localized in both time and frequency (Wirsing, 2020). There are many different families of wavelets named after their creators, the hair (Haar, 1910), the daubechies (Daubechies, 1992), the rieker (Riecker, 1940) etc., each baring different characteristics, intended for different types of time series data and complex systems. Among those families there is a wavelet extensively employed for conducting multiresolution analysis in real signals and that is the morlet wavelet[27], expressed in the following equation in a way that satisfies the admissibility condition[28] (Torrence and Combo, 1998; Grinsted ***et al.***, 2004; Aguiar-Conraria and Soares, 2014):

$$\psi_0(\mathrm{t}) = \pi^{-1/4} e^{i\omega_0 t} e^{-\frac{1}{2}t^2}, \qquad (5.2.1)$$

where t is the dimension-free time and ω_0 the frequency. Here in this study, we incorporate the morlet wavelet as the mother wavelet, with frequency of value 6, as proposed in adequate literature. Now, from the mother wavelet function (to be more precise with the scientific characterization), $\psi_{s,\tau}(\mathrm{t})$, we are able to create a family of wavelets is created, by just altering the scaling parameter, s, and the translation parameter, τ, that are its width and its shift in

[27] The morlet wavelet is the one that we incorporate for the CWT and the cross wavelet transform for retrieving the wavelet coherence figures for the pairs of VIX-Realized volatility, VIX-SP 500 etc.
[28] Guaranties that time series will be recovered from its wavelet transform-inversion of CWR.

time (Grinsted *et al*., 2004) and define the position of the wavelet in the frequency and in the time domain (Aguiar-Conraria and Soares, 2014):

$$\psi_{s,\tau}(\mathrm{t}) = \frac{1}{\sqrt{|s|}}\,\psi\left(\frac{t-\tau}{s}\right)dt, \qquad s,\tau\,\epsilon\mathbb{R}, s \neq 0 \tag{5.2.2}$$

Regarding now the CWT of a time series, X_t (with t=1, 2, …, T), it is defined as the convolution of X_t with the scaled and translated wavelet, $\psi_{s,\tau}(t)$, best portraited in Eq. (5.2.3) (Torrence and Combo, 1998):

$$W_t(s,\tau) = \int_{-\infty}^{\infty} X_t \frac{1}{\sqrt{|s|}}\psi * \left(\frac{t-\tau}{s}\right)dt, \tag{5.2.3}$$

where * denotes the convolution and $W_t(s,\tau)$ constitutes the local phase. The $|W_t(s,\tau)|^2$, is conceived as the local wavelet power spectrum, that is the local variance. At this point of CWT discussion, it is proper to mention that Torrence and Compo (1998) introduced the Cone of Influence (COI) in order to tackle with the edge effects created by the transformation process of a wavelet to the time domain. They define COI to be the area where wavelet power occurred by a discontinuity at the edge, drops to the e^{-2} of the value at the edge. They also define a statistical significance test for the wavelet power spectrum that quantifies the robustness. For a more integrated presentation of CWT, we will only refer to the equation provided by Grinsted et al. (2004) and will not entail deeper on the validation path. Eq. (5.4) just denotes the

probability of the wavelet power of a prespecified process with a given Fourier power spectrum, P_k[29] being greater than p is:

$$D\left(\frac{|W_t(s,\tau)|^2}{s_X^2} < p\right) = \frac{1}{2} P_k X_v^2(p)\,, \tag{5.2.4}$$

where v=1 when we real wavelets and v=2 when we have complex wavelet. The approximation of above equation is achieved through Monte Carlo simulations.

5.2.2 Wavelet coherence: The cross wavelet transform

For the cases we want to examine in depth the linkage and the strength of the co-movement between two time series, and desire to investigate the lead-lag effect, we resort to wavelet coherence provided we follow the appropriate steps. The first step is the cross-wavelet transform (XWT). XWT spots the moments where two time series have strong common power based on their CWTs. Then, the lead-lag info, that is the phase relationship between the two variables. Hudgins *et al.* (1993), denoted XWT between two time series, X(t) and Y(t) as:

$$W_{XY} = W_X W_Y^*\,, \tag{5.2.5}$$

where, W_X and W_Y are the CWTs of time series X(t) and Y(t), respectively and * stands for the complex conjugate. Their cross-wavelet power spectrum is defined as $|W_{XY}|$ and stands for the local covariance at each time-frequency period. Thus, the cross-wavelet power spectrum highlights the territories that

[29] P_k according to Allen and Smith (1996) is thought to be $P_k = \frac{1-a^2}{|1-ae^{-2ipk}|^2}$, with k being the Fourier frequency index.

attract high common power (Grinsted *et al.*, 2004). Again, as in CWT there is a statistical significance test distribution of the cross-wavelet power of two time series provided in Eq. (5.2.6), that is again approximated via Monte Carlo methods:

$$D\left(\frac{|W_t^X(s)\, W_t^{Y*}(s)|^2}{\sigma_X \sigma_Y} < p\right) = \frac{Z_v(p)}{v}\sqrt{P_k^X P_k^Y}, \tag{5.2.6}$$

where $Z_v(p)$ is the confidence level for probability p for a probability density function characterized by the square root of the product of two chi-square distributions (Torrence and Combo, 1998).

Having estimated cross-power spectrum one can proceed and test how coherent the conducted transformation between the time series, is. To do that we follow the functions found in Aguiar-Conraria and Soares (2014), who distinguish between the complex wavelet coherency of Eq. (5.2.7) and wavelet coherence as the absolute value of the complex wavelet coherency that appears in Eq. (5.2.8):

$$\rho_{XY} = \frac{S(W_{XY})}{[\,S(|W_X|^2)S(|W_Y|^2)]^{1/2}}, \tag{5.2.7}$$

And

$$R_{XY} = \frac{|S(W_{XY})|}{[\,S(|W_X|^2)S(|W_Y|^2)]^{1/2}}, \tag{5.2.8}$$

where R_{XY} is between $0 \leq R_{XY} \leq 1$ and S is a smoothing operator along time and frequency. According to Torrence and Webster (1998) a suitable smoothing operator is expressed through:

$$S(W) = S_{scale}\left(S_{time}\left(W_n(s)\right)\right), \tag{5.2.9}$$

where S denotes the smoothing operator, S_{scale} denotes smoothing along the wavelet scale axis and S_{time} denotes smoothing along time. Now since we used the morlet wavelet, we should also provide the relative smoothing operator again directed by Torrence and Webster (1998) and is given in Eq. (5.2.10):

$$S(W)|_S = \left(W_t(s) * c_1^{\frac{-t^2}{2s^2}}\right)|_2 , S_{time}(W)|_S = \left(W_t(s) * c_2 \Pi(0.6s)\right)|_t, \tag{5.2.10}$$

where c1 and c2 are normalization constants. The factor 0.6 is the scale of decorrelation length for the morlet wavelet (Torrence and Compo, 1998). Now, in order to investigate the lead-lag effect of time series X(t) to Y(t), we need to estimate the angle φ_{XY} of the complex coherency that returns the phase-difference of the two time series ($\varphi_{XY} = \varphi_X - \varphi_Y$) and is given by:

$$\varphi_{XY} = Arctan\left(\frac{\Im\left((W_{XY})\right)}{\Re\left((W_{XY})\right)}\right), \tag{5.2.11}$$

From the above we deduce that when $\varphi_{XY} = 0$ then the two time series co-move at the specified time-frequency domain. When $\varphi_{XY} \in$ (0, π/2), then X(t) and Y(t) are in phase and X(t) leads Y(t). When $\varphi_{XY} \in$ (−π/2, 0), then Y(t) leads X(t). When $\varphi_{XY} =$ π (or −π) then there exists an antiphase relation

(negative correlation). When $\varphi_{XY} \in (\pi/2, \pi)$, then Y(t) leads X(t). Finally, when $\varphi_{XY} \in (-\pi, -\pi/2)$ then X(t) leads Y(t). And clarifying a little more these relations, the following equation specifies how we convert φ_{XY} into the instantaneous time-lag between X and Y based on scale, s, and angular frequency ω(s):

$$(\Delta T)_{XY} = \frac{\varphi_{XY}}{\omega(s)} \tag{5.2.12}$$

Perhaps the above specifications are a little bit obsolete. Well, in order to clarify that, let us illustrate what the phase-difference actually returns. It returns the "coordinates" of the arrows that appear in the wavelet coherence figures and inform of the exact direction of which information flows in the system. It is the direction of the arrow inside the imaginary cycle. Simply put is the circular mean of the phase angles used to quantify the phase relationship between the time series under investigation (Grinsted *et al.*, 2004).

5.3 Data description

We conduct the multivariate spectral analysis on two implied-realized volatility pairs, on two implied volatility-relative market pair and also on a pair of implied volatilities. The first pair of implied-realized volatility consists of the two variables, that we are already familiar with from the previous chapters, the VIX volatility index and the realized volatility constructed out of 5min returns of S&P500 futures index. The second pair consists of the OVX index, the implied volatility of the WTI crude oil price, and its respective realized volatility constructed out of 5 min returns of the West Texas Intermediate (WTI) futures index. The choice of OVX index, along the respective realized volatility, it is all

but random. Many are the studies dedicated at the forecasting of crude oil volatility, oil prices etc. and incorporate information from other assets classes, volatilities indices, as the financialization of the crude oil market has raised the cross-dynamics in new heights (Degiannakis and Filis, 2018; 2022; Le Pen and Sevi, 2017). The third pair is comprised of the VIX volatility and the S&P500 stock market index returns. Since VIX is the major volatility index of the U.S stock market directly linked with S&P500 index, it is inevitable to refrain from its inclusion. Let us not forget that volatility indices are employed in many studies in order to enhance forecast accuracy for stock markets indices returns. The fourth pair is the combination of OVX volatility index with the WTI crude oil price index, a choice again justified by their direct linkage. Finally, the last pair consists of VIX index and OVX.

The sample of the first pair spans from 5th of January of 2010 up to 31st of August 2020, a total of 2686 trading days. The sample of the second pair spans from 15th of August 2011 up to 31st of August 2020, a total 2278 trading days. The sample for VIX-S&P500 starts at 5th of January 2010 and extends up to the end of 30th of November 2022. The sample of OVX-WTI index spans from 15th of August of 2011 and extents to 31st of August 2020 and so does the sample for the VIX-OVX. Samples' length is dictated by the availability of the realized volatilities. The realized measures were retrieved from TickData, while the data for the implied volatility indices were retrieved from CBOE. The rest indices are freely available. At this point we should mention that we

conducted all appropriate transformations in the time series, in order to apply the proposed methodologies[30].

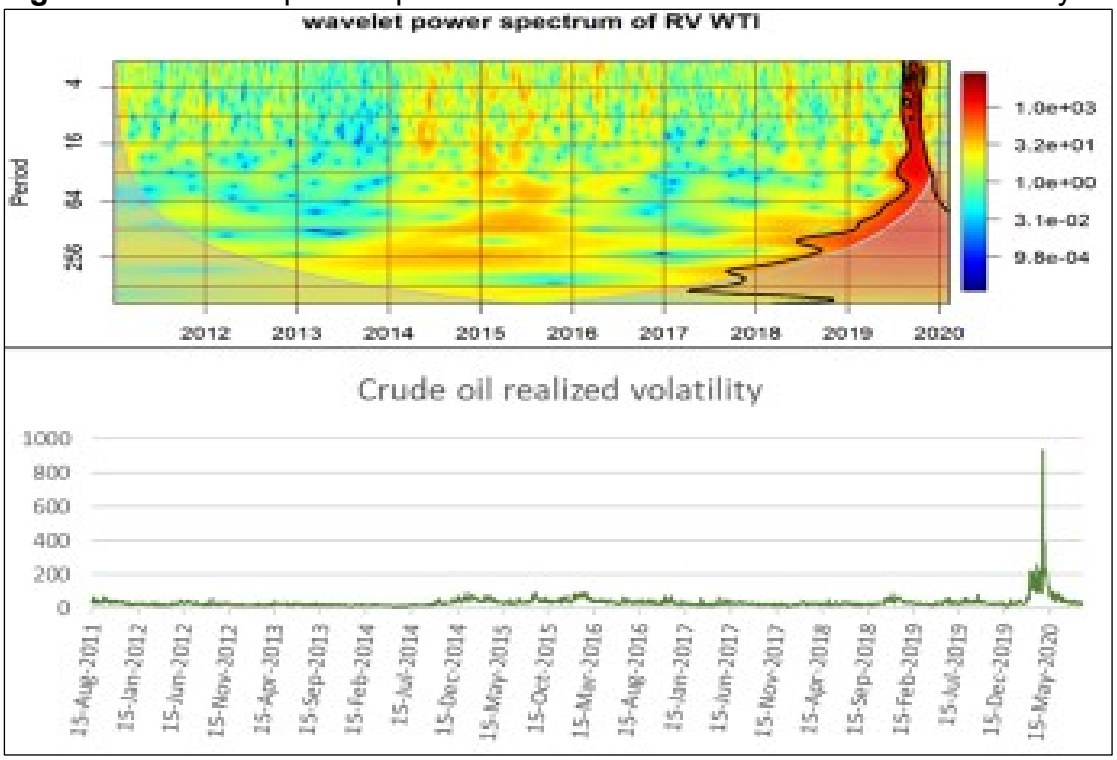
Since we conduct a wavelet analysis instead of presenting the descriptive statistics of the series, we thought it would have been better to include the wavelet power spectrums of each series. The wavelet power spectra estimates the strength of the cycles that exist in the series, at lower and higher frequencies (Crowley and Mayes, 2008). Fig. 5.1 to 5.4 present the series of implied and realized measures along with their wavelet power spectra, where we can see how proposed time series behaved in the time-frequency domain. The white line followed by the shaded area defines the COI, that is the regions where there are edge effects and not safe inference can possibly be drawn out of. Warmer colours indicate high power at specific range of frequencies and range of time.

By the examination of Fig. 5.1 we do observe statistically significant periods with high power. There are evident for only short periods of time, cycles of 10 and 22 days sporadically through entire sample, probably during some flash stock sales. What is more pronounced is the period of 2014 up to the end of 2016, the period of the oil price collapse, where there is a diffusion, there are small cycles with 5, 10-, 16-, 40- and 60-days periodicity, and there are two more extended with cycles of 160- and 252-days periodicity. Moreover, during 2018, the end of 2019 and the beginning of 2020 there is the highest power recorded with cycles extending to all frequency scales. It is the period that includes the cryptocurrency crash and the Covid-19 pandemic outburst that increased volatility at extreme heights and caused abrupt

[30] All estimations of proposed methodologies were conducted through the R programming language and the “biwavelet” package.

movements in many indices. Also, what is evident is the fact the cyclicalities recorded coincide with the way investors, risk management funds etc., reconstruct their positions, there are short time investors, investors that choose a longer exposition, and there are the institutional investors that may choose even longer periods.

Figure 5.1 Wavelet power spectrum of WTI futures returns realized volatility

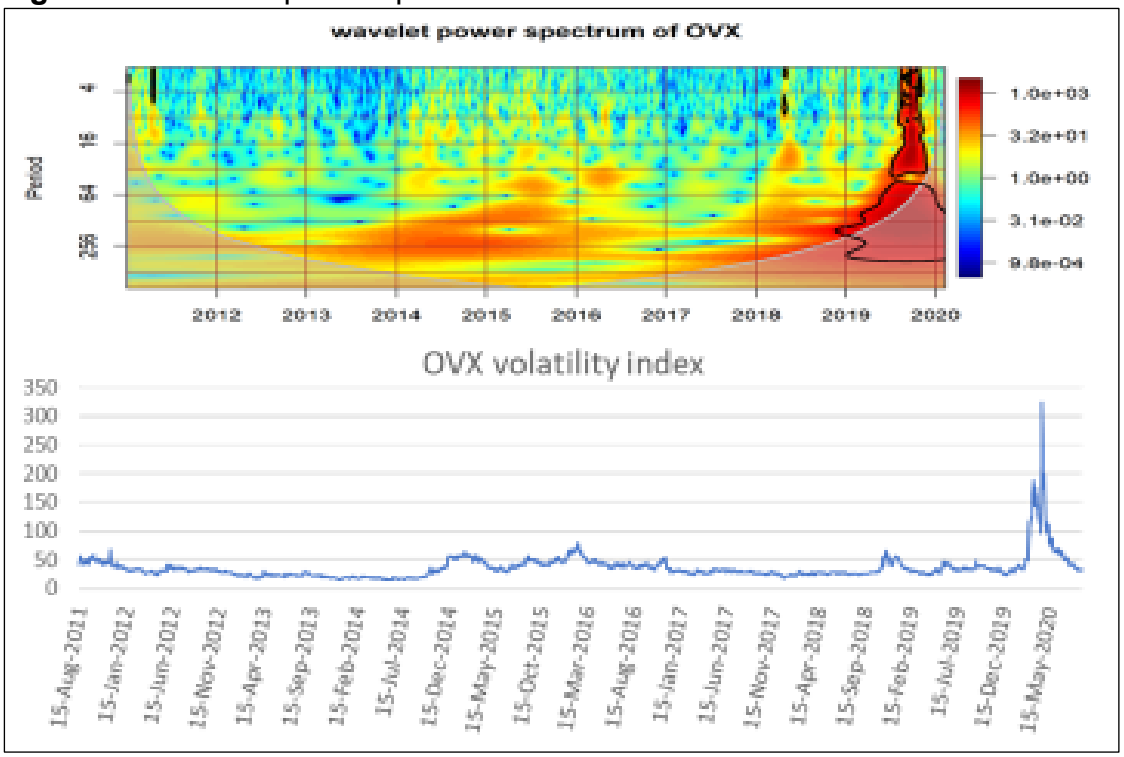


Note: Turbulent moments in the series coincide with density periods in the wavelet power spectra, where cycles of shorter and longer duration are spotted, especially during 2014-2015 period and 2018 to 2010.

Fig. 5.2 presents the power spectrum for the OVX index. Here the periods that attract the highest power are located in higher scales, although there are evident these short-term spots of higher power for cycles of 10- to 66-days for the same flash turmoil periods, we mentioned earlier. For the period starting from 2014, there is evident a cycle of 160- and 252 days up to the end of 2016 with significant high power, obviously due to the oil crisis. The power again starts to escalate at the mid of 2018 and peaks at the COVID-19 outbreak were there are small extent but of strong power cyclicalities covering

the entire frequency spectrum again compatible with the global crisis and the disruptions this condition entailed globally.

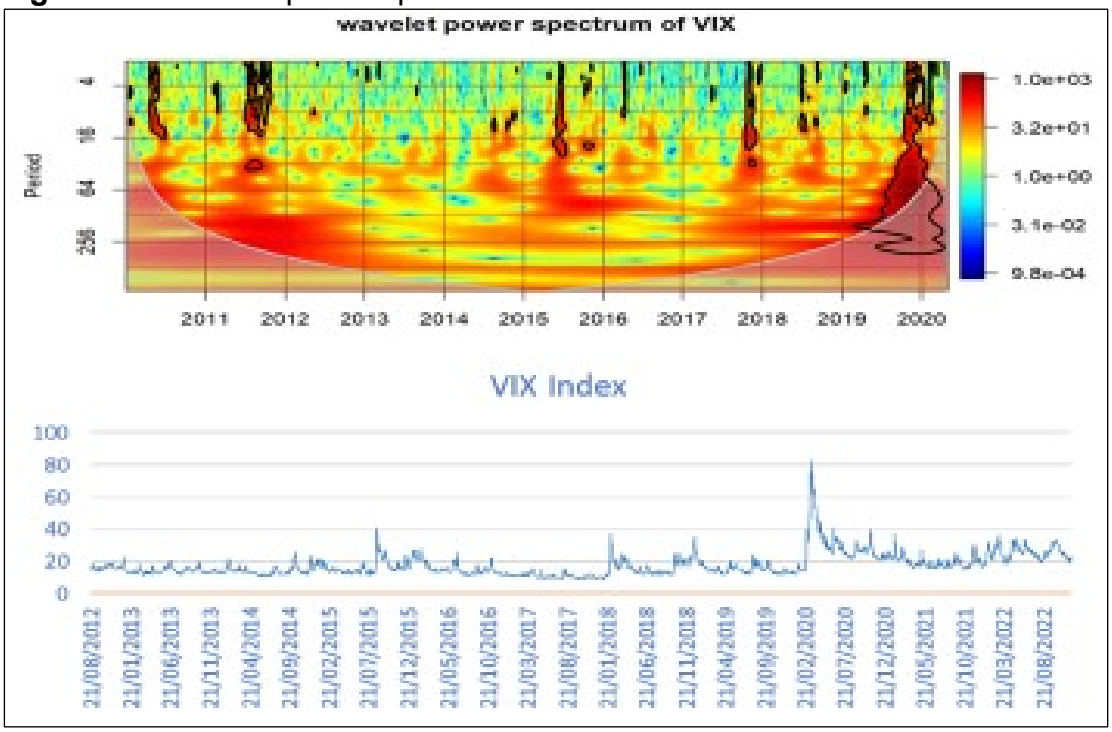
Figure 5.2 Wavelet power spectrum and series of OVX index



Note: Wavelet power spectra shows the strong annual and semi-annually cycles of the periods of the oil crise and the global pandemic outburst. For the period of the COVID-19 crisis there were recorded cyclicalities covering entire spectrum.

The power spectrum of VIX index is much more heated compared to previous power spectrums as Fig. 5.3 betrays. There are sporadically some small extent periods with daily, weekly, and biweekly frequencies with extreme high power, but there are periods of longer cycles of 44 days, 66 days, semi-annual and annual duration from 2011 to 2013, 2015 to 2016 and 2018 and on to the COVID-19 pandemic, where dynamics escalate at the extreme. VIX obviously captures all those shocks that cause markets' volatility to raise, and there quite a few during the last 10-12 years. Also, the longer cycles periods that extent probably coincide with the actual way markets operate and the movements conducted by markets participants in regular intervals through the trading year.

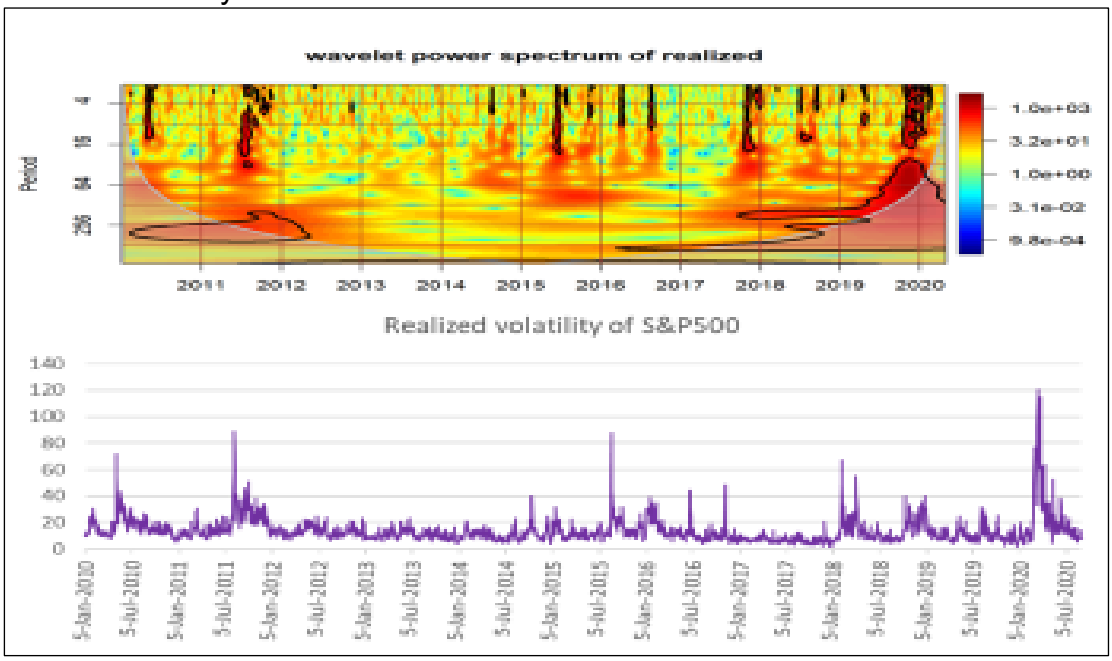
Figure 5.3 Wavelet power spectrum and series of VIX index



Note: the power spectrum of VIX is the most intense of all power spectra investigated up to this moment in our study. There are evident some strong daily, weekly, and so on, cycles at short periods but definitely there exist strong 44 days and on, cycles for most of the investigated period.

Finally, Fig. 5.4 depicts the power spectrum of S&P500 futures index realized volatility. The figure displays few similarities to VIX index, but definitely there are divergencies. The density is not so strong as is for the VIX and does not come with the continuation that VIX presents, although realized is much more persistent especially during the recorded crises, whether it was the Europe dept crisis, the oil crisis and of course the pandemic of COVID-19 outbreak. The black lines, define specific areas where semi-annual and annual cyclicalities are more profound. The power spectrum of realized volatility also reveals the periods when volatility of the markets started to build up due to troublesome conditions and when started to decline, that is the period between mid-2010 to 2013, 2014-2016 and 2018-2020 and on.

Figure 5.4 Wavelet power spectrum and series of S&P500 futures index realized volatility



Note: The power spectra of stock market realized volatility illustrates the density and the cycles recorded through the period under investigation where three crises were present.

5.4 Empirical findings

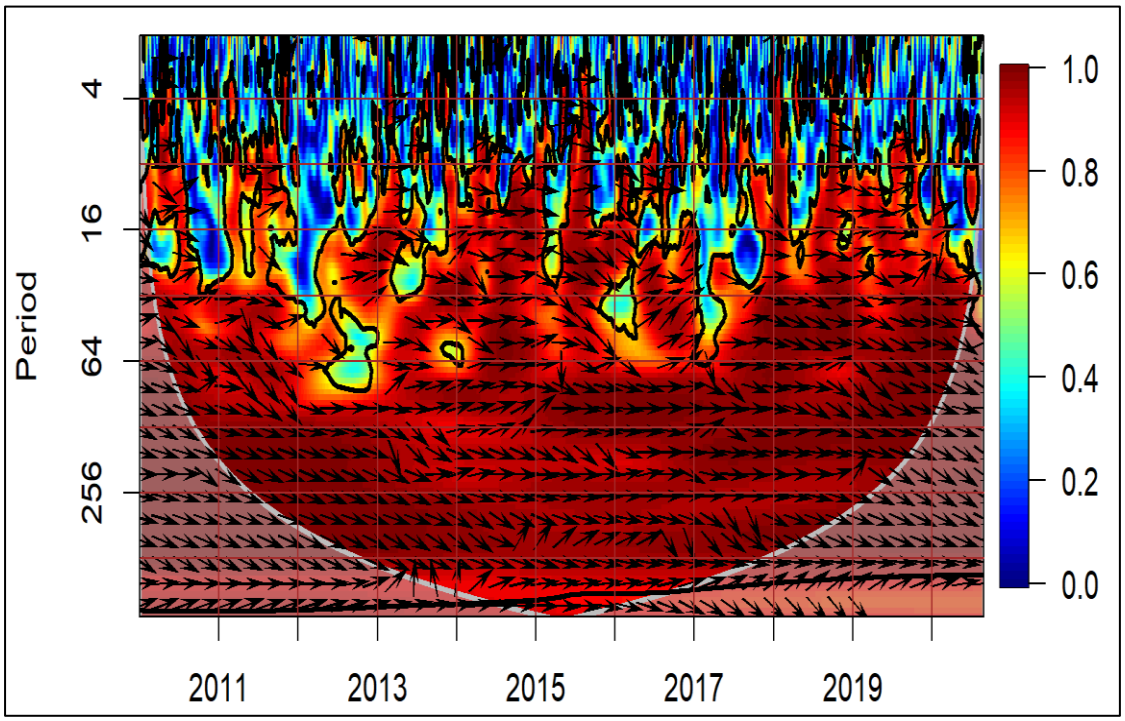
The examination of the wavelet power spectrum at least for the implied and realized volatilities of the stock and the energy market, revealed all those short and of longer duration periods where volatility raised significantly. This was a prime step for inspecting the cyclicalities of our time series prior to conducting the cross wavelet transform on the proposed pairs. In this section the analysis will be guided by the inspection of the wavelet coherence plots, that are also famed for their elaborated form as relation of the two series is diffused in the time frequency domain. Fig. 5.5 to 5.9 present the coherence of the five pairs. In the horizontal axis the time units are displayed, and along vertical axis appears the frequency in which co-movement was spotted. Since we have daily data, frequency scales are in days and record the periodicities observed in the data. The white line, as for the power spectra, delimits the COI, where information included in the shaded area cannot be used to draw

safe inference. Warm colors indicate strong co-movement, while the colder ones indicate periods of no significant coherence between the two time series. The arrows indicate the phase difference and inform of the lead-lag effect.

Our analysis starts from Fig. 5.5, that presents the coherence recorder for the first pair that is the VIX index along the realized volatility of S&P500 futures index. What is evident is their strong co-movement. For cycles of 10 days duration, co-movement is only spotted sporadically under the entire sample probably indicating correction movements of the markets due to flash crashes, stock markets fall, sale-offs etc. For the 44 days and 66 days cycles, up to the cycle of 252 days (trading year) co-movement is recorded in an almost continuous manner. Also, we observe, and this is the most critical part, that realized volatility leads VIX index for almost entire period under investigation. And being more precise from 2010 to 2012 realized volatility leads VIX for the 22, 44 and 66 days cycles and from 2011 to the end of the period under investigation for the 132 days cycle. Now, there are some sparingly moments, when implied volatility seamed to lead and that were recorded at the end of the years 2012 and 2013 for the 22 days cycle, at the end of year 2014 and the very start of 2015 for the 132 days cycle and finally, from mid-2016 until the end of the year for 22, 44 and 66 days cycles. The cycles of days 10, 22, 44, 66, and semi-annually and annually, coincide with the way investors, hedge funds, risk management funds etc., reconstruct their portfolios or their positions on short or long exposure, something noticeable from the previously investigated power spectra. Also, the short lived cycles coincide with market correction moves. Now, the fact that realized leads implied perhaps is not that surprising, since there exist studies that have

enhanced forecast accuracy of implied indices by the inclusion of realized measures, but still raises a point on how analysis and modelling should be conducted, when the two measures are present.

Figure 5.5 Wavelet Coherence: VIX vs. Realized volatility of S&P500 futures index

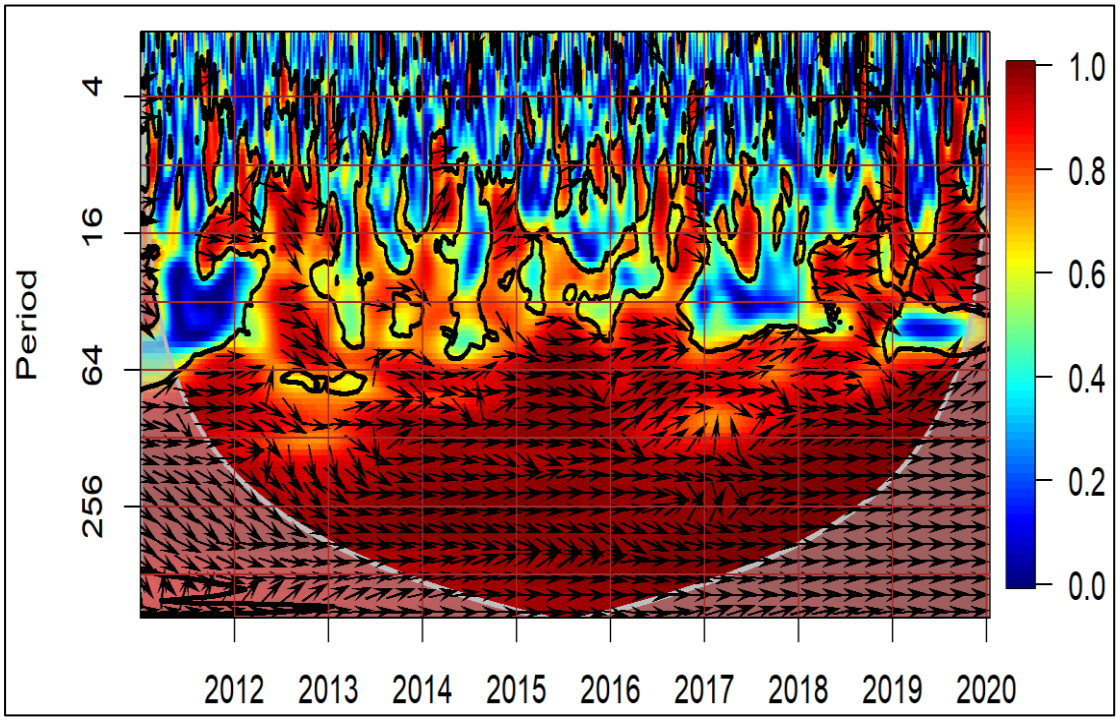


Note: The implied-realized volatility combination for the stock market reveals a strong positive co-movement with significant periodicities mainly becoming regular from the 22 days cycles and on. Definitely there are some evident strong bi-weekly relations altered with moments of no significant relation at all. Realized is leading implied in most cases.

The second pair of implied and realized volatilities, is the OVX index and the realized volatility of the WTI futures index. Here again we spot a very strong positive co-movement of the two variables from the 44 days cycle and up to the 252 days cycle for the entire period. There is also a periodicity of 22 days, throughout entire period but with discontinuities, moments where no significant relation is recorded. We also observe, the realized measure to lead the OVX index, starting from the 10 days cycle and on from the mid 2012 till the end of 2013, continuing to the 22, 44 days cycle and 66 from the mid 2018 up to 2020. There is definitely a 66-days cycle at years 2014-2015 that again

realized was leading implied. For the case of OVX, OVX seems to have led the 10 and 22 days cycle towards the end of 2014 and dominated at the 66 and 132 days cycle from the begging of 2016 until the end. This finding is illuminating of the driving powers between the two volatility measures.

Figure 5.6 Wavelet coherence: OVX vs. realized volatility of WTI futures index

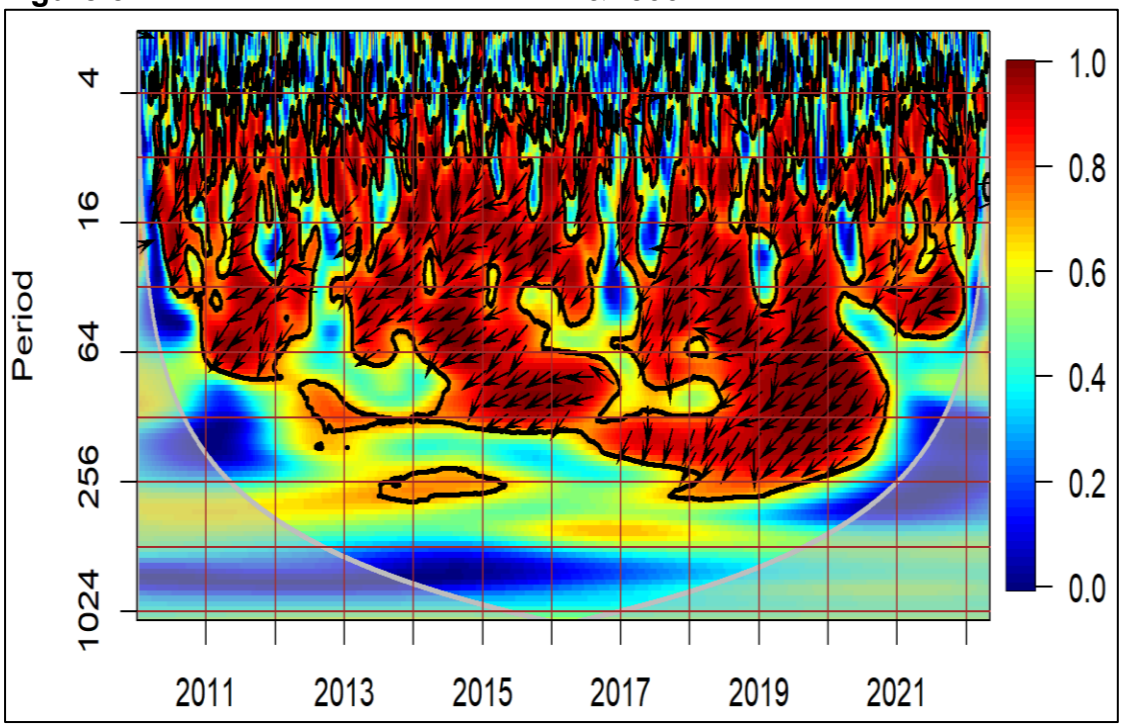


Note: The pair of implied-realized volatility for the energy market illustrates a leading realized volatility and extremely strong 66-days cycles and on. There are evident strong 22 days cycles and on for the period 2018-2020 where again realized leads implied volatility.

The next pair is the VIX index combined with the S&P500 stock index returns. Here we detect a strong co-movement, but in an anti-phase condition, that is an increase in VIX anticipates a decrease in stock index returns. VIX was leading in all phases. There were strong negative interrelations of 10, 22, 44, 66 days cycles, where VIX was leading S&P500 from 2011 up to mid-2012, from 2013 and up to almost the end of 2016. From 2017 up to almost the end of 2020 there is a cycle of 66 days to 252 days, where VIX once more leads. Findings somehow were expected as volatility drives most of market

movements, but what is really profound is the absence of any significant relation in a daily and weekly frequency something valid for all previous investigated relations.

Figure 5.7 Wavelet coherence: VIX vs. S&P500 returns

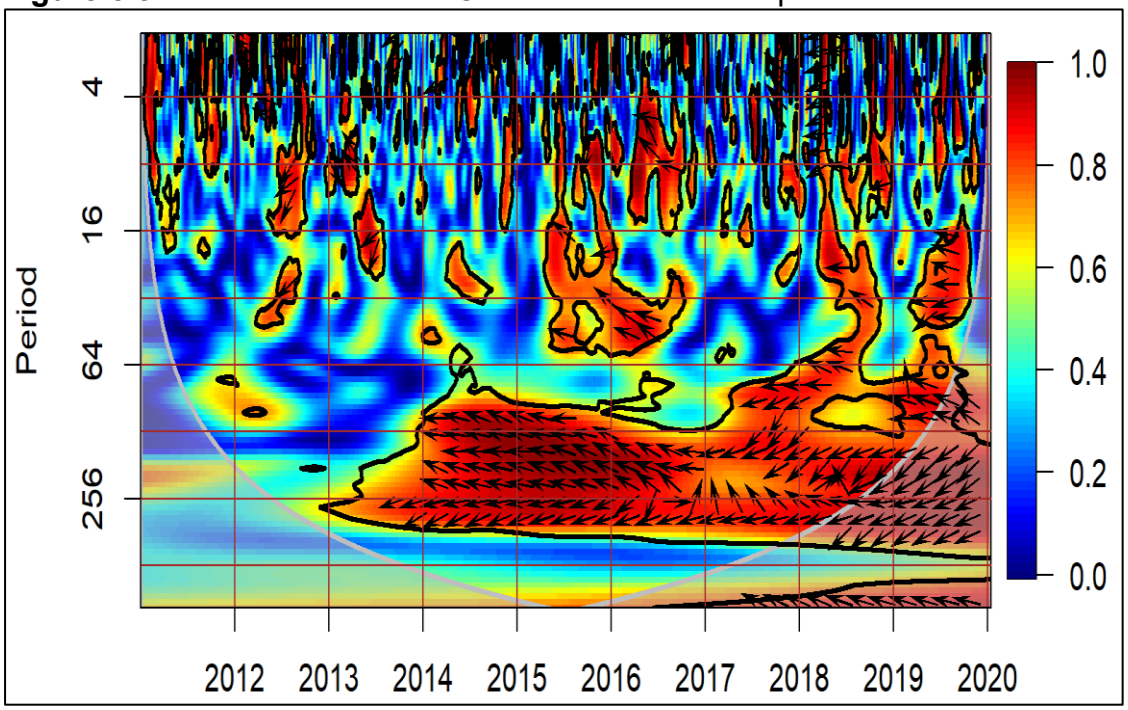


Note: The interrelation of VIX and S&P500 although strong is negative. VIX leads for all those moments and there is a periodicity spanning from 10 days cycles to 66 day cycles for the period 2011-2014 and from 2015-2020 relation is expanded to enclose the semi-annual and annual cycles.

The other pair is the OVX and the WTI crude oil price. Again, for this market we have a strong negative coherence but not in a continuous manner, as there were moments where no significant co-movement is recorded. From year 2014 up to the beginning of 2017 there exists a 132 to 252 days cycle where crude oil price guided the movements of OVX index, then there was an alteration in the leadership part and OVX led the anti-phase movements from 2018 to the rest of the remaining sample and for periodicities of 66 to 252 days. Also, oil price seems to have led the 10, 22 and 44 days cycle for a small fragment of time between 2015-2016 and the period in the beginning of COVID-19 pandemic. These findings clarify who relation alters between crises

of different origin, when there is an endogeneity and when there is global turbulence.

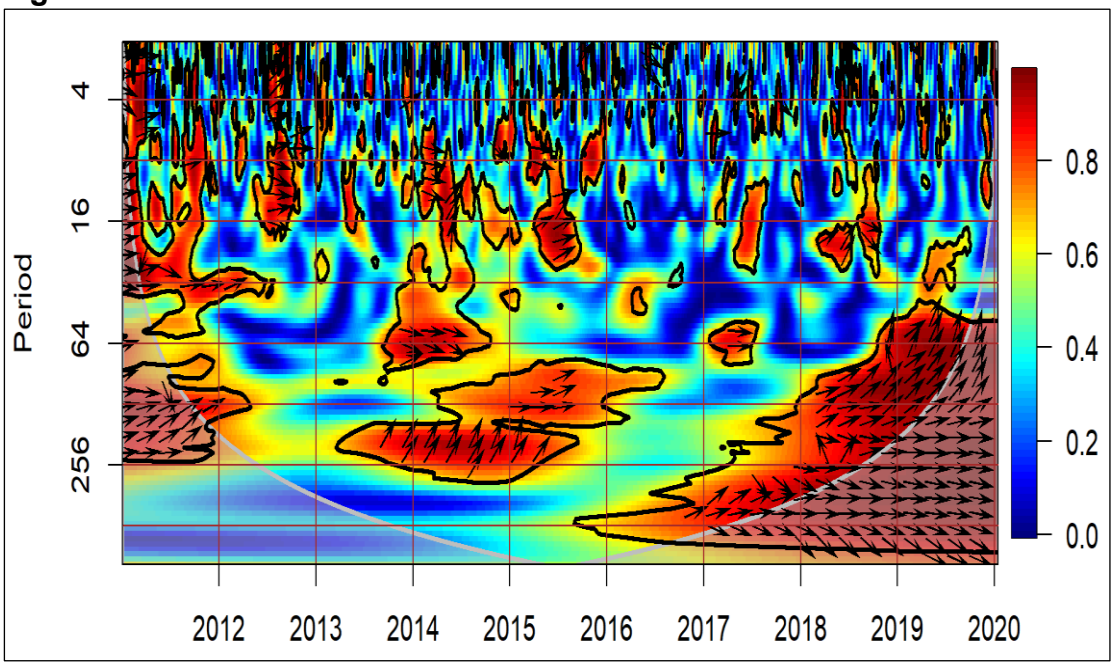
Figure 5.8 Wavelet coherence: OVX vs. WTI crude oil price



Note: The strong negative co-movement is being spotted for OVX and WTI crude oil price especially for the period of 2014-2020 for the semi-annual and annual cycles with altering lead-lag dependencies.

Finally, the last combination is VIX index paired with OVX index, a combination of two implied volatility indices that we frequently come across. There exists a strong positive co-movement especially during the periods where crises have burst out, especially during the oil crisis the years 2014-2015 and the COVID-19 pandemic. The relation is spotted in the frequencies of 60 to 252 days but with a difference. VIX index leads the OVX, for the semi-annual and annual cycles for the 2014-2015 period for the 60 days to 252 days for the 2019-2020 period, while OVX leads only for a small fragment of time during 2014 and for the 60 days cycle.

Figure 5.9 Wavelet coherence: VIX vs. OVX



Note: figure clearly depicts the periods where there is a strong and significant relation between the VIX and OVX, with VIX leading the annual and semi-annual cycle for the period 2014-2015 and 2018-2020.

The wavelet coherence analysis of all five pairs intended to uncover some hidden aspects concerning the interrelation of volatilities indices and of the adequate markets. One of these aspects was how strong these variables co-moved, especially under critical turmoil periods of the financial sector, the energy sector, or the global economy. There were such moments, where volatility raised to the extreme and their interrelation was strong enough. The other aspect was the direction of that co-movement, the lead-lag effect and that was quite illuminating. There was a common behavior in the realized volatility measures of the stock and the energy market, both measures led, while the implied volatilities indices followed. The third aspect was the frequencies, the cyclicalities in these relations, that was discovered to be in fully alignment with the regular operations that take place in the markets, it is the periodicities with which investors reconstruct their positions and markets go through correction moves in order to adjust imbalances during periods of

selloffs or flash jumps. Findings are illuminating because they provide a direct frame of volatility's dynamics direction, especially for the final pair, that comes in line with studies like the one from Delis *et al.* (2023), who find evidence on the predictive ability of other implied volatilities on the OVX index. Hence, wavelet analysis can contribute to highlighting the direction and bi-directional movements, the strength of the relation between two variables and the duration of this relationship. The direction could assist in choosing the best predictors, when engaging in forecasting but the duration could assist in both the forecasting practices, for choosing the best forecast horizon but also in the trading practices, depending on the relation, short-lived, long-lived, or persistent.

5.5 Conclusion

In this final chapter, our intention was to explore the ability, techniques adequate to the ones we incorporated in the main body of our study, have in extracting hidden features and patterns from time series. Afterall, all decomposition techniques presented in this thesis, were initially developed for feature extraction, able to assist in performing a more detailed analysis of complex, non-linear and non-stationary systems, that would be impossible to be performed by parametric frameworks. Thus, we took advantage of the pronounced ability of CWT and WC to decompose and analyze a time series, or two time series, in a continuous manner in the time and frequency domain, concurrently.

Throughout this thesis, we concentrated on two types of volatilities, the implied and the realized volatility and two variables as proxies, the VIX index, the major volatility index for the U.S stock market and the adequate realized

volatility of the S&P500 futures index. Here, we had to enhance our study by including another implied-realized volatility pair and the indices of the relative markets, in order to have a more integrated view of how their interconnections evolved under the chosen time fragment. The choice was not random. All these variables are key elements, as they sustain a tremendous network of assets, derivatives etc., directly developed on their account. Hence, we thought it would be interesting to develop an illustrative along a statistical point of view of how information flows through these channels. Findings justify this choice.

The first finding is that realized volatilities lead, and implied volatilities follow, with a periodicity higher than 44 days and on. Another fact is that there is, no significant relation, present, on a daily basis or even weekly, something valid for all pairs, with a few exclusions, when instantaneous short-lived movements were recorded probably due to flash troublesome market moments. Furthermore, there was the period 2014-2016, where crude oil price was leading OVX, while there were the years 2018-2020, where OVX led crude oil. That defines how important is the original source that results in a volatile environment creating pipeline reactions. In the first case was an oil crisis, an endogenous crisis, while in the second it was a crisis of a global flavor. Also, the relation between implied volatilities and their relative markets was negative for both pairs. VIX index led S&P500, but strong interrelation was spotted in a frequency of 10 days and on, while there were moments of no significant interrelation at all. Finally, we found VIX to lead OVX.

Thinking whether these findings are useful and why, we would claim that important implications were driven out of this study and are indeed useful

especially for those actively engaging in the markets for investing, hedging or diversifying purposes, for stakeholders, researchers, policy makers and forecasters. As for the later, findings inform for periodic movements, bidirectional positive or negative relations or the absence of any relation, a condition extremely important when developing multivariate modelling frameworks and when setting the forecast horizons.

Concluding, we should state that wavelet analysis offers advanced tools that allow for a careful and analytical inspection of all hidden features and peculiar behavior the financial instruments have, under specific time frames, when events of different origin occurred. That in turn adds in better understanding the invisible threads that bounds systems together and their transmission mechanisms. Here we covered only a few aspects of the underlying markets relations, motivated by the need to explore the tools other scientific fields have to offer. There are endless combinations and analysis knows no limits.

Chapter 6

Thesis conclusion

The present thesis highlighted the inclusion of decomposition techniques in the modelling and forecasting practice of financial time series and the gains they can offer. Here we concentrated on two substantial measures of risk, that dominate financial markets, the implied and realized volatility. Both measures are famed for their “difficult” behavioral characteristics, and the modelling rash that both have erupted. Hence, we thought that by incorporating disaggregation methods, appropriate for analyzing non-linear and non-stationary systems, in our study and decomposing these persistent measures into their intrinsic modes, that is into components of different frequencies where some discontinuities, periodicities, noise, etc., occurred we could possibly result in a more efficient modelling and forecasting procedure. Let us not forget that forecasts are the most critical elements for all those directly or indirectly engaging in stock or derivatives markets, whether they are investors, policy makers, risk managers, or stakeholders, in order to have an approximation of the imminent risk or the possible opportunities. Afterall, predicting the correct direction of an asset/market can be multi-beneficial. But a forecast can be properly quantified only if it is tested under the exact conditions markets operate and decisions are designed. Thus, evaluating forecasts following an actual trading strategy on futures contracts was the most appropriate way to reach correct conclusions.

For our study, after experimentation, we included the EMD, the EEMD, the SSA, the HVD, the EWT and the VMD techniques and the AR, the HAR, the HW and the LSTM frameworks. The set of the variables we peaked, were the VIX index and the high frequency returns of the S&P500 futures index, as proxies of implied and realized volatility for the studies of chapters 3 and 4, respectively. We peaked the VIX futures as proxy for the trading practice, where depending on the generated forecasts of implied or realized volatility received each day of the out-of-sample period, we would go long or short on the roll adjusted futures contracts. All indices were carefully chosen for their close proximity. The outcome was rewarding. For at least two out of the six decomposition-based models, we had astonishing returns on invested capital for both cases. For the case of implied volatility two of the best performing models were the EEMD-based and the EMD-based for all forecast horizons, with the best outcome returning 44 times the invested capital. For the case of realized volatility the best performing models for all forecast horizons were the EEMD-based and the HVD-based, with the best outcome returning 18 times the invested capital. Perhaps outcomes were not equally high for both measures but at least there were two models of the EEMD-based family, that worked for both and that were the EEMD-HW and the EEMD-HW-AR. Of course, we should state that these findings were only valid according to the cumulative returns earned out of the trading practice we incorporated. Under the classic statistical loss function no forecast gain could be quantified. Hence, in this thesis we also highlighted the contradictive outcome between different evaluation loss functions and how this should be carefully designed to align with the environment and the scope for which a forecast is being

generated. But there were another two findings that we almost neglected them. The one deals with the fact that the way individual components of each technique would be modelled and forecasted, matters for the final outcome. So, the combination plays a significant role in the level of quantified returns. The other deals with the roll of the futures contracts. We noticed that cumulative returns for all models, even for those that ill performed, were substantially higher when roll considered. Overall, chapters 2, 3 and 4 constituted an integrated study on the benefits of ensemble modelling, roll of futures, and model combinations, when forecasting financial time series. Thus, it worths becoming part of the trading practice provided objective-based evaluation criteria are considered.

The fifth chapter appeared to be isolated compared to the previous integrated analysis, but it is not so. On the contrary, the fifth chapter allowed to study other aspects of the decomposition methods and discover some hidden behaviors of implied and realized measures in the time-frequency domain. It allowed to study their interconnection and their cyclicalities. Thus, we included another one pair of implied-realized volatility for the crude oil price and two indices representative of the respective markets. Findings revealed that there were, strong 44, 66, 130 and 252 days cycles between implied and realized pairs. Realized was found to lead implied. Both implied volatilities and the relative markets had negative co-movements, while implied indices seemed to lead. Also, VIX was found to lead OVX at higher frequencies coinciding with the semi-annual and the annual cycle, at specific periods, such was the 2014-2015 period and the 2019-2020. There were and some sparingly moments that there seemed to be a directional relation in

lower frequencies but only for an infinitesimally lasting period, thus we cannot support that there was a specific periodic pattern. That said, indices found to not undergo any daily or weekly significant co-movement when tested in pairs for the period under investigation. Findings are important especially for forecasters who engage in multivariate modelling settings. The knowledge of when, how and in what frequency two variables co-moved, answers why there are significant forecast gains by including a specific exogenous variable that we thought to be strongly related dependent variable and why there is a void. Furthermore, it allows for a deeper understanding of how variables interacted during turbulent for the global economy moments, and bear in mind of how they may interact in a similar future moment.

Concluding, the present thesis raised some important implications on how to best incorporate methods, coming from other scientific fields, especially signal processing, in our advantage in order to optimally analyze and model financial time series and accomplish profitable forecasts. There are countless combinations in the vast and complex financial system and its underlying markets to experiment, uncover hidden relations, and set new trading rules.

References

Adrian, T., & J. Rosenberg. **(2008)** Stock Returns and Volatility: Pricing the Short-run and Long-run Components of Market Risk. *Journal of Finance, 63*, 2997–3030.

Aguiar-Conraria, L., & Soares, M.J. (2011) Business Cycle Synchronization and the Euro: a Wavelet Analysis. *Journal of Macroeconomics 33*, 477-489.

Aguiar-Conraria, L., & Soares, M.J. (2014) The Continuous Wavelet Transform: Moving Beyond Uni- and Bivariate Analysis. *ERN: Other Econometrics: Econometric Model Construction*.

Aït-Sahalia, Y., & Mancini, L. **(2008)** Out of sample forecasts of quadratic variation. *Journal of Econometrics, Volume 147(1),* 17-33.

Allen, R.M., & Smith, A.L. (1996) Monte Carlo detecting irregular oscillations in the presence of colored noise. *Journal of Climate 9(12)*.

Alexander, C., Korovilas, D. & Kapraun, **J. (2016)** Diversification with volatility products, *Journal of International Money, and Finance*, 65, 13-235.

Ali, M., Dost M. K., Alshanbari, H. M. & AAAH El-Bagoury. (2023) Prediction of Complex Stock Market Data Using an Improved Hybrid EMD-LSTM Model. *Applied Sciences, 13(3)*.

Alizadeh, S., Brandt, W.M., W. & Diebold, X. F. **(2001)** Range-Based Estimation of Stochastic Volatility Models. http://dx.doi.org/10.2139/ssrn.267788

Aloui, C., & Hkiri, B., (2014) Co-movements of GCC emerging stock markets: new evidence from wavelet coherence analysis. *Journal of Econometric Modelling 36*, 421–431.

Alqaralleh, H., & Canepa, A. (2021) Evidence of Stock Market Contagion during the COVID-19 Pandemic: A Wavelet-Copula-GARCH Approach. *Journal of Risk and Financial Management*, *14*(7), 329.

Andersen, G. T., & Benzoni, L. (2008) Realized Volatility. FRB of Chicago Working Paper No. 2008-14.

Andersen, T. G., & Bollerslev, T. (1998) Answering the skeptics: yes, standard volatility models do provide accurate forecasts, *International Economic Review*, 39, 885–905.

Andersen, T.G., T. Bollerslev, P.F. Christoffersen & Diebold F.X. (2005) Practical Volatility-82- and Correlation Modeling for Financial Market Risk Management, in M. Carey and R. Stulz (eds.), Risks of Financial Institutions. *Chicago: University of Chicago Press for National Bureau of Economic Research*.

Andersen, T., Bollerslev, T., Christoffersen, P., & Diebold, F. X. (2006) Volatility and correlation forecasting. In G. Elliott, C. W. J. Granger, & A. Timmermann (Eds.). *Handbook of economic forecasting*. Amsterdam: North Holland Press.

Andersen, T., Bollerslev, T., & Diebold, F. X. (2007) Roughing it up: Including jump components in the measurement, modelling and forecasting of return volatility. *The Review of Economics and Statistics*, 89(4), 701–720.

Andersen, T.G., Bollerslev, T., Diebold, F.X. & Labys, P. (2001) The distribution of realized exchange rate volatility. *Journal of the American Statistical Asso*ciation *96* (453), 42–55.

Andersen, T.G., T. Bollerslev, F.X. Diebold & Ebens, H. (2001) The Distribution of Stock Return Volatility, *Journal of Financial Economics,* 61, 43-76.

Angelidis, T., and & Degiannakis S. (2008) Volatility forecasting: Intra-day versus inter-day models. *Journal of International Financial Markets, Institutions & Money 18(5)*: 449–465. https://doi.org/10.1016/j.intfin.2007.07.001

Barndorff-Nielsen, O.E. & Shephard, N. (2002) Estimating Quadratic Variation Using Realized Variance, *Journal of Applied Econometrics*, 17, 457-477.

Barndorff-Nielsen, O. E., Hansen, P.R., Lunde, A. & Shephard, N. (2008) Designing realized kernels to measure the ex-post variation of equity prices in the presence of noise, *Econometrica,* 76 (6), 1481–1536.

Bekaert, G., & Hoerova, M. (2014) The VIX, the variance premium and stock market volatility. *Journal of Econometrics 183(2)*, 181–192.

Bengio, Y., Simard, P. & Frasconi, P. (1994) Learning long-term dependencies with gradient descent is difficult, *IEEE Transactions on Neural Networks*, 5(2), 157-166.

Black, F. & Scholes, M. (1973) The Pricing of Options and Corporate Liabilities, *Journal of Political Economy*, 81, 637-654

Bollerslev, T., Tauchen, G., & Zhou, H. (2009) Expected stock returns and variance risk premia, *Review of Financial Studies,* 22(11*),* 4463–4492.

Bollerslev T., Gibson M., & Zhou H. (2011) Dynamic Estimation of Volatility Risk Premia and Investor Risk Aversion from Option-Implied and Realized Volatilities, *CREATES Research Paper.*

Busch, T., Christensen, B.J., & Nielsen, M.Ø. (2011) The role of implied volatility in forecasting future realized volatility and jumps in foreign exchange, stock, and bond markets. *Journal of Economics 160 (1),* 48–57.

Canina, L., & Figlewski, S. (1993) The informational content of implied volatility. *Review of Financial Studies 6,* 659–681.

Chatfield, C. (1978) The Holt-Winters Forecasting Procedure. *Journal of the Royal Statistical Society, Series C (Applied Statistics)*, 27(3), 264–279.

Chatziantoniou, I., Degiannakis, S., Delis P., & Filis G. (2021) Forecasting oil price volatility using spillover effects from uncertainty indices. *Finance Research Letters 42*:101885. https://doi.org/10.1016/j.frl.2020.101885.

Chevalier, L., Laignel, B., Massei, N., Munier, S., Becker, M., Turki, I., ... & Cazenave, A. (2014) Hydrological variability of major French rivers over recent decades, assessed from gauging station and GRACE observations. *Hydrological Sciences Journal*, *59*(10), 1844-1855.

Christensen, B. J., & Prabhala, N. R. (1998) The relation between implied and realized volatility. *Journal of Financial Economics*, *50*, 125–150.

Civera, M. & Surace, C. (2021) A Comparative Analysis of Signal Decomposition Techniques for Structural Health Monitoring on an Experimental Benchmark. *Sensors*, 21, 1825.

Corsi, F. (2009) A simple approximate long-memory model of realized volatility. *Journal of. Financial Econometrics,* 7(2)*,* 174–196.

Corsi, F. & Reno, R. (2012) Discrete-time volatility forecasting with persistent leverage effect and the link with continuous-time volatility modeling, *Journal of Business and Economic Statistics,* 30(3), 368–380.

Daigler, R.T., & Rossi, L. (2006) A Portfolio of Stocks and Volatility. *Journal of Investing*, 99–106.

Daubechies, I. (1992) Ten Lectures on Wavelets, *Society for Industrial and Applied Mathematics*, Regional Conference Series in Applied Mathematics.

Degiannakis, S. (2008) Multiple days ahead realized volatility forecasting: Single, combined, and average forecasts. *Global Finance Journal 36,* 41-61.

Degiannakis, S. & Filis, G. (2017) Forecasting oil price realized volatility using information channels from other asset classes. *Journal of International Money and Finance,* 76(C), 28–49.

Degiannakis, S. & Filis, G. (2020) Oil price assumptions for macroeconomic policy, *MPRA Paper 100705, University Library of Munich*, Germany.

Degiannakis, S. & Filis, G. (2022) Oil price volatility forecasts: What do investors need to know? *Journal of International Money and Finance,* 18(5), 449–465.

Degiannakis, S., & Filis, G. (2018) Forecasting oil prices: High-frequency financial data are indeed useful. Energy Economics 76, 388–402.

Degiannakis, S., Filis, G. & Hassani, H. (2018) Forecasting global stock market implied volatility indices. *Journal of Empirical Finance* 46, 111–129.

Delis, P., Degiannakis, S. & Giannopoulos, C. (2023) What should be taken into consideration when Forecasting Oil Implied Volatility Index? *Energy Journal*, 44(5).

Dong, J, Dai, W., Tang, L., & Yu, L. (2019) Why do EMD-based methods improve prediction? A multiscale complexity perspective. *Journal of Forecasting*, 38, 714– 731.

Dragomiretskiy, K., & Zosso, D. (2014) Variational Mode Decomposition. *IEEE Transitions on Signal Processing*, 62, 531–544.

Eckstein, J. & Bertsekas. P.D. (1992) On the Douglas Rachford splitting method and the proximal point algorithm for maximal monotone operators. *Mathematical Programming*, 55(1-3), 293–318.

Engle, R.F., Hong, C.H., Kane, A., & Noh, J., (1993) Arbitrage valuation of variance forecasts with simulated options. *Advances in Futures and Options Research 6*, 393–415.

Engle, R.F., Kane, A., & Noh, J., (1996) Index-option pricing with stochastic volatility and the value of accurate variance forecasts. *Review of Derivatives Research 1 (2),* 139–157.

Feldman, M. (2006) Time-varying vibration decomposition and analysis based on the Hilbert transform. *Journal of Sound Vibrations*, 295, 518–530.

Feldman, M. (2011) Hilbert Transform Applications in Mechanical Vibration; John Wiley & Sons: Toronto, Canada.

Feldman, M. & Braun, S. (2017) Nonlinear vibrating system identification via Hilbert decomposition. *Mechanical Systems and Signal Processing*, 84(B), 65-96.

Fernandes, M., Medeiros, M.C. & Scharth, M. (2014) Modeling and predicting the CBOE market volatility index. *Journal of Banking Finance* 40, 1–10.

Flandrin, P., Rilling, G., and Goncalves, P. (2004) Empirical Mode Decomposition as a Filter Bank. *IEEE Signal Processing Letters*, 11, 112–114.

Forsberg, L. & Ghysels, E. (2006) Why Do Absolute Returns Predict Volatility so Well? http://dx.doi.org/10.2139/ssrn.929494

Gabay D. & Mercier, **B. (1976)** A dual algorithm for the solution of nonlinear variational problems via finite element approximation. *Computers & Mathematics with Applications,* 2(1), 17–40.

Gençay, R., Selçuk, F., & Whitcher, B. (2005) Multiscale systematic risk. *Journal of International Money and Finance*, *24*(1), 55-70.

Gilles, J. (2013) Empirical Wavelet Transform. *IEEE Transactions on Signal Processing*, 61, 3999-4010.

Golyandina, N. & Zhigljavsky, A. (2013) Singular Spectrum Analysis for Time Series. *Springer Briefs in Statistics.*

Golyandina, N., Nekrutkin, V. & Zhigljavsky, **A. (2001)** Analysis of Time Series Structure: SSA and Related Techniques. *Monographs on Statistics and Applied Probability*, 90.

Grinsted, A., Moore, J. C., & Jevrejeva, S. (2004) Application of the cross wavelet transform and wavelet coherence to geophysical time series. *Nonlinear processes in geophysics*, *11*(5/6), 561-566.

Grossmann, A., & Morlet, J. (1984) Decomposition of Hardy functions into square integrable wavelets of constant shape. *SIAM Journal on Mathematical Analysis*, *15*(4), 723-736.

Crowley, P. M., & Mayes, D. G. (2009) How fused is the Euro area core? An evaluation of growth cycle co-movement and synchronization using wavelet analysis. *OECD Journal: Journal of Business Cycle Measurement and Analysis*, *2008*(1), 63-95.

Haar A. (1910) Zur Theorie der orthogonalen Funktionensysteme*. Mathematische Annalen.69(3),* 331-371.

Hansen, P. R., Lunde, A. & Nason, J. M. (2011) The model confidence Set. *Econometrica,* 79(2), 453-497.

Hansen, P.R. (2005) A test for superior predictive ability. *Journal of Business Economic Statistics*, 23, 365–380.

Hansen, P.R., & Lunde, A. (2006) Realized variance and market microstructure noise. *Journal of Business and Economic Statistics, 24(2)*, 127–161.

Hansen, P.R., & Lunde, A., (2005) A Realized Variance for the Whole Day Based on Intermittent High-Frequency Data. Journal of Financial Econometrics 3 (4), 525–554.

Hassani, H., Kalantari, M. & Beneki, C. (2021) Comparative Assessment of Hierarchical Clustering Methods for Grouping in Singular Spectrum Analysis. *Applied Math,* 1, 18–36.

Hassani, H. (2007) Singular Spectrum Analysis: Methodology and Comparison. *Journal of Data Science,* 5,239–257.

Hassani, H. & Thomakos, D. (2010) A review on singular spectrum analysis for economic and financial time series. *Statistics and Its Interface,*3, 377–397.

Haugom, E., Langeland, H., Molnár, P., & Westgaard, S. (2014) Forecasting volatility of the US oil market. *Journal of Banking and Finance 47*, 1–14.

Hecht-Nielsen, R. (1992) Theory of the Backpropagation Neural Network**Based on non-indent by Robert Hecht-Nielsen, which appeared in Proceedings of the International Joint Conference on Neural Networks 1, 593–611, June 1989. *IEEE., Neural Networks for Perception*, 65-93.

Hochreiter, S. & Schmidhuber, J. (1997) Long short-term memory. *Neural Computation*, 9(8), 1735-80.

Huang, N. E., Shen, Z., Long, S. R., Wu, M. C., Shih, H. H., Zheng, Q. & Liu, H. H. (1998) The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 454, 903-995.

Huang, N.E., Shen, Z., & Long, S.R. (1999) A New View of Nonlinear Water Waves: The Hilbert Spectrum. *Annual Review of Fluid Mechanics*., 31, 417–457.

Huang, N.E., Wu, M.-L.C., Long, S.R., Shen, S.S.P., Qu, W., Gloersen, P. & Fan, K.L. (2003) A confidence limit for the empirical mode decomposition and Hilbert spectral analysis. *Proceedings of the Royal Society A, Mathematical, Physical and Engineering Sci*ences, 459, 2317–2345.

Huang, Y., Yan, C.J., & Xu, Q. (2012) On the difference between empirical mode decomposition and Hilbert vibration decomposition for earthquake motion records. *In Proceedings of the 15th World Conference on Earthquake Engineering*, Lisbon, Portugal.

Huang, X., & Tauchen, G., (2005) The relative contribution of jumps to total price variance. *Journal of Financial Econometrics 3*, 456–499

Hu, J. & Wang, J. (2015) Short-term wind speed prediction using empirical wavelet transform and Gaussian process regression, *Energy*, 93(2),1456-1466.

Hudgins, L.H., Hudgins, L.H., Friehe, C.A., Friehe, C.A., Mayer, M.E., & Mayer, M.E. (1993) Wavelet transforms and atmospheric turbulence. *Physical review letters, 71 20*, 3279-3282.

Hull, J. & White, A. (1987) The Pricing of Options on Assets with Stochastic Volatilities*, Journal of Finance, 42*, 281-300.

Huang, X., & Tauchen, G., (2005) The relative contribution of jumps to total price variance. J. Financial Econometrics 3, 456–499.

Jin, Z., Jin Y., & Chen, Z. (2022) Empirical mode decomposition using deep learning model for financial market forecasting. *Peer Journal of Computer Science, 8*.

Jorion, P. (1995) Predicting volatility in the foreign exchange market. *Journal of Finance 50*, 507–528.

Junsheng, C., **Dejie, Y., & Yu, Y. (2006)** A fault diagnosis approach for roller bearings based on EMD method and AR model. *Journal of Mechanical Systems and Signal Processing 20,* 350-362.

Kambouroudis, D. S., McMillan, D. G. & Tsakou, K. (2021) Forecasting realized volatility: The role of implied volatility, leverage effect, overnight returns, and volatility of realized volatility. *Futures markets,* 41(10),1618-1639.

Kambouroudis, D. S., McMillan D. G. & Tsakou, K. (2016) Forecasting stock return volatility: A comparison of GARCH implied volatility and realized volatility models. *Journal of Futures Markets*, 36(12), 1127-1163.

Konstantinidi, E., Skiadopoulos, G. & Tzagkaraki, E. (2008) Can the evolution of implied volatility be forecasted? Evidence from European and US implied volatility indices. *Journal of Banking and Finance,* 32(11), 2401-2411.

Labat, D. (2008) Wavelet analysis of the annual discharge records of the world's largest rivers. *Advances in Water Resources, 31*, 109-117.

Lahmiri, **S**. **(2016)** A variational mode decomposition approach for analysis and forecasting of economic and financial time series. *Expert Systems with Applications*, 55, 268-273.

Le Pen, Y., & Sévi, B., (2017) Futures trading and the excess co-movement of commodity prices. *Review of Finance. 22 (1)*, 381–418.

Li, Z., Wang, M., Wang, X., Liu, Z. & Shi, A. (2020) Oil Price Forecasting Based on Variational Mode Decomposition, Relative Entropy and LSTM Neural Network. *IOP Conference Series: Materials Science and Engineering*, 750.

Liu, M., Sun, X., Wang, Q. & Deng, S. **(2022)** Short-Term Load Forecasting Using EMD with Feature Selection and TCN-Based Deep Learning Model. *Energies*, 15, 7170.

Liu, W., Cao, S. & Chen, Y. **(2016)** Applications of variational mode decomposition in seismic time-frequency analysis. *Geophysics*, 81, 365–378.

Liu Y L., J A Patton, & Sheppard K. (2015) Does anything beat 5-minute RV? A comparison of realized measures across multiple asset classes, *Journal of Econometrics 187,* 293-311

Mallat, S. (1989) A theory for multiresolution signal decomposition: The wavelet representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence 11:7*, 674-93.

Meyer, Y. (1997) Wavelets, Vibrations and Scalings. *American Mathematical Society, Centre de Recherches Mathematiques*, *9*, 133.

Michańków, J., Sakowski, P. & Ślepaczuk R. **(2022)** LSTM in Algorithmic Investment Strategies on BTC and S&P500 Index. *Sensors (Basel)*. *22(3),* 917.

Moran, M. T. & Dash, S., (2007) "VIX Futures and Options: Pricing and Using Volatility Products to Manage Downside Risk and Improve Efficiency in Equity Portfolios. *Journal of Trading, 2 (3)*,96–105.

Noh, J., Engle, R.F., & Kane, A., (1994) Forecasting volatility and option prices of the S&P 500 index. *Journal of Derivatives 2 (1)*, 17–30.

Nomikos, N.K., & Pouliasis, P.K., (2011) Forecasting petroleum futures markets volatility: the role of regimes and market conditions. *Energy Economics, 33 (2),* 321–337.

Plakandaras, V., Papadimitriou, T., & Gogas, P. (2015) Forecasting Daily and Monthly Exchange Rates with Machine Learning Techniques. *Journal of Forecast*, *34*, 560–573.

Peel, M. C., Pegram, G. G. S. & McMahon, T. A. **(2007)** Empirical mode decomposition: improvement and application. *In Proceedings of*

International Congress Modelling Simulation, 1, 2996 – 3002. Modelling and Simulation Society of Australia, Canberra.

Ping Yu, Fang, J., Xu, Y. & Shi, Q. (2021) Application of Variational Mode Decomposition and Deep Learning in Short-Term Power Load Forecasting. *Journal of Physics: Conference Series, 1883.*

Prasad, A., & Bakhshi, P. (2022) Forecasting the Direction of Daily Changes in the India VIX Index Using Machine Learning. *Journal of Risk and Financial Management, 15.*

Ricker NH. (1940) The form and nature of seismic waves and the structure of seismograms. Geophysics. (4), 348-366.

Rilling, G. Flandrin, P. & Goncalves P. (2003) On empirical mode decomposition and its algorithms. *In Proceedings 6th IEEE-EURASIP Workshop on Nonlinear Signal and Image Processing*.

Rodríguez-Murillo, J.C., & Filella, M. (2020) Significance and Causality in Continuous Wavelet and Wavelet Coherence Spectra Applied to Hydrological Time Series. *Hydrology*.

Rua, A., & Nunes, L. C. (2009) International comovement of stock market returns: A wavelet analysis. *Journal of Empirical Finance*, *16*(4), 632-639.

Shankar, S., Dandapat S. & Barma, **S. (2021)** Classification of Epileptic Seizure from EEG Signal Based on Hilbert Vibration Decomposition and Deep Learning, *43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, Mexico, 2802-2805.

Shaari, M., Samsudin, R. & Shabri, A. (2018) Forecasting Drought Using Modified Empirical Wavelet Transform-ARIMA with Fuzzy C-Means Clustering. *Indonesian Journal of Electrical Engineering and Computer Science*, 11, 1152-1161.

Shi, P., Yang, W., Sheng, M. & Wang, M. (2017) An Enhanced Empirical Wavelet Transform for Features Extraction from Wind Turbine Condition Monitoring Signals. *Energies,* 10(7), 972.

Shu, W. & Gao, Q. (2020) Forecasting Stock Price Based on Frequency Components by EMD and Neural Networks. *In IEEE Access*, *8*, 206388-206395.

Singh, O., & Sunkaria, R. (2016) ECG signal denoising via empirical wavelet transform. *Australasian Physical & Engineering Sciences in Medicine*, 40,1-11.

Singh, M. J., Sharma L. N. & Dandapat, S. (2022) Hilbert Vibration Decomposition of Seismocardiogram for HR and HRV Estimation. *2022 IEEE International Conference on Signal Processing and Communications (SPCOM)*, Bangalore, India,1-5.

Sulandari, W., Subanar, M., Lee, H. & Rodrigues, P.C. (2020) Indonesian electricity load forecasting using singular spectrum analysis, fuzzy systems and neural networks, *Energy*, *190*.

Sun, S., Wang, S., Zhang, G., & Zheng, J. (2018) A decomposition–clustering–ensemble learning approach for solar radiation forecasting. *Solar Energy*, *163*, 189–199.

Szado, E. (2009) VIX Futures and Options: A Case Study of Portfolio Diversification during the 2008 Financial Crisis. *Journal of Alternative Investments, 12 (2),* 68–85.

Szado, E. (2018) The Distinctive Characteristics of VIX Futures and Options. *Working Paper*.

Szado, E. (2020) Selling VIX® Futures and Options for Portfolio Return Enhancement. *Working paper.*

Tang, L., Wu, Y., & Yu, L. (2018) A randomized-algorithm-based decomposition-ensemble learning methodology for energy price forecasting. *Energy*, *157,* 526–538.

Thomakos, D., Wang, T., & Wille, L. (2002) Modeling daily realized futures volatility using singular spectrum analysis. *Physica A: Statistical Mechanics and its Applications*, 312(3), 505–519.

Theodoridis, S. (2020) Machine Learning: A Bayesian and Optimization Perspective (Second Edition). *Academic Press*.

Torrence, C., & G. Compo. (1998). "A practical guide to wavelet analysis," *Bulletin of the American Meteorological Society 79:1*, pp. 61-78.

Torrence, C., & Webster, P.J. (1999) Interdecadal changes in the ENSO–monsoon system. *Journal of Climate*, *12*(8), 2679-2690.

Torrence, C., & Webster, P.J. (1997) The annual cycle of persistence in the El Nino-Southern Oscillation. *Quarterly Journal of Royal Meteorology Soc*iety.

Vacha, L. & Barunik, J. (2012). Co-movement of energy commodities revisited: evidence from wavelet coherence analysis*. Energy Economics 34(1):* 241–247.

Vrontos, S. D., Galakis J. & Vrontos, **I. D. (2021)** Implied volatility directional forecasting: a machine learning approach. *Quantitative Finance*, 21(10), 1687-1706.

Wu, Z., & Huang, N.E. (2009) Ensemble Empirical Mode Decomposition: A Noise-Assisted Data Analysis Method. *Advances in Adaptive Data Analysis*, 1, 1–41.

Winters, P.R. (1960). Forecasting sales by exponentially weighted moving averages. *Management Science,* 6(3), 324–342.

Yeh, J. R., Shieh, J. S., & Huang, N.E. (2010) Complementary ensemble empirical mode decomposition: A novel noise enhanced data analysis method. *Advances in Adaptive Data Analysis*, 2(02), 135-156.

Yu, L., Wang, S., & Lai, K.K. (2008) Forecasting crude oil price with an EMD-based neural network ensemble learning paradigm. *Energy Economics*, *30(5),* 2623–2635.

Yang, L., Cai X., Zhang, H., & Hamro, S. (2016) Interdependence of foreign exchange markets: A wavelet coherence analysis, *Economic Modelling (55),* 6-14.