Università degli studi di Milano-Bicocca

Dipartimento di Statistica e Metodi Quantitativi
Corso di Laurea Magistrale in Economia e Finanza

Long term guarantee measures: stabilità e sostenibilità finanziaria nelle compagnie di assicurazione

Relatore:
Prof.ssa Edit Rroji

Tesi di Laurea di:

Matteo Pastore
Matr. N. 870843

ANNO ACCADEMICO 2024/2025

Indice

Introduction **1**

1 Solvency II **2**

1.1 Requisiti quantitativi 3

2 Risk-Free Rate (RFR) curve **9**

2.1 Yield curve 9

2.1.1 Overnight Interest Rate Swap - OIS 9

2.1.2 Interest Rate Swap 12

2.1.3 Titoli di stato 12

2.2 Credit Risk Adjustment (CRA) 13

2.3 Il modello Nelson-Siegel-Svensson (1994) 15

2.4 Il modello Smith-Wilson (2000) 18

2.4.1 Hedging nel modello Smith-Wilson 27

2.4.2 Calibrazione del parametro α 29

2.4.3 Esempio Applicazione a OIS sull'ESTR 30

2.4.4 Esempio di Processo Gaussiano 34

3 Long-Term Guarantee Measures (LTG measures) **38**

3.1 Volatility Adjustment (VA) 39

3.1.1 VA in utilizzo 39

3.1.2 VA proposto 45

3.2 Matching Adjustment (MA) 47

3.3 Transitional Measures 48

3.3.1 Misura transitoria sui tassi d'interesse privi di rischio 48

3.3.2 Misura transitoria sulle riserve tecniche 49

4 Analisi empirica **51**

4.1 Analisi descrittiva 51

4.1.1 Selezione Input 53

4.2 Definizione della struttura del modello 55

4.2.1 Long Short Term Memory (LSTM) 55

4.2.2 Calcolo dei pesi ottimi 62

4.2.3 Nesterov-accelerated adaptive moment estimation (NADAM) 63

4.2.4 Boosting (applicazione di XGBoost a un LSTM) 65

4.2.5 Forecasting 71

4.3 Risultati finali 104

4.4 Conclusioni 125

5 Codice **126**

5.1 Codice Esempio Capitolo 2 126

5.2 Codice Capitolo 4 138

List of Figures **156**

Abstract

La presente tesi analizza le misure di lungo termine (Long-Term Guarantee measures, LTGM) introdotte dal regime prudenziale Solvency II con l'obiettivo di ridurre l'impatto della volatilità temporanea dei mercati sui requisiti minimi prudenziali. In particolare, l'attenzione si concentra sul volatility adjustment (VA), uno strumento regolamentare che modifica i tassi di sconto per la valutazione delle Best Estimate Liabilities (BEL), attenuando l'impatto di shock temporanei sui mercati finanziari. Dopo aver illustrato il quadro normativo e gli approcci di stima della curva risk-free, la tesi esamina il funzionamento delle LTG measures e del volatility adjustment come indicato da EIOPA e dalla regolamentazione in materia. La parte centrale del lavoro sviluppa un'analisi empirica: è stata costruita una rete neurale Long Short Term Memory (LSTM) con l'obiettivo di prevedere i valori futuri del VA per 18 paesi, utilizzando sia i lag storici del VA sia un insieme di regressori macro. Nel tentativo di migliorare la performance della rete, viene anche applicato un approccio di boosting ai residui della rete (Extreme Gradient Boosting) generando un approccio integrato. Il contributo della tesi si identifica principalmente nella progettazione, implementazione e validazione del modello predittivo. La procedura sviluppata integra metodologie econometriche, di machine learning e deep learning con tecniche di validazione out-of-sample, includendo sia previsioni iterative (basate su valori predetti con modelli econometrici) sia previsioni condizionate su valori reali, così da distinguere tra la capacità intrinseca della rete e la sua performance in uno scenario di previsione realistico.

1. Solvency II

Le compagnie assicurative rappresentano una parte fondamentale del sistema finanziario, secondo la relazione annuale IVASS relativa al 2023[1] i premi incassati dalle imprese assicurative ammontavano al 6,2% del PIL italiano, di cui il 4,4% ramo vita e l'1,8% ramo danni. L'elevata rilevanza sistemica di questi operatori è legata sia alle ingenti somme di denaro che gestiscono, sia alla peculiarità del loro modello operativo, caratterizzato da un ciclo finanziario inverso: le imprese incassano anticipatamente i premi, per poi adempiere alle obbligazioni contrattuali in un secondo momento. Queste specificità hanno portato negli anni allo sviluppo di normativa per ridurre il problema del moral hazard intrinseco all'attività. A livello europeo troviamo la direttiva framework Solvency II[2] che si pone lo scopo di garantire una prudente gestione dell'attività assicurativa al fine di limitare il rischio di default delle singole società che potrebbe causare danni agli assicurati, ridurre la liquidità nei mercati e portare a problemi di fiducia nel sistema finanziario. La direttiva è strutturata su tre pilastri: Il primo pilastro riguarda i requisiti quantitativi, in particolare la valutazione delle poste attive e passive di bilancio e i requisiti di solvibilità. Il secondo pilastro riguarda i requisiti di natura qualitativa, dunque la parte di corporate governance, l'istituzione dei tre livelli di controlli interni e le divisioni aziendali responsabili, nonché le caratteristiche che queste devono avere. Il terzo pilastro riguarda la trasparenza, quindi le modalità e la tipologia di informazioni che vanno comunicate all'autorità di vigilanza nazionale e agli altri stakeholders.

Per comprendere l'architettura di Solvency II è necessario chiarire in maniera semplice il funzionamento di un'impresa assicurativa, possiamo vedere il bilancio di questa come un processo stocastico a tempo discreto: $\tilde{A}_t - \tilde{L}_t$ con $\tilde{A}_t$ valore dell'attivo al tempo t e $\tilde{L}_t$ valore delle passività della compagnia al tempo t, questo a sua volta possiamo vederlo suddiviso in due parti $\tilde{X}_t + \tilde{L}_t^+$ con $\tilde{X}_t$ flussi di cassa in uscita tra $t-1$ e t (sia da contratti di assicurazione che spese diverse) ed $\tilde{L}_t^+$ esposizioni ancora in essere in t, per semplicità ipotizziamo che lo scostamento tra $t-1$ e t sia annuale e che la nostra società esemplificativa ponga in essere solo contratti annuali. Parten-

[1] IVASS, Relazione annuale 2023.

[2] Direttiva 2009/138/CE del Parlamento Europeo e del Consiglio.

do da $t = 0$ (tempo corrente) quello che intuitivamente la regolamentazione chiede di fare è far sì che $P\left[\tilde{A}_1 \geq \tilde{L}_1\right] = P\left[\tilde{L}_1 - \tilde{A}_1 \leq 0\right] >= 1 - p$ questa condizione implica in $t = 0$ una scelta di una allocazione sensata degli asset A_0 ed eventualmente l'aggiunta di capitale C_0 per aumentare il livello degli asset e garantire la condizione di "non rovina", ossia che la società sia solvibile con una probabilità molto elevata.
Questa condizione è spesso descritta attraverso l'utilizzo di una misura di rischio nel caso di Solvency II la misura indicata è il Value at Risk (VaR) calcolato al livello di confidenza 99.5%.

1.1 Requisiti quantitativi

Poiché la presente tesi si concentra prevalentemente sul primo pilastro di Solvency II, se ne delineano di seguito le principali componenti. La valutazione è volta a determinare il valore delle passività della compagnia di assicurazione (in particolare le *technical provisions*), alla determinazione del valore degli attivi a copertura e al calcolo del capitale minimo necessario per il calcolo del $VaR_{99.5\%}$, nell'esempio stiamo parlando di come determinare A_0, L_0 e C_0. Per quanto concerne gli attivi, il principio proposto all'articolo 75 di Solvency II prescrive: *le attività sono valutate all'importo al quale potrebbero essere scambiate tra parti consapevoli e consenzienti in un'operazione svolta alle normali condizioni di mercato* in particolare viene fatto riferimento ai principi IFRS (nello specifico IFRS 13) per la valutazione al fair value all'articolo 9 del regolamento delegato 2015/35[3] . In sintesi, questi principi prevedono di utilizzare come valore fair degli attivi i valori di mercato se l'attività è negoziata, se non disponibili utilizzare le quotazioni di strumenti "simili", se non ci sono nemmeno strumenti simili bisogna effettuare una valutazione del tipo valore attuale dei flussi di cassa futuri. Più complessa è la valutazione dei passivi che in Solvency II viene suddivisa in quattro macro blocchi: fondi propri, *best estimate liabilities* (BEL,ossia le riserve tecniche), *risk margin* e le altre passività. Guardando all'esempio le BEL consisterebbero sostanzialmente nella determinazione di L_0 che potremmo intuitivamente vedere come $\sum_{i=1}^{n} \frac{\mathbb{E}^{\mathbb{P}}\left[\tilde{X}_i\right]}{1+rf}$, il *risk margin* sarebbe una correzione applicata alle BEL che servirebbe a dare una remunerazione a un potenziale acquirente che comprasse la società assicurativa in caso di default della compagnia; la remunerazione

[3]Regolamento delegato (UE) 2015/35, Art. 75, e IFRS 13.

è dovuta al fatto che l'assunzione degli impegni dalla società in fallimento comporterebbe l'incremento del capitale richiesto alla società acquirente. I fondi propri rappresentano C_0 che dovrebbe far si che la condizione richiesta dalla misura di rischio sia rispettata, le altre passività sono invece ad esempio corporate bond emessi dalla compagnia.

Tornando a Solvency II, per quanto riguarda i fondi propri sono distinti in due gruppi: i fondi propri di base, ossia tutte quelle passività in bilancio che possono essere chiamate ad assorbire perdite, quindi sostanzialmente capitale e strumenti ibridi subordinati. I fondi propri ausiliari sono invece attività fuori bilancio che possono essere chiamate ad assorbire potenziali perdite (possono essere aggiunte al capitale regolamentare solo dopo l'approvazione dell'autorità di vigilanza); alcuni esempi sono lettere di credito, garanzie di cui dispone la società. A questo punto la normativa definisce un'ulteriore classificazione basata su tre tier, l'articolo 93 della direttiva Solvency II e gli articoli da 71 a 79 del regolamento delegato 2015/35[4] definiscono le modalità con cui suddividere le due tipologie sui tre *Tier*. Possiamo delineare in sintesi le caratteristiche rilevanti degli strumenti:

- l'elemento della piena e continua disponibilità dell'impresa.
- la capacità dell'elemento di essere chiamato ad assorbire le perdite.
- il grado di subordinazione (ossia lo strumento in caso di liquidazione deve essere pagato dopo tutte le altre obbligazioni, tra cui quelle verso gli assicurati e le società di riassicurazione).
- la durata dello strumento, se con scadenza definita, che questa scadenza sia congrua con la scadenza delle passività aziendali.
- se lo strumento è privo di obblighi/incentivi al rimborso.
- se l'elemento è privo di spese fisse obbligatorie (ad esempio assenza di dividendo fisso obbligatorio).
- se l'elemento è libero da gravami.

[4]Direttiva Solvency II, Art. 93 e Regolamento delegato 2015/35, Artt. 71-79.

A questo punto si può far rientrare nel Tier 1 solo gli elementi dei fondi propri di base che abbiano tutte le caratteristiche elencate in precedenza. Nel Tier 2 vanno i fondi propri di base che non rispettano alcune delle proprietà elencate (in particolare quella relativa alla disponibilità) e i fondi propri accessori che invece rispettano tutte le proprietà elencate. Nel Tier 3 rientrano sempre strumenti relativi ai fondi propri di base che non abbiano le caratteristiche necessarie per entrare nel Tier 1 o nel Tier 2, e i fondi propri accessori che non hanno tutte le caratteristiche elencate. Questa classificazione è fondamentale per la normativa in quanto questi tier andranno a comporre i requisiti di capitale relativi agli *Own Funds* (OF), infatti Solvency prevede che questi siano composti dal *Solvency Capital Requirement* (SCR) e il *Minimum Capital Reqirement* (MCR).
Il *solvency Capital Requirement* può essere composto da tutti e tre i livelli a patto che gli elementi ammessi al primo livello rappresentino almeno un terzo di tutti gli *own funds* e l'importo complessivo degli elementi ammessi al terzo livello sia invece inferiore a un terzo di tutti gli OF.
Il *Minimum Capital Requirement* è "contenuto" nel SCR e richiede un minimo livello di fondi propri a maggiore capacità di assorbimento delle perdite; infatti, possono rientrare nel conteggio del requisito solamente i fondi propri di base, e solamente quelli ammessi alla classificazione di primo e secondo livello, è inoltre necessario che almeno il 50% del MCR sia rappresentato da fondi propri di base ammessi al primo livello. Pensando all'esempio, stiamo parlando sostanzialmente della composizione dei C_0 dal punto di vista del passivo di bilancio, in particolare $SCR = C_0$. Come accennato, il calcolo di questo valore viene fatto calcolando il *Value At Risk* al livello del 99.5% di $\tilde{L}_1 - \tilde{A}_1$. Dal punto di vista tecnico, lo SCR (ossia: VaR al 99,5% dell'eccedenza delle passività sugli attivi) può essere calcolato attraverso la formula standard proposta dal regolatore con la seguente scomposizione[5]:

$$SCR = BSCR + SCR_{op} - Adj$$

con

1. il requisito patrimoniale di solvibilità di base (Basic Solvency Capital Requirement-BSCR)

2. il requisito patrimoniale per il rischio operativo (Solvency Capital Requirement for operational risk-SCR_{op})

[5]Cfr. EIOPA Guidelines on Solvency II, Standard Formula SCR.

3. l'aggiustamento per la capacità di assorbimento delle perdite delle riserve tecniche e delle imposte differite (adjustment loss-absorption -Adj).

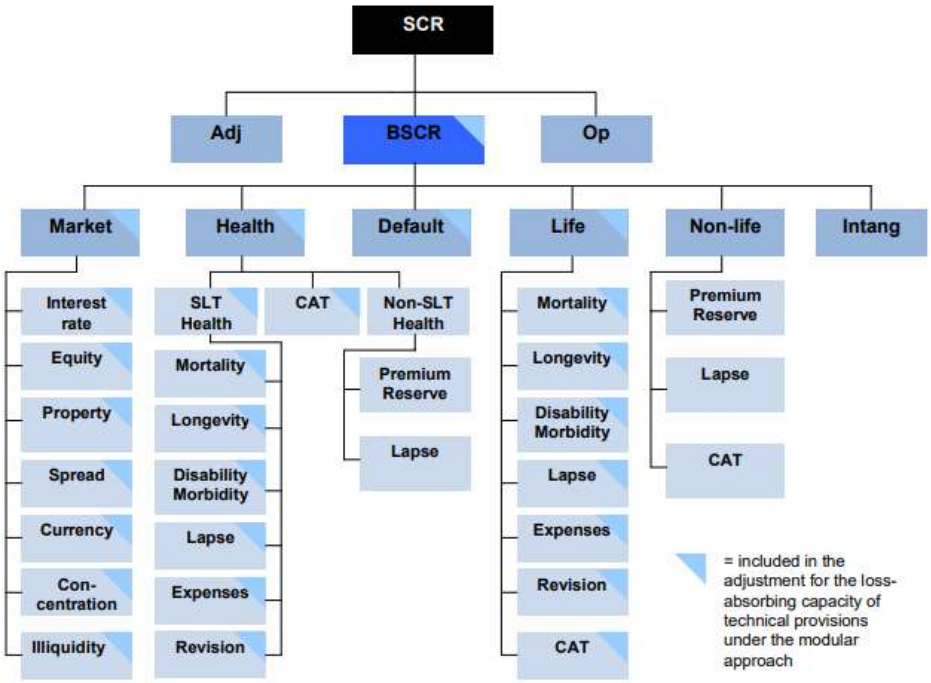

Figura 1.1: Struttura del Solvency Capital Requirement (SCR). Fonte: QIS5 Technical Specifications, pag. 90.

Il BSCR cattura in sé diverse fonti di rischio che compongono la variabile casuale $\tilde{L}_1 - \tilde{A}_1$: i rischi assicurativi, a loro volta suddivisi in tre componenti, la prima è la componente "non-life" e dovrebbe catturare oscillazioni inattese dei momenti di accadimento dei sinistri, entità dei sinistri, errori di tariffazione ecc. relativi ai contratti non-life.
seconda componente è relativa alla parte life insurance e dovrebbe catturare i rischi relativi all'andamento dei tassi di mortalità, volatilità dei tassi di revisione delle rendite, ecc.
La terza componente riguarderebbe la parte health insurance dove i fattori che generano la rischio-

sità sono simili alla categoria life, rilevante è ad esempio la volatilità delle spese mediche connesse ai contratti. I tre moduli hanno a loro volta dei sottomoduli di rischio. Altre componenti del BSCR sono il rischio di mercato che dovrebbe catturare tutte le oscillazioni principalmente della parte attiva dovute a movimenti dei diversi mercati: azionario, fixed income, currency ecc. Anche questa componente è composta da sottomoduli da calcolare. Il rischio di default della controparte che cattura ad esempio la possibilità che una società di riassicurazione peggiori la sua qualità creditizia o fallisca non adempindo alle proprie obbligazioni contrattuali. L'ultimo modulo è quello relativo alle attività immateriali la cui volatilità è associata al basso grado di liquidità e alla specificità del tipo di attività immateriale. Il BSCR secondo la standard formula è poi calcolato come:

$$BSCR = \sqrt{\sum_i \sum_j Corr_{i,j} SCR_i SCR_j} + SCR_{intangibles}$$

dove $Corr_{i,j}$ è la correlazione tra le sei fonti di rischio elencate precedentemente ed è fornita direttamente dal regolatore nel caso della formula standard. La seconda componente del SCR è il requisito patrimoniale per il rischio operativo, dal punto di vista del nostro esempio questi rientrano in quello che abbiamo definito come $\tilde{X}_t$ e sarebbero di fatto spese di conto economico legate alla gestione dell'impresa e diverse dalle obbligazioni contrattuali assicurative, quello che interessa adesso è l'aspetto della aleatorietà di queste spese e quindi il potenziale impatto sull'attivo di bilancio che dovrebbe coprire le obbligazioni assicurative. Dal punto di vista normativo esso copre il rischio di perdite derivanti ad esempio da disfunzioni di processi interni, spese legali inattese, eventi esogeni che generino danni inattesi alla società assicurativa, per il suo calcolo, relativamente ai contratti di assicurazione/riassicurazione dove il rischio di investimento è sopportato dalle controparti il calcolo del requisito viene fatto sull'importo delle spese annuali sostenute dal contraente in relazione a quelle esposizioni. Per quanto riguarda i contratti di assicurazione e riassicurazione diversi da quelli precedentemente indicati si calcola il requisito sul insieme dei premi più riserve tecniche detenute per quelle esposizioni, in questo secondo caso il requisito patrimoniale per il rischio operativo non può superare il 30% del BSCR calcolato sulle medesime posizioni.

L'ultima componente della sommatoria è *l'adjustment loss-absorption* riguarda la capacità di assorbimento delle perdite di riserve tecniche e imposte differite. Le riserve tecniche possono essere ridotte se ad esempio i contratti che stanno coprendo hanno clausole di partecipazione agli utili

(presenti ad esempio in alcune clausole dei contratti di riassicurazione come forme di incentivazione alla cessione di rischi ”buoni”). Anche le imposte differite possono essere utilizzate come forma di ”riduzione delle imposte future” e quindi utilizzate per coprire le passività dell’impresa e dunque essere fatte rientrare nel capitale di questa. Per il calcolo numerico dei diversi sotto-blocchi del VaR della standard formula vengono previsti due approcci da applicare a seconda del rischio:

- Factor Based Formula: è basata sull’applicazione di uno o più fattori a una misura di esposizione al rischio.
- Scenario Approach: consiste nella misurazione del ΔBOF, ossia della variazione del capitale regolamentare al seguito dell’applicazione di shock su alcune componenti di rischio; l’entità degli shock è stabilita da EIOPA.

La componente più rilevante ai fini della tesi è quella composta dalle *technical provisions*, in particolare la parte delle best estimate liabilities. Queste sono calcolate come valore attuale atteso delle obbligazioni contrattuali derivanti dai contratti di assicurazione in essere, quindi sostanzialmente

$$\frac{\mathbb{E}^{\mathbb{P}}\left[prestazioni - premi\right]}{1+rf} = \sum_{i=1}^{n} \frac{\mathbb{E}^{\mathbb{P}}\left[\tilde{X}_i\right]}{1+rf}.$$

Nel corso del presente elaborato analizzerò le modalità di calcolo del tasso risk-free al denominatore ed delle long term guarantee measures che consistono in degli *add-on* al tasso risk-free e hanno lo scopo di rendere più stabili i bilanci delle imprese di assicurazione.

2. Risk-Free Rate (RFR) curve

In questo capitolo descriverò come ricavare il tasso risk-free da utilizzare per il calcolo delle best estimate liabilities. Il calcolo consiste in quattro passaggi: 1) selezione degli strumenti finanziari da cui estrarre i tassi risk-free 2) estrapolazione dei tassi impliciti e costruzione della curva a scadenze discrete (Yield Curve) 3) eventuale aggiustamento per il rischio di credito residuo degli strumenti 4) interpolazione dei punti intermedi alle scadenze definite dagli strumenti del punto 1).

2.1 Yield curve

Come detto, per determinare i fattori di sconto risk-free per una specifica valuta impliciti nel mercato alla data odierna è necessario individuare degli strumenti finanziari quotati e sufficientemente liquidi da cui estrarre questa informazione. I punti di partenza che propongo sono tre: overnight interest rate swap, swap oppure government bond.

2.1.1 Overnight Interest Rate Swap - OIS

Gli Overnight Interest Rate Swap sono swap che hanno una gamba variabile costruita in questo modo: ipotizzando che r_t sia il tasso su cui è costruito il contratto, la gamba variabile paga

$$FL = nozionale \times \left[\prod_{i=1}^{d_n} \left(1 + \frac{n_i REF_i}{d_y} \right) - 1 \right].$$

- d_n numero di giorni lavorativi tra le date di pagamento delle gambe variabili dello swap
- $d_y$numero di giorni dell'anno considerato per la currency di riferimento del tasso
- n_i numero di giorni che ci sono tra due giorni di mercato consecutivi (ad esempio, se siamo a venerdì, il tasso si applicherà per 3 giorni)
- REF_i tasso di riferimento all'i-esimo giorno

La gamba fissa ha tipicamente scadenza annuale. Per ricavare il tasso risk-free per la valuta euro si utilizzano tipicamente gli Swap con gamba variabile indicizzata all'ESTR (Euro Short Term Rate). L'ESTR è costruito come il tasso medio a cui un insieme di banche ha effettuato operazioni overnight e valuta euro in un dato giorno; viene pubblicato alle 8:00 CET di ogni giorno lavorativo di Target2 (sistema di clearing tra banche). Come evidenziato in Bianchetti e Carlicchi 2013, queste forme di contratto sono le uniche per cui rimane valida la proprietà telescopica della floating leg. Affinché la proprietà telescopica sia valida è necessario che sia verificata la seguente proprietà:

$$D(0,t)\frac{1}{1+F(t,T)\tau(t,T)} = D(0,T). \tag{1}$$

l'equazione implica che investire per un primo periodo a un tasso noto da 0 e t, incassare l'investimento, e selezionare un successivo investimento tra t e T (che secondo il mercato in $t = 0$ ha un rendimento atteso pari a $F(t,T)$, tasso forward tra t e T) deve essere equivalente a bloccare l'investimento da 0 a T ad un tasso noto oggi. Il motivo per cui in tassi diversi da quelli *overnight*, come ad esempio l'Euribor, dove si ha sempre il merito creditizio elevato delle controparti ma non la natura overnight del tasso sottostante ai contratti, è dato dal differente grado di liquidità delle due posizioni. Ipotizzando di essere una banca appartenente alle banche che segnalano il proprio tasso per la costruzione dell'Euribor, ipotizzando ulteriormente di poter scegliere tra due strategie, la prima è prestare a un'altra banca del panel a 6 mesi e poi decidere se rinnovare il prestito, la seconda è prestare alla stessa banca per 12 mesi di fila. La prima strategia risulta più "sicura" perché nell'arco dei 6 mesi, se la banca a cui abbiamo prestato ha peggiorato il suo merito creditizio, possiamo decidere di non rinnovare il prestito ma prestare i soldi per i successivi sei mesi ad un'altra banca con rating ancora elevato, nella seconda strategia si è bloccati per 12 mesi, dunque si tenderà a chiedere un premio a livello di rendimento. Questo è il motivo per cui la proprietà rimane valida se parliamo di contratti indicizzati a tassi giornalieri, dove sostanzialmente la gamba variabile rappresenta una sorta di strategia di rinnovo giornaliero del contratto e dunque i rischi impliciti vengono "annullati" grazie alla brevissima durata dei contratti su cui i tassi sono costruiti, e all'elevata qualità creditizia delle controparti.

Dall'equazione (1) si ottengono i seguenti risultati:

$$F(t,T) = \frac{1}{\tau(t,T)}\left(\frac{D(0,t)}{D(0,T)} - 1\right).$$

$$F(t,T)\tau(t,T)D(0,T) = D(0,t) - D(0,T).$$

queste relazioni permettono di riscrivere il valore attuale atteso di mercato della gamba variabile come:

$$PV_{\text{Floating}} = \sum_{i=0}^{N-1} F(t_i,t_{i+1}) \cdot \tau(t_i,t_{i+1}) \cdot D(0,t_{i+1}) = \sum_{i=0}^{N-1} [D(0,t_i) - D(0,t_{i+1})] = D(0,t_0) - D(0,t_N).$$

Questa sarebbe la proprietà telescopica.

Ora ipotizzando di aggiungere un pagamento di nozionale di 100 a scadenza sulle due gambe otteniamo che il valore attuale della gamba variabile ad oggi è:

$$100PV_{\text{Floating}} + 100D(0,t_N) = 100\,[D(0,t_0) - D(0,t_N)] + 100D(0,t_N) = 100.$$

siccome il valore del contratto nel giorno della stipula deve essere pari a zero allora deve valere che:

$$100PV_{\text{Fixed}} + 100D(0,t_N) = 100.$$

questa condizione può essere sfruttata per ricavare i tassi risk-free impliciti negli OIS di mercato sull'ESTR.

Esempio: ipotizzando di avere la quotazione di uno swap con gamba fissa a 1 anno con scadenza a 1 anno al 1.5% e di utilizzare questo direttamente come primo tasso risk-free, inoltre, di avere un secondo contratto a scadenza 2 anni che quota $R_{ois2y} = 2\%$

$$2e^{-0.015\cdot 1} + 102e^{-Z_{2y}\cdot 2} = 100$$

$$D(0,2y) = e^{-Z_{2y}\cdot 2} = \frac{100 - 2e^{-0.015}}{102}.$$

per ricondursi alla notazione utilizzata in precedenza basta risolvere la seguente equazione in rf:

$$D(0,2y) = (1 + rf_{2yrs})^2.$$

al fine di procedere abbiamo bisogno di nuovi contratti, ipotizzando di avere un altro swap a scadenza 3 anni che quota un fair strike rate di $R_{ois3y} = 2.15\%$ avremo:

$$2.15e^{-0.015} + 2.15e^{-Z_{2y}\cdot 2} + 102.15e^{-Z_{3y}\cdot 3} = 100$$

però ora $e^{-Z_{2y}}$ è noto:

$$D(0,3y) = e^{-Z_{3y}\cdot 2} = \frac{100 - 2.15e^{-0.015} - 2.15e^{-Z_{2y}\cdot 2}}{102.15}.$$

$$= \frac{1 - 0.0215e^{-0.015} - 0.0215e^{-Z_{2y}\cdot 2}}{1 + 0.0215} = \frac{1 - R_{ois3y}\left[D(0,1y) + D(0,2y)\right]}{1 + R_{ois3y}}.$$

Il risultato di questa procedura è un vettore di tassi risk-free a scadenze discrete che verrà utilizzato nella fase successiva.

2.1.2 Interest Rate Swap

Un altro punto di partenza per la procedura di costruzione della curva risk-free nelle diverse valute può essere rappresentato dall'utilizzo di swap classici, nel caso dell'euro ad esempio dagli swap sull'euribor. Per quanto detto in precedenza, non è valida la proprietà telescopica. I punti di partenza sono rappresentati dagli swap fair rate dei contratti per le diverse scadenze, l'unica richiesta è che le gambe variabili dei diversi contratti abbiano stessa tenor. Illustrerò nella sezione successiva come aggiustare questi tassi per poterli inserire nella costruzione della curva.

2.1.3 Titoli di stato

Nel caso in cui gli swap in una certa valuta non siano abbastanza liquidi, possono essere utilizzati per la costruzione della curva risk-free direttamente titoli di debito pubblico nella valuta d'interesse. Assumendo di avere uno zero coupon bond a scadenza 3 mesi che quota 97.5€, uno ZCB a 6 mesi che quota 94.9€, uno ZCB a 1 anno che quota 90€ e un coupon bond a scadenza 1.5 anni con cedola semestrale di 4€ e quotazione 96€. Per estrarre i tassi impliciti è sufficiente fare:

$$100e^{-R_{3m}\cdot 0.25} = 97.5 \quad \Rightarrow \quad R_{3m} = 10.127\%.$$

$$100e^{-R_{6m}\cdot 0.5} = 94.9 \quad \Rightarrow \quad R_{6m} = 10.469\%.$$

$$100e^{-R_{1y}\cdot 1} = 90 \quad \Rightarrow \quad R_{1y} = 10.536\%.$$

$$4e^{-0.1046\cdot 0.5} + 4e^{-0.10536\cdot 1} + 104e^{-R_{1.5y}\cdot 1.5} = 96 \quad \Rightarrow \quad R_{1.5y} = 10.681\%$$

e aggiungendo ulteriori contratti, possiamo aumentare il numero di punti.
Anche in questo caso però è necessaria una correzione per ottenere il tasso risk-free in quanto questi tassi impliciti hanno una componente di rischio di credito.

2.2 Credit Risk Adjustment (CRA)

Questo aggiustamento va applicato al caso degli swap e dei government bond, l'aggiustamento si applica al fair swap rate nel caso degli swap, alla yield curve estratta nella sezione precedente nel caso dei bond. Illustrerò il metodo utilizzato da EIOPA in *RFR Technical Documentation*[1] per il calcolo del risk-free.
Come prescritto dall'articolo 45 del regolamento delegato UE del 2015: *L'aggiustamento per il rischio di credito di cui all'articolo 44, paragrafo 1, è determinato in modo trasparente, prudente, attendibile, obiettivo e coerente nel tempo. L'aggiustamento è determinato sulla base della differenza tra i tassi che tengono conto del rischio di credito riflesso nel tasso variabile dei contratti swap su tassi di interesse e i tassi su contratti swap indicizzati* overnight *aventi la stessa scadenza, qualora entrambi i tassi siano disponibili da mercati finanziari DLT. Il calcolo dell'aggiustamento è basato sul 50% della media di tale differenza in un arco temporale di un anno. L'aggiustamento non è inferiore a 10 punti base e non è superiore a 35 punti base.* L'individuazione del requisito DLT riguarda i mercati dove vengono scambiati questi strumenti che devono essere "deep, liquid, trasparent"; questa valutazione è fatta da EIOPA secondo alcune metriche presentate nel documento tecnico EIOPA 2024 , anche se rimane una parte di valutazione qualitativa. EIOPA propone tre situazioni.

Situazione 1

Per la valuta di riferimento, oltre agli swap, sono presenti degli OIS con gambe fisse con la stessa tenor della gamba variabile dello swap, inoltre EIOPA richiede che entrambe le classi di strumenti rispettino i requisiti DLT. Dopodiché si calcola lo spread tra fair strike rate dello swap e dell'OIS per ogni giorno di trading appartenente all'ultimo anno e per ogni scadenza presente nella curva

[1] EIOPA, Risk-Free Interest Rate Term Structures, Technical Documentation, 2024.

discreta costruita come descritto nella sezione precedente. Calcolati tutti gli spread se ne prende la media semplice complessiva e la si moltiplica per 0.5. Questo valore è ciò che va tolto ai fair strike rate degli swap, o ai tassi impliciti dei bond che abbiamo selezionato come base per la costruzione della curva. EIOPA afferma inoltre che se qualche quotazione nei giorni di trading dell'anno precedente dovesse mancare, allora si possono sostituire i dati mancanti con una regressione lineare sui prezzi. Qualora più del 20% dei dati di una delle due categorie di contratto dovesse mancare, allora si passerebbe all'utilizzo della metodologia per la situazione 3.

Situazione 2

Se non è possibile calcolare un CRA dedicato secondo la situazione 1 per alcune valute europee (es. NOK, CHF), si applica il valore di CRA dell'euro. Due casi specifici sono previsti per la corona norvegese e il franco svizzero.

Situazione 3

Nella situazione 3 rientrano tutte le valute per cui l'aggiustamento della situazione uno non è percorribile perché sostanzialmente mancano i requisiti DLT sui contratti necessari per il calcolo. Per ogni valuta in questa categoria viene calcolato un fattore di scala:

$$SF_x = \frac{\sum_{t \in JDLT} \left(r_t^x - \theta^{Euro} \right)}{\sum_{t \in JDLT} \left(r_t^{Euro} - \theta^{Euro} \right)}.$$

con:

$$\theta^{Euro} = min \left[0, r_t^{Euro} | r_t^{Euro} < 0 \right].$$

- $JDLT$ è l'insieme delle scadenze per cui c'è un titolo che rispetta il requisito DLT
- r_t^x sono i tassi calcolati nella sezione precedente per la currency x.
- r_t^{Euro} sono i tassi calcolati nella sezione precedente per la valuta euro.
- θ^{Euro} è il tasso input della zona euro più negativo, se non è presente nessun input negativo allora si mette zero.

L'idea è creare una sorta di misura di scostamento relativo rispetto a un punto comune (θ^{Euro}), se ad esempio, la valuta estera si discosta in media più della valuta euro da questo tasso allora il suo

aggiustamento sarà maggiore di 1. A questo punto si calcola lo scaling factor al percentile 25% tra tutti gli scaling factor delle valute nella terza situazione. Siccome difficilmente si avrà un quantile esatto, si utilizza un'interpolazione lineare sugli SF per estrarre il valore corrispondente al 25%. Infine si calcola:

$$CRA_x = \frac{SF^x}{SF^{25\%}} CRA^{25\%}$$

con $CRA^{25\%}$ fissato a 35bps[2].

In ogni caso, il CRA finale viene troncato a 10 o 35 bps se fuori da questo intervallo.

2.3 Il modello Nelson-Siegel-Svensson (1994)

Si discute ora uno dei metodi per la costruzione di una curva smooth che interpoli i punti discreti ottenuti dalle due sezioni precedenti. Il Modello di Nelson-Siegel-Svensson è una correzione del modello di Nelson e Siegel 1987. L'idea di partenza è molto semplice; trovare una funzione che sia in grado di assumere le forme tipiche delle yield curve, in particolare della forward rate curve istantanea, ossia monotona (crescente o decrescente), con una gobba o a forma di S. Il punto di partenza proposto nel paper è la scelta della soluzione di un'equazione differenziale del secondo ordine:

$$ay''(t) + by'(t) + cy(t) = 0.$$

scegliendo $y(t) = e^{rt}$ abbiamo che $y'(t) = re^{rt}$ e $y''(t) = r^2 e^{rt}$ da cui

$$ar^2 e^{rt} + bre^{rt} + ce^{rt} = 0.$$

$$e^{rt}\left(ar^2 + br + c\right) = 0.$$

siccome $e^{rt} \neq 0 \, \forall t \in \Re$ le soluzioni sono quelle per cui $\left(ar^2 + br + c\right) = 0$ risolvendo questa equazione caratteristica posso trovare al più due radici distinte (r_1, r_2) da cui la soluzione come combinazione:

$$y(t) = C_1 e^{r_1 t} + C_2 e^{r_2 t}.$$

[2] 1 bps (punto base) = $\frac{1}{10000}$

se indichiamo con $r(m)$ il forward rate istantaneo osservato oggi per il tempo m, Nelson e Siegel propongono la seguente soluzione:

$$r(m) = \beta_0 + \beta_1 e^{-\frac{m}{\tau_1}} + \beta_2 e^{-\frac{m}{\tau_2}}$$

con $r_1 = -1/\tau_1$, $r_2 = -1/\tau_2$ radici negative per non far andare a infinito il tasso quando m tende a infinito e β_0 costante di "lungo termine" in quanto se al tendere di m a infinito il forward rate istantaneo andrebbe a zero altrimenti. Nelson e Siegel osservano poi che in questa maniera il modello risulta sovraparametrizzato, il che genera problemi al momento del fitting, dunque per semplificare il modello scelgono la soluzione con radici negative identiche $(b^2 - 4ac = 0)$. La soluzione cambia leggermente; diventando:

$$r(m) = \beta_0 + \beta_1 e^{-\frac{m}{\tau}} + \beta_2 \left[\frac{m}{\tau} e^{-\frac{m}{\tau}}\right].$$

Integrando si ottiene l'equazione per la curva:

$$R(t) = \frac{1}{t}\int_0^t r(m)\mathrm{d}m = \beta_0 + (\beta_1 + \beta_2)\frac{\left[1 - e^{-\frac{t}{\tau}}\right]}{\frac{t}{\tau}} - \beta_2 e^{-\frac{t}{\tau}}.$$

In questo modello β_0 rappresenta la componente di lungo termine, β_1 è legata al breve termine e β_2 al medio termine. Svensson estende questo modello in Svensson 1994 introducendo un ulteriore termine di curvatura per migliorare il fitting:

$$r(m) = \beta_0 + \beta_1 e^{\frac{m}{\tau_1}} + \beta_2 \left[\frac{m}{\tau_1} e^{\frac{m}{\tau_1}}\right] + \beta_3 \left[\frac{m}{\tau_2} e^{\frac{m}{\tau_2}}\right].$$

Sempre integrando:

$$R(t) = \beta_0 + \beta_1 \frac{1 - e^{-\frac{t}{\tau_1}}}{\frac{t}{\tau_1}} + \beta_2 \left(\frac{1 - e^{-\frac{t}{\tau_1}}}{\frac{t}{\tau_1}} - e^{-\frac{t}{\tau_1}}\right) + \beta_3 \left(\frac{1 - e^{-\frac{t}{\tau_2}}}{\frac{t}{\tau_2}} - e^{-\frac{t}{\tau_2}}\right).$$

Per la stima dei parametri, Svensson propone di utilizzare la minimizzazione della somma degli scarti al quadrato tra yield osservati (ciò che abbiamo ricavato nelle sezioni precedenti) e il valore teorico $R(t)$ per t per cui abbiamo estratto lo yield discreto. Questo è equivalente a risolvere il seguente problema di massima verosimiglianza: Definiamo l'errore di stima per ogni osservazione come:

$$\varepsilon_i = R_i^{\text{oss}} - R_i^{\text{modello}}(\theta)$$

con $\theta = (\beta_0, \beta_1, \beta_2, \beta_3, \tau_1, \tau_2)$ Si assume che gli errori siano variabili casuali indipendenti e distribuiti secondo una normale standardizzata:

$$\varepsilon_i \sim \mathcal{N}(0, \sigma^2).$$

La funzione di densità associata a ciascun errore è quindi:

$$f(\varepsilon_i) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{\left(-\frac{1}{2\sigma^2}(R_i^{\text{oss}} - R_i^{\text{modello}}(\theta))^2\right)}.$$

Sotto l'ipotesi di indipendenza degli errori, la funzione di verosimiglianza totale diventa:

$$L(\theta, \sigma^2) = \prod_{i=1}^{n} f(\varepsilon_i)$$

ovvero:

$$L(\theta, \sigma^2) = \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n e^{\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(R_i^{\text{oss}} - R_i^{\text{modello}}(\theta))^2\right)}.$$

Applicando il logaritmo si ottiene la funzione di log-verosimiglianza:

$$l(\theta, \sigma^2) = -\frac{n}{2}\log(2\pi\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(R_i^{\text{oss}} - R_i^{\text{modello}}(\theta))^2.$$

Poiché il primo termine dipende solo da σ^2 e non da θ, massimizzare la log-verosimiglianza rispetto a θ equivale a minimizzare:

$$\sum_{i=1}^{n}(R_i^{\text{oss}} - R_i^{\text{modello}}(\theta))^2.$$

In breve, l'approccio adottato si basa sulla massimizzazione della verosimiglianza sotto ipotesi gaussiane degli errori, il che implica che la stima dei parametri θ è equivalente alla minimizzazione della somma dei quadrati delle deviazioni tra i tassi osservati e quelli modellizzati. Inoltre, l'utilizzo della massima verosimiglianza in una situazione di modello regolare determina una serie di proprietà vantaggiose dello stimatore $\hat{\theta}$ sotto l'ipotesi che l'assunzione sulla distribuzione degli errori sia coerente con i dati.

2.4 Il modello Smith-Wilson (2000)

E' necessario introdurre due dei suoi input che vanno predeterminati: l'*Ultimate Forward rate* (UFR) e il *Last Liquid Point* (LLP). *L'Ultimate Forward Rate* rappresenta il tasso a cui il modello convergerà dopo aver superato il *last liquid point*; dovrebbe rappresentare il tasso atteso di lungo periodo. EIOPA determina questo tasso a partire dai suoi driver principali, ossia il tasso di inflazione atteso per la specifica valuta e il tasso di interesse reale atteso, comune a tutte le valute.
Il tasso di interesse reale atteso è calcolato come:

$$\frac{1}{n}\sum_{i=1}^{n} r_{1960+i}$$

con r_i tasso di interesse reale calcolato come media dei tassi reali di Belgio, Germania, Francia, Italia, Olanda, Regno Unito, Stati Uniti. Per ognuno di questi paesi

$$r_i = \frac{\text{tasso nominale a breve termine-tasso di inflazione}}{1+\text{tasso di inflazione}}.$$

Il tasso di inflazione atteso dipende invece dagli obiettivi della banca centrale di riferimento per la valuta e risulta essere pari a:

- 1% se il target è minore o uguale a 1%
- 2% se il target è maggiore di 1% e minore di 3%
- 3% se il target è maggiore o uguale a 3% e minore di 4%
- 4% se il target è maggiore o uguale a 4%.

Per le banche centrali che non hanno annunciato un target ma si impegnano a tenere l'inflazione limitata tra bande si prende il punto medio delle bande. Per le banche centrali che non hanno annunciato nessuna forma di controllo si assume il 2%. L'Ultimate Forward Rate è calcolato come somma del tasso reale e del tasso di inflazione atteso. Nel tempo l'Ultimate Forward Rate è aggiornato solo se il cambiamento rispetto all'osservazione dell'anno precedente è superiore di 15 bps.
Il Last Liquid Point non è altro che la scadenza t per cui si ha l'ultimo tasso implicito nella yield

curve.

Un altro termine che verrà impiegato è denominato *Convergence Point* (CP) ossia il $\max(LLP + 40,60)$: oltre questo punto la curva deve convergere all'UFR. L'*intensity* α rappresenta la "velocità" a cui la curva converge all'UFR.

Il modello Smith-Wilson può essere interpretato come la ricerca di una funzione del prezzo dello zero coupon bond, fattore di sconto $p(0,t)$ in funzione della maturity t. Per la trattazione del modello verranno analizzate due prospettive, una puramente deterministica, che vede la ricerca della funzione $p(0,t)$ come soluzione di uno specifico problema di ottimizzazione, l'altra, stocastica, vede la determinazione della funzione $p(0,t)$ come calcolo di un valore atteso condizionato all'informazione osservabile sul mercato di uno specifico processo stocastico. Le due versioni sono utili nella ricerca dei punti problematici (e quindi soluzioni) del modello ed inoltre si complementano nel senso che la versione deterministica garantisce che i risultati trovati con la dimostrazione con componenti stocastiche abbia delle proprietà di ottimo lagate ad uno specifico problema di ottimizzazione.

Partendo dal problema deterministico, l'obiettivo è quello di determinare la funzione $p(0,t)$ in modo che sia regolare e che abbia un tasso di variazione contenuto. Per costruire tale funzione si parte dalla seguente parametrizzazione:

$$p(0,t) := (1 + g(t))\, e^{-w \cdot t}$$

dove:

- $w = \ln(1 + UFR)$ deriva dalla soluzione di $e^{-w \cdot 1} = (1 + UFR)^{-1}$ in w
- $g(t)$ è una funzione da determinare che in questo contesto rappresenta un aggiustamento al fattore di sconto determinato dall'UFR.

Per far sì che la funzione le fattore di sconto sia sufficientemente regolare la funzione $g(t)$ viene ricercata partendo da questo probelema di ottimizzazione:

$$\min_g L_\alpha[g] = \min_g \int_0^\infty g''(s)^2 + \alpha^2 g'(s)^2 \mathrm{d}s \tag{2}$$

con $\alpha > 0$ parametro che gestisce potenziali conflitti tra le due componenti nella ricerca del minimo, seguirà successivamente un'interpretazione. Con questi vincoli si sta chiedendo che la funzione di interpolazione dei fattori di sconto abbia derivata prima e seconda contenute. Come in precedenza, il punto di partenza deve essere determinato dall'informazione presente sul mercato. Pertanto, ipotizzando di avere un numero I di titoli, con insieme contenente tutte le scadenze diverse per cui i titoli quotati hanno flussi di cassa J possiamo definire il prezzo di un generico titolo quotato nel modo seguente:

$$p_i = \sum_{j \in J} c_{ij} p(0,t) = \sum_{j \in J} c_{ij} \left(1 + g(u_j)\right) e^{-w \cdot u_j} \quad \forall i = 1,..,I$$

- con p_i prezzo di mercato del titolo i-esimo
- u_j scadenza j-esima per cui si hanno cashflows dei titoli
- c_{ij} flusso di cassa del titolo i alla scadenza j-esima.

Le I funzioni vengono utilizzate come vincoli per la ricerca del punto di ottimo del problema (2). EIOPA individua la seguente funzione: $W(t,u) = \alpha \min(t,u) - e^{-\alpha \max(u,t)} \sinh(\alpha \min(t,u))$[3] la cui interpretazione è lasciata alla parte successiva. In ogni caso si può dimostrare che $g = W$ è componente della soluzione del problema (2) esposto in precedenza Kort e Vellekoop 2016. La soluzione va ricercata tra le funzioni appartenenti al seguente insieme:

$$\Omega = \{g \in C^2(\Re^+) : g(0) = a, g''(0) = 0, g' \in L^2(\Re^+), g'' \in L^2(\Re^+), \lim_{t \to \infty} g'(t) = 0, \lim_{t \to \infty} g''(t) = 0\}.$$

con $a \in \Re$.

La dimostrazione è svolta scrivendo la seguente lagrangiana con moltiplicatori scritti in questo modo: $\lambda_i = \alpha^3 \varepsilon_i$ allo scopo di semplificare la soluzione nei passaggi successivi. La soluzione del problema è ottenuta trovando la funzione g che minimizza la seguente lagrangiana:

$$\tilde{L}_\alpha[g] = \int_0^\infty g''(s)^2 + \alpha^2 g'(s)^2 \mathrm{d}s + \alpha^3 \sum_{i=1}^{I} \varepsilon_i \left[p_i - \sum_{j \in J} c_{ij} \left(1 + g(u_j)\right) e^{-w \cdot u_j} \right].$$

[3] $\sinh(x) = \frac{e^x - e^{-x}}{2}$

Se la soluzione del problema esposto esiste deve essere della forma:

$$g(t) = \sum_{i=1}^{I} \varepsilon_i \sum_{j \in J} c_{ij} e^{-w \cdot u_j} W(t, u_j)$$

da cui una funzione di sconto potrebbe essere la seguente

$$p(0,t) = e^{-w \cdot t} \left(1 + \sum_{i=1}^{I} \varepsilon_i \sum_{j \in J} c_{ij} e^{-w \cdot u_j} W(t, u_j) \right).$$

Le componenti che rimangono da specificare sono ε e α. Per ora si ipotizza di conoscere α la cui calibrazione verrà trattata successivamente e si procede con la calibrazione degli ε_i dai prezzi di mercato. Partiamo dalla specificazione del prezzo in questo framework

$$p_i = \sum_{k \in J} c_{ik} e^{-w \cdot u_k} \left(1 + \sum_{r=1}^{I} \varepsilon_r \sum_{j \in J} c_{rj} e^{-w \cdot u_j} W(u_k, u_j) \right) \ \forall i = 1, \ldots, I$$

da cui si ha che

$$p_i - \sum_{k \in J} c_{ik} e^{-w \cdot u_k} = \sum_{k \in J} \sum_{r=1}^{I} \sum_{j \in J} c_{ik} e^{-w \cdot u_k} W(u_k, u_j) c_{rj} e^{-w \cdot u_j} \varepsilon_r \ \forall i = 1, \ldots, I.$$

Definendo gli elementi dei vettori $\bar{\mathbf{P}}$ con $\bar{\mathbf{P}}_i := p_i - \sum_{k \in J} c_{ik} e^{-w \cdot u_k}$ e $\mathbf{Z}$ con $\mathbf{Z}_i := \varepsilon_i$; inoltre consideriamo le matrici $\mathbf{W}$ con elemento generico $\mathbf{W_{k,j}} = W(u_k, u_j)$ e $\mathbf{C}$ con elemento generico $\mathbf{C}_{ij} = c_{ij} e^{-w \cdot u_j}$ si può esprimere il sistema lineare degli I strumenti finanziari in forma matriciale:

$$\bar{\mathbf{P}} = \mathbf{CWC}^T \mathbf{Z}$$

e invertendo trovo gli ε_i elementi del vettore Z come segue:

$$\mathbf{Z} = (\mathbf{CWC}^T)^{-1} \bar{\mathbf{P}}.$$

Per semplificare la trattazione della componente stocastica del modello verrà ipotizzato che gli unici strumenti utilizzati per la calibrazione siano degli zero coupon bond, in questo modo la formula della curva diventa:

$$p(0,t) = e^{-w \cdot t} \left(1 + \sum_{i=1}^{I} \varepsilon_i e^{-w \cdot u_i} W(t, u_i) \right) = e^{-w \cdot t} + \sum_{i=1}^{I} \varepsilon_i e^{-w \cdot (t + u_i)} W(t, u_i). \tag{3}$$

Inoltre $\mathbf{C} = diag(e^{-w\cdot\mathbf{u}})$, con $\mathbf{u} := [u_1, \ldots, u_{\bar{J}}]^T$ dove $\bar{J}$ coincide con il numero degli elementi dell'insieme J, $\bar{\mathbf{P}}_\mathbf{i} = p_i - e^{-w\cdot u_i}$, da cui $\mathbf{Z} = \left(diag(e^{-w\cdot\mathbf{u}})\mathbf{W}diag(e^{-w\cdot\mathbf{u}})^T\right)^{-1}\bar{\mathbf{P}}$. E' possibile dimostrare che questa soluzione è collegata a uno specifico processo stocastico considerato in Lagerås e Lindholm 2016. In particolare partendo da un processo di Ornstein–Uhlenbeck definito dalla seguente equazione differenziale stocastica:

$$\mathrm{d}X_t = -\alpha X_t \mathrm{d}t + \alpha^{3/2}\mathrm{d}B_t$$

con B_t processo di Wiener, $\alpha > 0$ è il parametro di ritorno alla media; si assume inoltre che $X_0 \sim \mathcal{N}(0, \alpha^2) \perp\!\!\!\perp B_t \;\; \forall t > 0$. Definendo $\bar{X}_t = \int_0^t X_s \mathrm{d}s$ e $Y_t = e^{-w\cdot t}\left(1 + \bar{X}_t\right)$ è possibile dimostrare che questo processo stocastico è il processo che guida come i diversi fattori di sconto si manifestano per le diverse scadenze. In particolare conoscendo un insieme discreto di fattori di sconto è possibile determinare il valore atteso dei fattori di sconto per le scadenze non note:

$$\mathbb{E}^{\mathbb{P}}\left[Y_t | Y_{u_1} = p(0,u_1), \ldots, Y_{u_{\bar{J}}} = p(0,u_{\bar{J}})\right]$$

con $p(0,u_i)$ prezzi degli ZCB noti usati nella calibrazione. Si vuole dimostrare il seguente risultato

$$\mathbb{E}^{\mathbb{P}}\left[Y_t | Y_{u_1} = p(0,u_1), \ldots, Y_{u_{\bar{J}}} = p(0,u_{\bar{J}})\right] = p(0,t)$$

e che quindi il problema deterministico affrontato in (2) ha una sua esatta contropartita nel valore atteso condizionato del processo Y_t. Per la dimostrazione è necessario risolvere l'equazione differenziale stocastica $\mathrm{d}X_t$; sfruttando il lemma di Ito e la seguente trasformazione $f(t, X_t) = X_t e^{\alpha t}$

$$\begin{aligned}
\mathrm{d}\left[X_t e^{\alpha t}\right] &= X_t e^{\alpha t}\alpha \mathrm{d}t + e^{\alpha t}\mathrm{d}X_t \\
&= X_t e^{\alpha t}\alpha \mathrm{d}t + e^{\alpha t}\left(-\alpha X_t \mathrm{d}t + \alpha^{3/2}\mathrm{d}B_t\right) \\
&= \alpha^{3/2} e^{\alpha t}\mathrm{d}B_t.
\end{aligned}$$

Integrando:

$$\int_0^m \mathrm{d}\left[X_t e^{\alpha t}\right] = \int_0^m \alpha^{3/2} e^{\alpha t}\mathrm{d}B_t$$

da cui si ha che

$$X_m e^{\alpha m} - X_0 = \alpha^{3/2}\int_0^m e^{\alpha t}\mathrm{d}B_t$$

e quindi

$$X_m = X_0 e^{-\alpha m} + \alpha^{3/2} \int_0^m e^{-\alpha(m-t)} \mathrm{d}B_t.$$

da cui si deduce che

$$\mathbb{E}^{\mathbb{P}}[X_m] = \mathbb{E}^{\mathbb{P}}[X_0] e^{-\alpha m} + \alpha^{3/2} \int_0^m e^{-\alpha(m-t)} \mathbb{E}^{\mathbb{P}}[\mathrm{d}B_t] = 0.$$

Sfruttando l'indipendenza tra X_0 e B_t insieme alla proprietà di incrementi indipendenti del processo Wiener

$$\begin{aligned} V[X_m] &= V[X_0] e^{-2\alpha m} + \alpha^3 \int_0^m e^{-2\alpha(m-t)} V[\mathrm{d}B_t] \\ &= \alpha^2 e^{-2\alpha m} + \alpha^3 \int_0^m e^{-2\alpha(m-t)} \mathrm{d}t \\ &= \alpha^2 e^{-2\alpha m} + \frac{\alpha^2}{2}\left[1 - e^{-2\alpha m}\right] \\ &= \alpha^2 \left[e^{-2\alpha m} + \frac{1}{2} - \frac{e^{-2\alpha m}}{2}\right] \\ &= \alpha^2 \left[\frac{e^{-2\alpha m} + 1}{2}\right]. \end{aligned}$$

La distribuzione di X_m risulta quindi essere

$$X_m \sim \mathcal{N}(0, \alpha^2 \left[\frac{e^{-2\alpha m} + 1}{2}\right]) \tag{4}$$

inoltre ipotizzando $t > s$

$$Cov[X_s, X_t] = Cov\left[X_0 e^{-\alpha s} + \alpha^{3/2} \int_0^s e^{-\alpha(s-u)} \mathrm{d}B_u, X_0 e^{-\alpha t} + \alpha^{3/2} \int_0^t e^{-\alpha(t-u)} \mathrm{d}B_u\right].$$

Per la proprietà della covarianza di somme e l'indipendenza tra X_0 e B_t si ottiene

$$\begin{aligned} Cov[X_s, X_t] &= Cov[X_0, X_0] e^{-\alpha(t+s)} + Cov\left[\alpha^{3/2} \int_0^s e^{-\alpha(s-u)} \mathrm{d}B_u, \alpha^{3/2} \int_0^t e^{-\alpha(t-u)} \mathrm{d}B_u\right] \\ &= V[X_0] e^{-\alpha(t+s)} + \alpha^3 e^{-\alpha(t+s)} Cov\left[\int_0^s e^{\alpha u} \mathrm{d}B_u, \int_0^s e^{\alpha u} \mathrm{d}B_u + \int_s^t e^{\alpha u} \mathrm{d}B_u\right]. \end{aligned}$$

Sempre sfruttando l'indipendenza degli incrementi del processo di Wiener otteniamo il seguente risultato

$$\begin{aligned} Cov[X_s, X_t] &= \alpha^2 e^{-\alpha(t+s)} + \alpha^3 e^{-\alpha(t+s)} Cov\left[\int_0^s e^{\alpha u} \mathrm{d}B_u, \int_0^s e^{\alpha u} \mathrm{d}B_u\right] \\ &= \alpha^2 e^{-\alpha(t+s)} + \alpha^3 e^{-\alpha(t+s)} \int_0^s e^{2\alpha u} \mathrm{d}u \\ &= \alpha^2 e^{-\alpha t} \cosh(\alpha s) \quad se \quad t > s \end{aligned} \tag{5}$$

o più genericamente

$$Cov\left[X_s, X_t\right] = \alpha^2 e^{-\alpha \max(t,s)} \cosh(\alpha \min(t,s))^{4}.$$

Altra componente d'interesse è la covarianza tra due generici processi integrati $\bar{X}_s$,$\bar{X}_t$ sempre assumendo $t > s$ e sfruttando le proprietà della covarianza di somme si ha assumendo $0 \leq u \leq s$ e $0 \leq v \leq t$

$$Cov\left[\bar{X}_s, \bar{X}_t\right] = \int_0^t \int_0^s Cov\left[X_u, X_v\right] \mathrm{d}u\mathrm{d}v.$$

Utilizzando la proprietà di simmetria della covarianza e osservando che quando $0 \leq u \leq v \leq s$ o $0 \leq v \leq u \leq s$ gli integrali coincidono e sfruttando il teorema di Fubini, l'integrale doppio può essere riscritto nel seguente modo:

$$\begin{aligned} Cov\left[\bar{X}_s, \bar{X}_t\right] &= \int_0^s \int_0^v Cov\left[X_u, X_v\right] \mathrm{d}u\mathrm{d}v + \int_0^s \int_0^u Cov\left[X_u, X_v\right] \mathrm{d}v\mathrm{d}u + \int_s^t \int_0^s Cov\left[X_u, X_v\right] \mathrm{d}u\mathrm{d}v \\ &= 2\int_0^s \int_0^v Cov\left[X_u, X_v\right] \mathrm{d}u\mathrm{d}v + \int_s^t \int_0^s Cov\left[X_u, X_v\right] \mathrm{d}u\mathrm{d}v. \end{aligned}$$

Prendendo due volte il primo integrale per cui è valido che $0 \leq u \leq v \leq s$, è possibile sostituire (5) nella situazione $v > u$. Osserviamo inoltre che anche nel secondo integrale si ha che $0 \leq u \leq s \leq$

[4] $\cosh(x) = \frac{e^x + e^{-x}}{2}$

$v \leq t$ dunque anche in questo integrale si può sostituire il risultato per $v > u$ da cui

$$
\begin{aligned}
Cov\left[\bar{X}_s, \bar{X}_t\right] &= 2\int_0^s \int_0^v Cov\left[X_u, X_v\right] \mathrm{d}u\mathrm{d}v + \int_s^t \int_0^s Cov\left[X_u, X_v\right] \mathrm{d}u\mathrm{d}v \\
&= 2\int_0^s \int_0^v \alpha^2 e^{-\alpha v} \cosh(\alpha u) \mathrm{d}u\mathrm{d}v + \int_s^t \int_0^s \alpha^2 e^{-\alpha v} \cosh(\alpha u) \mathrm{d}u\mathrm{d}v \\
&= 2\int_0^s \alpha^2 e^{-\alpha v} \int_0^v \cosh(\alpha u) \mathrm{d}u\mathrm{d}v + \int_s^t \alpha^2 e^{-\alpha v} \mathrm{d}v \int_0^s \cosh(\alpha u) \mathrm{d}u \\
&= 2\int_0^s \alpha^2 e^{-\alpha v} \left[\frac{\sinh(\alpha u)}{\alpha}\right]_{u=0}^{v} \mathrm{d}v + \alpha \left[-\frac{1}{\alpha} e^{-\alpha v}\right]_{v=s}^{t} \left[\frac{\sinh(\alpha u)}{\alpha}\right]_{u=0}^{s} \\
&= 2\int_0^s \alpha e^{-\alpha v} \sinh(\alpha v) \mathrm{d}v - \left(e^{-\alpha t} - e^{-\alpha s}\right) \sinh(\alpha s) \\
&= 2\alpha \int_0^s e^{-\alpha v} \left[\frac{e^{\alpha v} - e^{-\alpha v}}{2}\right] \mathrm{d}v + \left(e^{-\alpha s} - e^{-\alpha t}\right) \sinh(\alpha s) \\
&= \alpha \int_0^s 1 - e^{-2\alpha v} \mathrm{d}v + \left(e^{-\alpha s} - e^{-\alpha t}\right) \sinh(\alpha s) \\
&= \alpha s - \alpha \int_0^s e^{-2\alpha v} \mathrm{d}v + \left(e^{-\alpha s} - e^{-\alpha t}\right) \frac{e^{\alpha s} - e^{-\alpha s}}{2} \\
&= \alpha s + \alpha \left[\frac{1}{2\alpha} e^{-2\alpha v}\right]_{v=0}^{s} + \frac{1 - e^{-\alpha s}}{2} - e^{-\alpha t} \sinh(\alpha s) \\
&= \alpha s + \left[\frac{e^{-2\alpha s}}{2} - \frac{1}{2}\right] + \frac{1 - e^{-\alpha s}}{2} - e^{-\alpha t} \sinh(\alpha s) \\
&= \alpha s - \frac{1 - e^{-\alpha s}}{2} + \frac{1 - e^{-\alpha s}}{2} - e^{-\alpha t} \sinh(\alpha s) \\
&= \alpha s - e^{-\alpha t} \sinh(\alpha s) \quad t > s.
\end{aligned}
$$

Più in generale si ha che

$$
Cov\left[\bar{X}_s, \bar{X}_t\right] = \alpha \min(s,t) - e^{-\alpha \max(s,t)} \sinh(\alpha \min(s,t)). \tag{6}
$$

Si può notare che (6) è esattamente la funzione W utilizzata in precedenza. Si ha inoltre che

$$
\mathbb{E}^{\mathbb{P}}\left[\bar{X}_t\right] = \int_0^t \mathbb{E}^{\mathbb{P}}\left[X_u\right] \mathrm{d}u = 0.
$$

A questo punto gli ultimi strumenti di interesse per la trattazione sono i seguenti

$$
\mathbb{E}^{\mathbb{P}}\left[Y_t\right] = e^{-w \cdot t}
$$

e

$$
\begin{aligned}
Cov\left[Y_s, Y_t\right] &= e^{-w(s+t)} Cov\left[\bar{X}_s, \bar{X}_t\right] \\
&= e^{-w(s+t)} \left[\alpha \min(s,t) - e^{-\alpha \max(s,t)} \sinh(\alpha \min(s,t))\right] \\
&= e^{-w(s+t)} W(s,t).
\end{aligned}
$$

Osservando che Y_t è una trasformazione lineare di $\bar{X}_t$ che a sua volta è trasformazione lineare di X_s che anch'esso è lineare in $\mathrm{d}B_t$ ossia un processo gaussiano, possiamo affermare che anche Y_t è un processo gaussiano, il che implica che il generico vettore $\mathbf{Y} = (Y_1, ..., Y_p) \sim \mathscr{N}_p,\ \forall p \in \mathbb{N}$ questo implica che

$$
\mathbb{E}^{\mathbb{P}}\left[Y_t | Y_{u_1} = p(0,u_1), \ldots, Y_{u_{\bar{J}}} = p(0,u_{\bar{J}})\right] = \mathbb{E}^{\mathbb{P}}\left[Y_t\right] + Cov\left[Y_t, \mathbf{Y}\right] Cov\left[\mathbf{Y}, \mathbf{Y}\right]^{-1} \left[p(0,\mathbf{u}) - \mathbb{E}^{\mathbb{P}}\left[\mathbf{Y}\right]\right].
$$

Con $\mathbf{Y}$ vettore contenente le variabili casuali per le scadenze $\mathbf{u}$ di cui si conosce la realizzazione. Sostituendo i risultati ricavati in precedenza si ottiene:

$$
\begin{aligned}
\mathbb{E}^{\mathbb{P}}\left[Y_t | Y_{u_1} = p(0,u_1), \ldots, Y_{u_{\bar{J}}} = p(0,u_{\bar{J}})\right] &= e^{-w\cdot t} + Cov\left[Y_t, \mathbf{Y}\right] Cov\left[\mathbf{Y}, \mathbf{Y}\right]^{-1} \left[p(0,\mathbf{u}) - e^{-w\cdot\mathbf{u}}\right] \\
&= e^{-w\cdot t} + Cov\left[Y_t, \mathbf{Y}\right] Cov\left[\mathbf{Y}, \mathbf{Y}\right]^{-1} \bar{\mathbf{P}} \\
&= e^{-w\cdot t} + Cov\left[Y_t, \mathbf{Y}\right] \left(diag(e^{-w\cdot\mathbf{u}}) \mathbf{W} diag(e^{-w\cdot\mathbf{u}})^T\right)^{-1} \bar{\mathbf{P}} \\
&= e^{-w\cdot t} + Cov\left[Y_t, \mathbf{Y}\right] \mathbf{Z} \\
&= e^{-w\cdot t} + \sum_{i=1}^{I} \varepsilon_i e^{-w\cdot(t+u_i)} W(t,u_i) \\
&= p(0,t).
\end{aligned}
$$

Come volevasi dimostrare, questo risultato è equivalente a quanto ottenuto risolvendo il problema (2). Vedere questo modello come un processo stocastico ci permette di fare analisi legate all'hedging. In particolare dall'equazione precedente, chiamando $\mathbf{D} := p(0,\mathbf{u})$ il vettore contenente i fattori di sconto discreti osservati, è possibile riscrivere:

$$
\begin{aligned}
\mathbb{E}^{\mathbb{P}}\left[Y_t | \mathbf{Y} = \mathbf{D}\right] &= \mathbb{E}^{\mathbb{P}}\left[Y_t\right] + Cov\left[Y_t, \mathbf{Y}\right] Cov\left[\mathbf{Y}, \mathbf{Y}\right]^{-1} \left[\mathbf{D} - \mathbb{E}^{\mathbb{P}}\left[\mathbf{Y}\right]\right] \\
&= \mathbb{E}^{\mathbb{P}}\left[Y_t\right] - Cov\left[Y_t, \mathbf{Y}\right] Cov\left[\mathbf{Y}, \mathbf{Y}\right]^{-1} \mathbb{E}^{\mathbb{P}}\left[\mathbf{Y}\right] + Cov\left[Y_t, \mathbf{Y}\right] Cov\left[\mathbf{Y}, \mathbf{Y}\right]^{-1} \mathbf{D}.
\end{aligned}
$$

Possiamo quindi scrivere

$$p(0,t) = \beta_0(t) + \beta^T(t)\mathbf{D} \tag{7}$$

con $\beta_0(t) := \mathbb{E}^{\mathbb{P}}[Y_t] - Cov[Y_t, \mathbf{Y}]\, Cov[\mathbf{Y}, \mathbf{Y}]^{-1}\, \mathbb{E}^{\mathbb{P}}[\mathbf{Y}]$ e vettore riga $\beta^T(t) := Cov[Y_t, \mathbf{Y}]\, Cov[\mathbf{Y}, \mathbf{Y}]^{-1}$.

2.4.1 Hedging nel modello Smith-Wilson

L'equazione (7) implica che nel modello di Smith-Wilson è possibile replicare un qualsiasi ZCB non quotato sul mercato attraverso l'utilizzo di una combinazione lineare degli ZCB quotati, in particolare è sufficiente detenere una quantità $\beta_0(t)$ liquida più delle quantità $\beta_i(t)$ degli strumenti quotati. Nel caso di portafoglio di passività composto da diversi flussi di cassa a diverse scadenze l'hedging può essere fatto nel seguente modo:

$$\begin{aligned} P^c &= \sum_t CF_t p(0,t) = \sum_t CF_t \left(\beta_0(t) + \beta(t)^T \mathbf{D}\right) \\ &= \sum_t CF_t \beta_0(t) + \sum_t CF_t \beta(t)^T \mathbf{D} \\ &= \bar{\beta}_0(t) + \bar{\beta}(t)^T \mathbf{D} \end{aligned}$$

con $\bar{\beta}_0(t) := \sum_t CF_t \beta_0(t)$ e $\bar{\beta}(t)^T := \sum_t CF_t \beta(t)^T$. I principali problemi del modello Smith-Wilson riguardano proprio le implicazioni di hedging. Il primo, meno rilevante, riguarda il fatto che per coprire ad esempio uno ZCB con scadenza t prima dell'LLP bisogni tendenzialmente prendere una posizione su tutti gli strumenti della curva; sarebbe preferibile invece, se la scadenza è intermedia alle scadenze di due strumenti utilizzati per la costruzione della curva, utilizzare soltanto questi per la copertura. Il problema più serio accade invece quando si parla di coperture di strumenti che hanno scadenza oltre il LLP. É infatti possibile dimostrare che la matrice $Cov[\mathbf{Y}, \mathbf{Y}]^{-1}$ implicata dallo specifico processo trasformato esposto in precedenza fa sì che per $p(0,t)$ con $t > LLP$ i segni dei $\beta_i(t)$ della regressione (7) avranno $sign(\beta_i(t)) = (-1)^{I-i}$ $i = 1, ..., I$ (Lagerås e Lindholm 2016) questo implica assumere sia posizioni lunghe che corte per replicare un singolo ZCB e questo a sua volta implica che la posizione di replicazione sarà, in valore assoluto, superiore alla posizione che si vuole coprire. Inoltre, col passare del tempo, questo tipo di oscillazione dei segni implica un turnover molto grande. Infatti, ipotizzando di aver usato per la calibrazione della curva 3 ZCB, il primo a scadenza 1 anno, il secondo a scadenza 2 anni, il terzo a scadenza 3 anni. La regola implica

che dopo un anno si rifarà la calibrazione, ma la posizione da assumere sulla prima scadenza sarà con segno positivo, quindi la posizione che si aveva assunto sullo ZCB a 2 (ora ZCB a 1 anno) anni diventa da negativa a positiva, mentre quella assunta sullo ZCB a 3 anni (ora ZCB a 2 anni) diventa negativa.

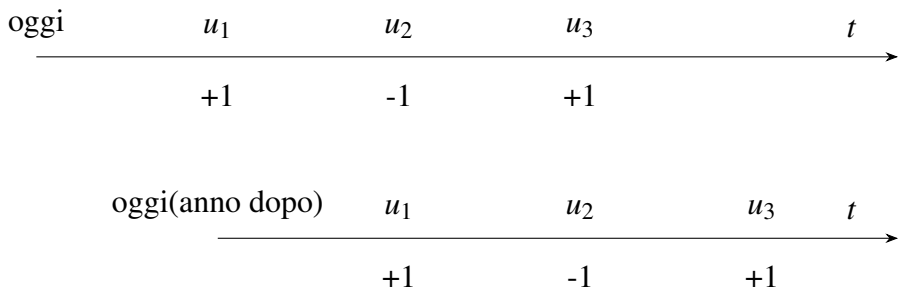

Figura 2.1: Esempio alternanza segni delle posizioni per hedging di uno ZCB con scadenza superiore al LLP.

Una potenziale soluzione a questo problema consiste nella selezione di un altro processo che abbia una matrice di covarianze costruita in modo tale che i coefficienti $\beta_i(t)$ abbiano tutti segno positivo. Per farlo è possibile sfruttare il risultato per cui i coefficienti di regressione β_i hanno lo stesso segno del coefficiente di covarianza parziale dei regressori (prezzi degli strumenti quotati $\mathbf{Y}$) sulla variabile Y_t dato da $Cov[Y_t, Y_{u_i}] - Cov[Y_t, \mathbf{Y}_{-u_i}] Cov[\mathbf{Y}_{-u_i}, \mathbf{Y}_{-u_i}]^{-1} Cov[\mathbf{Y}_{-u_i}, Y_t]$ con $\mathbf{Y}_{-u_i}$ $\mathbf{Y}$ senza la variabile al tempo u_i, è sufficiente scegliere un processo gaussiano la cui matrice varianze-covarianze sia tale da garantire che il segno di questo oggetto sia sempre positivo. Un tipo di matrice che ha questa proprietà è quella legata a un processo gaussiano con matrice varianze-covarianze la cui inversa sia una Matrice-M (Karlin e Rinott 1983). Una matrice A con $A_{ij} \leq 0 \, \forall i \neq j$ è una matrice-M se e solo se una delle seguenti condizioni è vera:

- Esiste un vettore $x \in \Re^n$ tale per cui $Ax > 0$.
- A è non singolare e tutti gli elementi di A^{-1} sono positivi.
- tutti i minori principali di A sono positivi.

Sotto queste condizioni è possibile dimostrare che se Σ è una matrice varianze covarianze di un processo gaussiano definita in questo modo:

$$\Sigma = \begin{bmatrix} \begin{bmatrix} V[Y_t] & Cov[Y_t, Y_{u_i}] \\ Cov[Y_t, Y_{u_i}] & V[Y_{u_i}] \end{bmatrix} & Cov[Y_t, \mathbf{Y}_{-u_i}] \\ Cov[Y_t, \mathbf{Y}_{-u_i}]^T & Cov[\mathbf{Y}_{-u_i}, \mathbf{Y}_{-u_i}] \end{bmatrix}$$

e la sua inversa è una matrice-M allora si ha che

$$Cov[Y_t, Y_{u_i}] - Cov[Y_t, \mathbf{Y}_{-u_i}] Cov[\mathbf{Y}_{-u_i}, \mathbf{Y}_{-u_i}]^{-1} Cov[\mathbf{Y}_{-u_i}, Y_t] \geq 0.$$ [5]

2.4.2 Calibrazione del parametro α

Fino ad ora si è considerato noto il valore dell'α. In pratica, per applicare il modello è necessario determinare questo valore. EIOPA propone di calibrare tale parametro a partire dal tasso forward istantaneo. Quest'ultimo è ricavato seguendo la procedura di seguito descritta.

Definendo $l(0,m)$ il tasso forward istantaneo osservato oggi per la scadenza che va da m a $m+\varepsilon$ si può identificare alla seguente identità con il fattore di sconto o prezzo di uno ZCB definito come segue:

$$p(0,t) = \exp\{-\int_0^t l(0,m)\mathrm{d}m\}$$

da cui:

$$\int_0^t l(0,m)\mathrm{d}m = -\ln(p(0,t)).$$

Derivando rispetto al tempo si ha che:

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_0^t l(0,m)\mathrm{d}m = -\frac{\mathrm{d}\ \ln(p(0,t))}{\mathrm{d}t}$$

ossia

$$l(0,t) = -\frac{\mathrm{d}\ \ln(p(0,t))}{\mathrm{d}t}.$$

Usando il modello Smith-Wilson si ottiene:

$$l(0,t) = -\frac{\mathrm{d}\, ln(e^{-w\cdot t} + \sum_{i=1}^{I} \varepsilon_i e^{-w\cdot(t+u_i)} W(t,u_i))}{\mathrm{d}t}.$$

[5] Karlin e Rinott 1983 p.426

$$= -\frac{1}{e^{-w\cdot t} + \sum_{i=1}^{I} \varepsilon_i e^{-w\cdot(t+u_i)} W(t,u_i)} \left[-we^{-w\cdot t} + \sum_{i=1}^{I} \left[-we^{-w\cdot(t+u_i)} W(t,u_i) + e^{-w\cdot(t+u_i)} \frac{dW(t,u_i)}{dt} \right] \varepsilon_i \right].$$

$$= w - \frac{1}{e^{-w\cdot t} + \sum_{i=1}^{I} \varepsilon_i e^{-w\cdot(t+u_i)} W(t,u_i)} \sum_{i=1}^{I} e^{-w\cdot(t+u_i)} \frac{dW(t,u_i)}{dt} \varepsilon_i.$$

Si nota che

$$\frac{\mathrm{d}W(t,u_i)}{\mathrm{d}t} = \begin{cases} \alpha - \alpha e^{-\alpha u_i} \cosh(\alpha t) & \text{se } t \leq u_i \\ \alpha e^{-\alpha t} \sinh(\alpha u_i) & \text{se } u_i \leq t \end{cases}$$

prendendo $t = CP$ tale per cui $CP > \max(J) = LLP$

$$l(0,CP) = w - \frac{\alpha}{e^{-w\cdot CP} + \sum_{i=1}^{I} \varepsilon_i e^{-w\cdot(CP+u_i)} W(CP,u_i)} \sum_{i=1}^{I} e^{-w\cdot(CP+u_i)} e^{-\alpha\cdot CP} \sinh(\alpha u_i) \varepsilon_i.$$

EIOPA propone di trovare α risolvendo il seguente problema:

$$\hat{\alpha} = \inf\{\alpha \geq 0.05 |\ g(\alpha) \leq 0.0001\} \tag{8}$$

con $g(\alpha) = |l(0,CP) - w|$. La precisa procedura utilizzata da EIOPA è disponibile aprendo le macro in VBA presenti in Smith e Wilson 2015.

2.4.3 Esempio Applicazione a OIS sull'ESTR

In questa sezione si applicherà tutto quanto visto in precedenza agli OIS sull'ESTR quotati alla data 07/01/2025. Le quotazioni utilizzate sono ricavate da Bloomberg ed in particolare solo il valore intermedio tra il bid e l'ask proposti dal provider. Le quotazioni sono le seguenti:

Tabella 2.1: Quotazioni Bloomberg degli OIS sull'ESTR alla data 07/01/2025.

Term	Unit	Ticker	Bid	Ask		mid
1	DY		2.917	2.917		2.917
1	WK	EESWE1Z	2.903259993	2.930239916		2.916749954
2	WK	EESWE2Z	2.911847353	2.923152685		2.917500019

Term	Unit	Ticker	Bid	Ask		mid
1	MO	EESWEA	2.865466595	2.895127773		2.880297184
2	MO	EESWEB	2.772936583	2.801663637		2.78730011
3	MO	EESWEC	2.660635471	2.677164555		2.668900013
4	MO	EESWED	2.578173161	2.584327221		2.581250191
5	MO	EESWEE	2.506432533	2.514167786		2.510300159
6	MO	EESWEF	2.440351963	2.446347713		2.443349838
7	MO	EESWEG	2.384613514	2.39048624		2.387549877
8	MO	EESWEH	2.344635963	2.350163937		2.34739995
9	MO	EESWEI	2.306516886	2.312483072		2.309499979
10	MO	EESWEJ	2.271858692	2.278341293		2.275099993
11	MO	EESWEK	2.244195938	2.250803947		2.247499943
1	YR	EESWE1	2.218995571	2.226004124		2.222499847
18	MO	EESWE1F	2.122331142	2.129668713		2.125999928
2	YR	EESWE2	2.101325274	2.107874632		2.104599953
3	YR	EESWE3	2.112565994	2.119434357		2.116000175
4	YR	EESWE4	2.14334631	2.14905405		2.14620018
5	YR	EESWE5	2.174256802	2.178943157		2.176599979
6	YR	EESWE6	2.207625628	2.212374449		2.210000038
7	YR	EESWE7	2.243067265	2.247832775		2.24545002
8	YR	EESWE8	2.274749279	2.279450893		2.277100086
9	YR	EESWE9	2.308026791	2.312673092		2.310349941
10	YR	EESWE10	2.34035635	2.345043659		2.342700005

Da questi input si procede come descritto nella sezione 2.1.1 e si ottengono i fattori di sconto. Questi verranno interpretati come se fossero prezzi di ZCB osservati alla data 07/01/2025 che pagano 1 alla data di scadenza dello swap da cui sono stati estratti. Si assume inoltre un $CP = 60$ e un $UFR = 4.5\%$ da cui deriva $w = 4.4\%$. A questo punto in Matlab si procede come indicato nella sezione 2.4.2. Per il calcolo dell'α ottimale viene utilizzato utilizzato il metodo della bisezione e viene individuato un valore di $\hat{\alpha} = 0.0992$ Una volta fissato α si calcola la formula (3) per $p(0,t)$ per diversi valori di t. Da questa viene estratta tramite la formula $r_t = -\frac{\ln p(0,t)}{t}$ la funzione dei tassi di sconto. Ripeto una procedura simile per il tasso forward istantaneo e ottengo i risultati descritti in seguito.

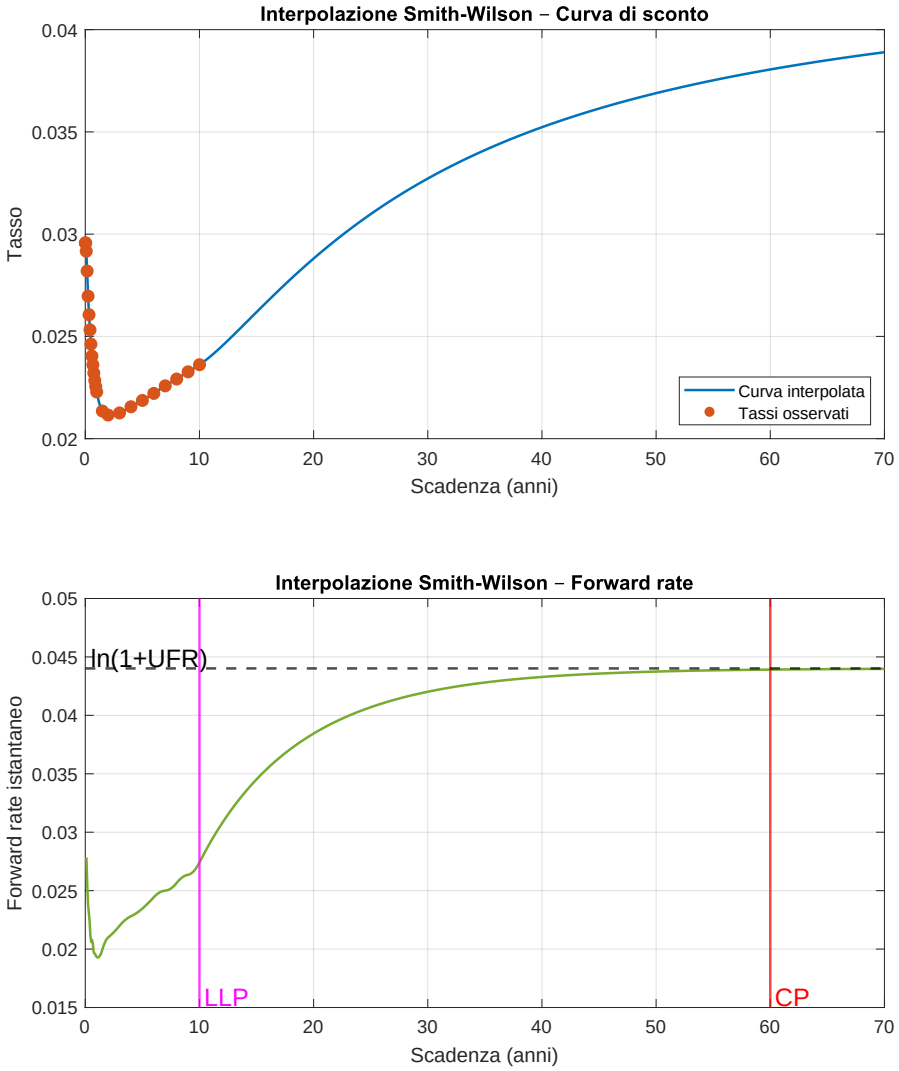

Figura 2.2: Grafico dei tassi di sconto interpolati (sopra) e grafico dei forward rate istantanei impliciti nel modello Smith-Wilson (sotto).

2.4.4 Esempio di Processo Gaussiano

Come esposto in precedenza, una soluzione potenziale al problema dell'alternanza dei pesi è data dalla scelta di un processo stocastico gaussiano con matrice varianze covarianze la cui inversa sia una matrice-M. Un esempio possibile è il seguente:

$$H_t = e^{-wt}(1 + X_t)$$

con X_t processo di Ornstein-Uhlenbeck descritto in precedenza. Dai risultati ricavati in precedenza (4), (5) è semplice dimostrare le seguenti proprietà:

$$\mathbb{E}^{\mathbb{P}}[H_t] = e^{-wt}$$

$$V(H_t) = e^{-2wt}\alpha^2\left[\frac{e^{-2\alpha t}+1}{2}\right]$$

$$cov(H_s, H_t) = e^{-w(t+s)}\alpha^2 e^{-\alpha t}\cosh(\alpha s) \text{ se} t > s$$

più genericamente:

$$\begin{aligned} cov(H_s, H_t) &= e^{-w(t+s)}\alpha^2 e^{-\alpha\max(t,s)}\cosh(\alpha\min(t,s)) \\ &= e^{-w(t+s)}OU(t,s) \end{aligned}$$

con $OU(t,s) := \alpha^2 e^{-\alpha\max(t,s)}\cosh(\alpha\min(t,s))$. Si osserva che anche $H_t \sim \mathcal{N}(e^{-wt}, e^{-2wt}\alpha^2\left[\frac{e^{-2\alpha t}+1}{2}\right])$ è un processo gaussiano in quanto trasformazione lineare di un processo gaussiano. Inoltre, la sua matrice varianze-covarianze è strettamente positiva per costruzione. Essa è anche semidefinita positiva in quanto tale e dunque invertibile. Questo porta alla conclusione che la sua inversa è una matrice-M. Definendo sempre **H** come vettore contenente le variabili casuali relative alle maturity

di cui si conosce la realizzazione, si procede come segue:

$$\begin{aligned}
\mathbb{E}^{\mathbb{P}}\left[H_t|H_{u_1}=p(0,u_1),\ldots,H_{u_{\bar{J}}}=p(0,u_{\bar{J}})\right] &= \mathbb{E}^{\mathbb{P}}\left[H_t\right]+Cov\left[H_t,\mathbf{H}\right]Cov\left[\mathbf{H},\mathbf{H}\right]^{-1}\left[p(0,\mathbf{u})-\mathbb{E}^{\mathbb{P}}\left[\mathbf{H}\right]\right] \\
&= e^{-w\cdot t}+Cov\left[H_t,\mathbf{H}\right]Cov\left[\mathbf{H},\mathbf{H}\right]^{-1}\left[p(0,\mathbf{u})-e^{-w\cdot\mathbf{u}}\right] \\
&= e^{-w\cdot t}+Cov\left[H_t,\mathbf{H}\right]Cov\left[\mathbf{H},\mathbf{H}\right]^{-1}\bar{\mathbf{P}} \\
&= e^{-w\cdot t}+Cov\left[H_t,\mathbf{H}\right]\left(diag(e^{-w\cdot\mathbf{u}})\mathbf{OU}diag(e^{-w\cdot\mathbf{u}})^T\right)^{-1}\bar{\mathbf{P}} \\
&= e^{-w\cdot t}+Cov\left[H_t,\mathbf{H}\right]\mathbf{Z} \\
&= e^{-w\cdot t}+\sum_{i=1}^{I}\varepsilon_i e^{-w\cdot(t+u_i)}OU(t,u_i) \\
&= p(0,t).
\end{aligned}$$

In questo caso si ha il vettore $\mathbf{Z}=\left(diag(e^{-w\cdot\mathbf{u}})OUdiag(e^{-w\cdot\mathbf{u}})^T\right)^{-1}\bar{\mathbf{P}}$,con $\mathbf{Z}_i=\varepsilon_i$, e la matrice $\mathbf{OU}$ tale per cui $\mathbf{OU_{k,j}}=OU(u_k,u_j)$. il forward rate istantaneo sotto queste ipotesi è ricavato sempre come:

$$l(0,t)=-\frac{\mathrm{d}\ \ln(p(0,t))}{\mathrm{d}t}.$$

Sotto le nuove ipotesi si ha

$$l(0,t)=-\frac{\mathrm{d}\,ln(e^{-w\cdot t}+\sum_{i=1}^{I}\varepsilon_i e^{-w\cdot(t+u_i)}OU(t,u_i))}{\mathrm{d}t}.$$

$$=-\frac{1}{e^{-w\cdot t}+\sum_{i=1}^{I}\varepsilon_i e^{-w\cdot(t+u_i)}OU(t,u_i)}\left[-we^{-w\cdot t}+\sum_{i=1}^{I}\left[-we^{-w\cdot(t+u_i)}OU(t,u_i)+e^{-w\cdot(t+u_i)}\frac{d\,OU(t,u_i)}{dt}\right]\varepsilon_i\right].$$

$$=w-\frac{1}{e^{-w\cdot t}+\sum_{i=1}^{I}\varepsilon_i e^{-w\cdot(t+u_i)}OU(t,u_i)}\sum_{i=1}^{I}e^{-w\cdot(t+u_i)}\frac{d\,OU(t,u_i)}{dt}\varepsilon_i.$$

Si nota che

$$\frac{\mathrm{d}\,OU(t,u_i)}{\mathrm{d}t}=\begin{cases}\alpha^3 e^{-\alpha u_i}\sinh(\alpha t) & \text{se } t\leq u_i\\ -\alpha^3 e^{-\alpha t}\cosh(\alpha u_i) & \text{se } u_i\leq t.\end{cases}$$

Anche in questo caso prendendo $t=CP$ tale per cui $CP>\max(J)=LLP$

$$l(0,CP)=w+\frac{\alpha}{e^{-w\cdot CP}+\sum_{i=1}^{I}\varepsilon_i e^{-w\cdot(CP+u_i)}OU(CP,u_i)}\sum_{i=1}^{I}e^{-w\cdot(CP+u_i)}\alpha^2 e^{-\alpha t}\cosh(\alpha u_i)\varepsilon_i.$$

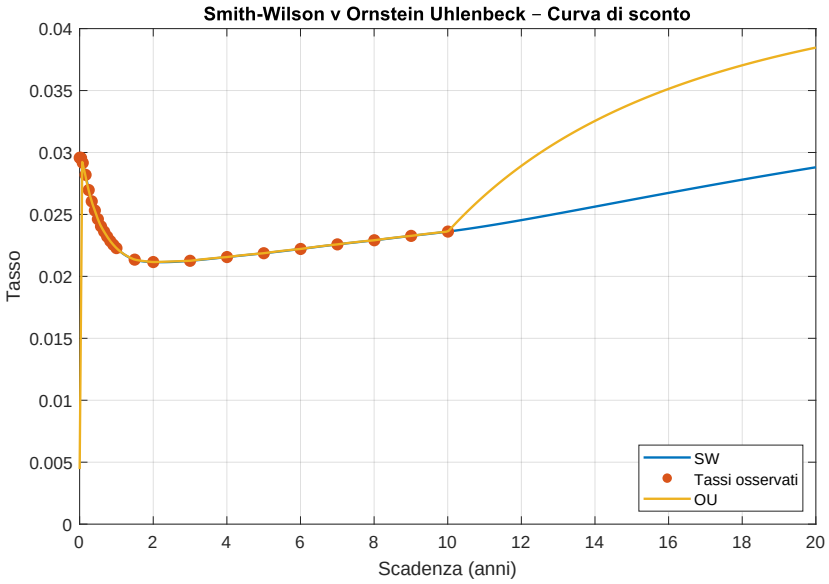

Figura 2.3: Grafico di confronto tra interpolazione con metodo Smith-Wilson e modello Ornstein-Uhlenbeck

Il livello di α ottimale è sempre ottenuto risolvendo (8). Applicando quanto esposto ai dati della tabella 2.1 si ottiene un $\alpha = 0.657$ e il seguente grafico di interpolazione: Come ci si aspetta, alla luce di quanto esposto nella risoluzione del problema deterministico, passando dal processo implicito nel modello SW al modello OU si sacrifica la regolarità della funzione interpolante in cambio della proprietà di hedging implicito nel modello con solo posizioni lunghe.

3. Long-Term Guarantee Measures (LTG measures)

Nel quadro regolamentare di Solvency II, le *Long-Term Guarantee Measures* rappresentano un insieme di strumenti volti a ridurre la volatilità artificiale dei bilanci delle imprese di assicurazione, in particolare nei rami vita, dove gli impegni contrattuali possono estendersi su orizzonti temporali di lungo periodo. Queste misure sono state introdotte per riflettere più accuratamente la natura a lungo termine del business assicurativo, riconoscendo che, in condizioni normali, le compagnie possono mantenere gli attivi fino a scadenza, senza doverli liquidare in periodi di stress di mercato. Le LTG measures sono state introdotte nel 2014 con la Direttiva 2014/51/EU (nota anche come *Omnibus II*[1]) e trovano fondamento legale nel Titolo I, Capo VI della Direttiva Solvency II. EIOPA, nel suo report annuale del 2020[2], evidenzia come circa l'80% delle riserve tecniche dell'Area Economica Europea faccia uso di almeno una delle misure previste.
Le principali LTG measures attualmente previste sono:

- **Volatility Adjustment (VA)**: un aggiustamento applicato al tasso risk-free per tener conto della variazione degli spread di mercato in periodi di stress, evitando reazioni eccessive nella valutazione delle riserve tecniche.

- **Matching Adjustment (MA)**: un'aggiunta al tasso risk-free applicabile su specifici portafogli (tipicamente annuity) quando gli attivi sono strettamente allineati ai passivi in termini di flussi futuri.

- **Transitional Measures**: misure transitorie su riserve tecniche e requisiti di capitale, con lo scopo di attenuare l'impatto del passaggio dal regime Solvency I a Solvency II.

Tutte queste misure sono soggette ad approvazione dell'autorità di vigilanza e devono essere applicate secondo criteri prudenziali, trasparenti e coerenti nel tempo. L'uso delle LTG measures impatta dunque le BEL, i requisiti patrimoniali (SCR) e la rappresentazione della solvibilità aziendale.

[1] Direttiva 2014/51/EU (Omnibus II), Titolo I, Capo VI.

[2] EIOPA, *Report on Long-Term Guarantees Measures and Measures on Equity Risk*, 2020

3.1 Volatility Adjustment (VA)

Il Volatility Adjustment è uno spread che viene aggiunto al tasso risk-free allo scopo di mitigare alcuni dei problemi generati dalla valutazione al fair value delle poste all'attivo. In particolare, quando i mercati sono in condizioni di stress, gli spread per i rischi di mercato possono non riflettere solo il rischio di insolvenza, ma anche altri elementi di rischio come il rischio di liquidità e fattori speculativi. La conseguenza è che il valore degli attivi potrebbe non riflettere per certi periodi l'effettiva redditività futura, soprattutto in situazioni dove la compagnia assicurativa deterrà le posizioni fino alla loro scadenza. L'intuizione sarebbe quella di far sì che i fondi propri non siano impattati da variazioni dei prezzi dovute a eventi "non-fondamentali" e temporanei e questo è giustificato a fronte di passività che sono a loro volta illiquide. Questo risultato si ottiene appunto facendo sì che il fattore di sconto delle technical provision (BEL) sia aggiustato per tenere conto di questa volatilità "extra". Per la determinazione della curva di sconto aggiustata per la volatilità EIOPA parte dal risultato dell'interpolazione Smith-Wilson e procede in questo modo:

1. Calcolo dell'aggiustamento per la volatilità.
2. Su tutte le scadenze discrete (annuali) fino al LLP si prendono i tassi impliciti e si aggiunge il VA.
3. Si riapplica il metodo Smith-Wilson su questi nuovi tassi aggiustati trattandoli come tassi di ZCB (sempre prendendo come prezzo 1 e nozionale a scadenza di 1), il resto degli input rimane invariato.

3.1.1 VA in utilizzo

Il volatility adjustment è attualmente calcolato per ogni paese in questo modo:

$$VA_{total} = 0.65 \left[S^{RC}_{currency} + \mathbf{1}_{S^{RC}_{country} > 0.85\%} \max(S^{RC}_{country} - 2S^{RC}_{currency}; 0) \right].$$

S^{RC} in generale indica uno "spread corretto per il rischio", l'interpretazione del funzionamento dello spread è: prima si calcola lo spread tra la curva risk-free e una qualche altra curva rappresentativa del rischio presente nell'attivo della società assicurativa. Questo spread racchiude in sé tutti i

rischi, per isolare il rischio di volatilità si calcola una risk correction che dovrebbe contenere tutti i rischi di tipo fondamentale, a questo punto il VA dovrebbe essere dato da $spread - risk\,correction$. Nella formula proposta da EIOPA questa differenza è scomposta in una componente generata dalla valuta e da una componente generata dal paese, le due componenti sono a loro volta separate in due parti, governativa e corporate.

Le componenti della formula sono definite come:

$S^{RC}_{currency} = S_{currency} - RC_{currency}$ come risk corrected spread per la valuta:

$$S_{currency} = w^{gov}_{currency} S^{gov+}_{currency} + w^{corp}_{currency} S^{corp+}_{currency}.$$

$$RC_{currency} = w^{gov}_{currency} RC^{gov+}_{currency} + w^{corp}_{currency} RC^{corp+}_{currency}.$$

con $S^{\cdot+}_{currency} = \max(S^{\cdot}_{currency}; 0)$ e $RC^{\cdot+}_{currency} = \max(RC^{\cdot}_{currency}; 0)$ con:

- $w^{gov}_{currency}$ è il peso dei government bond presenti nel portafoglio rappresentativo della specifica valuta.
- $w^{corp}_{currency}$ è il peso dei corporate bond presenti nel portafoglio rappresentativo della specifica valuta.
- $S^{gov}_{currency}$ è lo spread medio della componente fixed income governativa presente nel portafoglio rappresentativo della specifica valuta.
- $S^{gov}_{currency}$ è lo spread medio della componente fixed income diversa da quella governativa presente nel portafoglio rappresentativo della specifica valuta.
- $RC^{gov}_{currency}$ è la risk correction che dovrebbe contenere tutti i fattori di rischio fondamentali legati alla componente currency per la parte government bond.
- $RC^{corp}_{currency}$ è la risk correction che dovrebbe contenere tutti i fattori di rischio fondamentali legati alla componente currency per la parte corporate.
- $S^{RC}_{country}$ è costruito con la stessa logica, cambiano i portafogli di riferimento.

Resta da chiarire che cosa siano il portafoglio rappresentativo e il portafoglio di riferimento. Il portafoglio di riferimento è costruito a partire dal portafoglio rappresentativo. Questo portafoglio è interamente determinato da EIOPA sulla base di un attivo rappresentativo dell'impresa media. Ne vengono creati diversi, rappresentativi per ciascuna valuta (Es. euro, il portafoglio rappresenta l'attivo di un'impresa assicurativa media nella zona euro) e rappresentativi per ciascun paese (Es. Italia, il portafoglio rappresenta l'attivo di un'impresa assicurativa media italiana). Il portafoglio di riferimento è costruito a partire da questo attraverso l'utilizzo di indici di mercato sempre scelti da EIOPA relativi a yield curve su bond governativi e corporate; da questi si ottiene la curva necessaria per costruire S. Viene assunto inoltre che gli asset denominati in una determinata valuta coprano obbligazioni nella stessa valuta, mentre gli asset relativi a un determinato paese coprano obbligazioni del medesimo paese. I pesi delle componenti sono infatti forniti direttamente da EIOPA.

Costruzione dei portafogli di riferimento

Come già detto, per arrivare alla costruzione dei portafogli di riferimento è necessario costruire prima per ogni paese il portafoglio rappresentativo. Il portafoglio rappresentativo rappresenta una sorta di attivo di una società assicurativa media del paese per cui lo si sta valutando, è costruito da EIOPA a partire dalle seguenti informazioni:

Valore di mercato degli asset presenti nei portafogli. Da questa informazione si calcola:

- $w^{gov}_{currency}$,$w^{corp}_{currency}$, $w^{gov}_{country}$,$w^{corp}_{country}$.
- Le duration della componente fixed income, questa insieme ai pesi, permettono di individuare le duration della parte currency governativa, currency corporate, country governativa e country corporate.
- Il Credit Quality Step ossia un valore da una scala da 0 (solvibile) a 6 (rischioso) per la componente fixed income diversa da quella governativa.

Da ognuno di questi portafogli si estraggono poi quattro portafogli di yield market index.

Portafoglio di yield market index per valuta e componente governativa

Per ogni portafoglio rappresentativo di una valuta si controlla se ci sono titoli governativi, per ogni

emittente EIOPA individua uno yield market index, ossia una curva di sconto riferita allo specifico emittente, ed estrae il fattore di sconto con duration consistente alla duration dei titoli presenti nel portafoglio rappresentativo per ogni emittente. Lo spread per uno specifico emittente è calcolato come:

$$S^{gov}_{currency} = Ycurrency^{gov}_{duration} - Rf_{duration}.$$

$Ycurrency^{gov}_{duration}$ è calcolato tasso interno di rendimento di un portafoglio che rappresenta una media ponderata con pesi dati dai valori di mercato definiti dalle componenti della parte governativa nel portafoglio della specifica valuta, il portafoglio è costruito ipotizzando di investire una quantità determinata dal peso nello specifico yield market index per una specifica maturity definita dalla duration media (calcoalta sempre con questi pesi interni della componente governativa del portafolgio della specifica valuta) dei titoli governativi del portafoglio currency.

$$Rf_{duration}$$

rappresenta anch'esso un tasso interno di rendimento, questa volta calcolato ipotizzando di investire al tasso risk-free in un portafoglio sempre con ponderazioni indotte dai pesi con duration media equvalente alla duration del tir degli yield. Per il calcolo del tir si ipotizza un prezzo iniziale del portafoglio di 1

Portafoglio di yield market index per paese e componente governativa

La logica è simile a prima, però questa volta si guarda alla dimensione paese invece che alla valuta per determinare di quali emittenti estrarre lo yield market index relativamente al portafoglio rappresentativo. Il resto dei passaggi è uguale

$$S^{gov}_{country} = Ycountry^{gov}_{duration} - Rf_{duration}.$$

Viene calcolata per ogni emittente e poi viene calcolato lo spread medio utilizzando come pesi le ponderazioni a valori di mercato di ciascun emittente nella componente governativa della specifica valuta. **Portafoglio di yield market index per componente corporate**

Per la selezione, gli indici relativi alla parte corporate si tiene conto anche della distinzione emittente finanziario/non finanziario, del credit quality step (rating) e della valuta. Per l'identificazione

degli spread tutto quello già detto prima.

$$S^{corp}_{currency} = Ycurrency^{corp}_{duration} - Rf_{duration}.$$

$$S^{corp}_{country} = Ycountry^{corp}_{duration} - Rf_{duration}.$$

Risk correction per componente governativa

La logica dietro la risk correction sarebbe quella di catturare i rischi fondamentali, in particolare quelli di default e downgrading isolando la componente del VA. EIOPA propone la seguente formula:

$$30\%LTAS.$$

per esposizioni a governi locati nella european economic area;

$$35\%LTAS$$

per esposizioni a governi locati al di fuori della european economic area.

Risk correction per componente corporate

In questo caso la formula è:

$$\max\left(PD + CoD; 35\%LTAS\right).$$

LTAS rappresenta il long term average spread che è una media calcolata a 30 anni su tutti i trading days accessibili degli spread tra curva risk free e yield curve (definita dagli yield market index) per asset con stessa duration, credit quality e asset class. Tendenzialmente non si hanno 30 anni di storia in quanto il mercato degli OIS (necessari per il calcolo del CRA per la risk free curve) ha iniziato a svilupparsi a fine anni '90. EIOPA propone di utilizzare questo conto per i giorni di trading su cui è possibile eseguire il calcolo, per i giorni precedenti applicare direttamente la media calcolata sui giorni "affidabili".

CoD e *PD* sono calcolate sfruttando delle matrici di transizione. CoD sarebbe il costo di downgrade mentre PD il costo di default calcolato ipotizzando un tasso di recupero del 30%. Nel caso di uno ZCB ad esempio il CoD e la PD sono calcolate partendo da valore di mercato:

$$MV_M(m) = \frac{(1+r_M)^M}{(1+f_{m,M})^{M-m}}$$

dove M è la scadenza dello ZCB, m tempo futuro generico prima della scadenza, r_M è il tasso risk-free estratto con la metodologia precedentemente introdotta, f è il tasso forward osservato oggi per il periodo m-M. Questa uguaglianza vale per uno ZCB senza rischio di default; infatti sappiamo che in questa situazione è vero che

$$(1+r_m)^m(1+f_{m,M})^{M-m} = (1+r_M)^M.$$

da cui

$$MV_M(m) = (1+r_m)^m.$$

Questo è vero in assenza di rischi. Il valore di mercato di un titolo con rischi di default e downgrade è assunto essere

$$MV_{c,M}(m) = R_c^{\frac{M-m}{15}}(1+r_m)^m.$$

dove c rappresenta il credit quality step di partenza del titolo, R_c è una correzione specifica per lo step di riferimento, 15 rappresenta un'ipotesi di duration massima. Il cost of downgrade dalla classe X alla classe $Y < X$ al tempo m per uno ZCB con maturity M è dato da

$$CoD_{(X,Y),M}(m) = MV_{X,M}(m) - MV_{Y,M}(m) > 0.$$

Se $Y >= X$ il costo è uguale a zero.

Si procede costruendo la seguente matrice triangolare verso l'alto per la maturity M al tempo di valutazione m:

$$C_M^{(m)} = \begin{bmatrix} CoD_{(X,Y),M}(m)p_{X,Y} \text{ se } Y \neq \text{default} \\ 30\% MV_{X,M}(m)p_{X,Y} \text{ se } Y = \text{default} \end{bmatrix}_{X>Y\in CQS}.$$

serve inoltre questa matrice di transizione

$$Q = \begin{bmatrix} \sum_{k=1}^{n-1} p_{1,k} & 0 & \dots & p_{1,d} \\ p_{2,1} & \sum_{k=2}^{n-1} p_{2,k} & \dots & p_{2,d} \\ \vdots & \vdots & \ddots & \\ 0 & 0 & \dots & 1 \end{bmatrix}$$

Le somme sono date dal fatto che l'articolo 54 del Regolamento Delegato richiede di ipotizzare che dopo ogni downgrade intermedio il titolo debba essere sostituito con uno con credit quality

step almeno pari. L'ultima matrice da costruire è quella data dai valori alle scadenze discrete:

$$\begin{bmatrix} CoD_{primo\,step,M}(1) & \dots & CoD_{primo\,step,M}(M) \\ \vdots & \ddots & \vdots \\ CoD_{penultimo\,step,M}(1) & \dots & CoD_{penultimo\,step,M}(M) \\ CoD_{default\,step,M}(1)=0 & \dots & CoD_{default\,step,M}(M)=0 \end{bmatrix} = \bigcup_{m=1}^{M} \left(Q^{m-1} C_M^{(m)} \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 0 \end{bmatrix} \right).$$

dove l'unione va intesa in senso di concatenamento di vettori colonna uno a fianco all'altro.

Q^{m-1} implica un'ipotesi di stazionarietà per la matrice di transizione.

Il Cost of Default (in bps) per lo step legato al rating c maturity M denotato con $CoD_{c,M}^{(bp)}$ è calcolato risolvendo la seguente equazione:

$$\frac{1}{\left(1+r_m+CoD_{c,M}^{(bp)}\right)^M} = \frac{1}{(1+r_m)^M}\left(1-\sum_{m=1}^{M}\frac{CoD_{c,M}}{(1+r_m)^{m-0.5}}\right).$$

La probabilità di default è ottenuta nello stesso modo semplicemente sostituendo $\begin{bmatrix} 1 \\ 1 \\ \vdots \\ 0 \end{bmatrix}$ con $\begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}$ e alla matrice Q aggiustata per l'articolo 54 quella originale

$$T = \begin{bmatrix} p_{1,1} & p_{1,2} & \cdots & p_{1,d} \\ p_{2,1} & p_{2,2} & \cdots & p_{2,d} \\ \vdots & \vdots & \ddots & \\ p_{d,1} & p_{d,2} & \cdots & p_{d,d} \end{bmatrix}.$$

è comunque necessario calcolare sia $PD(T)$ che $PD(Q)$.

Il Cost of Downgrade finale per la classe c maturity M è dato da:

$$CoD_{c,M}^{(bp)} = \max\left[0; CoD_{c,M}^{(bp)} - (PD_{c,M}(T) - PD_{c,M}(Q))\right].$$

3.1.2 VA proposto

Il VA in utilizzo presenta sicuramente alcuni problemi, il primo dei quali risulta essere la mancata presa in considerazione delle caratteristiche specifiche dell'impresa assicurativa, ad esempio,

il duration miss-match che caratterizza la specifica società. I principali aggiustamenti necessari dovrebbero riguardare la possibilità per le singole imprese di utilizzare le proprie duration di portafoglio e i propri pesi di portafoglio relativamente alla componente governativa e di paese.
Nel 2020 è stata proposta una formula alternativa per il VA calcolata nel seguente modo

$$VA = VA_{perm} + VA_{macro}.$$

$$VA_{perm} = GAR \times AR_4 \times AR_5 \times scale_{currency} \times S^{RC}_{currency}.$$

GAR è il general application ratio fissato a 85%.
AR_4,AR_5 sono degli application ratios che sono specifici per compagnia assicurativa, AR_4 tiene conto del maturity missmatch tra attivo e passivo, AR_5 tiene conto dell'illiquidità degli attivi.
$scale_{currency} = \frac{1}{w^{gov}_{currency} + w^{corp}_{currency}}$ serve per far si che le scale sommino a 1, visto che i due pesi in $S^{RC}_{currency}$ non sommano necessariamente a 1.
La modifica sta nel calcolo di RC

$$RC^{gov+}_{currency} = 30\% \times \min(S^{gov+}_{currency}; LTAS^{+}) + 20\% \times \max(S^{gov+}_{currency} - LTAS^{+}; 0).$$

$$RC^{gov+}_{country} = 30\% \times \min(S^{gov+}_{country}; LTAS^{+}) + 20\% \times \max(S^{gov+}_{country} - LTAS^{+}; 0).$$

per la componente governativa nella european economic area.
Per la componente governativa al di fuori della zona euro si ha:

$$RC^{gov+}_{currency} = 50\% \times \min(S^{gov+}_{currency}; LTAS^{+}) + 40\% \times \max(S^{gov+}_{currency} - LTAS^{+}; 0).$$

$$RC^{gov+}_{country} = 50\% \times \min(S^{gov+}_{country}; LTAS^{+}) + 40\% \times \max(S^{gov+}_{country} - LTAS^{+}; 0).$$

Per la componente corporate si ha invece:

$$RC^{corp+}_{currency} = 50\% \times \min(S^{corp+}_{currency}; LTAS^{+}) + 40\% \times \max(S^{corp+}_{currency} - LTAS^{+}; 0).$$

$$RC^{corp+}_{country} = 50\% \times \min(S^{corp+}_{country}; LTAS^{+}) + 40\% \times \max(S^{corp+}_{country} - LTAS^{+}; 0).$$

si perde dunque la componente legata a PD e CoD.Inoltre, si ha:
$VA_{macro} = GAR \times AR_4 \times AR_5 \times w \times \max(scale_{country} \times S^{RC}_{country} - R \times Scale_{currency} \times S^{RC}_{currency}; 0)$

con: $scale_{country} = \frac{1}{w^{gov}_{country} + w^{corp}_{country}}$
$R = 1.3$ per facilitare l'attivazione della componente di paese rispetto alla formula in utilizzo.
$w = \min\left(1; \frac{S^{RC}_{country} - S^L}{S^H - S^L}\right) \times \mathbf{1}_{S^{RC}_{country} > S^L}$ con $S^L = 0.6\%$, $S^L = 0.9\%$ e $S^{RC}_{country}$ definito come quello per la currency modificando le risk correction rispetto alla formula in utilizzo.

3.2 Matching Adjustment (MA)

Il matching adjustment è previsto negli articoli 77 ter e 77 quater di Solvency II, è applicato a discrezione delle compagnie assicurative ma sostituisce il volatility adjustment sulle passività a cui è applicato. Il suo scopo dovrebbe essere quello di valutare meglio le passività assicurative replicabili. Infatti è previsto che per la sua applicazione la compagnia debba identificare un portafoglio di attivi che replichi i cashflow generati dalle passività a cui si vuole applicare il MA. Queste attività e passività dovrebbero essere gestite separatamente; l'idea è che siccome queste attività saranno detenute in modo che i cashflow coprano i flussi in uscita, il valore delle passività sarà equivalente. All'impresa di assicurazione rimangono comunque in capo alcuni rischi legati a queste attività che rientrano nello Spread Fondamentale. Per il calcolo del MA si procede calcolando due TIR. Il primo si ottiene risolvendo

$$\text{Valore attivi a copertura} = \sum_{t=1}^{n} \frac{\mathbb{E}^{\mathbb{P}}[CF_t]}{(1 + TIR_{asset})^t}.$$

il secondo

$$\sum_{t=1}^{n} \frac{\mathbb{E}^{\mathbb{P}}[CF_t]}{(1 + rf_t)^t} = \sum_{t=1}^{n} \frac{\mathbb{E}^{\mathbb{P}}[CF_t]}{(1 + TIR_{libilities})^t}$$

con rf_t tasso estratto con la procedura Smith-Wilson. Per quanto riguarda il fundamental spread si prende la media ponderata per il peso di portafoglio dei fundamental spread delle singole attività. Nel caso di attività governative FS è calcolato come:

$$FS_i = 30\% LTAS_i$$

per obbligazioni con stessa duration nella valuta di riferimento, 35% se le obbligazioni non sono dell'EEA. mentre nel caso di obbligazioni corporate

$$FS_i = \max(CoD + PD; 35\% LTAS_i).$$

La formula finale è:

$$MA = TIR_{asset} - TIR_{liabilities} - FS.$$

Questo aggiustamento va applicato come un add-on a ogni punto sulla curva risk-free di EIOPA (curva risk-free estratta da SW senza VA).

3.3 Transitional Measures

Le Transitional Measures sono strumenti introdotti dal regime Solvency II con l'obiettivo di facilitare il passaggio dal precedente quadro regolamentare (Solvency I) al nuovo sistema basato sul valore di mercato previsto da Solvency II. Sono pensate per mitigare gli impatti iniziali potenzialmente rilevanti sulla posizione patrimoniale delle compagnie assicurative, derivanti dal cambio di metodo di calcolo. L'utilizzo di queste misure è soggetto all'approvazione dell'autorità di vigilanza.

3.3.1 Misura transitoria sui tassi d'interesse privi di rischio

Prevista all'articolo 308 quater, prevede la possibilità per le compagnie di assicurazione di adattare un tasso aggiustato per calcolare le riserve Solvency relativamente ad alcune passività autorizzate. L'aggiustamento è infatti applicabile soltanto a:

- I contratti di assicurazione prodotti prima del 1 gennaio 2016 (data di entrata in vigore di Solvency II)
- alcune forme di riserve tecniche costituite secondo l'articolo 20 della direttiva 2002/83/CE (componente del quadro normativo di Solvency I abrogata con l'entrata in vigore di Solvency II).

L'applicazione della misura si sostanzia nell'aggiunta alla curva risk free (senza aggiustamenti per volatilità o matching) di uno spread, sulle diverse scadenze. Lo spread è definito nel seguente modo:

$$spread_t = \alpha_k \left(rf_{\text{Solvency I,t}} - TIR_{\text{Solvency II}} \right).$$

α_k è un parametro di ”decadenza” dell’aggiustamento nel senso che deve decrescere linearmente alla fine di ogni anno k, partendo da 100% l’1 gennaio 2016 e arrivare a zero l’1 gennaio 2032, riducendo nel tempo l’effetto dell’aggiustamento, permettendo la transizione.
$rf_{\text{Solvency I}}$ è la curva risk-free utilizzata dalla specifica compagnia secondo l’ordinamento normativo precedente (art. 20 direttiva 2002/83/CE).
$TIR_{\text{Solvency II}}$ è calcolato come tasso interno di rendimento che eguaglia il valore attuale delle passività a cui si vuole applicare l’aggiustamento. Usando la curva di EIOPA, è anche possibile utilizzare la curva corretta per il VA.

$$\sum_{t=1}^{n} \frac{CF_t}{\left(1+TIR_{\text{Solvency II}}\right)^t} = \sum_{t=1}^{n} \frac{CF_t}{\left(1+rf_t+VA_t\right)^t}.$$

La curva finale per le passività a cui è applicato l’aggiustamento sarà definita come:

$$rf_{t,Solvency\,II} + spread_t.$$

3.3.2 Misura transitoria sulle riserve tecniche

La misura transitoria sulle riserve tecniche è prevista all’articolo 308 quinquies di Solvency II. La logica di applicazione della misura è analoga a quella prevista per la curva dei tassi di interesse, solo che questa volta l’aggiustamento si applica direttamente al valore delle riserve tecniche secondo Solvency II (relativamente alle passività a cui si vuole applicare l’aggiustamento). Il calcolo è fatto nel seguente modo:

$$Aggiustamento_k = \alpha_k \left(TR_{\text{Solvency II}} - TR_{\text{Solvency I}}\right).$$

α_k in maniera simile a prima rappresenta un aggiustamento che deve decrescere linearmente alla fine di ogni anno k partendo da 100% l’1 gennaio 2016 e finendo a zero l’1 gennaio 2032.
$TR_{\text{Solvency II}}$ rappresentano le riserve tecniche calcolate secondo Solvency II l’1 gennaio 2016
$TR_{\text{Solvency I}}$ rappresentano le riserve tecniche calcolate secondo Solvency I il 31 dicembre 2015.
Il valore finale delle riserve tecniche all’anno k per cui si applica l’aggiustamento sarà dato da:

$$TR_k = TR_{\text{Solvency II},k} + Aggiustamento_k.$$

Su proposta dell'autorità di vigilanza o dopo la sua approvazione, $TR_{\text{Solvency II}}$ e $TR_{\text{Solvency I}}$ possono essere ricalcolate ogni 24 mesi o con frequenza maggiore in seguito a variazioni rilevanti del profilo di rischio dell'impresa.

4. Analisi empirica

4.1 Analisi descrittiva

In questo capitolo si procederà ad analizzare un data set di *Volatility Adjustment* pubblicato da EIOPA. Il dataset contiene osservazioni mensili dal 31/12/2015 al 30/04/2025 relativamente ai seguenti paesi: Austria, Belgio, Bulgaria, Repubblica Ceca, Germania, Spagna, Finlandia, Francia, Grecia, Ungheria, Irlanda, Italia, Olanda, Polonia, Portogallo, Svezia, Slovacchia, Regno Unito. Le serie storiche dei VA sono riportate nella Figura 4.1. Com'è possibile osservare dal grafico le serie di alcuni paesi sono molto simili tra loro (come ad esempio Austria, Belgio, Germania) mentre per altri paesi si osservano serie diverse; nella Grecia ad esempio si hanno livelli dell'aggiustamento più elevati rispetto agli altri paesi con un picco al 31/12/2015 dovuto alla crisi del debito sovrano (in particolare nel 2015 vennero previste ulteriori misure di austerità). Anche per l'Italia si osservano dei picchi nel 2018 in seguito a un anno caratterizzato da un deficit finanziario. In generale per tutti i paesi si osservano picchi nel 2020 generati dalla crisi COVID che vengono subito riassorbiti per poi avere un incremento nel 2022 dovuto all'invasione della Russia sul territorio dell'Ucraina. Per spiegare l'andamento del VA vengono selezionati i seguenti regressori:

- **Indici di paese**: serie storica di rendimenti del principale indice di mercato dei 18 paesi; questo regressore può essere visto come una proxy dei fondamentali dell'economia.
- **Paese**: numero intero tra 0 e 17 per indicare il paese di riferimento.
- **Yield 10 anni**: serie storica degli yield a 10 anni dei bond governativi del paese di riferimento.
- **Yield 1 anno**: serie storica degli yield a 1 anno dei bond governativi del paese di riferimento
- **Brent Crude Oil**: serie storica mensile del prezzo del petrolio Brent, rappresenta una proxy del costo dell'energia.

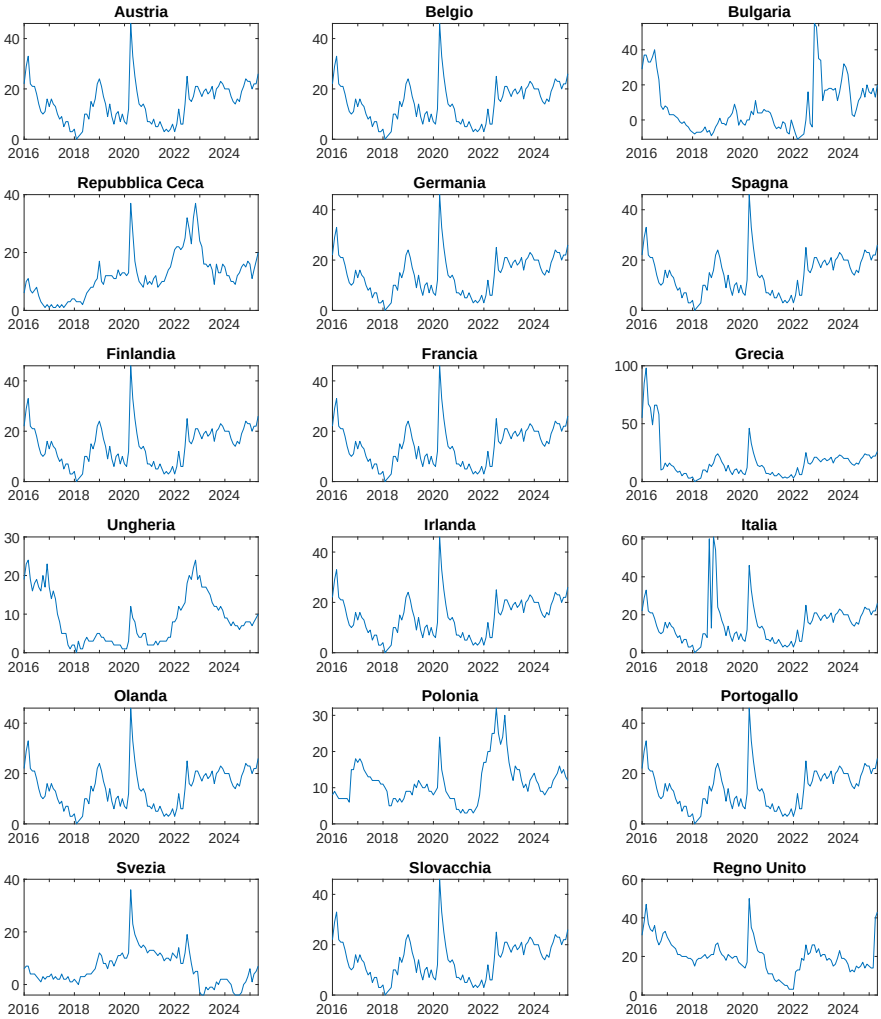

Figura 4.1: Serie storica dei Volatility Adjustment.

- **Gas Naturale**: serie storica mensile del prezzo del gas naturale Henry Hub. anche questa rappresenta una proxy del costo dell'energia.
- **HICP**: contiene la serie storica dell'inflazione annuale con osservazioni mensili per i 18 paesi.
- **S&P500**: serie storica dei rendimenti mensili dell'indice S&P500.
- **Spread 1-5 anni**: differenza tra yield dei titoli governativi a scadenza di 5 anni e di 1 anno.
- **Spread 1-10 anni**: differenza tra yield dei titoli governativi a scadenza di 10 anni e di 1 anno.
- **Spread 5-10 anni**: differenza tra yield dei titoli governativi a scadenza 10 anni e 5 anni.
- **Cambio Eur-Usd**: tasso di cambio euro dollaro.
- **VIX**: serie storica mensile dell'indice VIX calcolato da CBOE.
- **VSTOXX**: serie storica mensile del VSTOXX calcolato dall'EUREX.
- **iShares Core EUR Corp Bond UCITS ETF**: fondo contenente titoli di debito investment grade in euro, dovrebbe rappresentare una proxy della componente corporate del VA, vengono utilizzati i rendimenti mensili dell'indice come regressori.

Le fonti dei dati sono Bloomberg, Factset e il database di Eurostat. Viene svolta un'analisi delle correlazioni per determinare un primo aggiustamento agli input del modello.

4.1.1 Selezione Input

Per una selezione preliminare dei regressori sono analizzate le correlazioni lineari tra tutti i regressori indicati nella sezione precedente alle date di osservazione dei VA per ogni paese nella figura 4.2.

I regressori selezionati dopo questa breve analisi sono: VSTOXX, prezzo del Brent, prezzo del Gas Naturale, tasso di cambio euro dollaro, rendimenti dell'S&P500, il rendimento dell'indice

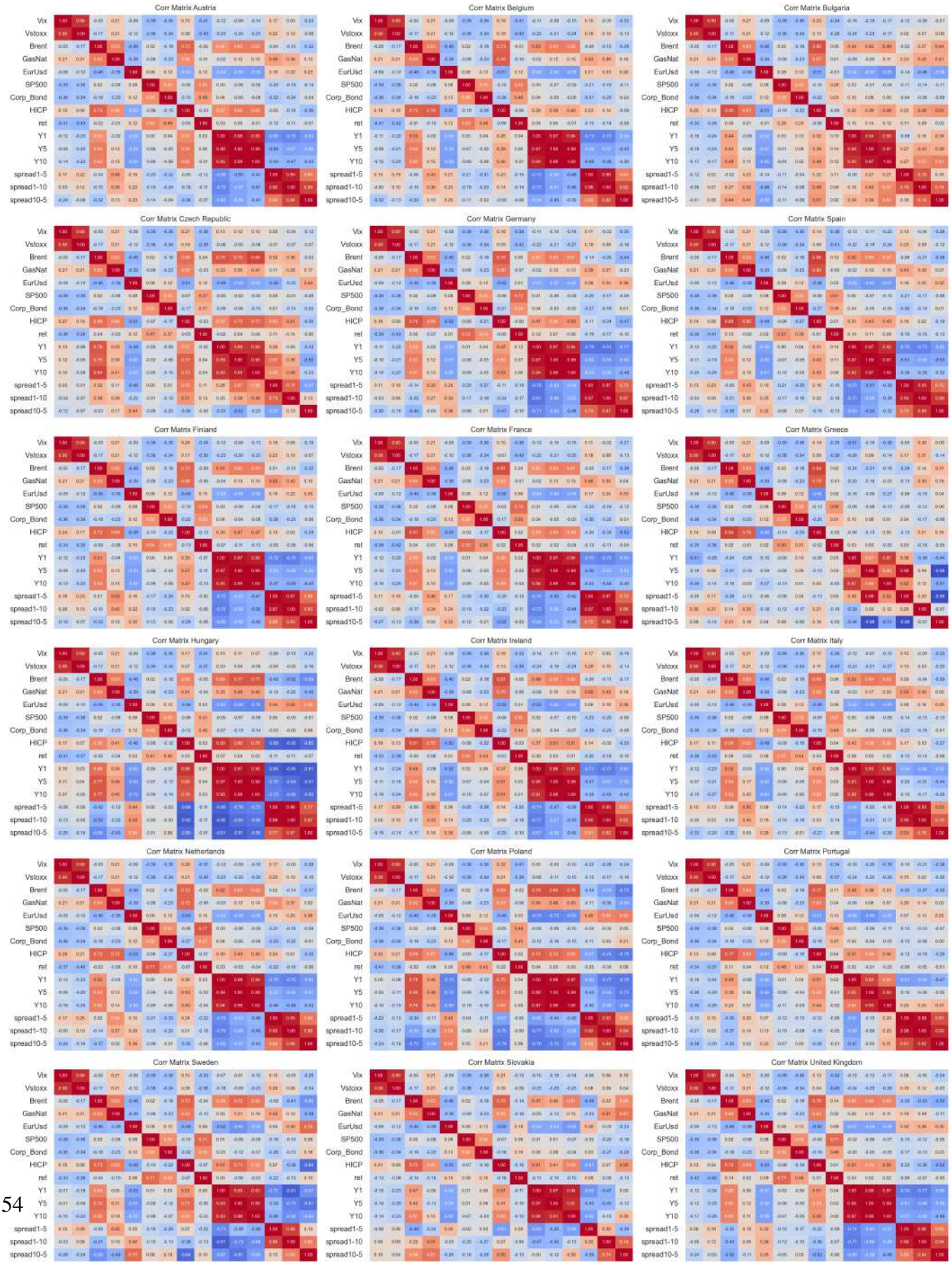

Figura 4.2: Matrici di correlazione

rappresentativo dei bond corporate, HICP (ossia il tasso d'inflazione per ogni paese), rendimenti dell'indice azionario di riferimento di ogni paese, lo yield a 5 anni di ogni paese, lo spread 1-5 degli yield dei titoli di stato di ogni paese, lo spread 5-10 degli yield dei titoli di stato di ogni paese e un embedding tridimensionale per indicare il paese di riferimento del VA. Tutti i regressori vengono processati venendo riscalati tra -1 e +1 per facilitare la convergenza del metodo di ottimizzazione che verrà discusso successivamente. La stessa procedura di riscalamento viene applicata ai VA di ogni paese.

4.2 Definizione della struttura del modello

4.2.1 Long Short Term Memory (LSTM)

Per analizzare i dati verrà utilizzata una struttura di rete neurale *Long Short Term Memory* che appartiene alla famiglia dei *Recurrent Network* (RN). La struttura di questi modelli rappresenta un'estensione dei *Feed Forward Network* (FFN) i quali possono essere definiti come una mappa:

$$\begin{aligned} \mathbf{z}^{(m)} : \{1\} \times \Re^{q_{m-1}} &\rightarrow \{1\} \times \Re^{q_m} \\ \mathbf{z} &\rightarrow \mathbf{z}^m(\mathbf{z}) = (1, \mathbf{z}_1^m, ..., \mathbf{z}_{q_m}^m(\mathbf{z})) \end{aligned}$$

con neuroni $1 \leq j \leq q_m$

$$\mathbf{z}_j^{(m)}(\mathbf{z}) = \phi\left(\sum_{l=0}^{q_{m-1}} w_{lj}^{(m)} z_l\right)$$

con:

- $\mathbf{z}^{(m)}$: rappresenta l'm-esimo layer del network ossia una funzione che va da $\Re^m$ a $\Re^n$ per cui nella prima dimensione di output si ha come risultato 1.

- q_m rappresenta il numero di dimensioni di output dell'm-esimo layer esclusa quella che restituisce sempre 1 (su queste dimensioni è applicata una funzione).

- $\mathbf{z}_j^{(m)}$ rappresenta il neurone ossia la funzione che restituisce il j-esimo output del layer m-esimo.

- ϕ rappresenta la funzione (detta funzione di attivazione) che caratterizza i neuroni del layer, queste funzioni vengono applicate al prodotto scalare $\mathbf{w}_j^{(m)T}\mathbf{z}^{(m-1)}$.
- $\mathbf{w}^{(m)T} := (w_{1j}^{(m)}, \ldots, w_{q_{m-1}j}^{(m)})$ vettore dei pesi del neurone j-esimo nel layer m-esimo.

Un feed forward network di profondità d può essere definito come:

$$\mathbf{x} \to \mathbf{z}^{(d:1)}(x) = \left(\mathbf{z}^{(d)} \circ \ldots \circ \mathbf{z}^{(1)}\right)(\mathbf{x}).$$

Nella figura 4.3 è fornita una rappresentazione grafica di un FFN[1].

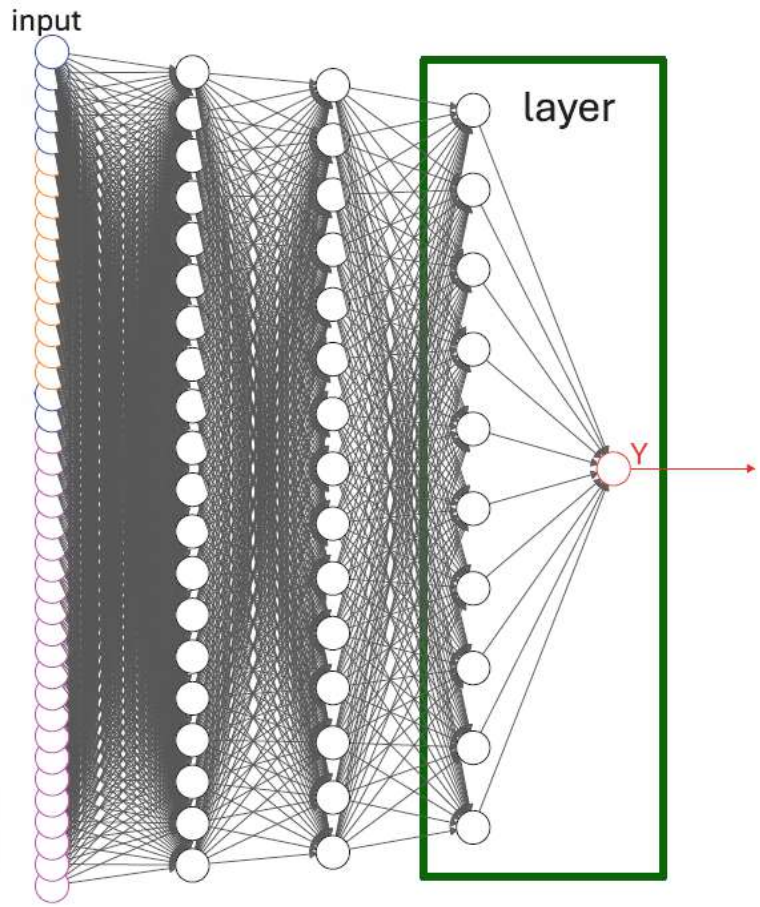

Figura 4.3: FFN rappresentazione grafica

Nel contesto delle serie temporali, i FFN presentano alcuni problemi, in primis se si vuole predire la variabile X_{t+1} utilizzando $(X_t, X_{t-1} \ldots, X_0)$ si avrebbe che la lunghezza degli input aumenta al

[1]Fonte figura: Wüthrich e Merz 2023

crescere di t, dunque il numero di parametri del primo layer aumenta complicando l'ottimizzazione. Una soluzione a questo problema è assumere la markovianità della serie (ossia ipotizzare che X_t sia influenzato solo al massimo da $X_{t-\tau}$). In generale il problema più grande è che in un FFN non viene dato alcun peso all'ordine di arrivo dei regressori. Per risolvere questo problema sono utilizzati i recurrent network. I RN sono definiti in maniera ricorsiva, partendo da un modello con un solo layer per $t \geq 1$ si ha che:

$$\begin{aligned} \mathbf{z}^{(1)} : \{1\} \times \Re^{q_0} \times \Re^{q_1} &\rightarrow \Re^{q_1} \\ (X_t, \mathbf{z_{t-1}}) &\rightarrow \mathbf{z_t} = \mathbf{z}^{(1)}(\mathbf{X_t}, \mathbf{z_{t-1}}) = \left(\mathbf{z}_1^{(1)}(X_t, \mathbf{z_{t-1}}), \ldots, \mathbf{z}_{q_1}^{(1)}(X_t, \mathbf{z_{t-1}})\right) \end{aligned}$$

con neuroni $1 \leq j \leq q_1$

$$\mathbf{z}_j^{[1]} = \mathbf{z}_j^{(1)}(X_t, \mathbf{z}) = \phi\left(\text{intercetta}_t^{(1)} + w_j^{(m)} X_t + \mathbf{u}_j^{(1)T} \mathbf{z_{t-1}}\right)$$

con:

- $\mathbf{u}_j^{(1)T}$: vettore di pesi associati ai lag degli output del modello al tempo precedente.
- q_0: numero di input, nell'esempio precedente è fissato a 1.

Per ottenere un modello di deep learning utilizzando i RN è sufficiente connettere più layer insieme; una rappresentazione grafica è fornita nella Figura: 4.4[2].
Per arrivare alle predizioni su X_{t+1} è comunque necessario connettere l'ultimo output del layer a un FFN. 4.5 I modelli long short term memory sono una forma di recurrent network a maggiore flessibilità, sono basati sulle funzioni di attivazione sigmoide $\phi_\sigma(x) := \frac{1}{1+e^{-x}} \in (0,1)$ e tangente iperbolica $\phi_{tanh}(x) := \frac{e^x - e^{-x}}{e^x + e^{-x}} \in (-1,1)$. Ognuno dei layer della figura 4.4 contiene quattro porte:

- *forghet gate*: modella quale informazione deve essere dimenticata per i periodi t succesivi

$$f_t^{[m]} = f_t^{(m)}(\mathbf{z_t}^{[m-1]}, \mathbf{z_{t-1}}^{[m]}) = \phi_\sigma\left(\mathbf{w_f}^{(m)T} \mathbf{z_t}^{[m-1]} + \mathbf{u_f}^{(m)T} \mathbf{z_{t-1}}^{[m]}\right) \in (0,1)^{q_m}$$

$\mathbf{w_f}^{(m)} \in \Re^{(q_{m-1}+1) \times q_m}$ e $\mathbf{u_f}^{(m)} \in \Re^{q_m \times q_m}$

[2]Fonte della Figura: Wüthrich e Merz 2023

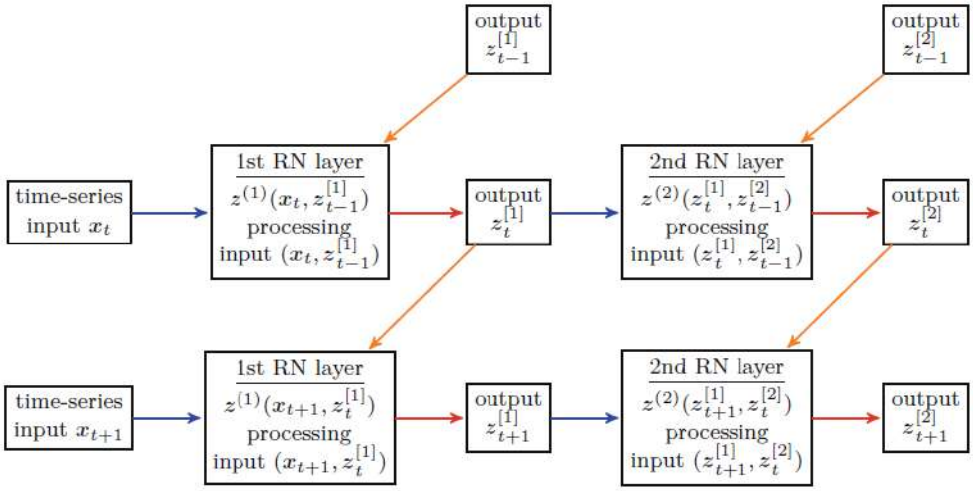

Figura 4.4: Deep RN con profondità di due layer

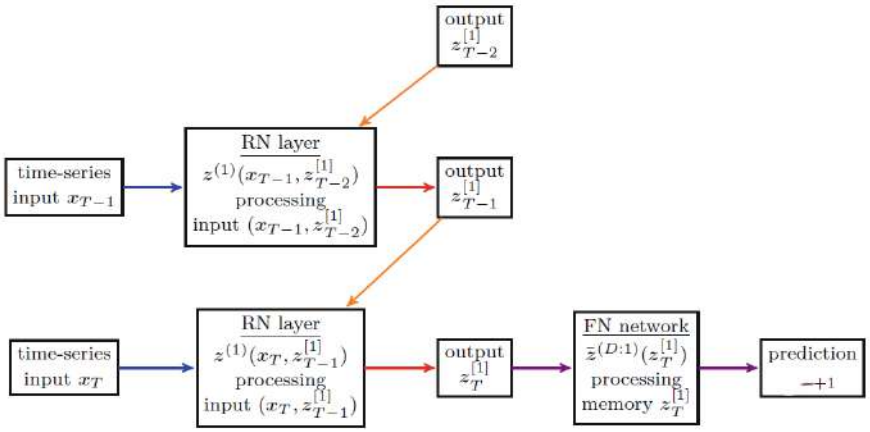

Figura 4.5: RN con profondità di 1 layer con FFN

- *input gate* modella come aggiornare l'informazione da portare nei periodi successivi

$$i_t^{[m]} = i_t^{(m)}(\mathbf{z_t}^{[m-1]}, \mathbf{z_{t-1}}^{[m]}) = \phi_\sigma\left(\mathbf{w_i}^{(m)T}\mathbf{z_t}^{[m-1]} + \mathbf{u_i}^{(m)T}\mathbf{z_{t-1}}^{[m]}\right) \in (0,1)^{q_m}$$

$\mathbf{w_i}^{(m)} \in \Re^{(q_{m-1}+1)\times q_m}$ e $\mathbf{u_i}^{(m)} \in \Re^{q_m \times q_m}$

- *output gate* modella che informazione utilizzare nel periodo t

$$o_t^{[m]} = o_t^{(m)}(\mathbf{z_t}^{[m-1]}, \mathbf{z_{t-1}}^{[m]}) = \phi_\sigma\left(\mathbf{w_o}^{(m)T}\mathbf{z_t}^{[m-1]} + \mathbf{u_o}^{(m)T}\mathbf{z_{t-1}}^{[m]}\right) \in (0,1)^{q_m}$$

$\mathbf{w_o}^{(m)} \in \Re^{(q_{m-1}+1)\times q_m}$ e $\mathbf{u_o}^{(m)} \in \Re^{q_m \times q_m}$

- *cell state process* salva l'informazione da portare nei periodi successivi

$$\begin{aligned} c_t^{[m]} &= c^{(m)}(\mathbf{z_t}^{[m-1]}, \mathbf{z_{t-1}}^{[m]}, c_{t-1}^{[m]}) \\ &= f_t^{[m]} \odot c_{t-1}^{[m]} + i_t^{[m]} \odot \phi_{tanh}\left(\mathbf{w_c}^{(m)T}\mathbf{z_t}^{[m-1]} + \mathbf{u_c}^{(m)T}\mathbf{z_{t-1}}^{[m]}\right) \in \Re^{q_m} \end{aligned}$$

$\mathbf{w_c}^{(m)} \in \Re^{(q_{m-1}+1)\times q_m}$, $\mathbf{u_c}^{(m)} \in \Re^{q_m \times q_m}$ e $\odot$ prodotto elemento per elemento.

l'output dell'm-esimo layer al tempo t è calcolato come:

$$\mathbf{z}_t^{[m]} = z_t^{(m)}(\mathbf{z_t}^{[m-1]}, \mathbf{z_{t-1}}^{[m]}, c_t^{[m]}) = o_t \odot \phi(c_t^{[m]}) \in \Re^{q_m}$$

con ϕ funzione di attivazione generica. Viene anche utilizzato un embeddig layer per i codici dei paesi, l'embedding layer consiste nel moltiplicare una matrice di variabili dummy per ogni classe (in questo caso i 18 paesi) per una matrice di pesi di embedding, lo scopo è quello di passare da uno spazio in $\Re^n$ definito dalle variabili dummy dove le osservazioni distano tutte l'una dall'altra $\sqrt{2}$ in termini di norma euclidea ad uno spazio dove le diverse categorie possono essere più o meno vicine a seconda dell'informazione apportata all'output, il layer è definito nel seguente modo:

$$e_r = \mathbf{oh}'_{d_0}\mathbf{E}_r \in \Re^r$$

dove:

- e_r è un vettore di dimensionalità r (numero di dimensioni in cui si convertono le dummy) in seguito si passerà da 18 paesi a 3 dimensioni di embedding.

- $\mathbf{oh_{d_0}}$ matrice $d_0 \times n_{classi}$ di variabili dummy per ogni classe (d_0 numero di unità statistiche)

- $\mathbf{E}_r$ matrice $n_{classi} \times r$ di pesi di embedding che vanno stimati.

Sostanzialmente gli output di questa procedura vengono trattati come veri input del modello al posto delle classiche variabili dummy.

Per quanto riguarda il caso in esame, la struttura più performante osservata è basata su un LSTM shallow e come attivazione ϕ_{tanh} per la componente relativa ai VA passati con un periodo di *look-back* di 20 e una componente di forward network per integrare i restanti regressori 4.6. I regressori vengono integrati insieme agli output dell'ultimo livello della componente LSTM in un forward layer e tanh come funzione di attivazione. Per finire gli input vengono collegati all'output tramite un ultimo forward layer con attivazione lineare. La loss function da ottimizzare per trovare i pesi ottimi selezionata è l'errore quadratico medio. Com'è noto la minimizzazione dell'errore quadratico medio è connesso all'assunzione di una distribuzione normale della variabile dipendente, questo permette di esprimere la forma distributiva del VA_t dell'i-esimo paese nel seguente modo:

$$VA_t = g(VA_{t-1}, ..., VA_{t-20}, X) + \varepsilon_t \tag{1}$$

con X vettore contenente i regressori diversi dal VA, $\varepsilon_t \sim \mathcal{N}(0, \sigma^2)$ e σ^2 costante in ogni t e g funzione che caratterizza la rete neurale e che rappresenta il valore atteso condizionato all'informazione presente al tempo $t-1$ sotto la misura fisica.

Il dataset originale è suddiviso in tre componenti: training set che contiene tutte le osservazioni 31-08-2017 al 31-01-2024 (vengono escluse le prime 20 osservazioni che fanno parte dei regressori) per un totale di 78 osservazioni per ognuno dei 18 paesi, il validation set contiene 10 osservazioni per ognuno dei 18 paesi che vanno dal 29-02-2024 al 29-11-2024, queste osservazioni servono per valutare la performance della rete per selezionare i pesi ottimi, infine un test set composto da 5 osservazioni per ognuno dei 18 paesi dal 31-12-2024 al 30-04-2025.

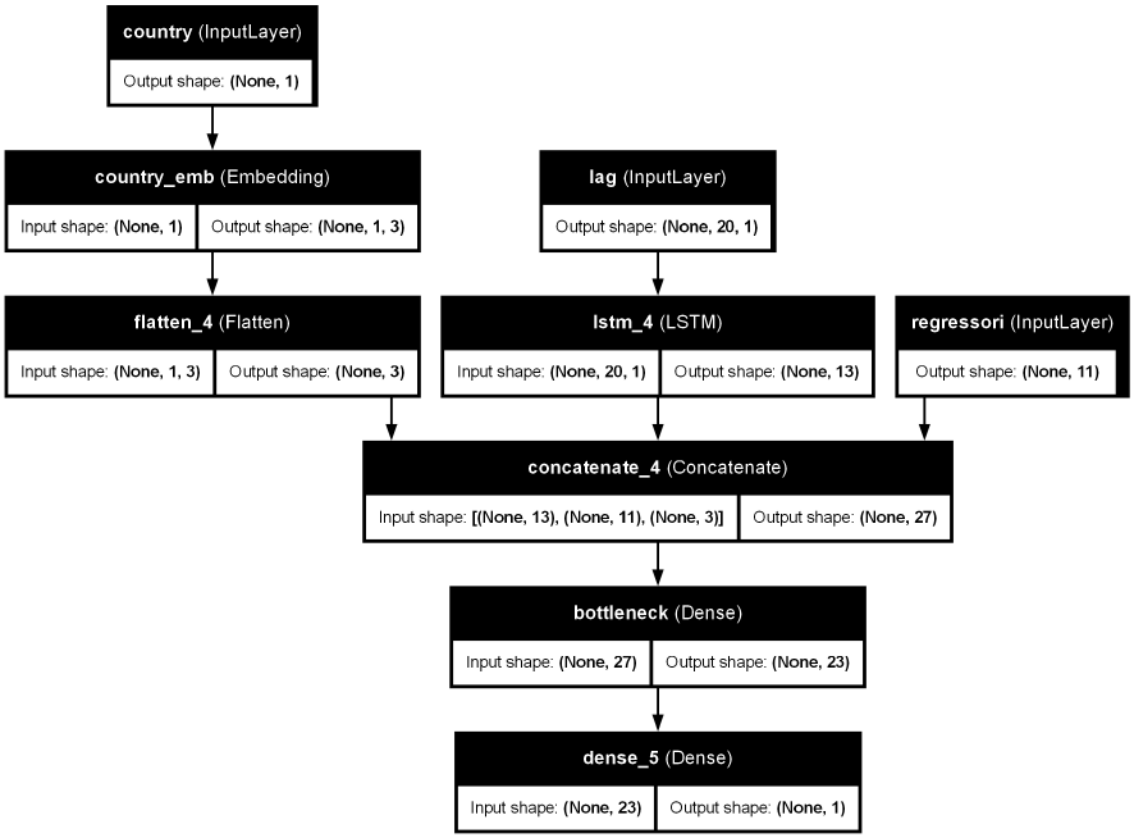

Figura 4.6: Struttura di network utilizzata per il VA.

4.2.2 Calcolo dei pesi ottimi

Come già esposto per trovare i valori ottimi è necessario ottimizzare la funzione obbiettivo, che in questo caso è rappresentata nel seguente modo:

$$\min_{\mathbf{w}} \sum_{i=1}^{18} \sum_{t=0}^{T} (VA_t - g(VA_{t-1}, ..., VA_{t-20}, X))^2. \tag{2}$$

Sussiste però un problema dettato dal teorema di universalità di approssimazione degli FFN dimostrato da Hornik, Stinchcombe e White 1989, il risultato rilevante è che la famiglia di funzioni definita nel seguente modo:

$$A^{q_0} = \{A : \{1\} \times \Re^{q_0} \to \Re; \mathbf{x} \to \mathbf{A}(x) = \mathbf{wx}, \mathbf{w} \in \Re^{q_0+1}\}$$

famiglia contenente tutte le funzioni affini, se su questa si definisce la seguente famiglia di funzioni:

$$\Sigma^{q_0}(\psi) = \{f : \{1\} \times \Re^{q_0} \to \Re; \mathbf{x} \to f(x) = \sum_{j=0}^{q_1} \beta_j \psi(A_j(x)),\ A_j \in A^{q_0}, \beta_j \in \Re, q_1 \in N\}$$

insieme di tutti gli shallow FFN con ψ funzione *squashing*, ossia una funzione tale per cui

$$\lim_{x \to -\infty} \psi(x) = 0, \lim_{x \to +\infty} \psi(x) = 1.$$

Questa famiglia di funzioni contiene qualsiasi funzione continua reale definita su compatti in $\{1\} \times \Re^{q_0}$. Osservando che $(\tanh(x) + 1)/2$ è una funzione *squashing* e che gli ultimi due livelli della Figura 4.6 hanno esattamente la struttura delle funzioni nella famiglia $\Sigma^{q_0}(\psi)$, possiamo concludere dal teorema che la struttura esposta è in grado di approssimare qualsiasi funzione definita su un compatto. Questo è in realtà un grosso problema perchè si sta evidenziando che la rete è in grado di portare la distanza dell'equazione (2) a zero. Da un punto di vista statistico questo porta a problemi di overfitting, ossia l'incapacità del modello di generalizzarsi ai dati al di fuori del training set. Per risolvere questo problema si è infatti suddiviso il dataset in training set e validation set. Si procede sempre minimizzando la (2) sul training set, ma vengono scelti i pesi che minimizzano la perdita sul validation set. Per la selezione degli iperparametri che definiscono la rete, ossia numero di neuroni relativi al layer LSTM e al layer FFN e il learning rate λ da utilizzare nella procedura di ottimizzazione, è stato utilizzato un comando di ricerca su una griglia di possibili valori. Sono stati selezionati gli iperparametri che hanno riportato una performance migliore sul validation set.

4.2.3 Nesterov-accelerated adaptive moment estimation (NADAM)

Per minimizzare la funzione obbiettivo viene utilizzato il metodo NADAM una variante della classica *gradient descent*, questa tecnica è basata sull'assunzione che la funzione obbiettivo sia differenziabile sul dominio definito dai pesi (e che dunque il vettore gradiente esista in tutti i punti di interesse) a questo punto inizializzando il pesi del modello in un punto casuale $\mathbf{w_{start}}$ è sufficiente calcolare il vettore gradiente della funzione di perdita $\nabla f(\mathbf{w_{start}})$ e sfruttare il fatto che il vettore gradiente punta nelle direzione di massima crescita (e quindi spostarsi nella direzione opposta porta nella direzione di massima decrescita per la differenziabilità) ed aggiustare i pesi: $\mathbf{w_{step1}} = \mathbf{w_{start}} - \lambda \nabla f(\mathbf{w_{start}})$ e continuare a eseguire questa procedura fino a che non si raggiunge il punto che trova una validation loss sufficientemente piccola, λ rappresenta un numero piccolo a piacere, più è piccolo più la convergenza è lenta, più è grande più è facile saltare le vicinanze di un potenziale punto di minimo. Il vettore gradiente viene calcolato normalmente usando il concetto di *backpropagation* che consiste nell'applicazione della regola della catena per calcolare le derivate presenti nel vettore gradiente come prodotto delle derivate delle diverse funzioni che compongono la rete; questo concetto è principalmente rilevante dal punto di vista dell'efficienza del codice. Un primo modo per migliorare il metodo esposto in precedenza è aggiungere il *momentum*, per farlo si definisce un parametro μ compreso tra 0 e 1, prendendo il vettore gradiente in un generico step $\nabla f(\mathbf{w_{step_{t-1}}})$ si procede calcolando il momentum come $m_{step_t} = \mu m_{step_{t-1}} + \nabla f(\mathbf{w_{step_{t-1}}})$ e a questo punto si calcola il nuovo set di pesi come $w_{step_t} = w_{step_{t-1}} - \lambda m_{step_t}$ in questo modo nelle direzioni dove si procede nello stesso senso la dimensione dello step aumenta (velocizzando la convergenza), mentre nelle dimensioni dove gli step sono in direzioni contrastanti (ad esempio perchè vicine ad un potenziale minimo) la dimensione dello spostamento si riduce. Il *Nesterov accelerated gradient* consiste nella stessa procedura ma invece che calcolare il gradiente nella posizione dei pesi precedenti si calcola il vettore gradiente in $\nabla f(\mathbf{w_{step_{t-1}}} + \lambda \mu m_{step_{t-1}})$ ossia si sposta ulteriormente il vettore dei pesi nella direzione indotta dalla somma dei vettori gradienti passati, questo dovrebbe migliorare la convergenza. Il resto della procedura è identico al momentum classico, quindi si procede col calcolo del momentum come $m_{step_t} = \mu m_{step_{t-1}} + \nabla f(\mathbf{w_{step_{t-1}}} + \lambda \mu m_{step_{t-1}})$ e $w_{step_t} = w_{step_{t-1}} - \lambda m_{step_t}$. Questo tipo di ottimizzazione è la prima componente di NADAM. La

seconda componente deriva dagli algorimtmi basati su norma L_2, il primo metodo si chiama *adaptive subgradient descent (ADAgrad)* e consiste nello scalare la dimensione di λ in modo che questo diventi più piccolo man mano che ci si avvicina ad un potenziale punto di minimo in modo da ridurre la dimensione degli step. Si calcola $n_{step_t} = n_{step_{t-1}} + \nabla f(\mathbf{w_{step_{t-1}}})^2$ come somma dei valori al quadrato dei vettori (l'operazione di elevamento al quadrato va intesa come elemento per elemento). A questo punto i pesi vengono aggiustati nel seguente modo: $w_{step_t} = w_{step_{t-1}} - \lambda \frac{\nabla f(\mathbf{w_{step_{t-1}}})}{\sqrt{n_{step_t}}+\varepsilon}$ dove l'operazione di frazione va intesa elemento per elemento, ε rappresenta una costante vicina a zero per evitare che il vettore somma dei gradienti al quadrato vada a zero in alcune o tutte le sue componenti. Per come è impostato l'aggiustamento si nota che l'impatto di λ è ridotto tanto più la somma dei vettori gradienti al quadrato è elevata, ossia tanto più quella dimensione è cambiata rispetto al punto di partenza. Il problema di questo metodo è che la somma dei gradienti al quadrato è strettamente crescente e quindi all'aumentare del numero di step inevitabilmente le ampiezze degli spostamenti diventano comunque prossime a zero su tutte le dimensioni, questo blocca la procedura di ottimizzazione. Per risolvere questo problema viene introdotta *RMSProp* simile ad ADAgrad ma con $n_{step_t} = vn_{step_{t-1}} + (1-v)\nabla f(\mathbf{w_{step_{t-1}}})^2$ con v compreso tra 0 e 1, in questo modo si evita che la somma dei vettori gradienti al quadrato collassi a zero introducendo una sorta di meccanismo di memoria. L'aggiustamento dei pesi viene fatto allo stesso modo di ADAgrad $w_{step_t} = w_{step_{t-1}} - \lambda \frac{\nabla f(\mathbf{w_{step_{t-1}}})}{\sqrt{n_{step_t}}+\varepsilon}$. Unendo RMSProp e il momentum classico si ottiene la procedura *ADAM*, questa prevede il calcolo del momentum come $m_{step_t} = \mu m_{step_{t-1}} + (1-\mu)\nabla f(\mathbf{w_{step_{t-1}}})$ e il termine di scala L_2 sempre come $n_{step_t} = vn_{step_{t-1}} + (1-v)\nabla f(\mathbf{w_{step_{t-1}}})^2$ in più si definiscono le seguenti versioni con bias del momentum e del termine di scala come $\hat{m}_{step_t} = \frac{m_{step_t}}{1-\mu^t}$ e $\hat{n}_{step_t} = \frac{n_{step_t}}{1-v^t}$ ma questi aggiustamenti servono solo per evitare che a inizio procedura il termine di scala e il termine di momentum siano troppo piccoli, essendo μ e v compresi tra 0 e 1 il bias sparisce all'aumentare del numero degli step che vengono fatti. I pesi vengono aggiustati come $w_{step_t} = w_{step_{t-1}} - \lambda \frac{\hat{m}_{step_t}}{\sqrt{\hat{n}_{step_t}}+\varepsilon}$. Unendo il Nesterov accelerated gradient e ADAM si ottiene NADAM una procedura definita nel seguente modo, si calcolano sempre il momentum e la componente di scala come in precedenza:$m_{step_t} = \mu m_{step_{t-1}} + (1-\mu)\nabla f(\mathbf{w_{step_{t-1}}})$, $n_{step_t} = vn_{step_{t-1}} + (1-v)\nabla f(\mathbf{w_{step_{t-1}}})^2$ e le loro versioni scalate $\hat{m}_{step_t} = \frac{m_{step_t}}{1-\mu^t}$ e $\hat{n}_{step_t} = \frac{n_{step_t}}{1-v^t}$, si ese-

gue un passaggio simile anche per il gradiente $\hat{g}_{step_t} = \frac{\nabla f(\mathbf{w_{step_{t-1}}})}{1-\mu^t}$ e a questo punto si può calcolare $\bar{m} = (1-\mu)\hat{g}_{step_t} + \mu \hat{m}_{step_t}$. Questo termine serve per calcolare il vettore gradiente al prossimo step come definito nel Nesterov momentum. I pesi vengono aggiustati $w_{step_t} = w_{step_{t-1}} - \lambda \frac{\bar{m}_{step_t}}{\sqrt{\hat{n}_{step_t}}+\varepsilon}$; i riferimenti relativi alle procedure possono essere trovati in Dozat 2016. Il metodo NADAM è stato scelto sulla base dei risultati proposti in Ferrario, Noll e Wuthrich 2020. Nello specifico viene utilizzata la versione stocastica di NADAM che consiste nel calcolare la direzione di spostamento usando un sottocampione del training set, questo viene definito batch, ogni epoca di spostamento è terminata quando dei batch di cui la dimensione totale è pari a quella del campione originale sono usati per calcolare i diversi aggiustamenti dei pesi. Nel caso specifico è stata usata una batchsize di 64 osservazioni. Lo scopo di questa "stima" del vero vettore gradiente è quello di velocizzare la procedura di addestramento. Al termine di ogni epoca viene valutata la funzione di perdita (2) sul validation set. Quando la funzione di perdita non si riduce per 50 epoche successive, vengono selezionati i pesi associati alla perdita sul validation set minima.

4.2.4 Boosting (applicazione di XGBoost a un LSTM)

In questa sezione si prova ad applicare il concetto di boosting a una rete LSTM con lo scopo di provare a migliorarne la performance out of sample. Il boosting in generale consiste nel migliorare la performance di uno stimatore debole (ossia uno stimatore che risulta leggermente meglio di una previsione randomica) sommandolo ad altri stimatori deboli con lo scopo di ottenere uno stimatore efficace. Il boosting solitamente si esegue utilizzando gli alberi di regressione; si procede dunque a delinearne gli aspetti principali. Gli alberi di regressione rappresentano una sorta di metodo di classificazione, ipotizzando di avere uno spazio di caratteristiche per ogni osservazione definito come $\mathbf{x} := (x_1, ..., x_q) \in X$ con X spazio di tutte le potenziali caratteristiche (per semplicità in $\Re$), e una funzione di perdita come obbiettivo nella quale è presente qualche parametro da stimare, negli alberi di regressione si procede suddividendo il campione originale (equivalentemente X) in modo da ottenere una partizione composta da due componenti che minimizzi il più possibile la funzione di perdita. Lo split viene eseguito sulla base di una delle q variabili dello spazio campionario, scegliendo il livello c di quella variabile che suddivide lo spazio campionario in

modo che la somma delle perdite calcolate sui due nuovi campioni sia la più piccola possibile (e più piccola della perdita calcolata sul livello precedente, altrimenti non conviene fare lo split). L'albero è accresciuto ripetendo questa procedura per ogni foglia che viene generata. Per esprimere in maniera matematica ciò che viene fatto ad ogni step:

$$\min_{1\leq l\leq q}\min_{c\in\Re}\left[\min_{\beta_0}\sum_{x_i\in X_t, x_{i,l}\leq c} L(\beta_0)+\min_{\beta_1}\sum_{x_i\in X_t, x_{i,l}>c} L(\beta_1)\right]$$

con:

- q numero totale di caratteristiche
- c livello a cui eseguire lo split sulla caratteristica prescelta
- x_i i-esima unità statistica
- $x_{i,l}$ l-esima caratteristica dell'i-esima unità statistica
- L funzione di perdita obbiettivo
- β_0, β_1 parametri da stimare
- t sequenza di 0 e 1, 0 sta ad indicare foglia destra, 1 sta ad indicare foglia sinistra, rappresenta dunque il percorso seguito per arrivare ad una specifica foglia.
- X_t spazio delle caratteristiche (o unità statistiche) presente nella foglia precedente

L'output di questo split sarà dato da X_{t0}, X_{t0} ossia dalle unità statistiche della foglia sinistra e della foglia destra e dalle stime $\hat{\beta}_0$, $\hat{\beta}_1$. Bisognerebbe definire anche una regola di stop in quanto è possibile dimostrare che la funzione di perdita si va sempre a ridurre fino a che non viene ottenuto un numero di foglie pari al numero di unità statistiche (Wuthrich e Buser 2023), ma ai fini del boosting questo concetto è di minore importanza. Con questo strumento in mente un modo semplice per boostare la rete neurale della figura 4.6 è quello di prendere gli output del modello che si ottengo dal layer bottleneck (23 valori per ogni osservazione) e trattarli come se fossero delle rielaborazioni degli input. A questo punto nell'ultimo layer alla luce della funzione di perdita espressa in (2) è

eseguita una regressione lineare classica dei VA_t su queste rielaborazioni degli input. Prendendo le stime della rete neurale prima del boosting come weak estimator di partenza, è possibile calcolare i residui della rete neurale. Calcolati i residui, è possibile applicare la procedura di boosting a questi residui usando come regressori i 23 output del layer bottleneck della figura 4.6. La procedura consiste nell'applicare degli alberi di regressione ai residui, generalmente questi alberi dovranno dare origine a poche foglie in modo che siano dei *weak learner* (ad esempio si esegue un unico split) e si procede a calcolare le predizioni basate sullo split costruito sugli 8 regressori. Una volta calcolate queste previsioni si ricalcolano i residui tra nuove previsioni e vecchi residui e si ripete la procedura per un numero r di volte. Lo stimatore finale sarà dunque dato dallo stimatore della rete neurale sommato alla somma di tutti gli stimatori della procedura di boosting. Una componente rilevante che differenzia gli alberi di regressione è come vengono calcolate le previsioni. La procedura di boosting selezionata è definita *Extreme gradient boosting (XGBoost)* che appartiene alla famiglia delle tecniche di *gradient boosting* e consiste in una sorta di problema di ottimizzazione funzionale[3]. Ipotizzando di avere una struttura di stimatori definita nel seguente modo:

$$\begin{aligned}
\hat{y}_i^0 &= \hat{VA}_{rete} \\
\hat{y}_i^1 &= \hat{y}_i^0 + f_1(x_i) \\
&\vdots \\
\hat{y}_i^r &= \sum_{k=1}^{r} f_k(x_i) + \hat{VA}_{rete} = \hat{y}_i^{r-1} + f_r(x_i)
\end{aligned}$$

con:

- r: r-esimo passo di boosting.
- x_i: i-esima unità statistica composta nel caso in esame da 23 regressori.
- $f_k(x_i)$: stimatore del residuo al k-esimo passo di boosting.

Si definisce una generica funzione obbiettivo:

$$obj^r = \sum_{i=1}^{n} l(VA_i, \hat{y}_i^r) + \sum_{k=1}^{r} \omega(f_k) \; r = 1, \ldots \tag{3}$$

con:

[3] ricerca della funzione che lega i regressori alla variabile obbiettivo

- l: generica funzione di perdita (nel caso in esame, errore quadratico medio $\frac{1}{n}(VA_i - \hat{y}_i^r)^2$). Si osserva che per $r = 1$ $\frac{1}{n}(VA_i - \hat{y}_i^1)^2 = \frac{1}{n}(VA_i - \hat{VA}_{rete} - f_1(x_i))^2 = \frac{1}{n}(\text{residui rete} - f_1(x_i))^2$

- ω: rappresenta un termine di penalizzazione della complessità dell'albero.

Ipotizzando ora di essere al r-esimo passo di boosting, vengono assunti noti tutti i valori nei r-1 passi precedenti ed è possibile riscrivere la (3) in modo da esplicitare le incognite da stimare al t-esimo passo:

$$obj^r = \sum_{i=1}^{n} l(VA_i, \hat{y}_i^{r-1} + f_r(x_i)) + \omega(f_r) + costante \tag{4}$$

A questo punto prendendo l'espansione di Taylor al secondo ordine di (4) in $\hat{y}_i^{r-1}$ possiamo ottenere un'equazione quadratica da ottimizzare per calcolare $f_r(x_i)$ ottimo:

$$obj^r \approx \sum_{i=1}^{n} \left[l(VA_i, \hat{y}_i^{r-1}) + g_i f_r(x_i) + \frac{1}{2} h_i f_r(x_i)^2 \right] + \omega(f_r) + costante \tag{5}$$

con:

- $g_i = \partial_{\hat{y}_i^{r-1}} l(VA_i, \hat{y}_i^{r-1})$

- $h_i = \partial^2_{\hat{y}_i^{r-1}} l(VA_i, \hat{y}_i^{r-1})$

Si è assunto che la funzione di perdita sia derivabile due volte e che f_r sia un numero di dimensione non eccessivamente grande. Rimuovendo le costanti che non hanno impatto sul risultato di ottimizzazione possiamo riscrivere la funzione obbiettivo (5) come:

$$obj^r \approx \sum_{i=1}^{n} \left[g_i f_r(x_i) + \frac{1}{2} h_i f_r(x_i)^2 \right] + \omega(f_r) \tag{6}$$

Definendo la stima dell'albero di regressione $f_r = w_{q(x)}$, $w \in \Re$, $q : \Re^{23} \to 1, ..., T$ con $q(x)$ funzione che definisce in che foglia terminale finisce la generica unità statistica, T numero complessivo di foglie finali. w definisce il valore reale da assegnare alla foglia assegnata da $q(x)$. In XGBoost

il termine di complessità è definito come $\omega(f) = \gamma T + \frac{1}{2}\lambda \sum_{j=1}^{T} w_j^2$[4] da cui la (6) diventa:

$$obj^r \approx \sum_{i=1}^{n} \left[g_i w_{q(x_i)} + \frac{1}{2} h_i w_{q(x_i)}^2 \right] + \gamma T + \frac{1}{2}\lambda \sum_{j=1}^{T} w_j^2. \tag{7}$$

Indicando con $I_j = \{i | q(x_i) = j\}$ ossia insieme di unità statistiche assegnate alla stessa foglia, così facendo la (7) si può riscrivere come:

$$obj^r \approx \sum_{j=1}^{T} \left[w_j \sum_{i \in I_j} g_i + \frac{1}{2} w_j^2 \left(\lambda + \sum_{i \in I_j} h_i \right) \right] + \gamma T. \tag{8}$$

Ora è possibile definire il valore ottimale della foglia ottimizzando la (8)

$$\begin{aligned} \partial_{w_j} obj^r &= 0 \\ \sum_{i \in I_j} g_i + w_j \left(\lambda + \sum_{i \in I_j} h_i \right) &= 0 \\ G_j + w_j \left(\lambda + H_j \right) &= 0 \\ w_j^* &= -\frac{G_j}{\left(\lambda + H_j \right)} \forall j \in 1, \ldots, T \end{aligned}$$

dove si è posto $G_j := \sum_{i \in I_j} g_i$, $H_j := \sum_{i \in I_j} h_i$. Da cui la (8):

$$obj^{r*} \approx -\frac{1}{2} \sum_{j=1}^{T} \frac{G_j^2}{\left(H_j + \lambda \right)} + \gamma T. \tag{9}$$

A questo punto la variabile (nel caso in esame tra le 23) e il livello della variabile vanno scelti a ogni passo dell'albero in modo tale da ridurre il più possibile la (9), oppure, più semplicemente, basta controllare il valore del miglioramento di perdita ottenuto eseguendo lo split. Ipotizzando di essere in una foglia generica t il valore della perdita contenuto nella foglia secondo la (9) sarà $obj_t^{r*} \approx -\frac{1}{2}\frac{G^2}{(H+\lambda)} + \gamma$ dopo uno split la funzione di perdita del passo t si dividerà in $t0$,$t1$ con una loss complessiva di: $obj_{t0+t1}^{r*} \approx -\frac{1}{2}\left(\frac{G_{sinistra}^2}{(H_{sinistra}+\lambda)} + \frac{G_{destra}^2}{(H_{destra}+\lambda)} \right) + \gamma 2$ a questo punto si calcola il *gain*

[4] Il termine di complessità è legato al tema di selezione dell'albero ottimale in quanto, come evidenziato in Wuthrich e Buser 2023, la funzione di perdita si ridurrebbe sempre ad ogni split, lo scopo di questo termine è porre un limite al numero di split.

come:

$$
\begin{aligned}
Gain &= -(obj^{r*}_{t0+t1} - obj^{r*}_{t}) = \frac{1}{2}\left(\frac{G^2_{sinistra}}{(H_{sinistra}+\lambda)} + \frac{G^2_{destra}}{(H_{destra}+\lambda)} - \frac{G^2}{(H+\lambda)}\right) - \gamma \quad (10)\\
&= \frac{1}{2}\left(\frac{G^2_{sinistra}}{(H_{sinistra}+\lambda)} + \frac{G^2_{destra}}{(H_{destra}+\lambda)} - \frac{(G_{destra}+G_{sinistra})^2}{((H_{destra}+H_{sinistra})+\lambda)}\right) - \gamma. \quad (11)
\end{aligned}
$$

Se il Gain è positivo, si esegue lo split, se è negativo, allora non conviene eseguire lo split e l'albero smette di crescere su quella foglia. Nel caso in esame è stata utilizzata una procedura di ottimizzazione di alcuni iperparametri connessi al boosting, in particolare si costruisce una griglia con i seguenti valori:

- nestimators: [100, 200, 500], ossia si provano a combinare 100, 200 o 500 stimatori.
- max depth: [2, 3, 4, 5], ossia la massima profondità del singolo albero di boosting, ad esempio una profondità massima di 2 implica che al più ci saranno 2 split con al più quattro foglie per ogni albero costruito durante la procedura di boosting.
- learning rate: [0.01, 0.05, 0.1], è il tasso di apprendimento η. La sua applicazione consiste nel moltiplicare gli stimatori di ogni passaggio di boosting per questa costante prima di sommarli alla stima precedente e di calcolare i nuovi residui da inserire nel passaggio di boosting successivo.

  $$
  \begin{aligned}
  \hat{y}^0_i &= \hat{VA}_{rete}\\
  \hat{y}^1_i &= \hat{y}^0_i + \eta f_1(x_i)\\
  &\vdots\\
  \hat{y}^r_i &= \sum_{k=1}^{r} \eta f_k(x_i) + \hat{VA}_{rete} = \hat{y}^{r-1}_i + \eta f_r(x_i)
  \end{aligned}
  $$

  Lo scopo è quello di indebolire gli stimatori.
- subsample: [0.7, 0.9, 1.0], che percentuale del training set selezionare randomicamente ad ogni passo di boosting (per ogni albero si pesca un campione), se $= 1$ si utilizza tutto il campione. Lo scopo è quello di indebolire ulteriormente lo stimatore proveniente dal singolo albero.

- colsample bytree: [0.7, 0.9, 1.0], rappresenta la percentuale di regressori da utilizzare nello specifico passo di boosting, ad ogni passo si pescano randomicamente dei regressori, se $= 1$ si usano sempre tutti i regressori. Lo scopo è ancora una volta quello di rendere gli stimatori deboli ed evitare l'overfitting.

- early stopping rounds: fissato a 50 e rappresenta il parametro per cui se non si ha un miglioramento della funzione di perdita sul validation set per 50 passaggi di boost consecutivi, allora la procedura di boosting viene terminata e vengono sommati solo gli stimatori fino all'ultimo per cui c'è stato un miglioramento della funzione di perdita sul validation set. In ogni caso il numero massimo di stimatori sommabili è definito dalla max depth.

I migliori parametri vengono selezionati attraverso la 3k cross-validation, ossia il trianing set viene suddiviso in 3 blocchi di dimensione identica e i parametri vengono provati sul 2 blocchi e validati sul terzo, una volta che questa procedura è fatta utilizzando come set di valutazione ognuno dei tre blocchi si ottiene una perdita complpessiva associata al set di parametri come media delle 3 perdite calcolate. Si sceglie il set di parametri con la perdita media più piccola. Il validation set durante il fit di ogni iperparametro su ogni combinazione di blocchi è usato per valutare l'early stopping della procedura allo scopo sempre di evitare l'overfitting.

4.2.5 Forecasting

L'ultima componente riguarda il calcolo del valore atteso condizionato nelle date future. Per poter eseguire il forecasting delle osservazioni nelle cinque date di test è necessario eseguire il forecasting anche dei regressori selezionati. Per i seguenti regressori: VSTOXX, cambio euro dollaro, gas naturale, Brent e rendimenti corporate sono selezionati dei modelli specifici scegliendo tra strutture ARMA, ARCH, GARCH. La scelta del miglior modello è eseguita prendendo il modello con AIC e BIC più bassi. Per quanto riguarda i rendimenti degli indici azionari dei paesi e l'S&P500 è usato un modello LSTM insieme al layer di Embedding per indicare il paese di riferimento e cercare di estrarre dell'informazione crossezionale. Anche per HICP, yield a 5 anni dei titoli governativi, spread 1-5 e spread 5-10 vengono utilizzati delle strutture di LSTM con inserimento dell'embedding layer per tenere conto del paese di riferimento, per la stima degli iperparametri viene utilizzata

la stessa procedura discussa in precedenza basata sulla selezione dei parametri con minor perdita sul validation set. La scelta di usare dei modelli di predizione basati su rete è stata dettata dalla piccola dimensione del dataset a disposizione del singolo paese preso in considerazione. Relativamente ai modelli dove non si sono usate reti neurali si è aumentata la dimensione del dataset. Prima di analizzare i diversi modelli si procede con una breve descrizione delle caratteristiche dei principali modelli classici utilizzati.

ARMA(p,q)

I modelli ARMA(p,q) possono essere espressi nella seguente formula:

$$X_t = \mu + \sum_{i=1}^{p} \phi_i X_{t-i} + \sum_{j=1}^{q} \theta_j \varepsilon_{t-j} + \varepsilon_t \tag{12}$$

con ε_t IID normali standard. Una soluzione unica e stazionaria all'equazione esiste se e solo se le radici dell'equazione $1 - \sum_{i=1}^{p} \phi_i z^p = 0$ si trovano al di fuori del cerchio unitario **Pelagatti2015UC**.

ARCH(p)

I modelli ARCH servono per catturare le dinamiche di serie storiche dove la varianza evolve nel tempo, un problema presente molto frequentemente nelle serie storiche finanziarie.

$$X_t = \mu + \sigma_t \varepsilon_t \tag{13}$$

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{p} \alpha_i X_{t-i}^2 \tag{14}$$

con $\alpha_0 > 0$ e $\alpha_i > 0 \ \ \forall i = 1, ..., p$.

GARCH(p,q)

I modelli GARCH servono per catturare le dinamiche di serie storiche dove la varianza evolve nel tempo e presenta anche del clustering.

$$X_t = \mu + \sigma_t \varepsilon_t \tag{15}$$

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{p} \alpha_i X_{t-i}^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2 \tag{16}$$

con $\alpha_0 > 0$ e $\alpha_i > 0 \forall i = 1, ..., p$ e $\beta_j > 0 \forall j = 1, ..., q$, vengono anche controllati modelli misti del tipo AR-ARCH/GARCH dove μ ha una dinamica del tipo $\mu_t = w + \phi X_{t-1}$. Per tutti questi modelli la stima dei parametri è eseguita utilizzando la massima verosimiglianza:

$$f_{X_0,...,X_n}(X_0, ..., X_n) = f_{X_0}(X_0) \prod_{i=1}^{n} f_{X_t|X_{t-1},...,X_0}(X_t|X_{t-1}, ..., X_0).$$

Viene sfruttata l'assunzione di normalità di ε per definire la struttura di f, n rappresenta la dimensione del campione. Nel caso della stima di massima verosimiglianza del modello GARCH è necessario anche definire un valore per σ_0^2, nei casi in esame viene utilizzata la varianza empirica calcolata sul campione $\sigma_0^2 := \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2$.

VSTOXX

Nel prevedere il VSTOXX si estende il dataset per la stima prendendo le osservazioni mensili da partire dal 04-01-1999 e si usano come training set le osservazioni fino al 29-11-2024. Vengono utilizzati l'*Augemented dickey fuller test* (ADF)[5] e il test per l'eteroschedasticità di Engle (HET)[6], il primo per vedere se è necessario integrare la serie per ottenere la stazionarietà in media, il secondo per vedere se c'è stazionarietà in varianza (ipotesi nulla). Se il test viene rifiutato, allora sarà necessario utilizzare un modello ARCH o GARCH per modellare la varianza condizionale. Nel caso in esame è rifiutata sia la nulla dell'ADF sia la nulla dell'HET, questo implica stazionarietà in media e necessità di utilizzare modelli per catturare la varianza condizionale. Per cercare di capire la potenziale struttura della media vengono utilizzati il grafico della funzione di autocorrelazione 4.8 e il grafico della funzione di autocorrelazione parziale 4.9, i due grafici suggeriscono una struttura AR 1 per la media, ma vengono comunque testati altri modelli.
Vengono comunque testati altri modelli e valutati guardando l'AIC e il BIC. Alla luce della tabella

[5]adfuller della libreria statsmodels in python esegue il test per identificare presenza di radice unitaria di Dickey Fuller su $\Delta y_t = \psi y_{t-1} + \sum_{i=1}^{p} \alpha_i \Delta y_{t-i} + u_t$ con $\Delta y_t = y_t - y_{t-1}$ e $p = \text{int}\left(12(\frac{n}{100})^{\frac{1}{4}}\right)$

[6]het_arch della libreria statsmodels in python, esegue il test F sui coefficienti angolari della regressione $y_t^2 = \alpha_0 + \alpha_1 y_{t-1}^2 + ... + \alpha_q y_{t-q}^2 + \varepsilon_t$ sotto le ipotesi classiche OLS; q è scelto come $q = \min(\text{int}\left(12\left(\frac{n}{100}\right)^{\frac{1}{4}}\right), \text{int}\left(\frac{n-1}{2}\right))$ con n lunghezza del dataset in input.

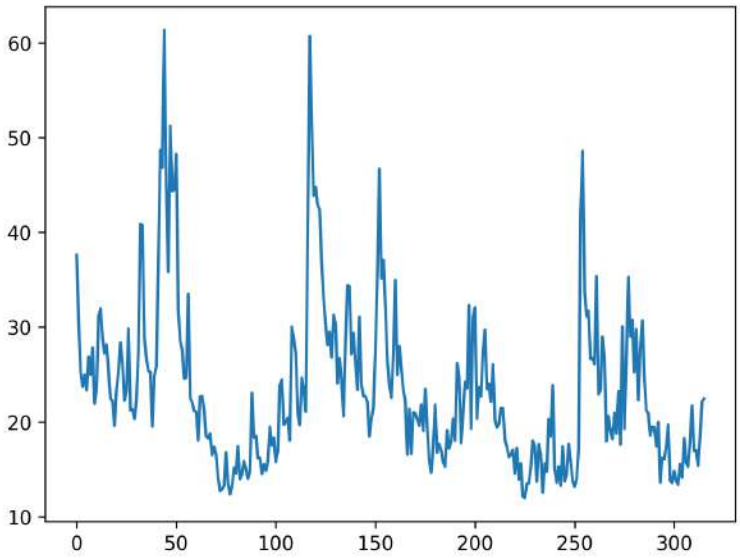

Figura 4.7: Serie storica del VSTOXX mensile

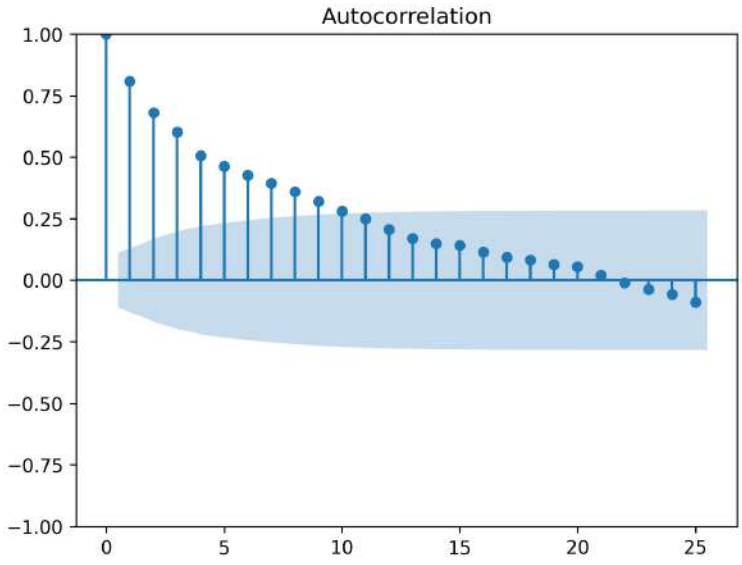

Figura 4.8: ACF del VSTOXX

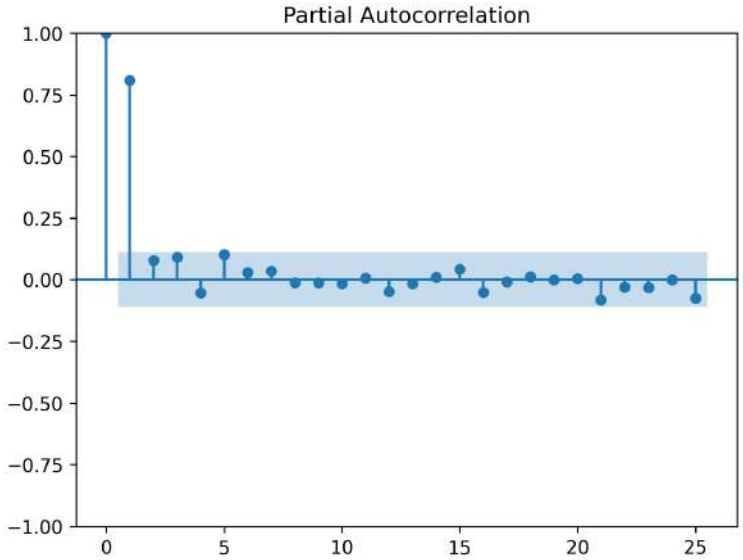

Figura 4.9: PACF del VSTOXX

Tabella 4.1: Confronto modelli: AIC e BIC per VSTOXX

Modelli	AIC	BIC
AR(1) ARCH(1)	1849.519747	1864.466036
AR(1) GARCH(1,1)	1843.656589	1862.339451
ARCH(1,1)	2066.091031	2077.310410
AR(1)	1891.261512	1902.480891
ARMA(1,1)	1891.049751	1906.008923

4.7[7] il miglior modello è il GARCH con media AR1. I risultati del modello sono riportati nella tabella 4.2.

Tabella 4.2: Summary del modello AR(1) GARCH(1,1) per la serie del VSTOXX

```
                           AR - GARCH Model Results
==============================================================================
Dep. Variable:                      y   R-squared:                       0.656
Mean Model:                        AR   Adj. R-squared:                  0.655
Vol Model:                      GARCH   Log-Likelihood:               -916.828
Distribution:                  Normal   AIC:                           1843.66
Method:            Maximum Likelihood   BIC:                           1862.34
                                        No. Observations:                  310
Date:                Sun, Aug 17 2025   Df Residuals:                      308
Time:                        21:55:55   Df Model:                            2
                              Mean Model
=====================================================================
                 coef    std err          t      P>|t|  95.0% Conf. Int.
---------------------------------------------------------------------
Const          3.9112      0.851      4.594  4.356e-06 [  2.242,  5.580]
y[1]           0.8130  3.630e-02     22.396 4.345e-111 [  0.742,  0.884]
                             Volatility Model
==========================================================================
                 coef    std err          t      P>|t|     95.0% Conf. Int.
--------------------------------------------------------------------------
omega         10.6517      3.888      2.740  6.152e-03    [  3.031, 18.272]
alpha[1]       0.3044      0.118      2.580  9.869e-03  [7.319e-02,  0.536]
beta[1]        0.2773      0.149      1.863  6.244e-02 [-1.441e-02,  0.569]
==========================================================================

Covariance estimator: robust
```

Tutti i parametri stimati sono significativi, si procede con il forecasting per le 5 date successive al 29/11/2024 4.10. Il valore atteso condizionato del VSTOXX verrà utilizzato per fare il forecasting del volatility adjustment nei diversi paesi.

[7] Per ARCH(1,1) si intende un modello ARCH con media costante.

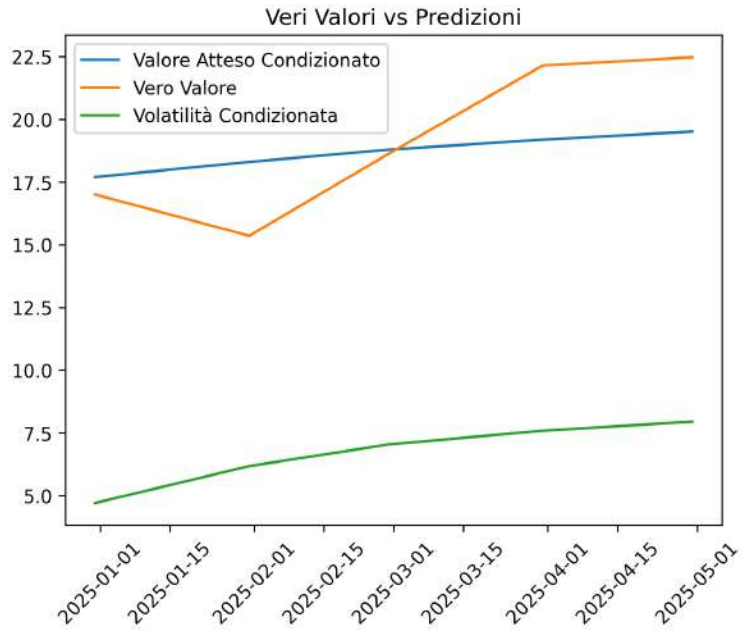

Figura 4.10: Previsioni del VSTOXX con frequenza mensile

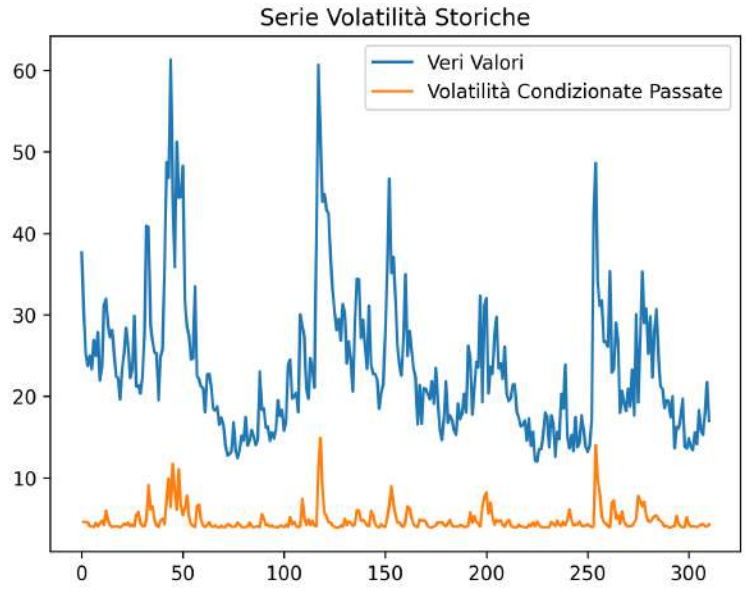

Figura 4.11: Grafico con volatilità condizionata storica implicata dal modello AR(1)-GARCH(1,1) (linea arancione) e serie storica del VSTOXX (linea blu)

Brent

Per quanto riguarda il Brent viene seguita una procedura simile. Il dataset parte dal 30/06/2004 e finisce al 30/04/2025, ancora una volta il training set si estende fino al 29/11/2024. La serie storica è riportata in Figura 4.13. Eseguendo l'ADF risulta che la serie è stazionaria ma il p-value è del 2%,

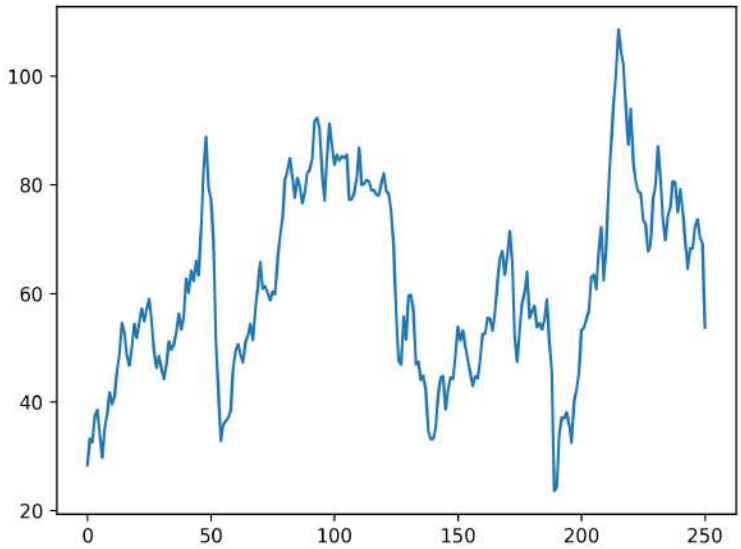

Figura 4.12: Serie storica prezzo Brent in euro

si decide di integrarla lo stesso e operare con la serie integrata. eseguendo nuovamente l'ADF e l'HET entrambi rifiutano l'ipotesi nulla, dunque si ha una serie stazionaria in media ma che richiede l'utilizzo di modelli ARCH/GARCH per modellare la varianza condizionale. Si controlla anche in questo caso l'ACF e il PACF 4.14, 4.15, i due grafici suggeriscono una struttura ARMA(1,1) per la media. Come in precedenza vengono testati diversi modelli e selezionato il migliore secondo il criterio dell'AIC e BIC 4.3. Si sceglie di utilizzare il modello GARCH(1,1) con media costante, in quanto i coefficienti ARMA(1,1) non risultano significativamente diversi da zero; i parametri associati al modello sono riportati in Figura: 4.4. Anche se il β_1 non è significativo, si è comunque deciso di tenere questo modello. Le predizioni delle differenze saranno tutte identiche e uguali a μ. Per ottenere i valori originali si è presa l'osservazione al 29/11/2024 e si è sommata alla prima differenza ottenendo il valore attuale, ripetendo questa procedura per le successive osservazioni si ottengono le predizioni del livello del Brent. Sui rendimenti del Brent è successivamente applicato

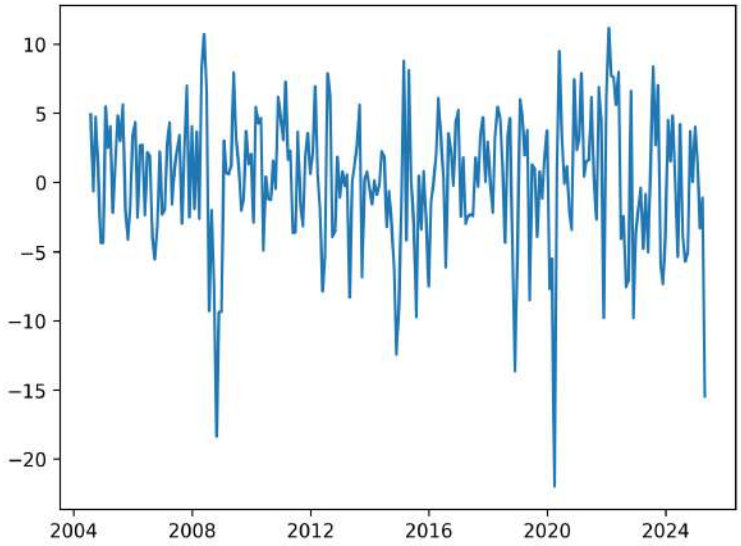

Figura 4.13: Serie storica della differenza di prezzo del Brent in dollari

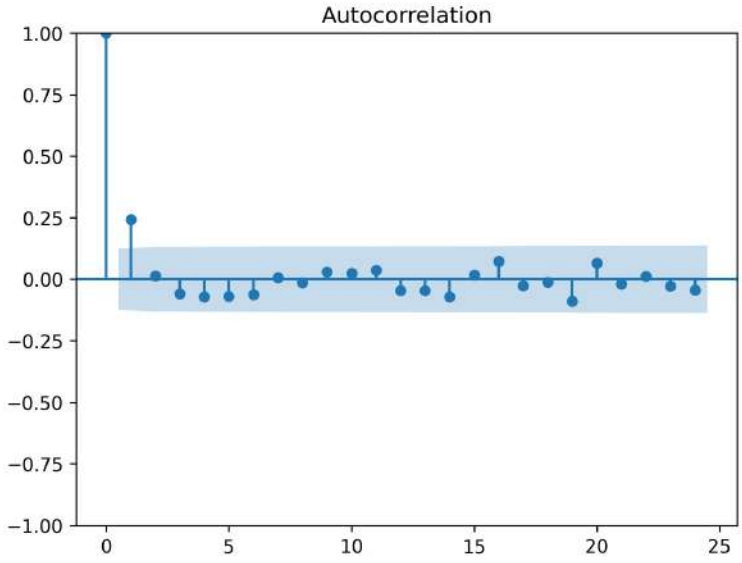

Figura 4.14: ACF delle differenze del prezzo del Brent

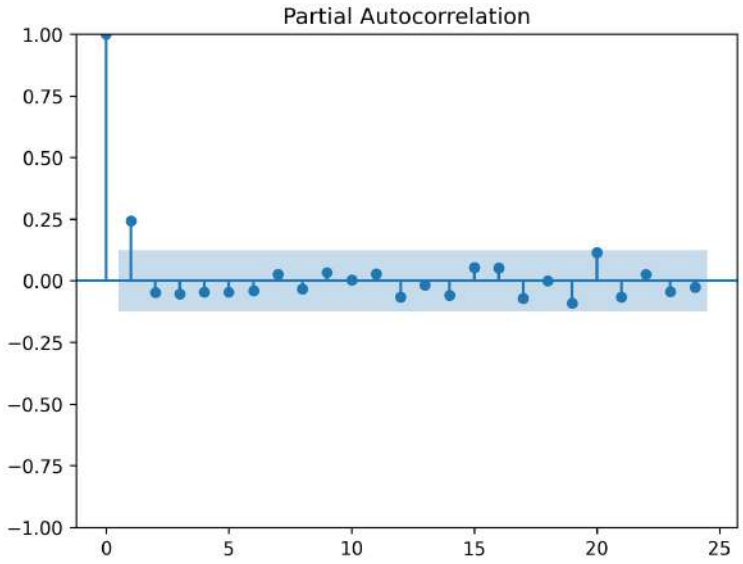

Figura 4.15: PACF delle differenze del prezzo del Brent

Tabella 4.3: Confronto modelli: AIC e BIC

Modelli	AIC	BIC
ARMA(1,1)	1464.085583	1478.090616
AR(1)	1462.783923	1473.287697
ARCH(1)	1459.036288	1469.540063
GARCH(1,1)	1455.653772	1469.658805

Tabella 4.4: Summary del modello GARCH(1,1)

```
                     Constant Mean - GARCH Model Results
==============================================================================
Dep. Variable:                      y   R-squared:                       0.000
Mean Model:             Constant Mean   Adj. R-squared:                  0.000
Vol Model:                      GARCH   Log-Likelihood:               -723.827
Distribution:                  Normal   AIC:                           1455.65
Method:            Maximum Likelihood   BIC:                           1469.66
                                        No. Observations:                  245
Date:                Mon, Aug 18 2025   Df Residuals:                      244
Time:                        09:25:20   Df Model:                            1
                                Mean Model
========================================================================
                 coef    std err          t      P>|t|     95.0% Conf. Int.
------------------------------------------------------------------------
mu             0.4995      0.278      1.798  7.212e-02 [-4.489e-02,  1.044]
                              Volatility Model
========================================================================
                 coef    std err          t      P>|t|     95.0% Conf. Int.
------------------------------------------------------------------------
omega          9.0919      2.381      3.819  1.342e-04    [  4.425, 13.758]
alpha[1]       0.4346      0.193      2.251  2.439e-02  [5.617e-02,  0.813]
beta[1]        0.2311      0.154      1.502      0.133 [-7.053e-02,  0.533]
========================================================================

Covariance estimator: robust
```

un modello GARCH(1,1) per predire la volatilità dei rendimenti e utilizzare questi valori in un secondo momento per creare una versione alternativa della rete; i risultati sono riportati nelle Figure: 4.16, 4.17.

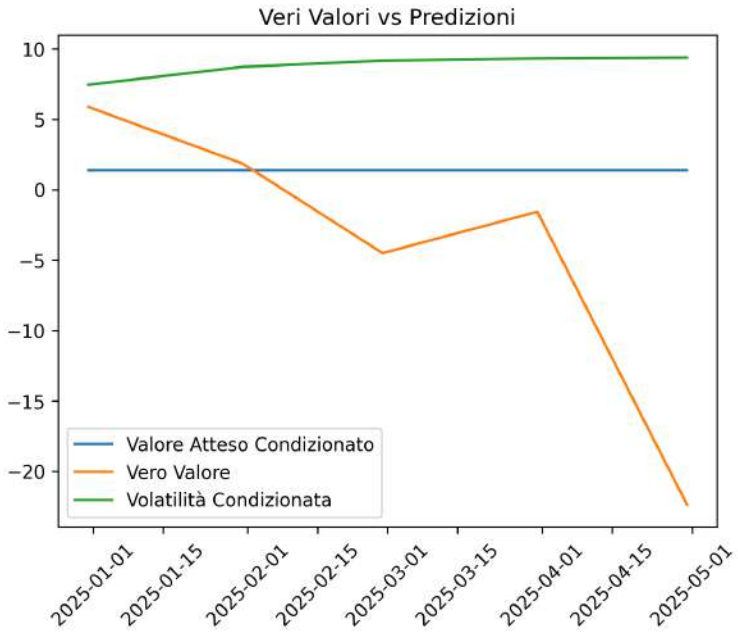

Figura 4.16: Previsioni della media (che è costante, linea blu) e volatilità (linea verde) dei rendimenti del Brent

Gas Naturale

Anche in questo caso si aumenta il dataset partendo dal 31/07/2000 al 30/04/2025 con training set fino al 29/11/2024; la serie è riportata in Figura:4.18. Anche in questo caso viene fatto l'ADF test e risulta che la serie non è stazionaria con un p-value del 7% superiore al 5%. Si procede dunque integrando la serie 4.19. Ancora una volta si ricalcolano l'ADF e l'HET entrambi rifiutano l'ipotesi nulla al 5% il che implica stazionarietà e la necessità di utilizzare un modello ARCH/GARCH per modellare la varianza condizionale. Si procede ancora una volta con i grafici dell'ACF e PACF. Guardando i grafici 4.20 e 4.21 non sembrerebbe essere suggerita nessuna struttura relativamente alla media condizionale, è presente qualche correlazione significativa ma leggermente al di sopra delle bande di significatività, al più una struttura ARMA(1,1) o ARMA(2,2). Si procede con il

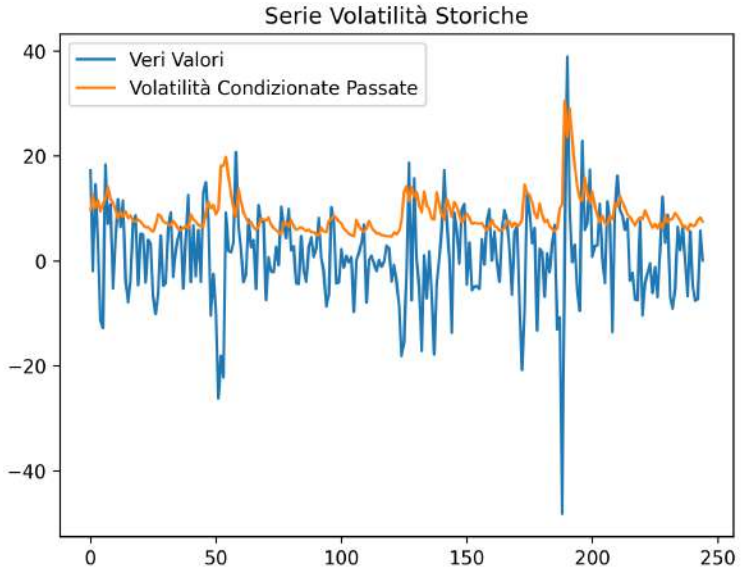

Figura 4.17: Volatilità storiche dei rendimenti del Brent secondo modello GARCH(1,1)

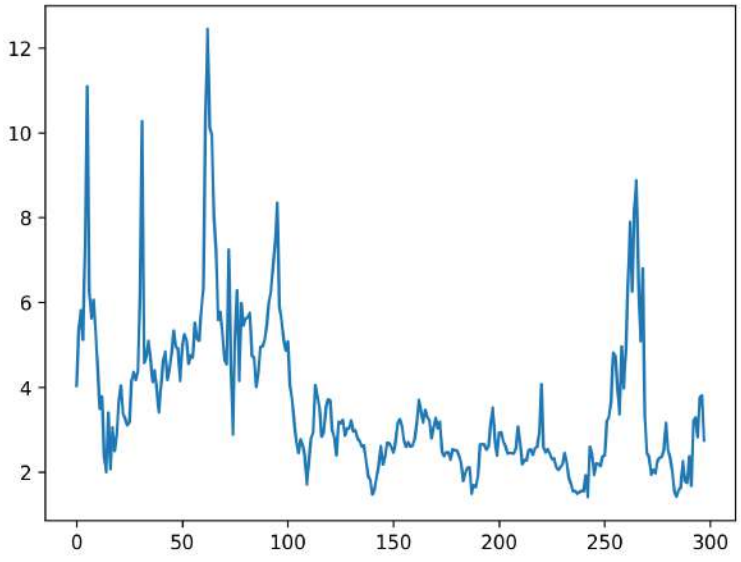

Figura 4.18: Serie storica del prezzo del gas naturale in dollari

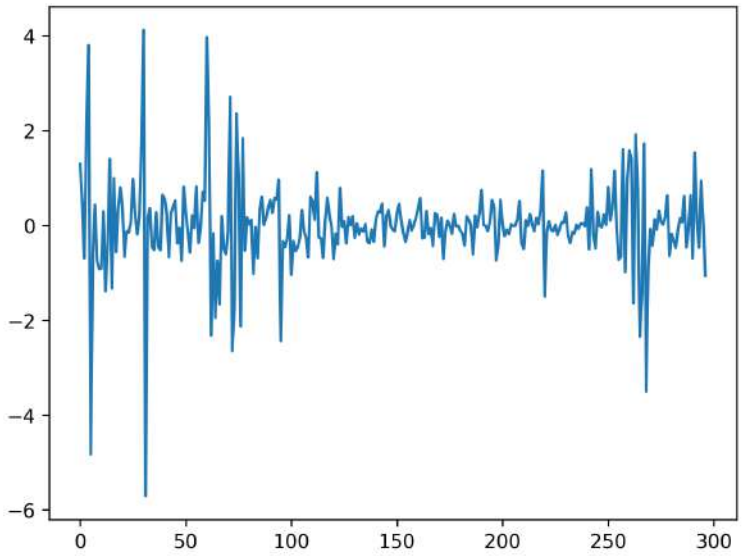

Figura 4.19: Serie storica delle differenze di prezzo del gas naturale

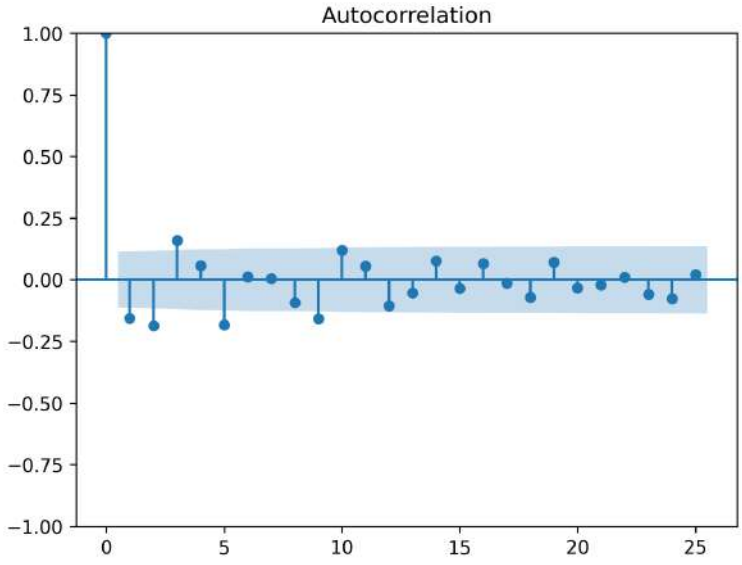

Figura 4.20: ACF delle differenze del prezzo del gas naturale

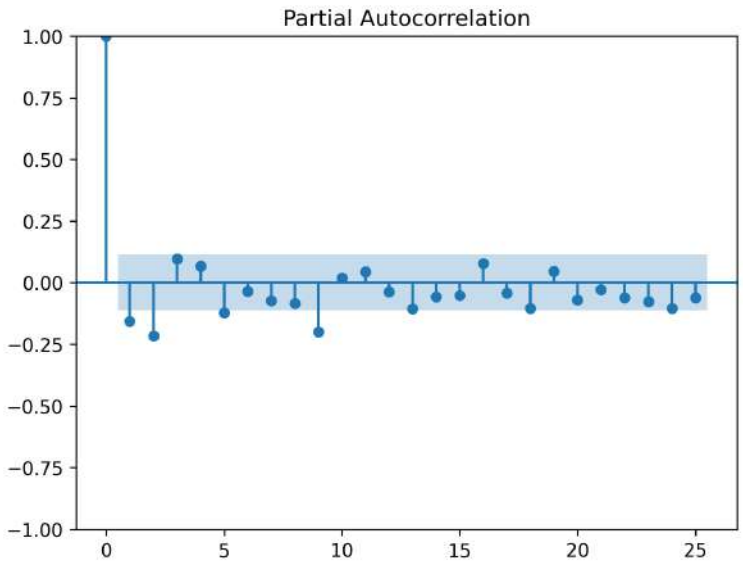

Figura 4.21: PACF delle differenze del prezzo del gas naturale

controllo su diversi modelli 4.5 Si decide di utilizzare il modello AR(1) GARCH(1,1) di cui il i

Tabella 4.5: Confronto modelli: AIC e BIC differenze di prezzo del Brent

Modelli	AIC	BIC
GARCH(1,1)	593.026318	607.733334
ARCH(1)	616.957016	627.987277
Zero GARCH(1,1)	591.390802	602.421064
Zero ARCH(1)	615.467092	622.820600
AR1 GARCH(1,1)	584.629054	602.995670
AR1 ARCH(1)	596.598871	611.292164
ARMA(1,1)	780.442433	795.149448

risultati sono riportati in Figura 4.6. Si procede eseguendo le previsioni per le 5 date successive come in precedenza nella Figura: 4.22. Come nel caso del Brent, i livelli di prezzo vengono estratti sommando l'ultima osservazione del training set alla prima predizione per ottenere il primo livello; si ripete ciclicamente la procedura per le 5 date. Anche in questo caso ai rendimenti del Brent viene applicato un modello Zero ARCH(1) per estrarre le volatilità dei rendimenti storici e le predizioni delle volatilità da usare successivamente 4.16, 4.24.

Tabella 4.6: Summary del modello AR(1) GARCH(1,1) per differenze di prezzo gas naturale

```
                          AR - GARCH Model Results
==============================================================================
Dep. Variable:                      y   R-squared:                       0.024
Mean Model:                        AR   Adj. R-squared:                  0.021
Vol Model:                     GARCH   Log-Likelihood:               -287.315
Distribution:                 Normal   AIC:                           584.629
Method:           Maximum Likelihood   BIC:                           602.996
                                       No. Observations:                  291
Date:               Mon, Aug 18 2025   Df Residuals:                      289
Time:                       10:31:18   Df Model:                            2
                                 Mean Model
============================================================================
                 coef    std err          t      P>|t|       95.0% Conf. Int.
----------------------------------------------------------------------------
Const         -0.0160  3.692e-02     -0.434      0.665 [-8.836e-02,5.635e-02]
y[1]          -0.1655  7.254e-02     -2.282  2.250e-02   [ -0.308,-2.334e-02]
                             Volatility Model
==========================================================================
                 coef    std err          t      P>|t|     95.0% Conf. Int.
--------------------------------------------------------------------------
omega          0.0514  5.210e-02      0.987      0.324 [-5.071e-02,  0.154]
alpha[1]       0.5287      0.183      2.896  3.779e-03    [  0.171,  0.886]
beta[1]        0.4713      0.213      2.214  2.680e-02  [5.417e-02,  0.889]
==========================================================================

Covariance estimator: robust
```

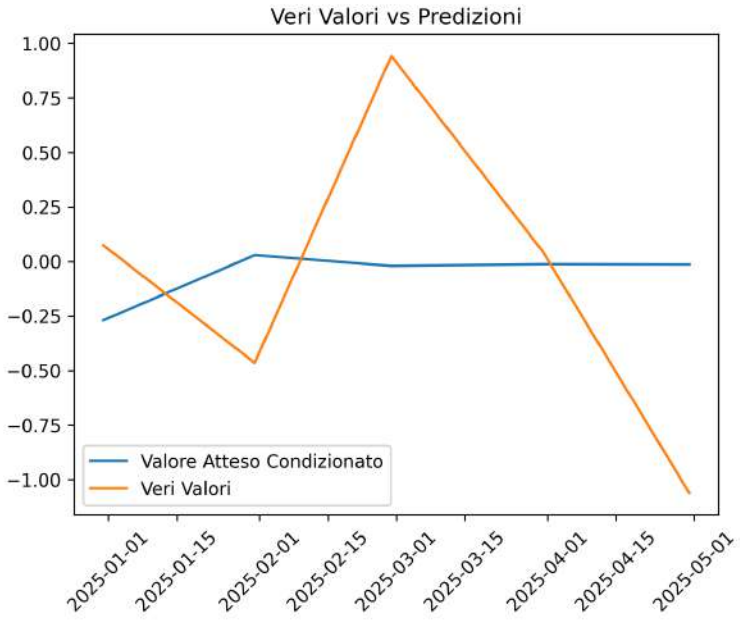

Figura 4.22: Previsioni delle differenze di prezzo del gas naturale

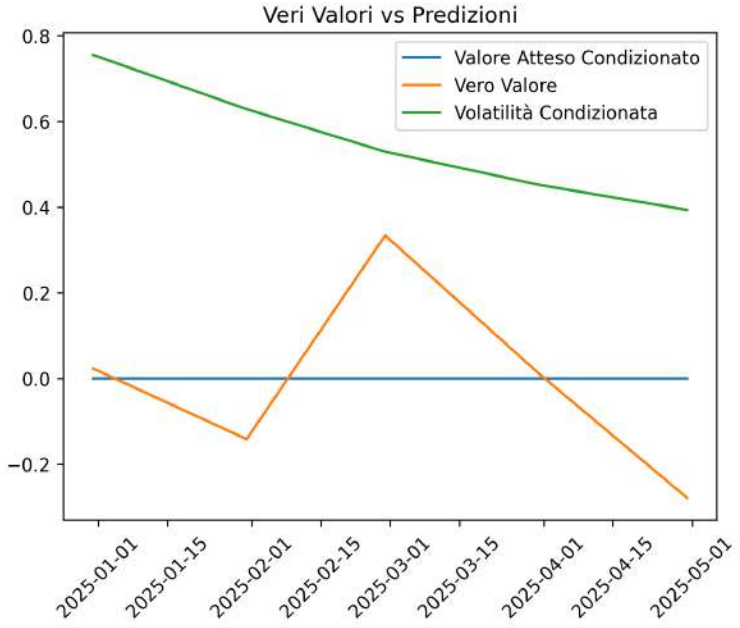

Figura 4.23: Previsioni dei rendimenti (linea blu) e delle volatilità (linea verde) dei rendimenti del gas naturale secondo il modello Zero ARCH(1).

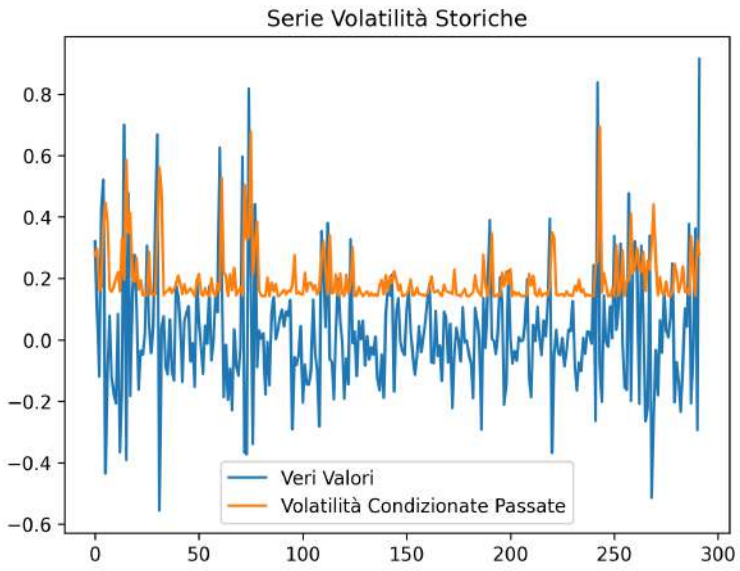

Figura 4.24: Grafico volatilità storiche previste dal modello Zero ARCH(1) per i rendimenti del gas naturale.

Cambio Euro-Dollaro

Il cambio tiene conto di quanti dollari può comprare 1 euro ogni mese. La serie storica inizia dal 31/07/2000 e termina sempre il 30/04/2025; il training set è sempre definito fino al 29/11/2024. La serie storica è riportata in Figura: 4.25 Ancora una volta con l'ADF si ottiene un p-value del 19%, non si rifiuta l'ipotesi nulla dunque non si può rifiutare l'ipotesi che la serie non sia stazionaria, si procede a integrarla, la serie integrata è riportata in Figura: 4.26. Rifacendo il test ADF e HET si rifiuta in entrambi i casi l'ipotesi nulla, dunque di nuovo si ha una serie stazionaria con necessità di utilizzo di modelli ARCH/GARCH per modellare la varianza condizionale. Si procede con l'analisi dei grafici ACF 4.27 e PACF 4.28. Osservando i grafici non sembrerebbe esserci nessuna struttura nella media condizionale, si procede fittando modelli con media nulla. Questo implica predizioni costanti del livello del cambio. I livelli di AIC e BIC sono riportati nella tabella 4.7, viene selezionato il modello GARCH(1,1) con media zero, i risultati sono riportati in 4.8. Le predizioni sulle 5 date successive risulteranno costanti a zero. Per ottenere i livelli del cambio euro dollaro si ripete la medesima procedura utilizzata nei casi precedenti. Anche in questo caso si è applicato un modello Zero ARCH(2) ai rendimenti per estrarre le volatilità passate e prevederle

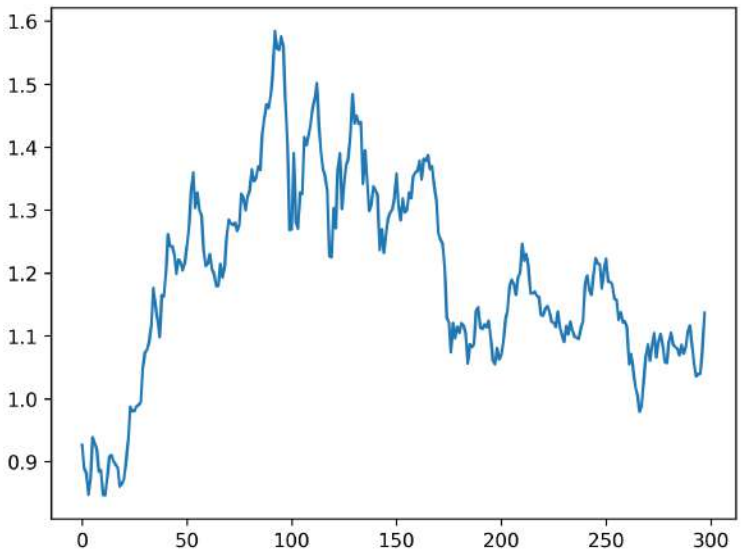

Figura 4.25: Serie storica del cambio euro-dollaro

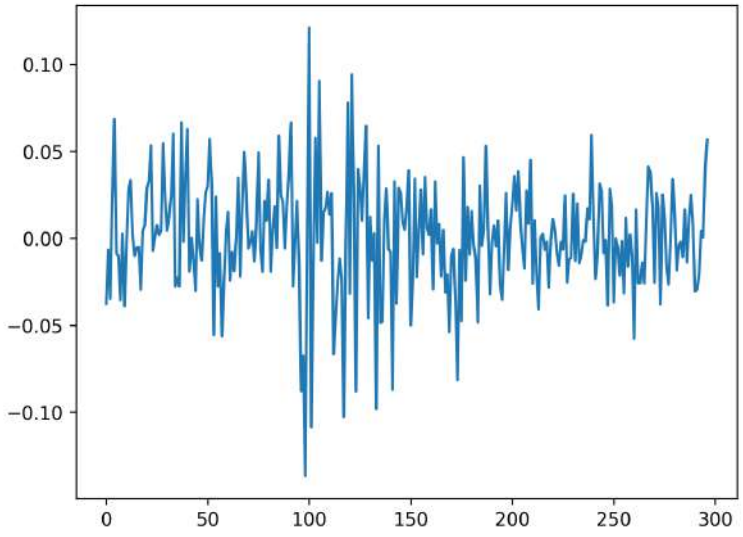

Figura 4.26: Serie storica delle differenze nel cambio euro-dollaro

Tabella 4.7: Confronto modelli: AIC e BIC per differenza di cambio euro-dollaro

modelli	AIC	BIC
Zero GARCH(2,2)	-1196.838613	-1178.454844
Zero ARCH(2)	-1196.838613	-1178.454844
Zero GARCH(1,1)	-1199.752426	-1188.722165
Zero GARCH(1,2)	-1198.412393	-1183.705378

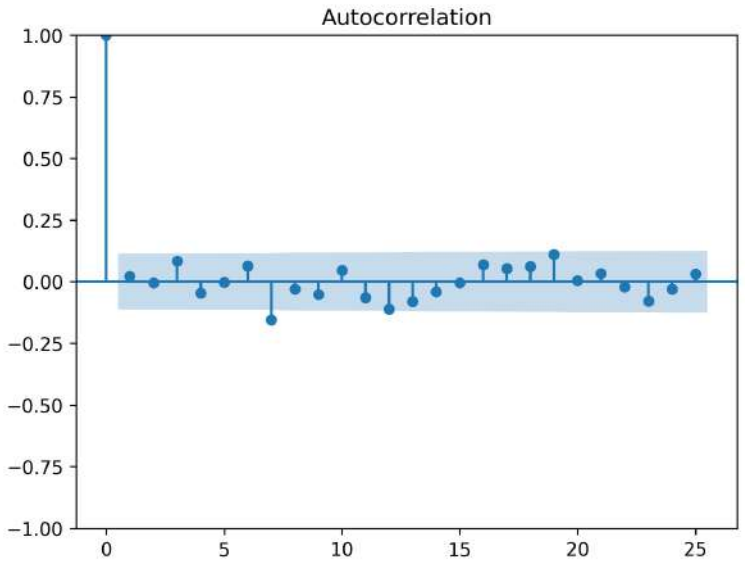

Figura 4.27: ACF delle differenze nel cambio euro-dollaro

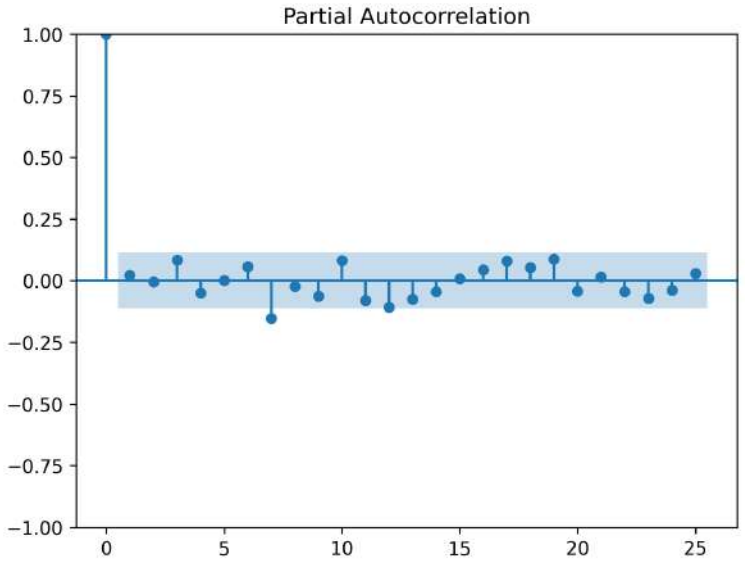

Figura 4.28: PACF delle differenze nel cambio euro-dollaro

Tabella 4.8: Summary del modello Zero GARCH(1,1) per cambio euro-dollaro

```
                       Zero Mean - GARCH Model Results
==============================================================================
Dep. Variable:                      y   R-squared:                       0.000
Mean Model:                 Zero Mean   Adj. R-squared:                  0.003
Vol Model:                      GARCH   Log-Likelihood:                602.876
Distribution:                  Normal   AIC:                          -1199.75
Method:            Maximum Likelihood   BIC:                          -1188.72
                                        No. Observations:                  292
Date:                Mon, Aug 18 2025   Df Residuals:                      292
Time:                        11:26:11   Df Model:                            0
                             Volatility Model
============================================================================
                 coef    std err          t      P>|t|      95.0% Conf. Int.
----------------------------------------------------------------------------
omega      1.9355e-05  6.491e-06      2.982  2.865e-03 [6.633e-06,3.208e-05]
alpha[1]       0.0885  4.065e-02      2.178  2.941e-02   [8.859e-03,  0.168]
beta[1]        0.8937  3.104e-02     28.791 2.778e-182     [  0.833,  0.955]
============================================================================

Covariance estimator: robust
```

per le 5 date successive 4.29, 4.30.

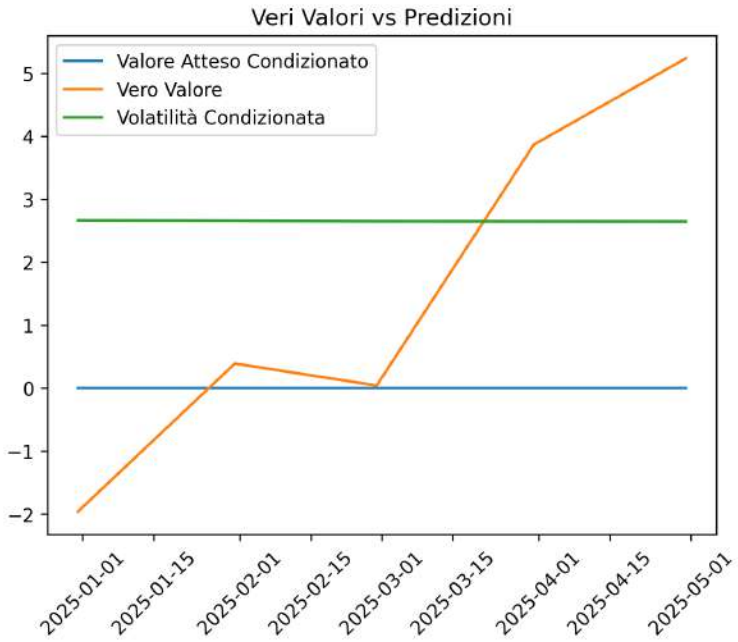

Figura 4.29: Previsioni del valore atteso (linea blu) e della volatilità condizionata (linea verde) nei rendimenti del cambio euro-dollaro

Rendimenti Corporate

La serie storica che va dal 30/10/2009 al 30/04/2025 è riportata nella figura 4.31, il training set si estende fino al 29/11/2024 Sia l'ADF che l'HET rifiutano l'ipotesi nulla suggerendo nuovamente stazionarietà in media e non stazionarietà in varianza. I grafici di ACF 4.32 e PACF 4.33 non suggeriscono nessun tipo di struttura sulla media, ancora una volta si procede a verificare AIC e BIC di modelli GARCH e ARCH con media zero o media costante. Gli AIC e BIC sono riportati in figura 4.9 Viene scelto il modello con BIC più basso ossia lo Zero GARCH(1,1), questo perchè la media del modello a media costante e i parametri del modello AR(1) GARCH(1,1) non risultavano significativamente diversi da zero. I risultati del modello sono riportati nella Figura: 4.10. Le predizioni e le volatilità previste sono riportate in figura 4.34, mentre nella figura 4.35 , sono riportate le volatilità passate implicite nel modello Zero GARCH(1,1)

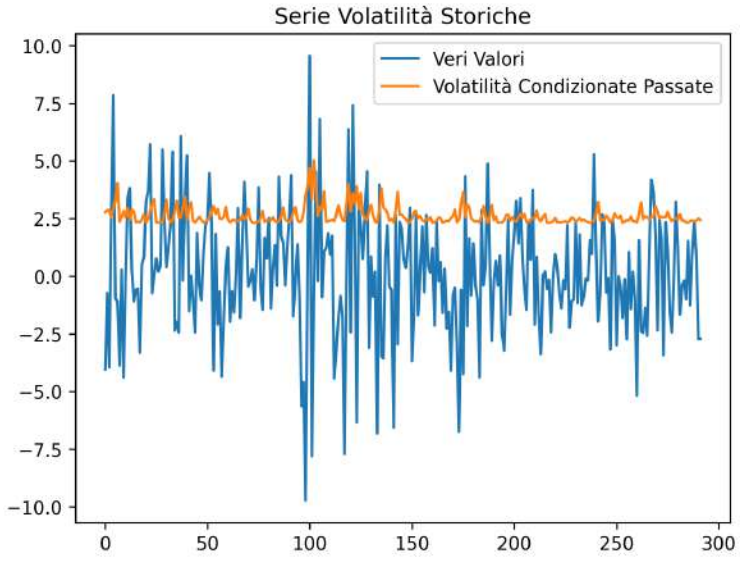

Figura 4.30: Volatilità storiche (linea arancione) implicite nella serie dei rendimenti del cambio Eur-Usd (linea blu)

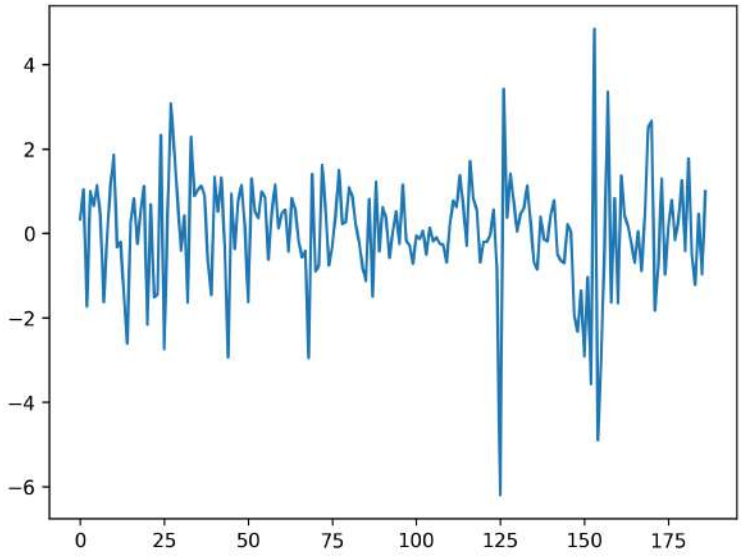

Figura 4.31: Rendimenti mensili del fondo iShares Core EUR Corp Bond.

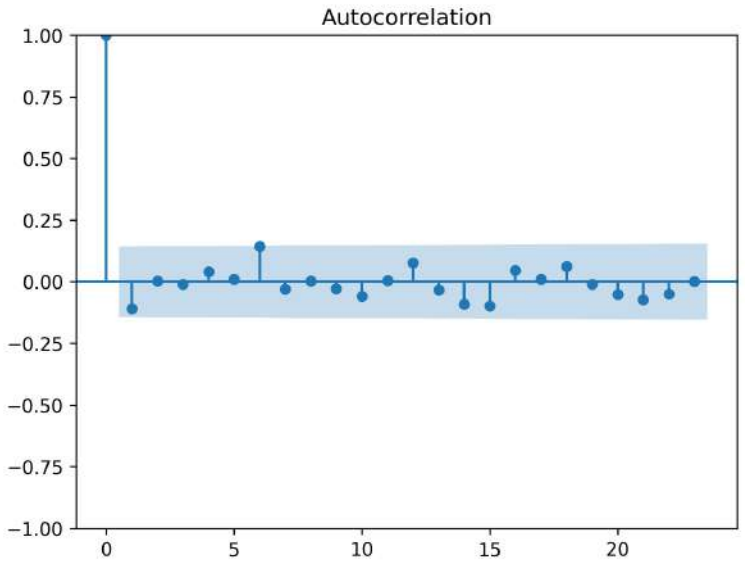

Figura 4.32: ACF dei rendimenti corporate.

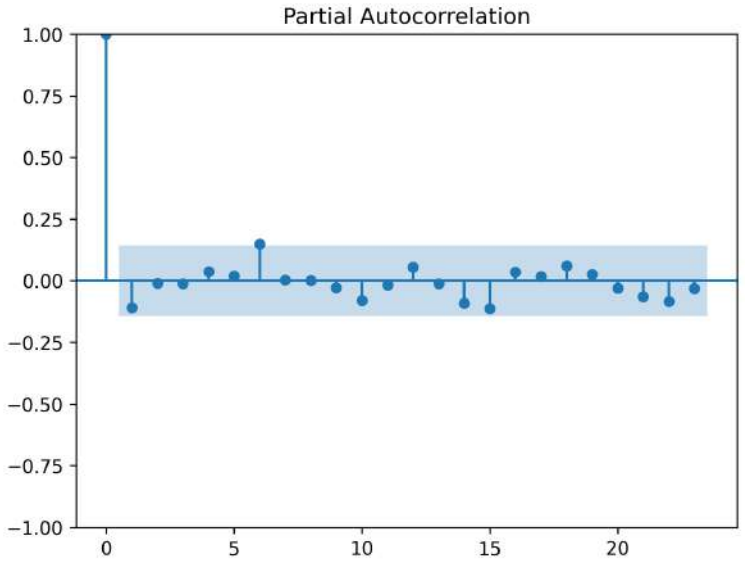

Figura 4.33: PACF dei rendimenti corporate.

Tabella 4.9: Confronto modelli: AIC e BIC

Modelli	AIC	BIC
AR(1)_GARCH(1,1)	601.154387	617.146872
GARCH(1,1)	601.767436	614.583463
Zero_GARCH(1,1)	602.521880	612.133900

Tabella 4.10: Summary del modello Zero GARCH(1,1) per i rendimenti corporate

```
                       Zero Mean - GARCH Model Results
==============================================================================
Dep. Variable:                      y   R-squared:                       0.000
Mean Model:                 Zero Mean   Adj. R-squared:                  0.005
Vol Model:                      GARCH   Log-Likelihood:               -298.261
Distribution:                  Normal   AIC:                           602.522
Method:            Maximum Likelihood   BIC:                           612.134
                                        No. Observations:                  182
Date:                Mon, Sep 01 2025   Df Residuals:                      182
Time:                        08:37:08   Df Model:                            0
                             Volatility Model
=====================================================================
                 coef    std err          t      P>|t|    95.0% Conf. Int.
---------------------------------------------------------------------
omega          0.6625      0.308      2.148  3.170e-02 [5.806e-02,  1.267]
alpha[1]       0.3724      0.135      2.763  5.736e-03   [  0.108,  0.637]
beta[1]        0.2882      0.116      2.474  1.337e-02 [5.985e-02,  0.516]
=====================================================================

Covariance estimator: robust
```

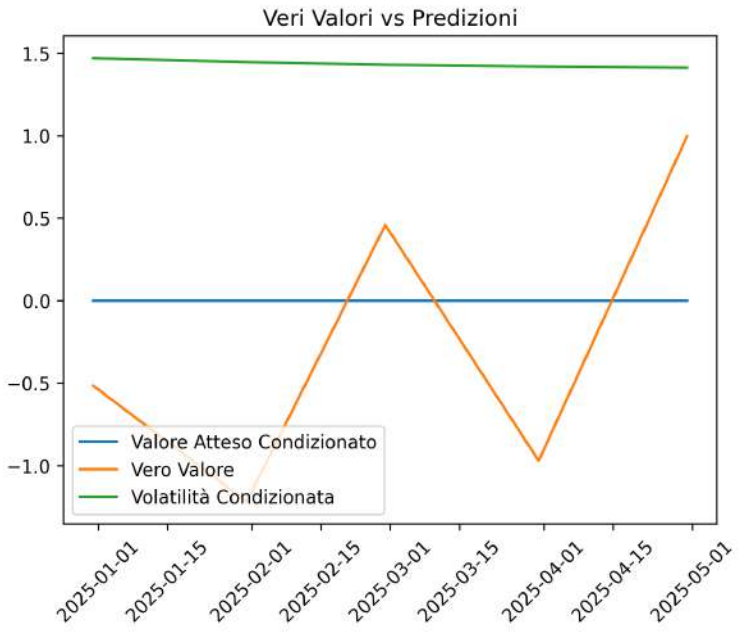

Figura 4.34: Forecast della media e della volatilità dei rendimenti corporate.

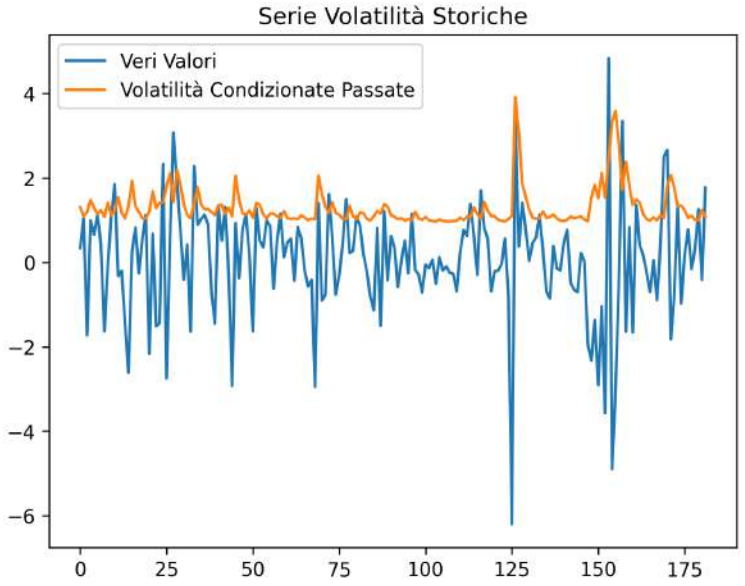

Figura 4.35: Volatilità storiche dei rendimenti corporate.

HICP

Come già anticipato per l'HICP dei 18 paesi verrà utilizzata una rete LSTM con utilizzo di embedding per la componente paese. Il modello migliore ottenuto controllando diversi iperparametri ha la struttura nella figura 4.36, il metodo di ottimizzazione utilizzato è lo stesso descritto in precedenza e utilizzato per il VA. La performance complessiva sul test set dei pesi selezionati è data da MSE di 0.0158, un MAE di 0.0961, un R^2 di 0.1743, un r di Pearson di 0.6726. La performance solo sulle prime predizioni complessiva è di MSE: 0.0035, MAE: 0.0445, r di Pearson: 0.9284 Le predizioni della rete sono riportate in Figura: 4.37 Come per i VA anche in questo caso gli input sono stati riscalati tra -1 e 1 per facilitare l'apprendimento della rete. Per ottenere le osservazioni nella scala originale è sufficiente invertire la scala.

Rendimenti indici + rendimenti S&P500

Anche per i rendimenti si esegue una procedura molto simile a quanto esposto in precedenza, ma si aggiunge anche l'S&P500 ai diversi rendimenti, rappresentativo del paese USA, in questo modo si crea un set di predizioni anche per questo regressore. La struttura del modello è riportata in figura 4.38. La performance dei pesi selezionati sul test set complessivo è un MSE di 0.0622, un MAE di 0.1944, un R^2 di 0.1577 e un r di Pearson di 0.4438. La performance complessiva sulle prime predizioni è MSE:0.0130, MAE:0.0964 e r di Pearson: 0.8522 .I valori predetti dalla rete per i 18 paesi sul training, validation e forecastati sul test set sono riportati in figura 4.39, i risultati per S&P500 invece in Figura: 4.40. Anche in questo caso i valori sono stati scalati tra -1 e 1.

Yield 5 anni

In questa sezione viene eseguita la medesima procedura per gli yield a 5 anni. Il modello selezionato è quello di figura 4.41 Le metriche sul test set complessivo delle predizioni sono: MSE di 0.0386, MAE di 0.1699, R^2 di 0.7367 e r di Pearson di 0.9630. La performance sulle prime predizioni dei 18 paesi è di MSE: 0.0178, MAE:0.1184, r di Pearson: 0.9828. Il grafico delle previsioni del modello riportato in Figura: 4.42.

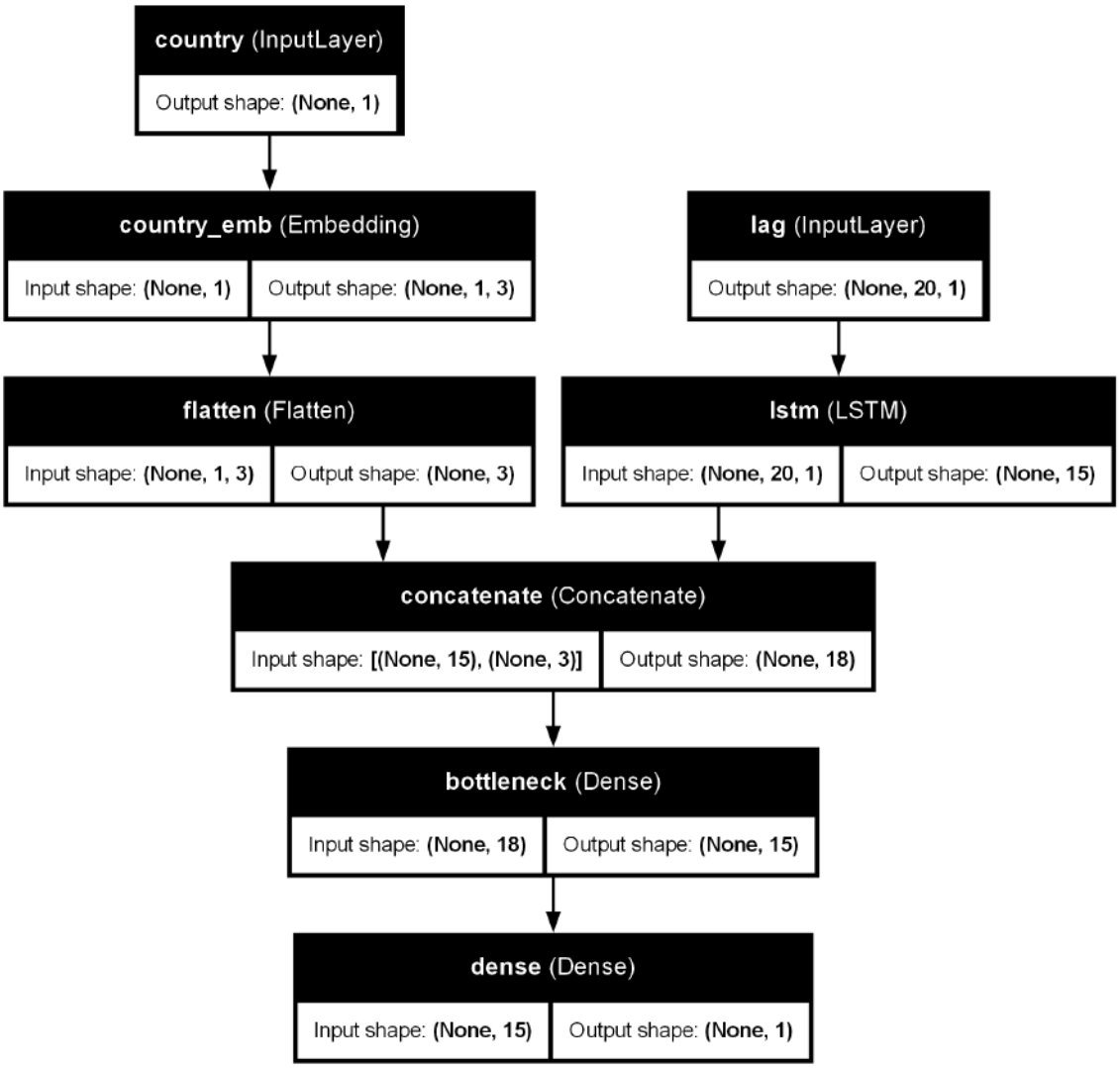

Figura 4.36: Struttura Rete HICP

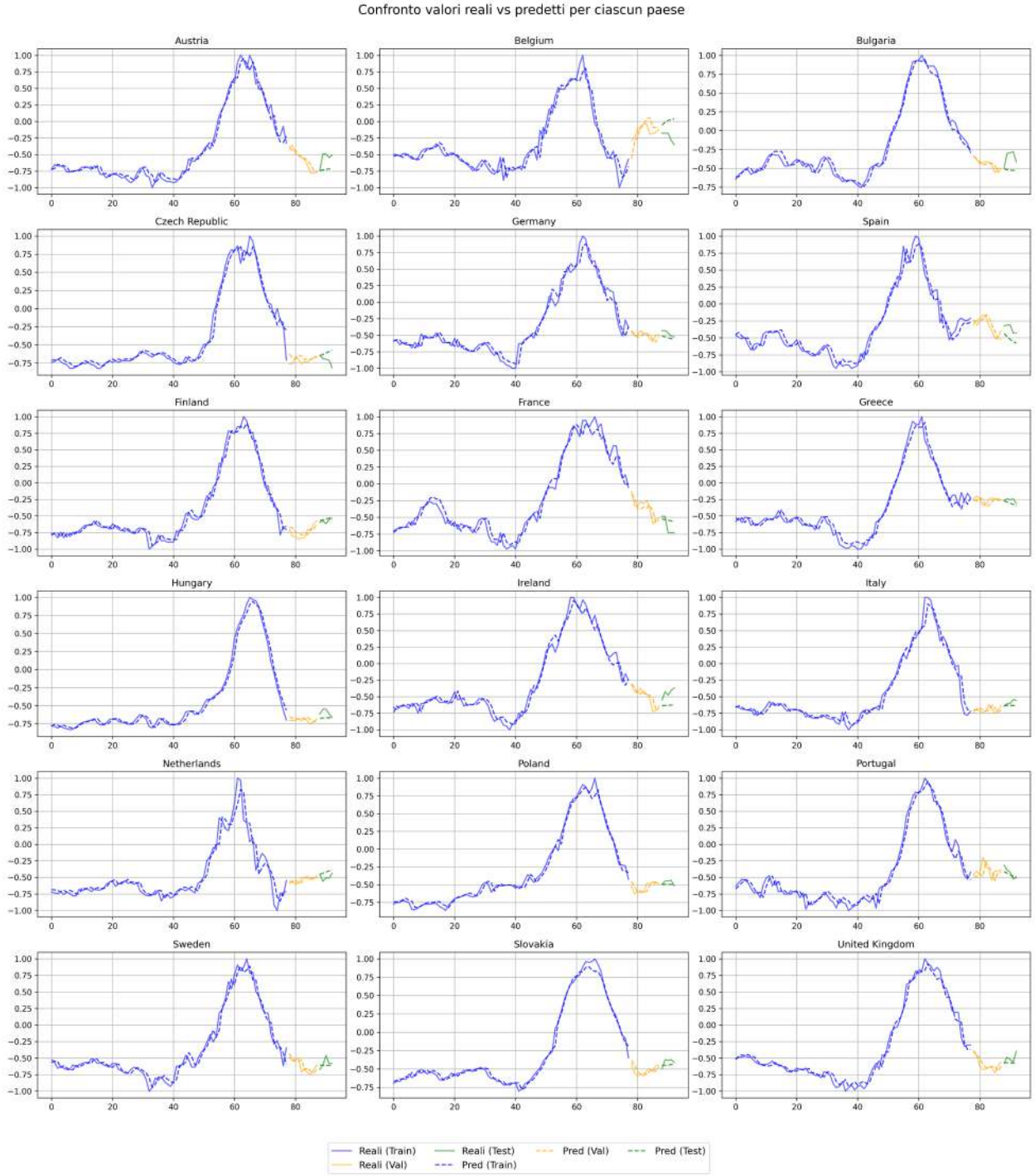

Figura 4.37: Predizioni rete HICP

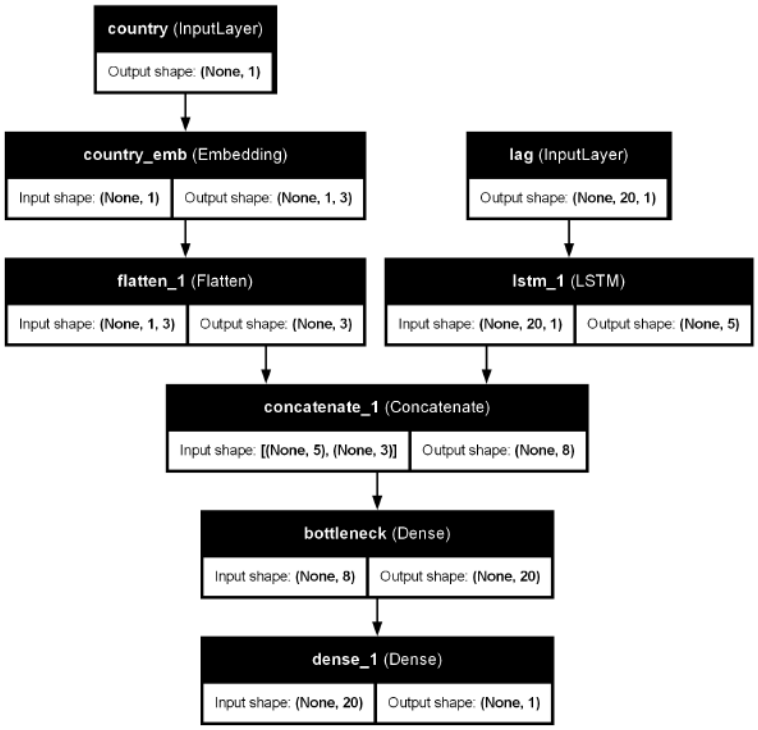

Figura 4.38: Modello per i rendimenti degli indici azionari.

spread 1-5

Per lo spread 1-5 il modello migliore selezionato risulta 4.43. La performance sul test set complessivo è un MSE di 0.0634, un MAE di 0.1835, un R^2 di 0.5758 e un r di Pearson di 0.8330. Le metriche solo sulle prime predizioni sono MSE: 0.0189, MAE: 0.1158 e r di Pearson: 0.9430 Le previsioni sono riportate in Figura: 4.44

spread5-10

In questo caso la struttura di modello scelta è quella 4.36, MSE di 0.1011, MAE di 0.2387, R^2 di 0.2082 e r di Pearson di 0.8254. Le metriche solo sulle prime predizioni sono MSE: 0.0184, MAE:0.1117, r di Pearson: 0.9751. Le previsioni del modello sono riportate in Figura: 4.46.

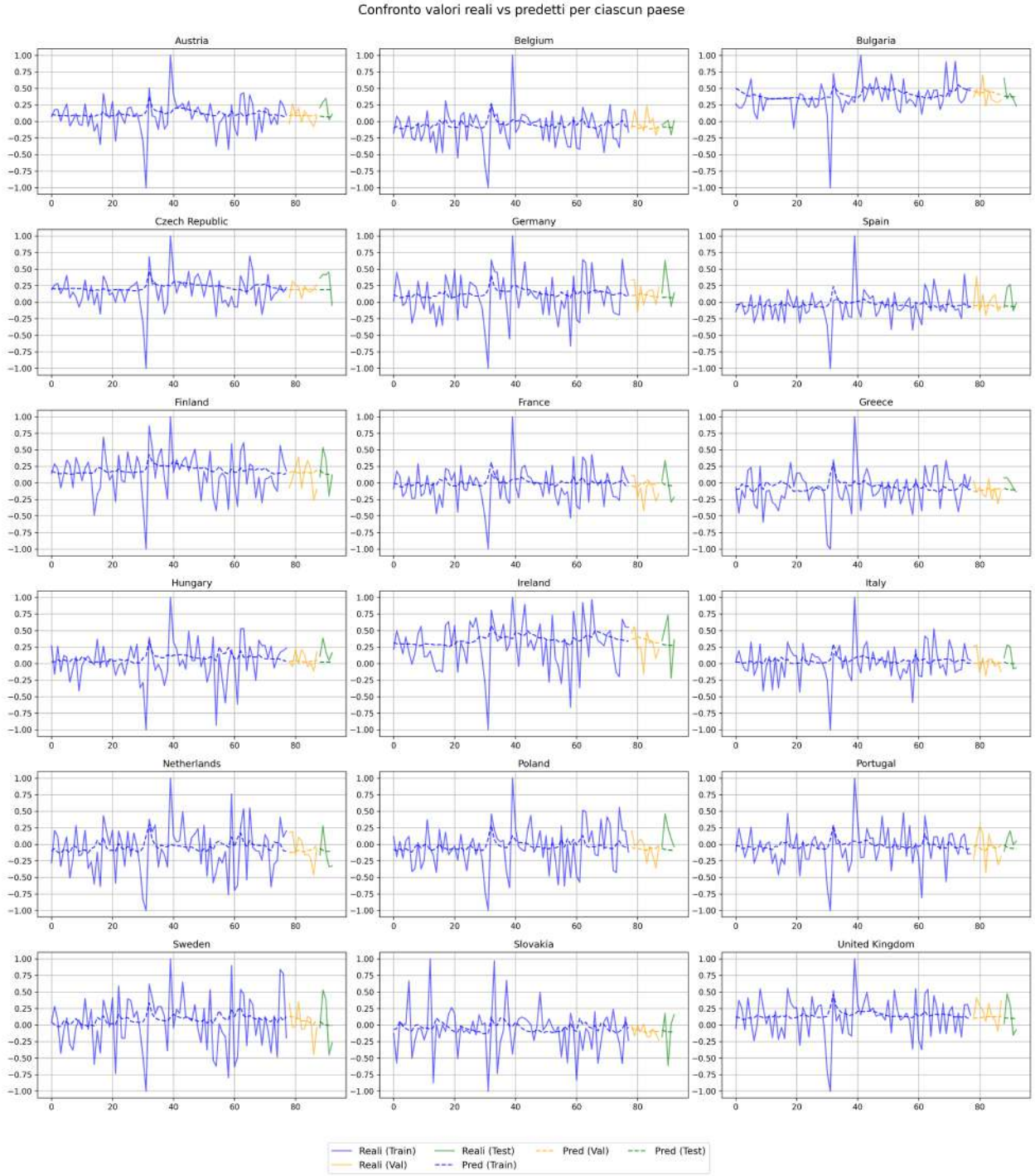

Figura 4.39: Predizioni modello per indici azionari

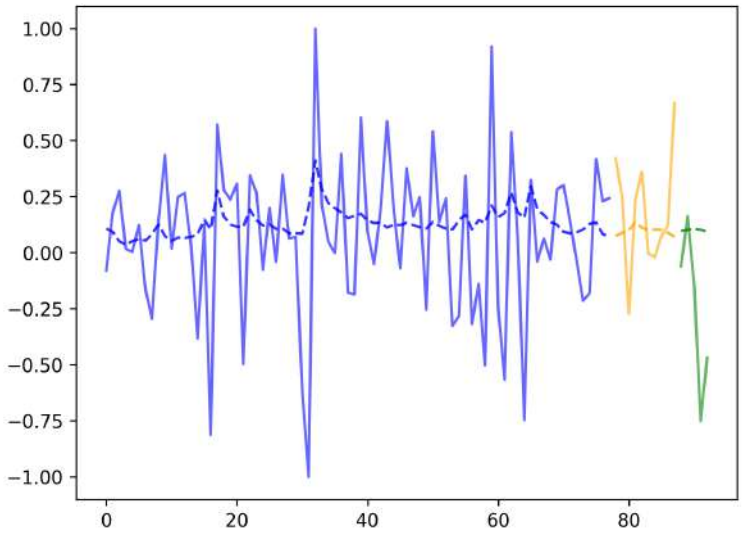

Figura 4.40: Predizioni modello per rendimenti S&P500.

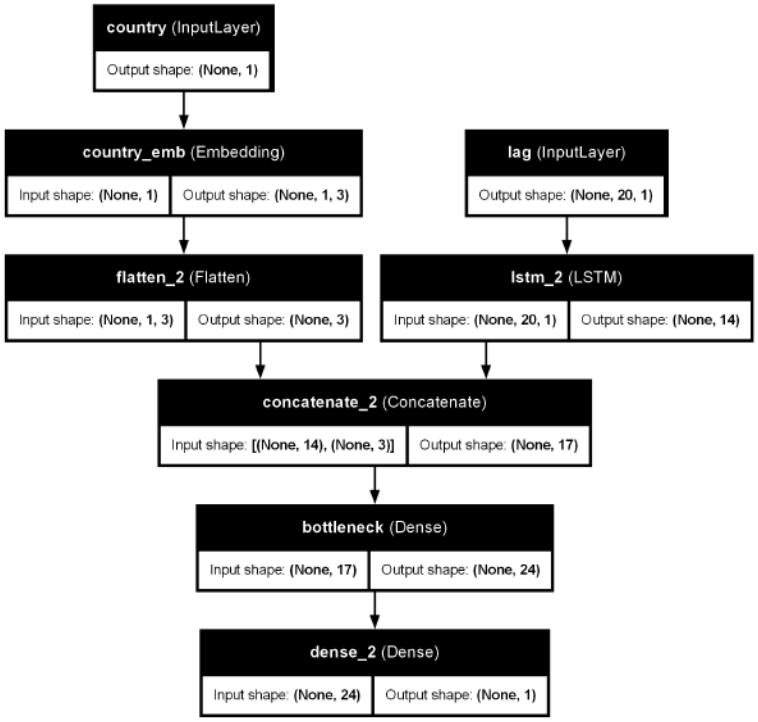

Figura 4.41: Modello per gli yield a 5 anni dei titoli di stato dei diversi paesi

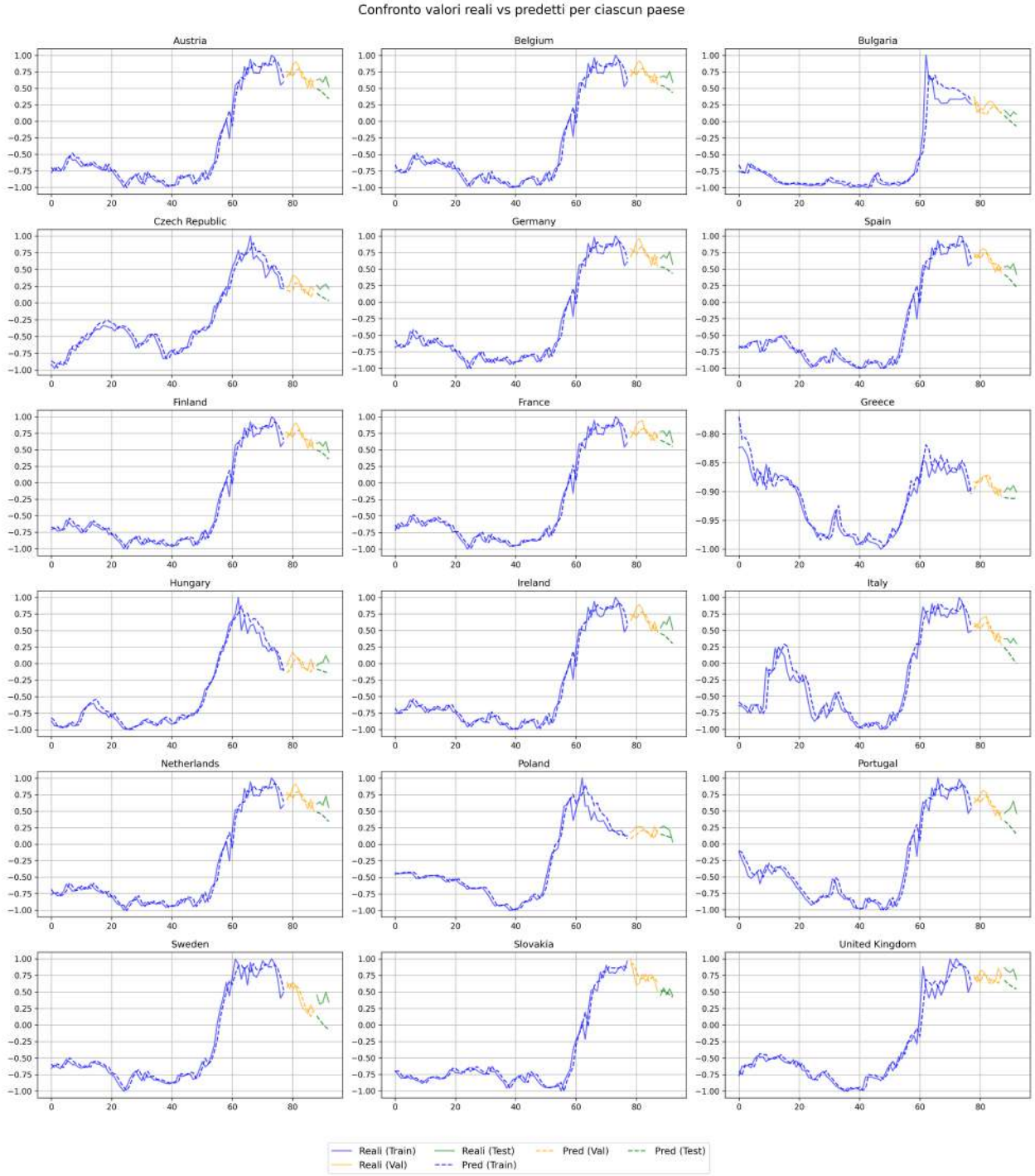

Figura 4.42: previsioni dello yield a 5 anni dei diversi paesi

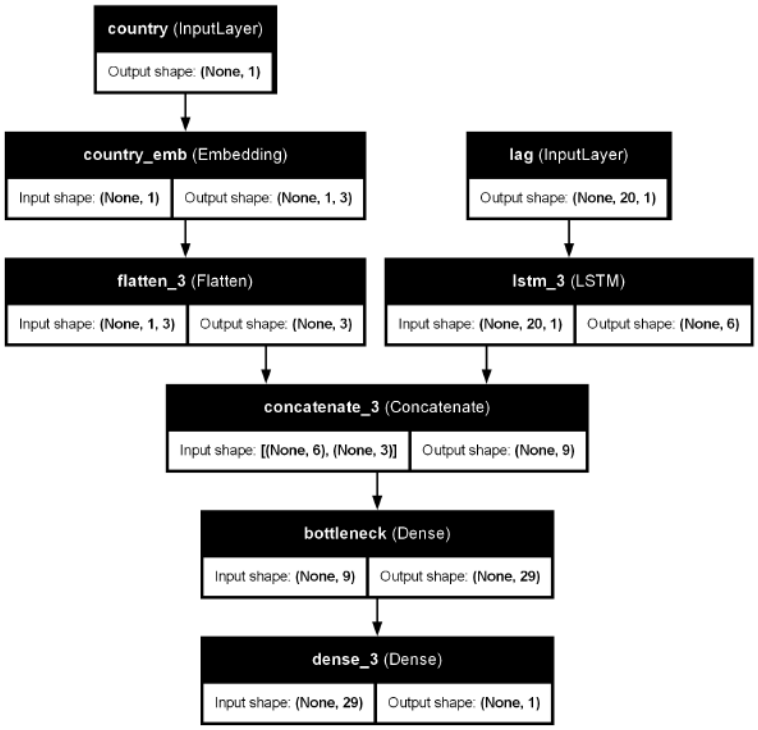

Figura 4.43: Modello per gli yield a 5 anni dei titoli di stato dei diversi paesi

4.3 Risultati finali

In questa sezione si riportano i risultati finali della rete applicati al volatility adjustment dei diversi paesi utilizzando per il forecasting i risultati della sezione e utilizzando la metodologia di boosting descritta in precedenza. La performance sul test set complessivo prima del boost è la seguente, MSE[8]: 0.0324, MAE[9]: 0.1165 R^2[10]: 0.4255 e r di Pearson[11]: 0.6796. Dopo il boost, MSE: 0.0.0313 , MAE: 0.1154, R^2: 0.4449 e r di Pearson: 0.6796. La metriche pre-boost e post-boost solo sulle prime previsioni del test set complessivamente sui 18 paesi sono: MSE: 0.0048, MAE: 0.0506 e r di Pearson: 0.9483 passano dopo il boost a MSE: 0.0041, MAE: 0.0472 e r di Pearson: 0.9483. Per verificare su quali paesi la procedura di boosting fornisce dei risultati più validi si è

[8] Errore quadratico medio

[9] Errore assoluto medio

[10] Indice di determinazione $\frac{\text{varianza spiegata}}{\text{varianza totale}}$

[11] $\frac{Cov(y,\hat{y})}{\sqrt{Var(y)Var(\hat{y})}}$

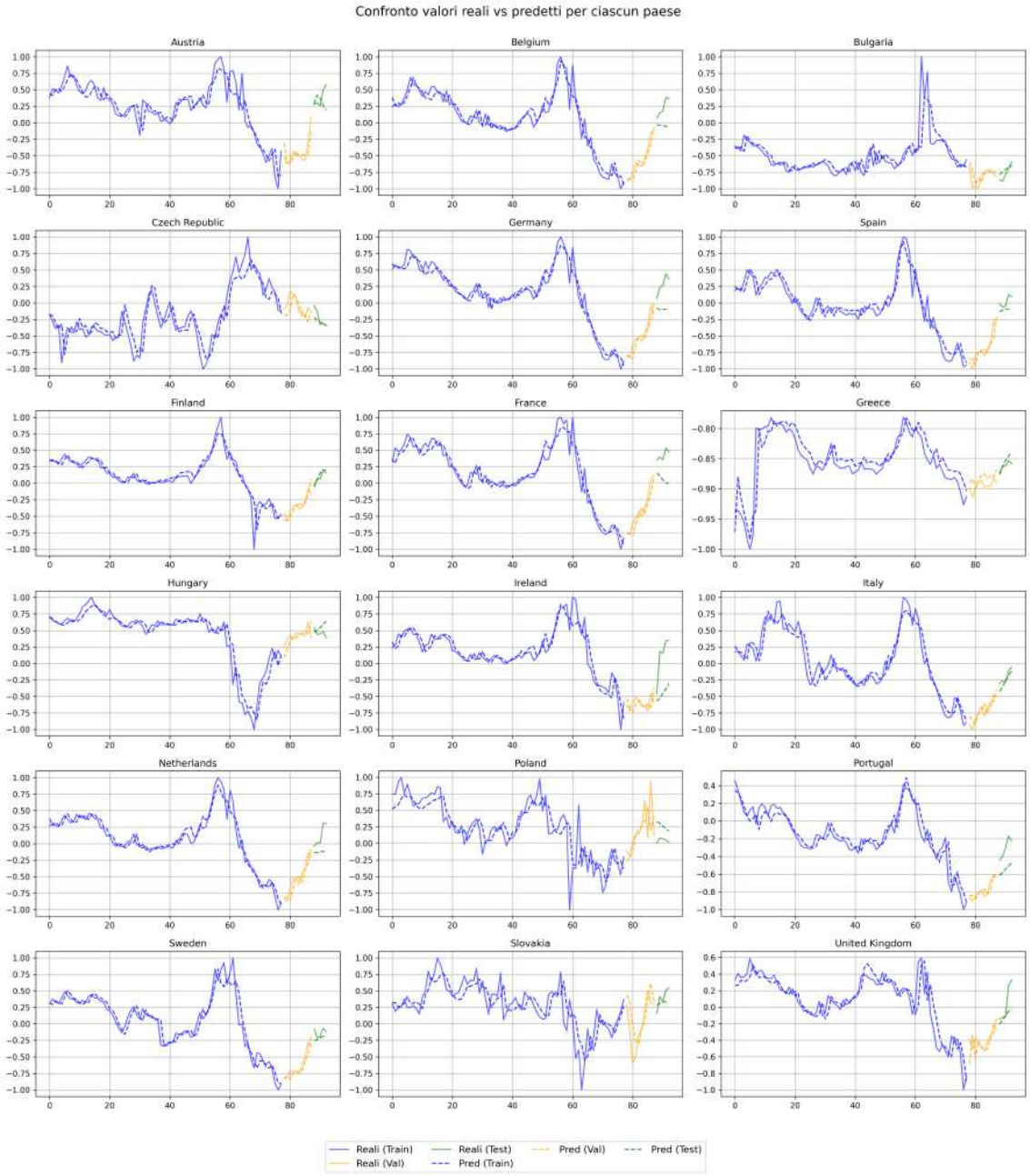

Figura 4.44: previsioni della rete dello spread 1-5 anni dei diversi paesi

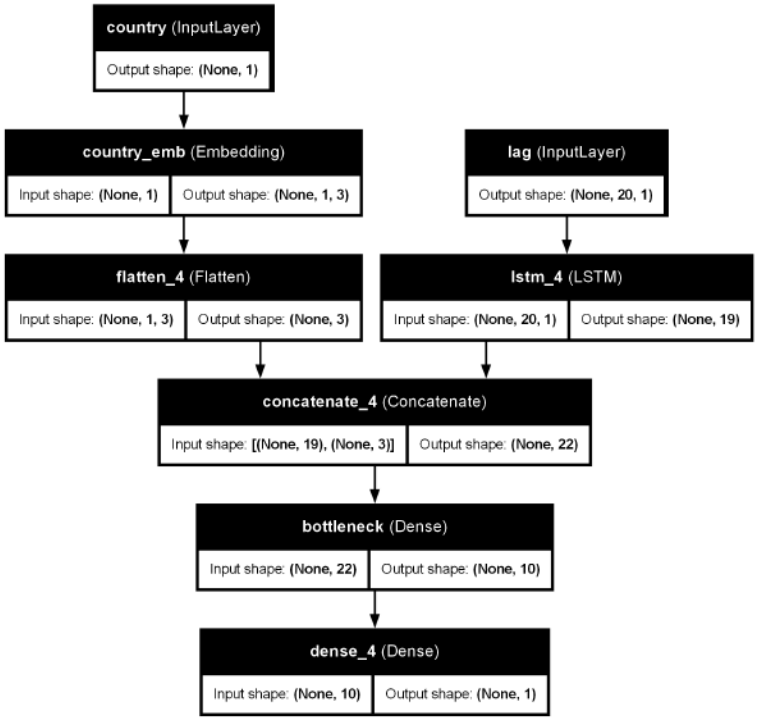

Figura 4.45: Modello per gli yield a 5 anni dei titoli di stato dei diversi paesi

calcolata la performance pre-boost e post-boost su ogni paese sul training set 4.11, sul validation set 4.12 e sul test set 4.13 attraverso MSE, MAE.

Per validare la qualità della rete si controlla come cambiano le performance della rete se al posto dei valori forecastati vengono inseriti i veri valori realizzati dei regressori. Prima del boosting le metriche riportano miglioramento: MSE: 0.0218, MAE: 0.1032, R^2: 0.6140 e r di Pearson: 0.8020 complessivamente sul test set. I risultati sono riportati nelle Figure: 4.51 e 4.52.

Modello con volatilità implicite

Siccome il forecasting di alcuni regressori si dimostra complesso, si prova ad aggiungere ai regressori le volatilità previste dai modelli ARCH e GARCH utilizzati nella sezione forecasting. Vengono inoltre rimossi il prezzo del Brent, il prezzo del gas naturale, i rendimenti dell'SP500 e i rendimenti degli indici specifici di paese, questi vengono sostituiti dalla volatilità del VSTOXX, la volatilità dei rendimenti del Brent, la volatilità dei rendimenti del gas naturale, la volatilità del cambio euro dollaro e la volatilità dei rendimenti corporate. Viene ricontrollata la matrice di cor-

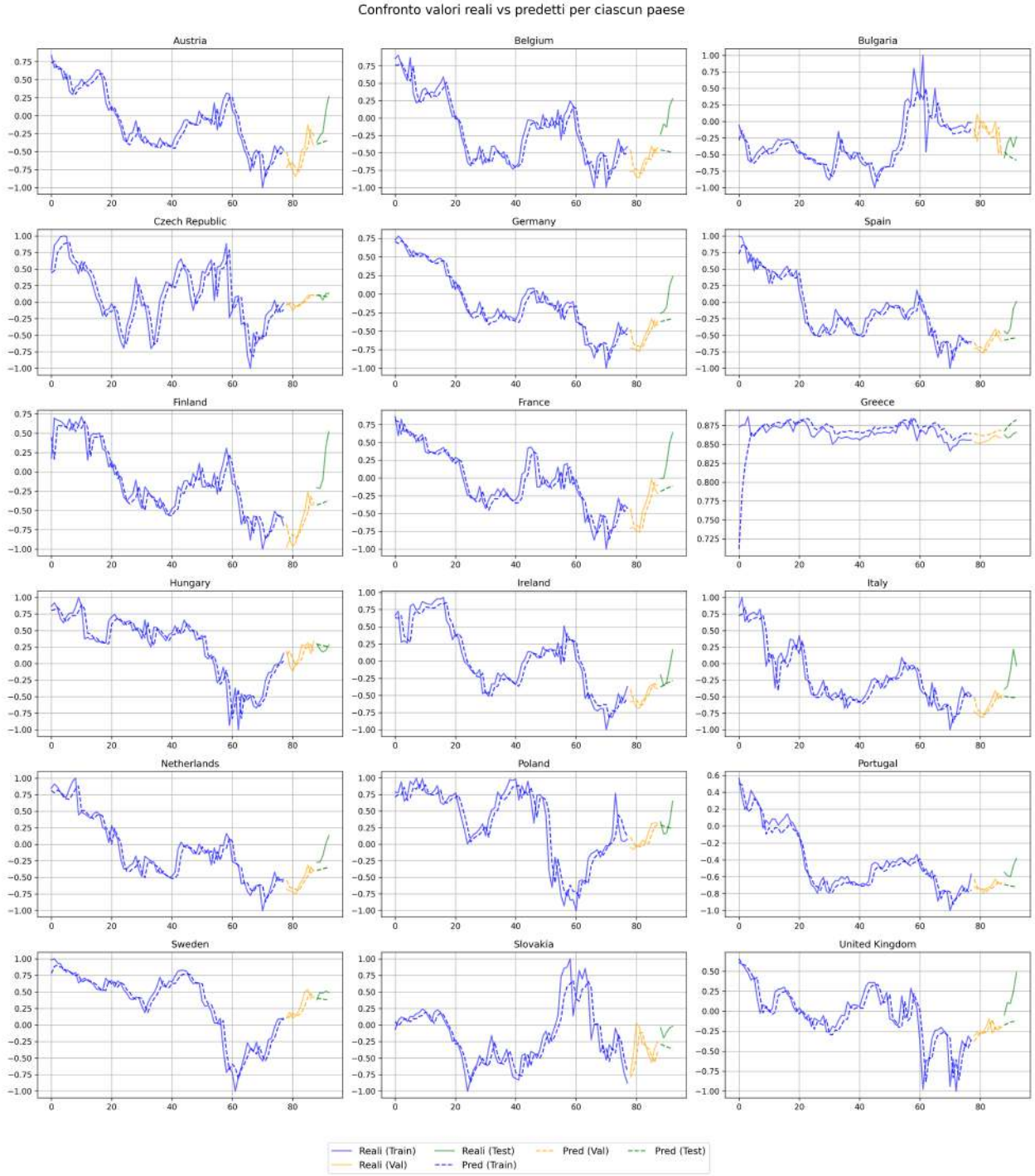

Figura 4.46: previsioni dello spread 5-10 anni dei diversi paesi

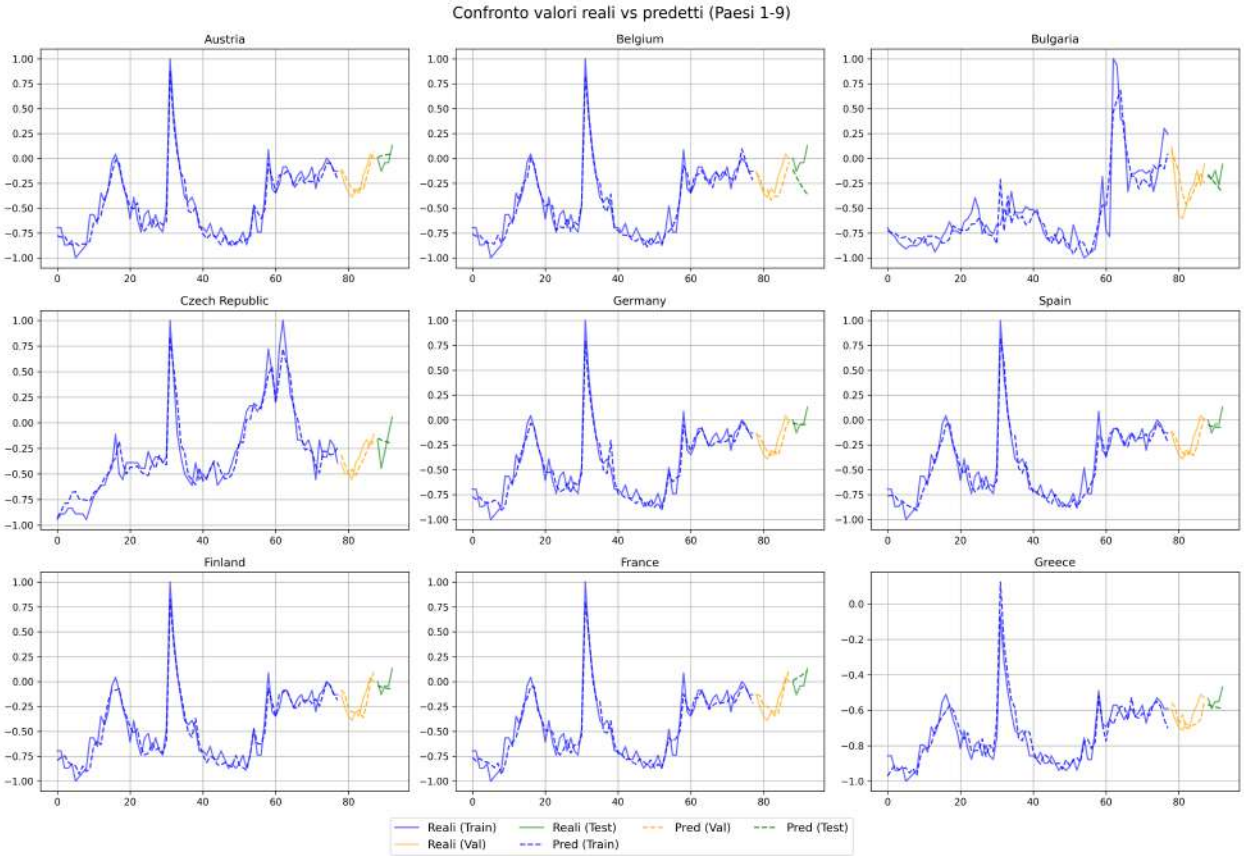

Figura 4.47: Previsioni rete pre-boost primi 9 paesi (modello 1)

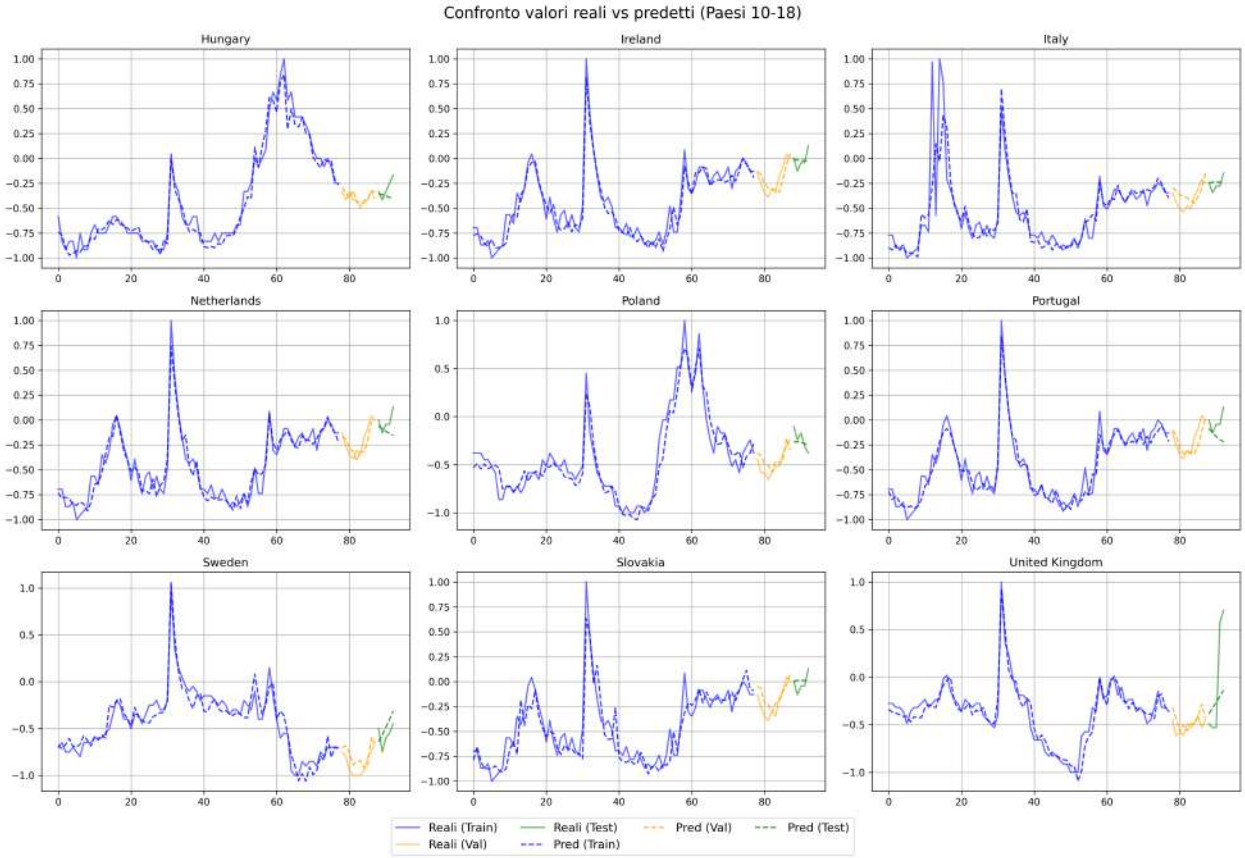

Figura 4.48: Previsioni rete pre-boost restanti 9 paesi (modello 1)

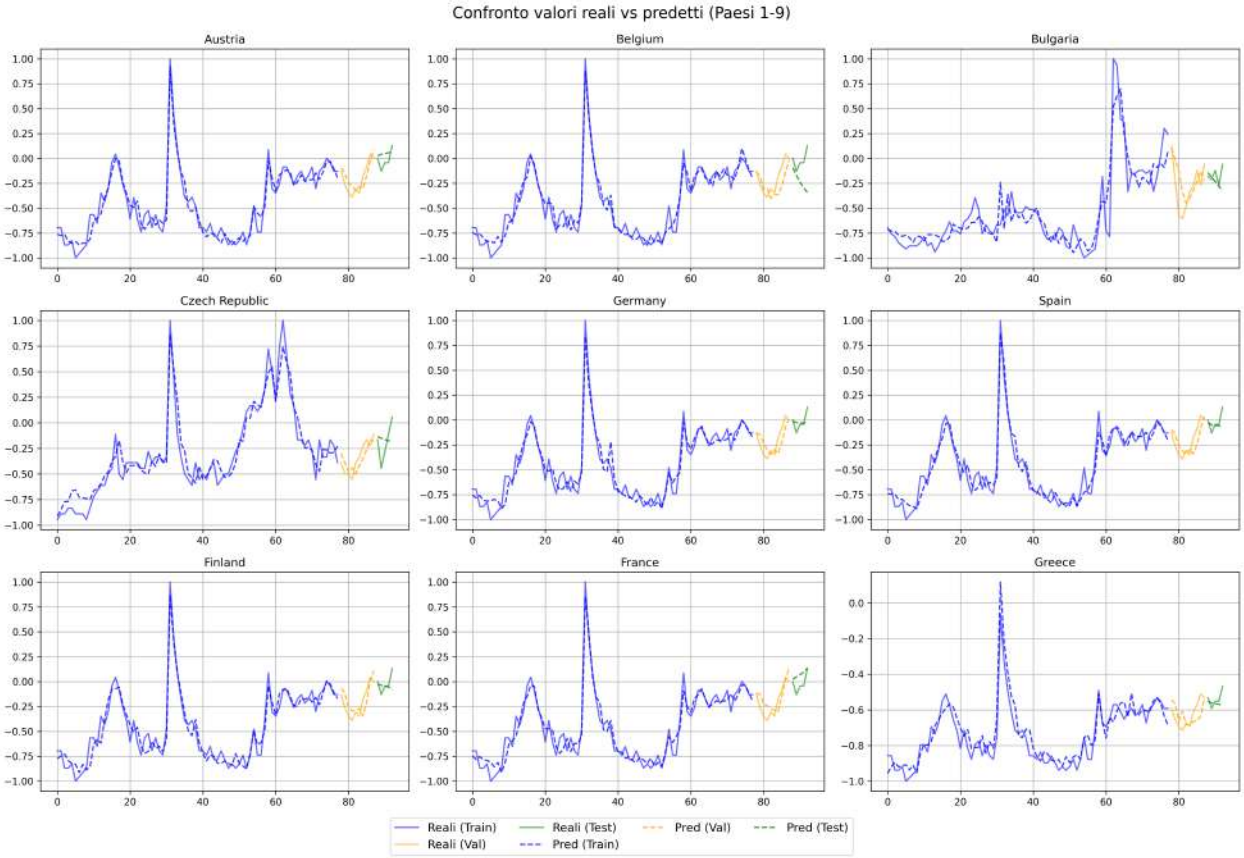

Figura 4.49: Previsioni rete + XGboost primi 9 paesi (modello 1)

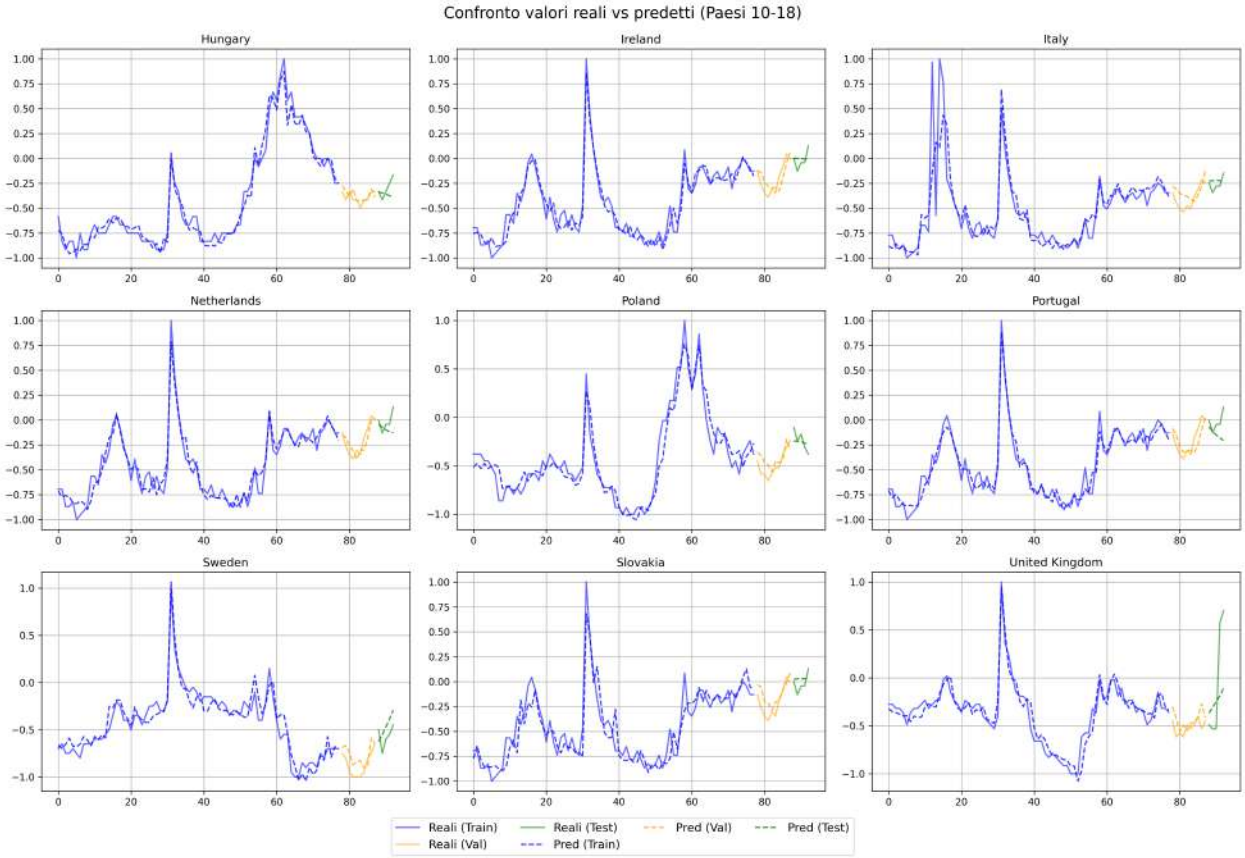

Figura 4.50: Previsioni rete + XGboost restanti 9 paesi (modello 1)

Tabella 4.11: MSE, MAE e pre-boost e post-boost per il singolo paese sul training set (modello 1)

Paese	MSE_pre	MSE_post	MAE_pre	MAE_post
Austria	0.0094	0.0084	0.0790	0.0756
Belgium	0.0092	0.0083	0.0765	0.0726
Bulgaria	0.0319	0.0297	0.1289	0.1272
Czech Republic	0.0173	0.0173	0.1010	0.1017
Germany	0.0106	0.0096	0.0817	0.0780
Spain	0.0108	0.0097	0.0831	0.0805
Finland	0.0100	0.0092	0.0772	0.0754
France	0.0097	0.0088	0.0796	0.0766
Greece	0.0041	0.0044	0.0521	0.0535
Hungary	0.0086	0.0076	0.0762	0.0722
Ireland	0.0108	0.0097	0.0823	0.0789
Italy	0.0521	0.0448	0.1123	0.1084
Netherlands	0.0102	0.0095	0.0795	0.0757
Poland	0.0183	0.0170	0.1037	0.0990
Portugal	0.0090	0.0080	0.0744	0.0708
Sweden	0.0118	0.0110	0.0844	0.0820
Slovakia	0.0233	0.0213	0.1136	0.1097
United Kingdom	0.0096	0.0093	0.0763	0.0762

Tabella 4.12: MSE, MAE e pre-boost e post-boost per il singolo paese sul validation set (modello 1)

Paese	MSE_pre	MSE_post	MAE_pre	MAE_post
Austria	0.0073	0.0078	0.0787	0.0819
Belgium	0.0151	0.0132	0.0980	0.0912
Bulgaria	0.0370	0.0390	0.1517	0.1517
Czech Republic	0.0093	0.0087	0.0883	0.0883
Germany	0.0101	0.0096	0.0845	0.0827
Spain	0.0110	0.0105	0.0875	0.0908
Finland	0.0101	0.0107	0.0858	0.0921
France	0.0089	0.0099	0.0803	0.0859
Greece	0.0051	0.0049	0.0575	0.0561
Hungary	0.0030	0.0036	0.0473	0.0505
Ireland	0.0088	0.0089	0.0793	0.0828
Italy	0.0105	0.0122	0.0972	0.1030
Netherlands	0.0068	0.0063	0.0718	0.0717
Poland	0.0098	0.0113	0.0861	0.0942
Portugal	0.0103	0.0097	0.0793	0.0825
Sweden	0.0147	0.0172	0.1061	0.1162
Slovakia	0.0117	0.0141	0.0902	0.0961
United Kingdom	0.0132	0.0154	0.0833	0.0925

Tabella 4.13: MSE, MAE pre-boost e post-boost per il singolo paese sul test set (modello 1)

Paese	MSE_pre	MSE_post	MAE_pre	MAE_post
Austria	0.0089	0.0107	0.0824	0.0920
Belgium	0.0750	0.0679	0.2281	0.2120
Bulgaria	0.0195	0.0171	0.0927	0.0895
Czech Republic	0.0321	0.0327	0.1461	0.1494
Germany	0.0077	0.0072	0.0596	0.0604
Spain	0.0097	0.0084	0.0742	0.0646
Finland	0.0102	0.0091	0.0735	0.0638
France	0.0101	0.0127	0.0823	0.0920
Greece	0.0037	0.0026	0.0478	0.0381
Hungary	0.0165	0.0142	0.0995	0.0898
Ireland	0.0078	0.0079	0.0658	0.0697
Italy	0.0048	0.0056	0.0568	0.0664
Netherlands	0.0200	0.0173	0.1096	0.1000
Poland	0.0085	0.0077	0.0776	0.0744
Portugal	0.0336	0.0293	0.1420	0.1323
Sweden	0.0212	0.0242	0.1439	0.1535
Slovakia	0.0085	0.0096	0.0770	0.0867
United Kingdom	0.2864	0.2800	0.4387	0.4419

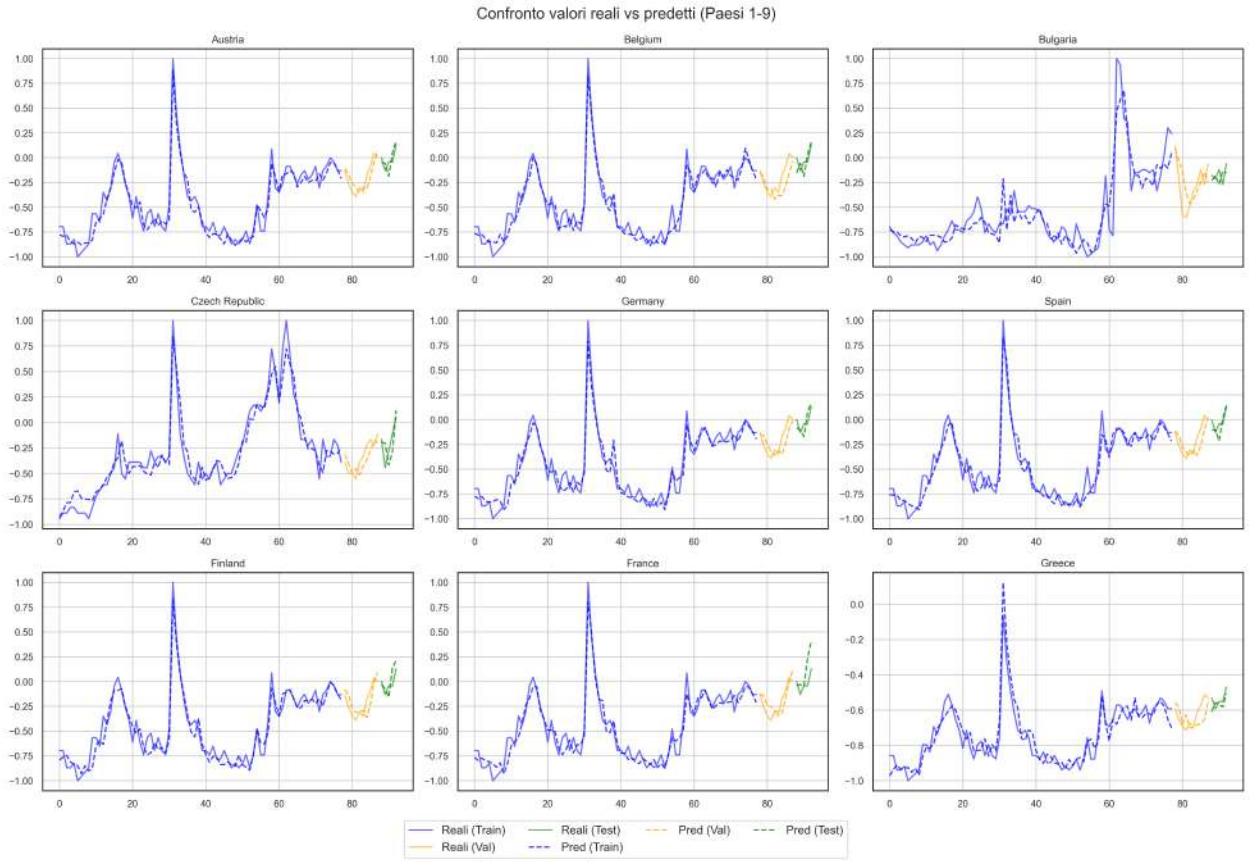

Figura 4.51: Previsioni rete primi 9 paesi utilizzando le vere realizzazioni dei regressori (modello 1)

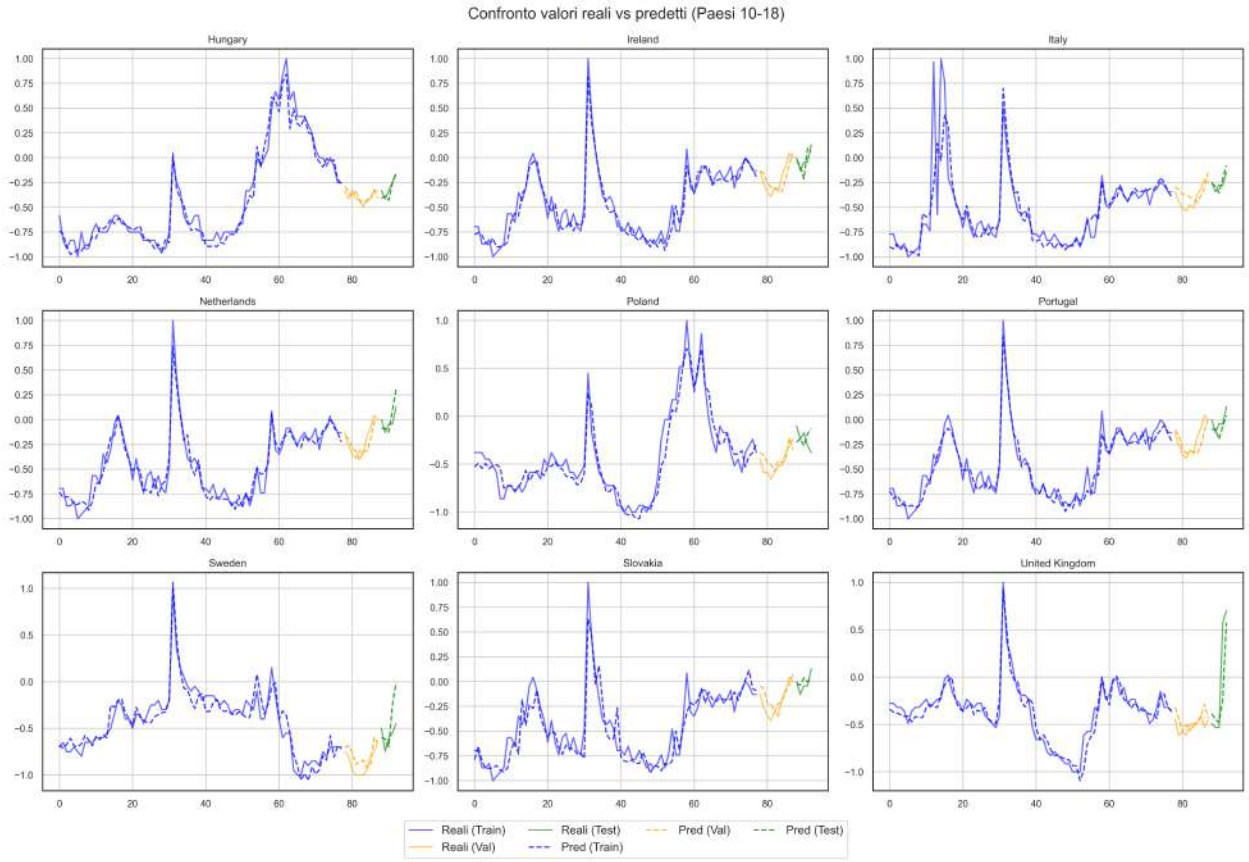

Figura 4.52: Previsioni rete dei restanti 9 paesi utilizzando le vere realizzazioni dei regressori (modello 1)

relazioni per vedere se questi nuovi regressori hanno forti relazioni lineari tra di loro e con altri regressori. Dalla figura 4.53 non sembra ci siano alte correlazioni lineari tra i nuovi regressori. Vengono quindi aggiunti tutti.

La performance pre-boost del modello 2 è MSE: 0.0318, MAE: 0.1197, R^2: 0.4361 e r di Pearson: 0.6743. Sulle prime predizioni è di MSE: 0.0047, MAE: 0.0525 e r di Pearson: 0.9397. Dopo il boost la performance è di MSE: 0.0302 MAE: 0.1230, R^2: 0.46401 e r di Pearson: 0.6942. Sulle prime predizioni MSE: 0.0079, MAE: 0.0750 e r di Pearson: 0.8872. Le performance suddivise per paese sono riportate nelle Figure: 4.14, 4.15, 4.16.

Tabella 4.14: MSE, MAE pre-boost e post-boost per il singolo paese sul training set (Modello 2)

Paese	MSE_pre	MSE_post	MAE_pre	MAE_post
Austria	0.0282	0.0050	0.1170	0.0549
Belgium	0.0276	0.0047	0.1154	0.0534
Bulgaria	0.0497	0.0136	0.1381	0.0866
Czech Republic	0.0246	0.0084	0.1053	0.0713
Germany	0.0292	0.0049	0.1175	0.0538
Spain	0.0265	0.0041	0.1112	0.0496
Finland	0.0294	0.0047	0.1171	0.0513
France	0.0285	0.0046	0.1184	0.0542
Greece	0.0091	0.0036	0.0705	0.0479
Hungary	0.0203	0.0076	0.1088	0.0686
Ireland	0.0274	0.0050	0.1142	0.0538
Italy	0.0749	0.0096	0.1526	0.0648
Netherlands	0.0281	0.0054	0.1135	0.0563
Poland	0.0293	0.0097	0.1179	0.0733
Portugal	0.0265	0.0081	0.1125	0.0668
Sweden	0.0259	0.0078	0.1120	0.0662
Slovakia	0.0293	0.0072	0.1145	0.0628
United Kingdom	0.0238	0.0069	0.1051	0.0662

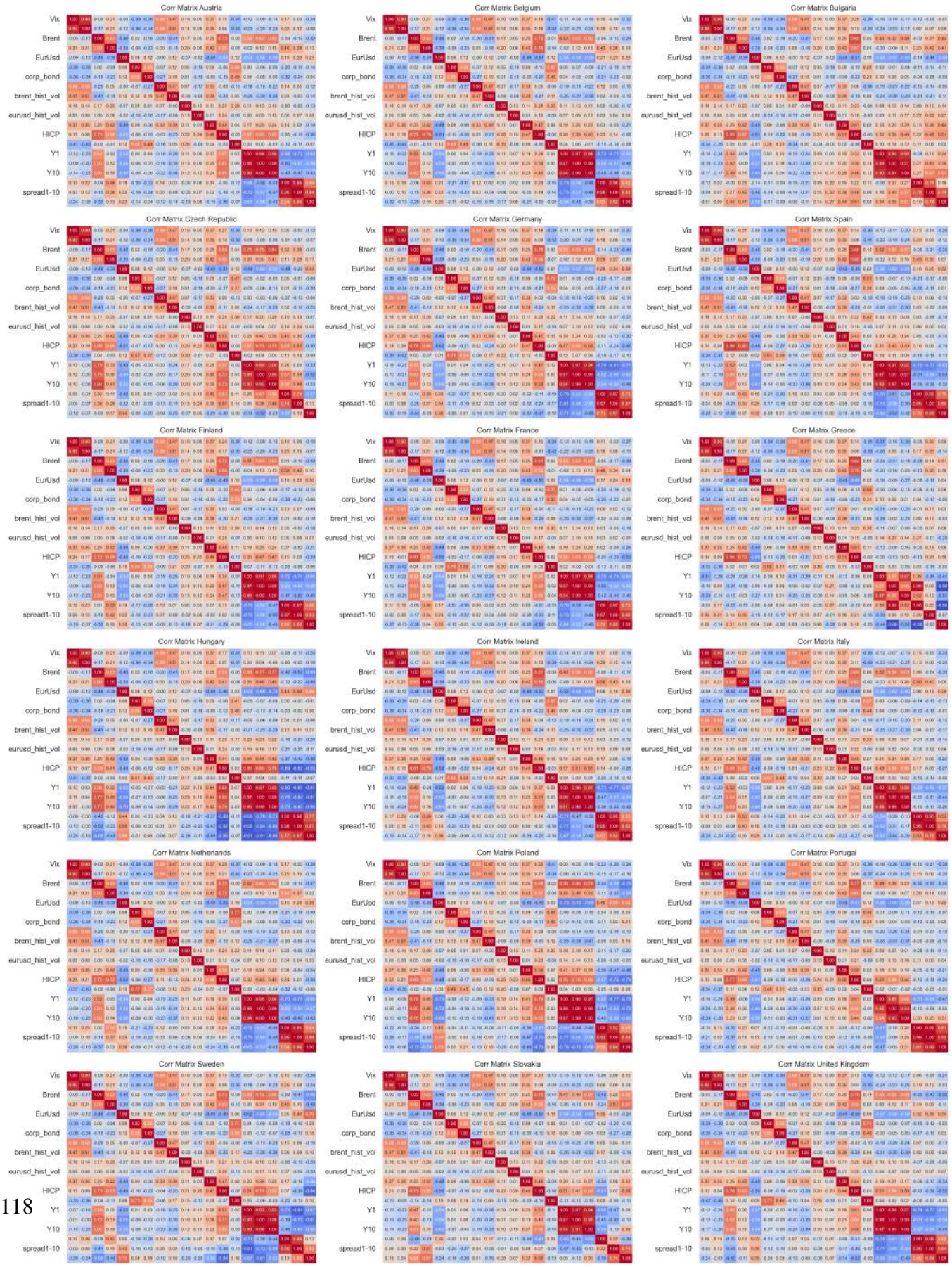

Figura 4.53: Matrici di correlazione

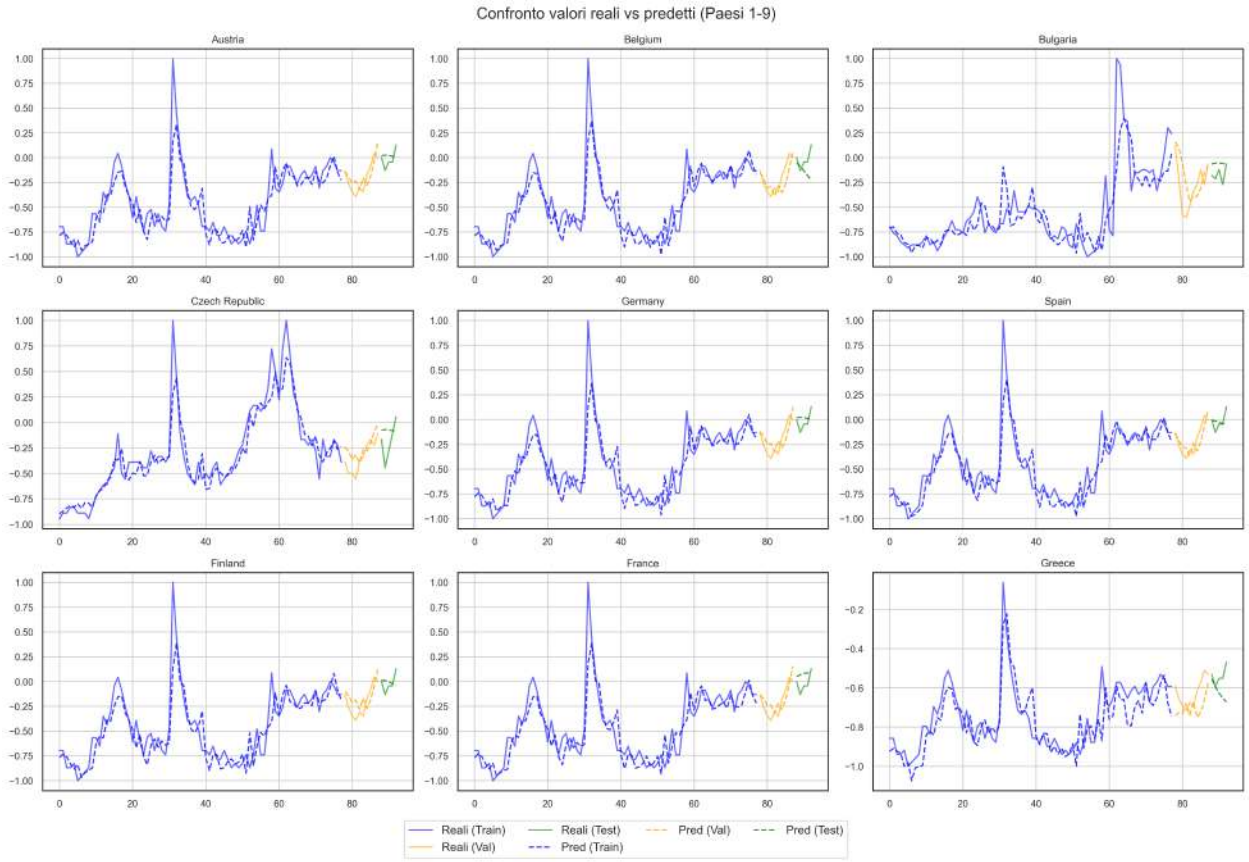

Figura 4.54: Previsioni rete pre-boost primi 9 paesi (Modello 2)

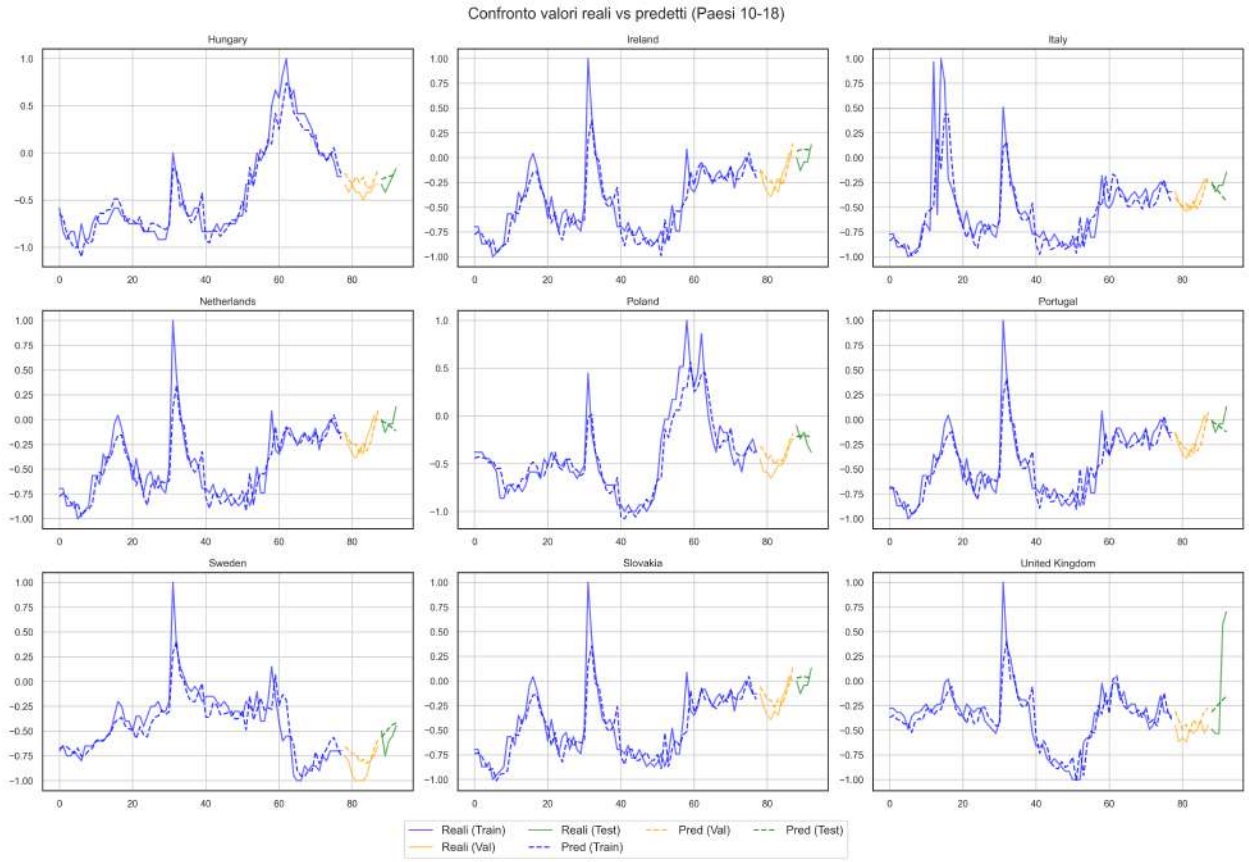

Figura 4.55: Previsioni rete pre-boost restanti 9 paesi (Modello 2)

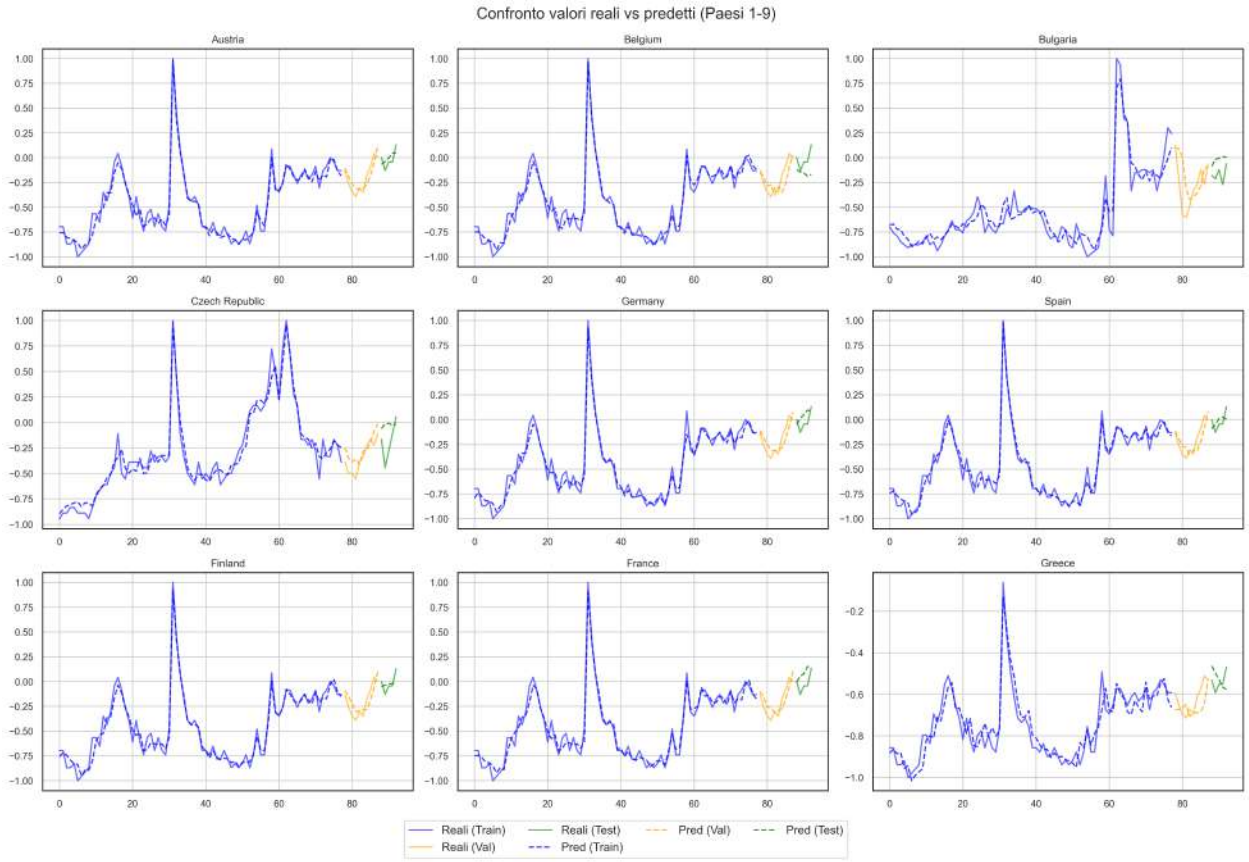

Figura 4.56: Previsioni rete + XGboost primi 9 paesi (Modello 2)

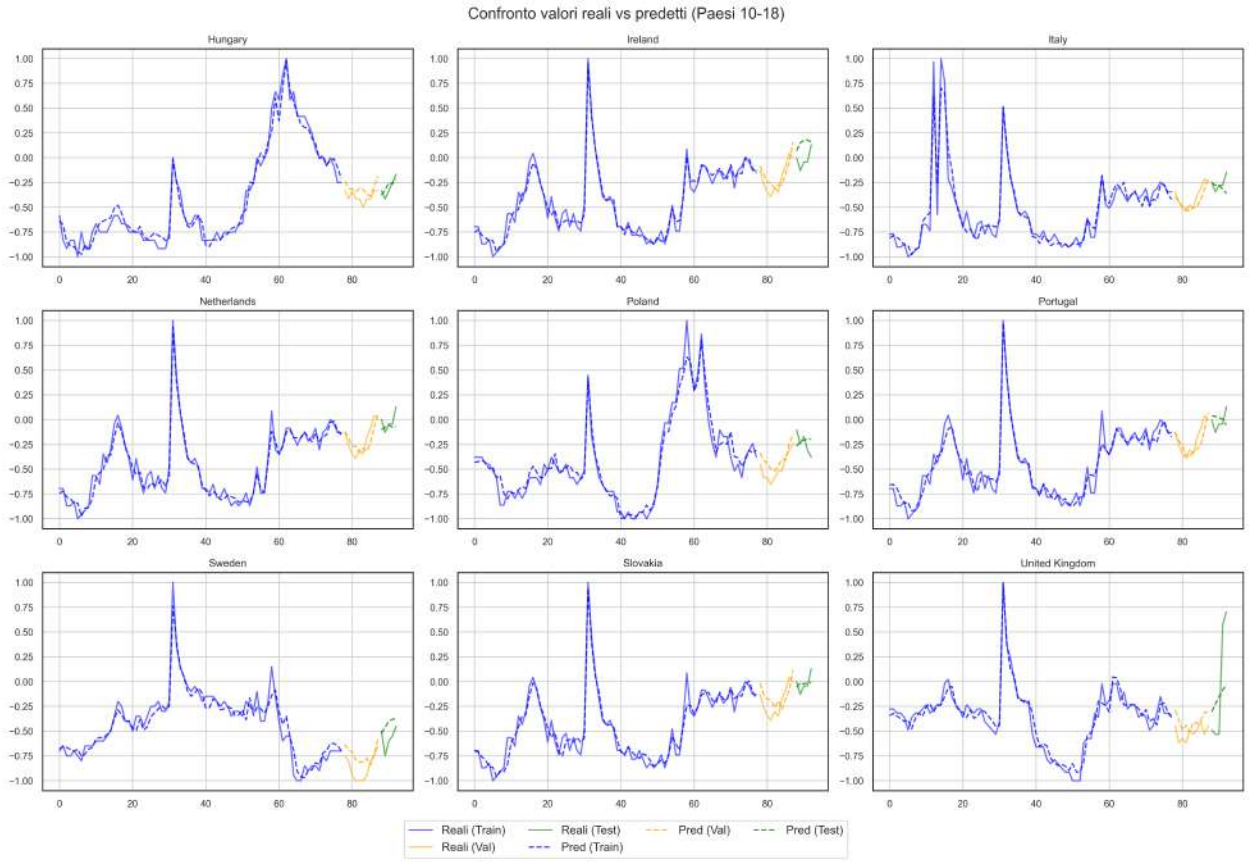

Figura 4.57: Previsioni rete + XGboost restanti 9 paesi (Modello 2)

Tabella 4.15: MSE, MAE pre-boost e post-boost per il singolo paese sul validation set (Modello 2)

Paese	MSE_pre	MSE_post	MAE_pre	MAE_post
Austria	0.0112	0.0065	0.0931	0.0733
Belgium	0.0114	0.0109	0.0921	0.0887
Bulgaria	0.0545	0.0581	0.1755	0.1755
Czech Republic	0.0181	0.0131	0.1140	0.0990
Germany	0.0112	0.0068	0.0960	0.0738
Spain	0.0096	0.0086	0.0896	0.0837
Finland	0.0151	0.0082	0.1106	0.0845
France	0.0110	0.0079	0.0927	0.0809
Greece	0.0102	0.0051	0.0865	0.0565
Hungary	0.0161	0.0099	0.1074	0.0858
Ireland	0.0116	0.0093	0.0950	0.0874
Italy	0.0067	0.0047	0.0709	0.0540
Netherlands	0.0112	0.0089	0.0951	0.0829
Poland	0.0168	0.0124	0.1101	0.1012
Portugal	0.0110	0.0086	0.0943	0.0771
Sweden	0.0251	0.0236	0.1366	0.1353
Slovakia	0.0154	0.0178	0.1121	0.1262
United Kingdom	0.0212	0.0201	0.1237	0.1221

Tabella 4.16: MSE, MAE pre-boost e post-boost per il singolo paese sul test set (Modello 2)

Paese	MSE_pre	MSE_post	MAE_pre	MAE_post
Austria	0.0098	0.0060	0.0859	0.0750
Belgium	0.0315	0.0277	0.1346	0.1380
Bulgaria	0.0179	0.0289	0.1131	0.1511
Czech Republic	0.0422	0.0561	0.1685	0.1913
Germany	0.0094	0.0108	0.0836	0.0953
Spain	0.0096	0.0078	0.0658	0.0848
Finland	0.0104	0.0074	0.0784	0.0696
France	0.0155	0.0176	0.1093	0.1026
Greece	0.0114	0.0051	0.0807	0.0610
Hungary	0.0077	0.0038	0.0743	0.0554
Ireland	0.0166	0.0363	0.1162	0.1641
Italy	0.0211	0.0103	0.1076	0.0655
Netherlands	0.0145	0.0109	0.0824	0.0806
Poland	0.0101	0.0153	0.0880	0.1048
Portugal	0.0151	0.0136	0.0870	0.0987
Sweden	0.0177	0.0314	0.1137	0.1504
Slovakia	0.0110	0.0077	0.0941	0.0737
United Kingdom	0.3015	0.2479	0.4713	0.4519

4.4 Conclusioni

Nel capitolo 2 è stato introdotto un metodo per interpolare i nodi della curva di sconto (metodo Smith-Wilson) con l'introduzione di concetti come l'Ultimate Forward Rate che permettono di interpolare la curva di sconto risk-free anche per scadenze oltre il Last Liquid Point. Questa procedura ha implicazioni rilevanti soprattutto per la determinazione del valore attuale delle passività assicurative nel ramo vita che possono avere scadenze oltre il LLP e rappresentare parti significative del portafoglio delle passività della compagnia assicurativa. Per quanto riguarda il capitolo 4, si è provato a determinare un modello in grado di spiegare e prevedere il volatility adjustment proposto da EIOPA, per farlo sono state create due reti neurali e si è anche provato a migliorare ulteriormente le previsioni delle reti utilizzando una procedura integrata tramite l'utilizzo della tecnica di Extreme Gradient Boosting. Per quanto riguarda i risultati della prima rete, si possono osservare dei risultati intermedi; in alcuni casi la rete prevede bene solo per alcuni paesi. Guardando alla procedura di boosting, questa sembra funzionare bene sul complesso delle prime previsioni per ogni paese, mentre su tutto il test set non porta miglioramenti su tutti i paesi. Per il primo modello si è anche provato a sostituire i valori dei regressori previsti dai diversi modelli con i veri valori osservati; la rete migliora la performance e anche la procedura di boosting sembra apportare dei miglioramenti. Questi risultati fanno pensare che quanto meno la rete riesca a catturare una buona parte del comportamento di fondo del VA e che i regressori selezionati siano coerenti. Nella seconda rete si prova a sostituire dei regressori la cui predizione risultava difficile con alcune volatilità implicite nei modelli utilizzati nella parte di regressione. Questa rete sembra performare leggermente meglio per quanto riguarda le previsioni sul test set, anche in questo caso la procedura di boosting migliora la performance sulle prime previsioni nel complesso di tutti i paesi ma quando si va a vedere la performance sulle 5 date scomposta per paese i risultati sono misti. In conclusione, la rete del modello 1 potrebbe essere usata dalla compagnia assicurativa per calcolare un VA dinamico durante il calcolo delle BEL, nello specifico prendendo l'output della rete come certo, è sufficiente modificare uno degli input introducendo uno shock per avere il volatility adjustment sotto shock e utilizzare questa versione come spread aggiunto al tasso risk-free.

5. Codice

5.1 Codice Esempio Capitolo 2

Per questo esempio è stato utilizzato Matlab

Main esempio capitolo 2

```
settlement ="07/01/2025";
formatDate ="dd/mm/yyyy";
NumberWorkDays=[1,5,10];
NumberMonths=[1:12,18,(2:10)*12];
dateois = [datewrkdy(datetime(2025,01,07), NumberWorkDays, 1),...
           datemnth(datetime(2025,01,07),NumberMonths)]';
dates.settlement =datenum(settlement,formatDate);
dates.OIS = datenum(dateois);
%quotazioni OIS su ESTR al 07/01/2025
R.OIS=[2.91700,
2.916749954,
2.917500019,
2.880297184,
2.78730011,
2.668900013,
2.581250191,
2.510300159,
2.443349838,
2.387549877,
2.34739995,
2.309499979,
2.275099993,
2.247499943,
2.222499847,
2.125999928,
2.104599953,
```

```
2.116000175,
2.14620018,
2.176599979,
2.210000038,
2.24545002,
2.277100086,
2.310349941,
2.342700005]/100;
%estrazione dei fattori di sconto e prezzi impliciti a partire dalle
%quotazioni come discusso nel capitolo 2
discCurve=OISbootstrap(dates,R);


% prezzi di ZCB impliciti
pt = discCurve.discounts(2:end);
%tenors
U=yearfrac(dates.settlement,dates.OIS,3);
%si ipotizza un UFR di 0.045
UFR = 0.045;
w = log(1+UFR);
%si ipotizza un convergence point di 60 anni
CP = 60;
%costruzione della matrice C
C = eye(size(U,1)).*exp(-w.*U);

massimi = max(U, U');
minimi  = min(U, U');

%ricerca dell'alpha ottimo
alpha_opt = FindMinAlphaTol(pt,w,U,C,massimi,minimi,CP,0.0001);
%tol fissata a 0.0001 come da regolamentazione EIOPA.

%si plotta il grafico per vedere la forma funzionale che si sta
%ottimizzando, si osserva che decresce strettamente in alpha.
```

```
x = linspace(0,5,1000); %valori di alpha
y = zeros(size(x));
for i = 1:1000
 y(i) = (abs(fwd_rateSW_OPT(C, pt, w, U, massimi, minimi, x(i), CP) - w));
end

plot(x,y)



%calcolo di Z input della funzione di di sconoto
Z = (C * SW(massimi, minimi, alpha_opt) * C') \ (pt - exp(-w .* U));

%calcolo della funzione di sconto
t = linspace(0,70,1000);
P = zeros(size(t));

for i = 1:1000
    P(i) = exp(-w * t(i)) + (exp(-w.*(t(i) + U)).*W(t(i),U,alpha_opt))' *Z
        ;
end
%calcolo della funzione del tasso forward istantaneo
tassi_forward = zeros(size(t));
for i = 1:1000
    tassi_forward(i) = fwd_rateSW(C,pt,w,U,massimi,minimi,alpha_opt,t(i));
end


%plot della funzione di sconto e della funzione dela tasso forward
    istantaneo.
% punti noti a priori
tassi_a_priori = -log(pt)./U;
%punti a posteriori
tassi_SW = -log(P)./t;
```

```
figure(1)
set(gcf, "Position", [100, 100, 1200, 450])

subplot(2,1,1)
plot(t, tassi_SW, "Color", [0 0.45 0.74], "LineWidth", 1.5)
hold on
scatter(U, tassi_a_priori, 60, [0.85 0.33 0.10], "filled")
xlabel("Scadenza (anni)", "FontSize", 12)
ylabel("Tasso", "FontSize", 12)
title("Interpolazione Smith-Wilson - Curva di sconto", "FontSize", 13)
grid on
legend("Curva interpolata", "Tassi osservati", "Location", "southeast")
set(gca, "FontSize", 11)
hold off

subplot(2,1,2)
plot(t, tassi_forward, "Color", [0.47 0.67 0.19], "LineWidth", 1.5)
xlabel("Scadenza (anni)", "FontSize", 12)
ylabel("Forward rate istantaneo", "FontSize", 12)
title("Interpolazione Smith-Wilson - Forward rate", "FontSize", 13)
grid on
set(gca, "FontSize", 11)

%%%%%%%%%%%%%%%Esempio processo gaussiano altrenativo
    %%%%%%%%%%%%%%%%%%%%%%%

%plot della forma funzionale del processo hornstein hulenback
x = linspace(0,5,1000); %valori di alpha
y = zeros(size(x));
for i = 1:1000
 y(i) = (abs(fwd_rateOU_OPT(C, pt, w, U, massimi, minimi, x(i), CP) - w));
end

```

```
plot(x,y)

%calibrazione dell'alpha
alpha_optOU = FindMinAlphaTolOU(pt,w,U,C,massimi,minimi,CP,0.0001);

%calolo dello Z componente della funzione del fattore di sconto
Z_OU = (C * OU(massimi, minimi, alpha_optOU) * C') \ (pt - exp(-w .* U));

%costruzione della funzione del fattore di sconto nel caso del processo
%alternativo
t = linspace(0.001,70,1000);
P_OU = zeros(size(t));

for i = 1:1000
    P_OU(i) = exp(-w * t(i)) + (exp(-w.*(t(i) + U)).*O(t(i),U,alpha_optOU)
        )' *Z_OU;
end

tassi_forward_OU= zeros(size(t));
for i = 1:1000
    tassi_forward_OU(i) = fwd_rateOU(C,pt,w,U,massimi,minimi,alpha_optOU,t
        (i));
end


%punti a posteriori OU
%grafico dell'interpolazione dei tassi
tassi_OU = -log(P_OU)./t;
plot(t,tassi_OU)
hold on
scatter(U, tassi_a_priori, 60, [0.85 0.33 0.10], "filled")
hold off
%grafico del forward rate istantaneo nel caso del processo alternativo
plot(t,tassi_forward_OU)
```

```
%grafico di confronto dei due processi
figure(2)
set(gcf, "Position", [100, 100, 1200, 450])
plot(t, tassi_SW, "Color", [0 0.45 0.74], "LineWidth", 1.5)
xlim([0,20])
hold on
scatter(U, tassi_a_priori, 60, [0.85 0.33 0.10], "filled")
plot(t,tassi_OU,"LineWidth",1.5)
xlabel("Scadenza (anni)", "FontSize", 12)
ylabel("Tasso", "FontSize", 12)
title("Smith-Wilson v Ornstein Uhlenbeck - Curva di sconto", "FontSize",
    13)
grid on
legend("SW", "Tassi osservati", "OU", "Location", "southeast")
set(gca, "FontSize", 11)
hold off
```

Funzione: OISbootsrap

```
function discCurve=OISbootstrap(dates,R)
%Funzione estre fattori di sconto impliciti nelle quotazioni OIS
%dates = struct contenente settlement (data di calcolo curva) e OIS (
    maturities dei contratti selezionati)
%R = struct contenente OIS (fair strike dei contratti selezionati)
%Dati iniziali
DF=zeros(size(dates.OIS));
maturities=yearfrac(dates.settlement,dates.OIS,3);
idx=find(maturities<=1,1,"last");
%Procedura esposta nel capitolo 2
%Tassi OIS con maturity inferiore o uguale a 1 anno
yearF=yearfrac(dates.settlement,dates.OIS(1:idx),2);% 2 = act/360
DF(1:idx)=1./(1+yearF.*R.OIS(1:idx));
%%Tassi OIS con maturity superiore a 1 anno
periodYF=[yearF(end);yearfrac(dates.OIS(idx:end-1),dates.OIS(idx+1:end),2)
```

```
    ];
for i=idx+1:length(dates.OIS)
    DF(i)=(1-R.OIS(i)*sum(periodYF(1:i-idx).*DF(idx:i-1)))/(1+periodYF(i+1-idx)*R.OIS(i));
end
%curve relative ai contratti OIS
discCurve.knots=[dates.settlement;dates.OIS];
discCurve.discounts=[1;DF];
discCurve.rates=[0;-log(DF)./maturities];
end
```

Funzione: FindMinAlphaTol

```

function alpha_min = FindMinAlphaTol(pt, w, U, C, massimi, minimi, CP, tol)
%Funzione che Ricerca il piu piccolo alpha tale che g(alpha) <= tol e che alpha >= 0.05
%pt, w, U, C, massimi, minimi,CP = input della funzione fwd_rate necessaria
%per il calcolo di g(alpha)
%tol= soglia sul g(alpha) fissata a 0.0001 da EIOPA
%g(alpha) = abs(fwd_rate(C, pt, w, U, massimi, minimi, alpha_high, CP) - w)

% Intervallo iniziale di ricerca
alpha_low = 0.00001;
alpha_high = 1;

%Controllo che la condizione sia soddisfatta almeno in alpha_high
%se risulta soddisfatta allora l'alpha minimo sara minore o uguale all' alpha_high fissato a 1
%in questo caso
if abs(fwd_rateSW_OPT(C, pt, w, U, massimi, minimi, alpha_high, CP) - w) > tol
```

```
    error("Nessun valore di alpha nell'intervallo soddisfa la condizione g
        (alpha) <= tol");
end

%Utilizzo metodo della bisezione per trovare l'alpha minimo per cui la
%condizione g(alpha) <= tol risulta soddisfatta
while (alpha_high - alpha_low) > 1e-6
    alpha_mid = (alpha_low + alpha_high)/2;
    g_mid = abs(fwd_rateSW_OPT(C, pt, w, U, massimi, minimi, alpha_mid, CP
        ) - w);

    if g_mid <= tol
        alpha_high = alpha_mid;
    else
        alpha_low = alpha_mid;
    end
end

alpha_min = alpha_high;  % primo valore che soddisfa la condizione
if alpha_min <= 0.05
    alpha_min = 0.05; %valore minimo proposto da EIOPA
end
```

Funzione: fwd_rateSW_OPT

```

function [fwd] = fwd_rateSW_OPT(C, prezzi, w, U, massimi, minimi, alpha, t
    )
% Calcola il forward rate al tempo CP nel metodo Smith-Wilson, questa
% funzione serve solo per l'ottimizzazione.
% C = matrice diagonale con exp(-w*U)
% prezzi = vector dei discount factors osservati
% U = vettore delle maturity
% t = CP

```

```
%calcolo vettore dei pesi Z da formula
Z = (C * SW(massimi, minimi, alpha) * C') \ (prezzi - exp(-w .* U));

numeratore = 1 + alpha * sum(U .* diag(C) .* Z);
denominatore = sum(sinh(alpha .* U) .* diag(C) .* Z);
kappa = numeratore / denominatore;

fwd = w + alpha / (1 - kappa * exp(alpha * t));
end
```

Funzione: SW

```
function [sw] = SW(massimi, minimi, alpha)
% Calcola Wilson function per calcolo della Wilson matrix
%massimi = matrice quadrata che contiene i massimi delle maturity
%per cui si hanno osservazioni
%minimi = matrice quadrata che contiene i minimi delle maturity
%per cui si hanno osservazioni
%alpha = parametro di convergenza da stimare, o implicito nel modello
sw = alpha .* minimi - exp(-alpha .* massimi) .* sinh(alpha .* minimi);
end
```

Funzione: W

```
function [wi] = W(t,u, alpha)
%Wilson function per calcolo scalare-vettore
wi = alpha .* min(t,u)-exp(-alpha.* max(t,u)).*sinh(alpha.*min(t,u));
end
```

Funzione: fwd_rateSW

```
function [res] = fwd_rateSW(C, prezzi, w, U, massimi, minimi, alpha, t)
% Calcola il forward rate al tempo t nel metodo Smith-Wilson (l(0,t))
% C = matrice diagonale con exp(-w*U)
```

```
% prezzi = vettore dei discount factors osservati
% U = vettore delle maturity per cui risulta presente uno strumento quotato
% t = maturity d'interesse
derivate = zeros([size(U,1),1]);
for i = 1:size(U,1)
    derivate(i) = alpha*exp(-alpha * t)*sinh(alpha*U(i));
    if t < U(i)
    derivate(i) = alpha -alpha*exp(-alpha*U(i))*cosh(alpha*t);
    end
end

Z = (C * SW(massimi, minimi, alpha) * C') \ (prezzi - exp(-w .* U));
numeratore = (exp(-w*(t+U)).*Z)' * derivate;
denominatore = exp(-w*t)+(exp(-w*(t+U)).*Z)' * W(t,U,alpha);
res = w- numeratore/denominatore;
end
```

Funzione: FindMinAlphaTolOU

```
function alpha_min = FindMinAlphaTolOU(pt, w, U, C, massimi, minimi, CP, tol)
% Ricerca il pi  piccolo alpha tale che g(alpha) <= tol nel caso del
% processo ornstein uhlenback
%pt, w, U, C, massimi, minimi,CP = input della funzione fwd_rate necessaria
%per il calcolo di g(alpha)
%tol= soglia sul g(alpha) fissata a 0.0001 da EIOPA
% g(alpha) = abs(fwd_rate(C, pt, w, U, massimi, minimi, alpha_high, CP) - w)

% Intervallo iniziale di ricerca
alpha_low = 0.00001;
alpha_high = 1;

```

```
%Controllo che la condizione sia soddisfatta almeno in alpha_high
%se risulta soddisfatta allora l'alpha minimo sara minore o uguale all'
    alpha_high fissato a 1
%in questo caso
if abs(fwd_rateOU_OPT(C, pt, w, U, massimi, minimi, alpha_high, CP) - w) >
     tol
    error("Nessun valore di alpha nell'intervallo soddisfa la condizione g
        (alpha) <= tol");
end

%Utilizzo metodo della bisezione per trovare l'alpha minimo per cui la
%condizione g(alpha) <= tol risulta soddisfatta
while (alpha_high - alpha_low) > 1e-6
    alpha_mid = (alpha_low + alpha_high)/2;
    g_mid = abs(fwd_rateOU_OPT(C, pt, w, U, massimi, minimi, alpha_mid, CP
        ) - w);

    if g_mid <= tol
        alpha_high = alpha_mid;
    else
        alpha_low = alpha_mid;
    end
end

alpha_min = alpha_high;  % primo valore che soddisfa la condizione
if alpha_min <= 0.05
    alpha_min = 0.05;
end
```

Funzione: fwd_rateOU_OPT

```
function [fwd] = fwd_rateOU_OPT(C, prezzi, w, U, massimi, minimi, alpha, t
    )
% Calcola il forward rate al tempo CP per l'ottimizzazione col methodo OU
% C = matrice diagonale con exp(-w*U)
```

```
% prezzi = vector dei discount factors osservati
% U = vettore delle maturity
% t = CP
Z = (C * OU(massimi, minimi, alpha) * C') \ (prezzi - exp(-w .* U));
numeratore = exp(-w*t);
denominatore = (Z.*exp( -alpha.*(t+U)))' * O(t,U,alpha);
k = numeratore/denominatore;
fwd = w - alpha/(1+k) ;
end
```

Funzione: OU

```
function [hu] = OU(massimi, minimi, alpha)
% Calcola HU funtion per calcolo della HU matrix
%massimi = matrice quadrata che contiene i massimi delle maturity
%per cui si hanno osservazioni
%minimi =  matrice quadrata che contiene i minimi delle maturity
%per cui si hanno osservazioni
%alpha = parametro di convergenza da stimare, o implicito nel modello
hu = alpha^2 .* exp(-alpha .* massimi) .* cosh(alpha .* minimi);
end
```

Funzione: O

```
function [h] = O(t,u, alpha)
%HU function per calcolo scalare-vettore
h = alpha^2.* exp(-alpha.* max(t,u)).*cosh(alpha.*min(t,u));
end
```

Funzione: fwd_rateOU

```
function [res] = fwd_rateOU(C, prezzi, w, U, massimi, minimi, alpha, t)
% Calcola il forward rate al tempo t nel metodo Ornstein Hulenbeck (l(0,t)
    )
% C = matrice diagonale con exp(-w*U)
% prezzi = vettore dei discount factors osservati
```

```
% U = vettore delle maturity per cui si ha uno strumento quotato
% t = maturity d'interesse
derivate = zeros([size(U,1),1]);
for i = 1:size(U,1)
    derivate(i) = -alpha^3 * exp(-alpha*U(i))*cosh(alpha*U(i));
    if t < U(i)
    derivate(i) = alpha^3 *exp(-alpha * U(i))*sinh(alpha*t);
    end
end

Z = (C * OU(massimi, minimi, alpha) * C') \ (prezzi - exp(-w .* U));
numeratore = (exp(-w*(t+U)).*Z)' * derivate;
denominatore = exp(-w*t)+(exp(-w*(t+U)).*Z)' * O(t,U,alpha);
res = w- numeratore/denominatore;
end
```

5.2 Codice Capitolo 4

Viene indicata solo la pipeline principale per questioni di spazio.

Pipeline Rete + XGBoost

```
import random
import matplotlib
matplotlib.use('Qt5Agg')
import matplotlib.pyplot as plt
import numpy as np
import pandas as pd
from tensorflow.keras.layers import Input, Dense, Multiply, LSTM,
    Embedding, Flatten, Concatenate, Dropout, BatchNormalization
from tensorflow.keras.models import Model
from tensorflow.keras.callbacks import ModelCheckpoint, EarlyStopping
from tensorflow.keras.optimizers import Nadam
from sklearn.preprocessing import MinMaxScaler, StandardScaler
```

```
from sklearn.metrics import mean_squared_error, mean_absolute_error,
    r2_score, mean_absolute_percentage_error
from tensorflow.keras.utils import plot_model
import seaborn as sns
import datetime
from IPython.display import Image
import pydot
import xgboost as xgb
from  sklearn.tree import DecisionTreeRegressor, plot_tree
import arch
from statsmodels.graphics.tsaplots import month_plot, quarter_plot,
    plot_acf, plot_pacf
from statsmodels.tsa.seasonal import seasonal_decompose
from statsmodels.tsa.stattools import adfuller
from pmdarima import auto_arima, ARIMA, model_selection
from sklearn.model_selection import ParameterGrid
from keras import regularizers
from statsmodels.stats.diagnostic import het_arch
import statsmodels.api as sm
from statsmodels.tsa.arima.model import ARIMA
from scipy.stats import pearsonr
import keras_tuner as kt
import json
from sklearn.model_selection import RandomizedSearchCV

"""
Viene saltata la parte di preparazione dei dati per contenere lo spazio,
    come esposto tutti i dati vengono scalati tra -1 e 1.
"""

early_stop = EarlyStopping(monitor='val_loss', patience=50,
    restore_best_weights=True, verbose=1)


```

```
#costruzione del tuner degli iperparametri
#build model e una funzione definita da ma che contiene la struttura principale del modello

tuner = kt.RandomSearch(build_model, objective="val_loss", executions_per_trial=1, overwrite=True)

#funzione che testa le diverse combinazioni di iperparametri
tuner.search(
    [train_hist_va_train, input_mod1.values, input_mod1_codes.values],
    train_output_va_train,
    validation_data=([train_hist_va_val, input_mod1_val.values, input_mod1_val_codes.values], train_output_va_val),
    epochs=200,
    batch_size=64,
    callbacks=[early_stop],
    verbose=1
)

best_model = tuner.get_best_models(1)[0]
best_model.save_weights("best_model.weights.h5")
# best_model.load_weights("best_model_weights.h5")
best_hps = tuner.get_best_hyperparameters(1)[0]

#salvataggio iperparametri migliori
with open("best_hps.json", "w") as f:
    json.dump(best_hps.values, f)
#apertura iperparametri migliori
with open("best_hps.json", "r") as f:
    best_params = json.load(f)
#costruzione del modello a partire dai migliori parametri selezionati
best_model = build_fixed_model(best_params)
plot_model(best_model,to_file = "model.png", show_shapes = True, show_layer_names = True)
```

```
best_model.load_weights("best_model.weights.h5")

#estrazione predizioni della rete
predictions_train = best_model.predict(x = [train_hist_va_train,input_mod1
    , input_mod1_codes])
predictions_val = best_model.predict(x = [train_hist_va_val,
    input_mod1_val, input_mod1_val_codes])
pred_vere = best_model.predict(x = [test_hist_va, input_mod1_test.values,
    input_mod1_test_codes.values])
#predizioni iterative sul test set
predizioni = [[] for i in range(len(paesi))]
va_input_rolling = test_hist_va[list(range(0, len(test_output_va),
    len_test)),:,:]
for i in range(len_test):
    codes_input_rolling = input_mod1_test_codes.iloc[np.array(list(range
        (0, len(test_output_va), len_test)))+i].values
    X_input_rolling = input_mod1_test_predicted.iloc[np.array(list(range
        (0, len(test_output_va), len_test)))+i,:].values
    predictions_test = best_model.predict(x=[va_input_rolling,
        X_input_rolling, codes_input_rolling])
    for j in range(len(paesi)):
        va_input_rolling[j,0:lookback-1,:] = va_input_rolling[j,1:,:]
        predizioni[j].append(predictions_test[j][0])
        va_input_rolling[j, lookback-1, :] = predictions_test[j]

predizioni = pd.DataFrame(np.concatenate([np.array(predizioni[i]).reshape
    (-1,1) for i in range(len(paesi))],axis=1 ), columns=paesi)
#predizioni = pd.DataFrame(np.concatenate([pred_vere[np.arange(0,len_test)
     + i * len_test ,:] for i in range(len(paesi))], axis = 1) , columns =
    paesi )

print("MSE:", mean_squared_error(test_output_va, np.concatenate([
    predizioni.iloc[:,i] for i in range(len(paesi))],axis=0 )) )
print("MAE:", mean_absolute_error(test_output_va,  np.concatenate([
```

```
    predizioni.iloc[:,i] for i in range(len(paesi))],axis=0 )))
print("R^2:", r2_score(test_output_va, np.concatenate([predizioni.iloc[:,i
    ] for i in range(len(paesi))],axis=0 ) ))
print("Rpearson:", pearsonr(test_output_va.reshape(-1), np.concatenate([
    predizioni.iloc[:,i] for i in range(len(paesi))],axis=0 ))[0] )

#estrazione predizioni per ogni singolo paese
predictions_train = pd.concat([pd.DataFrame(predictions_train[
    indici_train1+len(indici_train1)*i], columns=[paesi[i]]) for i in range
    (len(paesi))], axis=1 )
predictions_val = pd.concat([pd.DataFrame(predictions_val[np.array(list(
    range(0,len(indici_val))))+len(indici_val)*i], columns=[paesi[i]]) for
    i in range(len(paesi))], axis=1 )
reali_train = pd.concat([pd.DataFrame(train_output_va_train[indici_train1+
    len(indici_train1)*i], columns=[paesi[i]]) for i in range(len(paesi))],
     axis=1 )
reali_val = pd.concat([pd.DataFrame(train_output_va_val[np.array(list(
    range(0,len(indici_val))))+len(indici_val)*i], columns=[paesi[i]]) for
    i in range(len(paesi))], axis=1 )

indici_pred = np.linspace(start=0,stop=len(indici_test)-1, num= len(
    indici_test), dtype=int)
reali_test = pd.concat([pd.DataFrame(test_output_va[indici_pred+len(
    indici_pred)*i], columns=[paesi[i]]) for i in range(len(paesi))], axis
    =1 )

#plot predizioni
plot_rvp(paesi,reali_train,predictions_train,reali_val, predictions_val,
    reali_test, predizioni)

#analisi performance
len_train = 113-len_test-len_val-lookback
MSE = []
MAE = []
```

```
r_pearson = []
for i in range(len(paesi)):
    MSE.append(mean_squared_error(train_output_va_train[len_train*i:len_train*i + len_train],predictions_train.iloc[:,i].values ))
    MAE.append(mean_absolute_error(train_output_va_train[len_train * i:len_train * i + len_train], predictions_train.iloc[:,i].values))
    r_pearson.append(pearsonr(train_output_va_train[len_train * i:len_train * i + len_train].reshape(-1), predictions_train.iloc[:, i].values)[0])
tab1train = pd.DataFrame(np.concatenate([np.array(paesi).reshape(-1,1),np.array(MSE).reshape(-1,1),np.array(MAE).reshape(-1,1),np.array(r_pearson).reshape(-1,1)], axis = 1), columns = ["Paese","MSE_pre", "MAE_pre","r_pear"] )
print(tab1train)
MSE = []
MAE = []
r_pearson = []
for i in range(len(paesi)):
    MSE.append(mean_squared_error(train_output_va_val[len_val*i:len_val*i + len_val],predictions_val.iloc[:,i].values ))
    MAE.append(mean_absolute_error(train_output_va_val[len_val * i:len_val * i + len_val], predictions_val.iloc[:,i].values))
    r_pearson.append(pearsonr(train_output_va_val[len_val * i:len_val * i + len_val].reshape(-1), predictions_val.iloc[:, i].values)[0])
tab1val = pd.DataFrame(np.concatenate([np.array(paesi).reshape(-1,1),np.array(MSE).reshape(-1,1),np.array(MAE).reshape(-1,1),np.array(r_pearson).reshape(-1,1)], axis = 1), columns = ["Paese","MSE_pre", "MAE_pre","r_pear"] )
print(tab1val)
MSE = []
MAE = []
r_pearson = []
for i in range(len(paesi)):
    MSE.append(mean_squared_error(test_output_va[len_test*i:len_test*i +
```

```
        len_test],predizioni.iloc[:,i].values ))
    MAE.append(mean_absolute_error(test_output_va[len_test * i:len_test *
        i + len_test], predizioni.iloc[:, i].values))
    r_pearson.append(pearsonr(test_output_va[len_test * i:len_test * i +
        len_test].reshape(-1),predizioni.iloc[:, i].values )[0] )
tab1 = pd.DataFrame(np.concatenate([np.array(paesi).reshape(-1,1),np.array
    (MSE).reshape(-1,1),np.array(MAE).reshape(-1,1),np.array(r_pearson).
    reshape(-1,1)] , axis = 1), columns = ["Paese","MSE_pre", "MAE_pre","
    r_pear"] )
print(tab1)
(tab1.iloc[:,-1].apply(pd.to_numeric, errors="coerce") >0).sum()

#controllo solo sulle predizioni targhet
print(mean_squared_error(test_output_va[[i*len_test for i in range(len(
    paesi))]]   ,predizioni.iloc[0,:] ))
print(mean_absolute_error(test_output_va[[i*len_test for i in range(len(
    paesi))]]   ,predizioni.iloc[0,:] ))
print(pearsonr(test_output_va[[i*len_test for i in range(len(paesi))]].
    reshape(-1)   ,predizioni.iloc[0,:].values.reshape(-1)))
#collocazione dei paesi nello spazio tridimensionale individuato dal
    embedding layer.
embedding_model = Model(inputs = best_model.get_layer("country").output,
    outputs = best_model.get_layer("country_emb").output )
posizioni = embedding_model.predict(codes_input_rolling)
posizioni = posizioni.squeeze()
x, y, z = posizioni[:,0], posizioni[:,1], posizioni[:,2]
fig = plt.figure(figsize=(10,7))
ax = fig.add_subplot(111, projection="3d")
ax.scatter(x, y, z, c="b", marker="o")
for i, country in enumerate(paesi):
    ax.text(x[i], y[i], z[i], country, fontsize=8)
ax.set_xlabel("X")
ax.set_ylabel("Y")
ax.set_zlabel("Z")
```

```
plt.title("Risultato Dell'Embedding Layer")
plt.savefig("Embedding.png")
plt.show(block = True)



#Tentativo di applicazione del boosting
#si estraggono le predizioni al layer bottleneck come indicato nella tesi per il trainig, validation e test set.
intermediate_model = Model(inputs=[best_model.get_layer("lag").output, best_model.get_layer("regressori").output, best_model.get_layer("country").output], outputs=best_model.get_layer("bottleneck").output)
bottleneck_train = intermediate_model.predict([train_hist_va_train, input_mod1.values, input_mod1_codes.values])
bottleneck_val = intermediate_model.predict([train_hist_va_val, input_mod1_val.values, input_mod1_val_codes.values])
bottleneck_test = predittore_bottleneck2(paesi=paesi,test_hist_va=test_hist_va ,test_output_va=test_output_va,
                                input_mod = input_mod1_test_predicted, input_mod_country = input_mod1_test_codes, len_test=len_test, model=intermediate_model, axis = 0, pred = predizioni)

#estrazione dei residui sul trainig e valdiation set
previsioni = best_model.predict(x=[train_hist_va_train,input_mod1, input_mod1_codes])
previsioni_val = best_model.predict(x=[train_hist_va_val, input_mod1_val, input_mod1_val_codes])
previsioni_test = np.concatenate([np.array(predizioni.iloc[:,i]) for i in range(len(paesi))], axis=0).reshape(-1,1)
residui_train = train_output_va_train - previsioni
residui_val = train_output_va_val - previsioni_val

```

```
#costruzione della grid di parametri per il tuning del booster
param_grid = {
    "n_estimators": [100, 200, 500],
    "max_depth": [2, 3, 4, 5],
    "learning_rate": [0.01, 0.05, 0.1],
    "subsample": [0.7, 0.9, 1.0],
    "colsample_bytree": [0.7, 0.9, 1.0]
}

#costruzione del generico modello di boosting che va inserito in
    RandomizedSearchCV di #sklearn
xgb_model = xgb.XGBRegressor(objective="reg:squarederror",
    early_stopping_rounds=50)


search = RandomizedSearchCV(
    estimator=xgb_model,
    param_distributions=param_grid,
    n_iter=20,
    scoring="neg_mean_squared_error",
    cv=3,
    verbose=1,
    n_jobs=-1
)

search.fit(bottleneck_train, residui_train,
           eval_set=[(bottleneck_val, residui_val)],
           verbose=0)

print("Best␣params:", search.best_params_)

#caricamento stimatore dei residui con parametri che realizzano la miglior
    performance secondo la 3fold cross-valdiation
# e tenuto conto del validation set per l'early stop
```

```
best_xgb = search.best_estimator_

#predizioni di boosting sui residui
boost_train = best_xgb.predict(bottleneck_train).reshape(-1,1)
boost_val   = best_xgb.predict(bottleneck_val).reshape(-1,1)
boost_test  = best_xgb.predict(bottleneck_test).reshape(-1,1)

#predizioni finali = rete + boosting
pred_finali_train = previsioni + boost_train
pred_finali_val   = previsioni_val + boost_val
pred_finali  = previsioni_test + boost_test

#costruzione dei dataframe finali
predictions_train = pd.concat([pd.DataFrame(pred_finali_train[
    indici_train1+len(indici_train1)*i], columns=[paesi[i]]) for i in range
    (len(paesi))], axis=1 )
predictions_val = pd.concat([pd.DataFrame(pred_finali_val[np.array(list(
    range(0,len(indici_val))))+len(indici_val)*i], columns=[paesi[i]]) for
    i in range(len(paesi))], axis=1 )
reali_train = pd.concat([pd.DataFrame(train_output_va_train[indici_train1+
    len(indici_train1)*i], columns=[paesi[i]]) for i in range(len(paesi))],
     axis=1 )
reali_val = pd.concat([pd.DataFrame(train_output_va_val[np.array(list(
    range(0,len(indici_val))))+len(indici_val)*i], columns=[paesi[i]]) for
    i in range(len(paesi))], axis=1 )

indici_pred = np.linspace(start=0,stop=len(indici_test)-1, num= len(
    indici_test), dtype=int)
predictions_test = pd.concat([pd.DataFrame(pred_finali[indici_pred+len(
    indici_pred)*i], columns=[paesi[i]]) for i in range(len(paesi))], axis
    =1 )
reali_test = pd.concat([pd.DataFrame(test_output_va[indici_pred+len(
    indici_pred)*i], columns=[paesi[i]]) for i in range(len(paesi))], axis
    =1 )
```

```
#plot predizioni
plot_rvp(paesi,reali_train,predictions_train,reali_val, predictions_val, reali_test, predictions_test)

#analisi performance
print("MSE:", mean_squared_error(test_output_va, np.concatenate([predictions_test.iloc[:,i] for i in range(len(paesi))],axis=0 )) )
print("MAE:", mean_absolute_error(test_output_va,  np.concatenate([predictions_test.iloc[:,i] for i in range(len(paesi))],axis=0 )))
print("R2:", r2_score(test_output_va, np.concatenate([predictions_test.iloc[:,i] for i in range(len(paesi))],axis=0 )))
print("Rpearson:", pearsonr(test_output_va.reshape(-1), np.concatenate([predictions_test.iloc[:,i] for i in range(len(paesi))],axis=0 ))[0] )

MSE = []
MAE = []
r_pearson = []
for i in range(len(paesi)):
    MSE.append(mean_squared_error(train_output_va_train[len_train*i: len_train*i + len_train],predictions_train.iloc[:,i].values ))
    MAE.append(mean_absolute_error(train_output_va_train[len_train * i: len_train * i + len_train], predictions_train.iloc[:, i].values))
    r_pearson.append(pearsonr(train_output_va_train[len_train * i: len_train * i + len_train].reshape(-1),predictions_train.iloc[:, i].values)[0])
tab2train = pd.DataFrame(np.concatenate([np.array(paesi).reshape(-1,1),np.array(MSE).reshape(-1,1),np.array(MAE).reshape(-1,1), np.array(r_pearson).reshape(-1,1)], axis = 1), columns = ["Paese","MSE_post", "MAE_post","r_pear_post"] )
print(tab2train)

tab_train = pd.concat([tab1train["Paese"],tab1train["MSE_pre"],tab2train["MSE_post"],tab1train["MAE_pre"],tab2train["MAE_post"],tab1train["r_pear
```

```
    "],tab2train["r_pear_post"]],axis=1 )
with open("tab_train.tex", "w") as f:
    f.write(tab_train.to_latex(index=False))
print((tab1train.iloc[:,[1,2]].apply(pd.to_numeric, errors="coerce").
    values -tab2train.iloc[:,[1,2]].apply(pd.to_numeric, errors="coerce")).
    values >0)

MSE = []
MAE = []
r_pearson = []
for i in range(len(paesi)):
    MSE.append(mean_squared_error(train_output_va_val[len_val*i:len_val*i
        + len_val],predictions_val.iloc[:,i].values ))
    MAE.append(mean_absolute_error(train_output_va_val[len_val * i:len_val
         * i + len_val], predictions_val.iloc[:, i].values))
    r_pearson.append(pearsonr(train_output_va_val[len_val * i:len_val * i
        + len_val].reshape(-1),predictions_val.iloc[:, i].values)[0])
tab2val = pd.DataFrame(np.concatenate([np.array(paesi).reshape(-1,1),np.
    array(MSE).reshape(-1,1),np.array(MAE).reshape(-1,1), np.array(
    r_pearson).reshape(-1,1)], axis = 1), columns = ["Paese","MSE_post", "
    MAE_post","r_pear_post"] )
print(tab2val)

print((tab1val.iloc[:,[1,2]].apply(pd.to_numeric, errors="coerce").values
    -tab2val.iloc[:,[1,2]].apply(pd.to_numeric, errors="coerce")).values
    >0)
tab_val = pd.concat([tab1val["Paese"],tab1val["MSE_pre"],tab2val["MSE_post
    "],tab1val["MAE_pre"],tab2val["MAE_post"],tab1val["r_pear"],tab2val["
    r_pear_post"]],axis=1 )
with open("tab_val.tex", "w") as f:
    f.write(tab_val.to_latex(index=False))

MSE = []
MAE = []
```

```
r_pearson = []
for i in range(len(paesi)):
    MSE.append(mean_squared_error(test_output_va[len_test*i:len_test*i + len_test],predictions_test.iloc[:,i].values ))
    MAE.append(mean_absolute_error(test_output_va[len_test * i:len_test * i + len_test], predictions_test.iloc[:, i].values))
    r_pearson.append(pearsonr(test_output_va[len_test * i:len_test * i + len_test].reshape(-1), predictions_test.iloc[:, i].values)[0])
tab2 = pd.DataFrame(np.concatenate([np.array(paesi).reshape(-1,1),np.array(MSE).reshape(-1,1),np.array(MAE).reshape(-1,1), np.array(r_pearson).reshape(-1,1)], axis = 1), columns = ["Paese","MSE_post", "MAE_post","r_pear_post"] )
print(tab2)
print(mean_squared_error(test_output_va[[i*len_test for i in range(len(paesi))]],predictions_test.iloc[0,:] ))
print(mean_absolute_error(test_output_va[[i*len_test for i in range(len(paesi))]],predictions_test.iloc[0,:] ))
print(pearsonr(test_output_va[[i*len_test for i in range(len(paesi))]].reshape(-1),predictions_test.iloc[0,:].values.reshape(-1)))
print((tab1.iloc[:,[1,2]].apply(pd.to_numeric, errors="coerce").values - tab2.iloc[:,[1,2]].apply(pd.to_numeric, errors="coerce").values) >0)

tab_test = pd.concat([tab1["Paese"],tab1["MSE_pre"],tab2["MSE_post"],tab1["MAE_pre"],tab2["MAE_post"],tab1["r_pear"],tab2["r_pear_post"]],axis=1)
with open("tab.tex", "w") as f:
    f.write(tab_test.to_latex(index=False))
```

Funzione: build_model

```
def build_model(hp):
    """
    Questa funzione va utilizzata con tuner allo scopo di individuare gli iperparametri ottimali
    :param hp: iperparametri da individuare
```

```
    :return: struttura rete
    """
    va_input = Input(shape=(lookback,1), name="lag")
    Reg_input = Input(shape=(input_mod1.shape[1],), name="regressori")
    Country_input = Input(shape= (1,), name = "country")
    Country_emb = Embedding(input_dim = len(paesi), output_dim = 3, name =
        "country_emb" )(Country_input)
    Country_emb = Flatten()(Country_emb)
    # dimensione LSTM come iperparametro
    x = LSTM(hp.Int("units_lstm", min_value=2, max_value=20, step=1))(
       va_input)
    x = Concatenate()([x, Reg_input,Country_emb])
    x = Dense(hp.Int("units_dense", min_value=2, max_value=32, step=1),
              activation="tanh")(x)
    output_layer = Dense(1)(x)

    model = Model(inputs=[va_input, Reg_input, Country_input], outputs=
       output_layer)
    model.compile(
        loss="mse",
        optimizer=Nadam(learning_rate=hp.Choice("lr", [0.001, 0.005,
           0.01]))
    )
    return model
```

Funzione: build_fixed_model

```
def build_fixed_model(params):
    """
    Questa funzione serve per costruire la rete coi parametri ottimali
       individuati da tuner search
    :param params: Parametri ottimimi individuati con tuner search
    :return: struttura rete con parametri inseriti
    """
    va_input = Input(shape=(lookback,1), name="lag")
```

```
    Reg_input = Input(shape=(input_mod1.shape[1],), name="regressori")

    Country_input = Input(shape=(1,), name="country")
    Country_emb = Embedding(input_dim=len(paesi), output_dim=3, name="
        country_emb")(Country_input)
    Country_emb = Flatten()(Country_emb)
    x = LSTM(params["units_lstm"])(va_input)
    x = Concatenate()([x, Reg_input, Country_emb])
    x = Dense(params["units_dense"], activation="tanh", name = "bottleneck
        ")(x)
    output_layer = Dense(1)(x)

    model = Model(inputs=[va_input, Reg_input, Country_input], outputs=
        output_layer)
    model.compile(
        loss="mse",
        optimizer=Nadam(learning_rate=params["lr"])
    )
    return model
```

Funzione: predittore_bottleneck_2

```
def predittore_bottleneck2(paesi,test_hist_va,test_output_va, input_mod,
    input_mod_country ,len_test,model, axis, pred):
    """
    funzione per fare la predizione del botteleneck sul test set in
        maniera iterativa
    :param paesi: lista dei paesi
    :param test_hist_va: VA storici
    :param test_output_va: VA in output
    :param input_mod: input diversi dal VA
    :param len_test: lunghezza del test set
    :param model: modello di previsione
    :param axis: asse su cui collegare i gruppi di paesi (0,1)
    :param pred: predizioni gia esistenti
```

```
    :return: df di predizioni per ogni paese
    """
    predizioni = [[] for i in range(len(paesi))]
    va_input_rolling = test_hist_va[list(range(0, len(test_output_va),
        len_test)),:,:]
    for i in range(len_test):
        country_input_rolling = input_mod_country.iloc[np.array(list(range
            (0, len(test_output_va), len_test)))+i]
        X_input_rolling = input_mod.iloc[np.array(list(range(0, len(
            test_output_va), len_test)))+i,:].values
        predictions_test = model.predict(x=[va_input_rolling,
            X_input_rolling, country_input_rolling])
        for j in range(len(paesi)):
            va_input_rolling[j,0:lookback-1,:] = va_input_rolling[j,1:,:]
            predizioni[j].append(predictions_test[j,:])
            va_input_rolling[j, lookback-1, :] = pred.iloc[i,j]
    predizioni = pd.DataFrame(
        np.concatenate([np.array(predizioni[i]) for i in range(len(paesi))
            ], axis=axis))

    return predizioni
```

Elenco delle figure

1.1 Struttura del Solvency Capital Requirement (SCR). Fonte: QIS5 Technical Specifications, pag. 90. 6

2.1 Esempio alternanza segni delle posizioni per hedging di uno ZCB con scadenza superiore al LLP. 28

2.2 Grafico dei tassi di sconto interpolati (sopra) e grafico dei forward rate istantanei impliciti nel modello Smith-Wilson (sotto). 33

2.3 Grafico di confronto tra interpolazione con metodo Smith-Wilson e modello Ornstein-Uhlenbeck 36

4.1 Serie storica dei Volatility Adjustment. 52

4.2 Matrici di correlazione 54

4.3 FFN rappresentazione grafica 56

4.4 Deep RN con profondità di due layer 58

4.5 RN con profondità di 1 layer con FFN 58

4.6 Struttura di network utilizzata per il VA. 61

4.7 Serie storica del VSTOXX mensile 74

4.8 ACF del VSTOXX 74

4.9 PACF del VSTOXX 75

4.10 Previsioni del VSTOXX con frequenza mensile 77

4.11 Grafico con volatilità condizionata storica implicata dal modello AR(1)-GARCH(1,1) (linea arancione) e serie storica del VSTOXX (linea blu) 77

4.12 Serie storica prezzo Brent in euro 78

4.13 Serie storica della differenza di prezzo del Brent in dollari 79

4.14 ACF delle differenze del prezzo del Brent 79

4.15 PACF delle differenze del prezzo del Brent 80

4.16 Previsioni della media (che è costante, linea blu) e volatilità (linea verde) dei rendimenti del Brent 82
4.17 Volatilità storiche dei rendimenti del Brent secondo modello GARCH(1,1) 83
4.18 Serie storica del prezzo del gas naturale in dollari 83
4.19 Serie storica delle differenze di prezzo del gas naturale 84
4.20 ACF delle differenze del prezzo del gas naturale 84
4.21 PACF delle differenze del prezzo del gas naturale 85
4.22 Previsioni delle differenze di prezzo del gas naturale 87
4.23 Previsioni dei rendimenti (linea blu) e delle volatilità (linea verde) dei rendimenti del gas naturale secondo il modello Zero ARCH(1). 87
4.24 Grafico volatilità storiche previste dal modello Zero ARCH(1) per i rendimenti del gas naturale. 88
4.25 Serie storica del cambio euro-dollaro 89
4.26 Serie storica delle differenze nel cambio euro-dollaro 89
4.27 ACF delle differenze nel cambio euro-dollaro 90
4.28 PACF delle differenze nel cambio euro-dollaro 90
4.29 Previsioni del valore atteso (linea blu) e della volatilità condizionata (linea verde) nei rendimenti del cambio euro-dollaro 92
4.30 Volatilità storiche (linea arancione) implicite nella serie dei rendimenti del cambio Eur-Usd (linea blu) 93
4.31 Rendimenti mensili del fondo iShares Core EUR Corp Bond. 93
4.32 ACF dei rendimenti corporate. 94
4.33 PACF dei rendimenti corporate. 94
4.34 Forecast della media e della volatilità dei rendimenti corporate. 96
4.35 Volatilità storiche dei rendimenti corporate. 96
4.36 Struttura Rete HICP 98
4.37 Predizioni rete HICP 99
4.38 Modello per i rendimenti degli indici azionari. 100

4.39 Predizioni modello per indici azionari 101
4.40 Predizioni modello per rendimenti S&P500. 102
4.41 Modello per gli yield a 5 anni dei titoli di stato dei diversi paesi 102
4.42 previsioni dello yield a 5 anni dei diversi paesi 103
4.43 Modello per gli yield a 5 anni dei titoli di stato dei diversi paesi 104
4.44 previsioni della rete dello spread 1-5 anni dei diversi paesi 105
4.45 Modello per gli yield a 5 anni dei titoli di stato dei diversi paesi 106
4.46 previsioni dello spread 5-10 anni dei diversi paesi 107
4.47 Previsioni rete pre-boost primi 9 paesi (modello 1) 108
4.48 Previsioni rete pre-boost restanti 9 paesi (modello 1) 109
4.49 Previsioni rete + XGboost primi 9 paesi (modello 1) 110
4.50 Previsioni rete + XGboost restanti 9 paesi (modello 1) 111
4.51 Previsioni rete primi 9 paesi utilizzando le vere realizzazioni dei regressori (modello 1) 115
4.52 Previsioni rete dei restanti 9 paesi utilizzando le vere realizzazioni dei regressori (modello 1) 116
4.53 Matrici di correlazione 118
4.54 Previsioni rete pre-boost primi 9 paesi (Modello 2) 119
4.55 Previsioni rete pre-boost restanti 9 paesi (Modello 2) 120
4.56 Previsioni rete + XGboost primi 9 paesi (Modello 2) 121
4.57 Previsioni rete + XGboost restanti 9 paesi (Modello 2) 122

Bibliografia

Barucci, Emilio, Daniele Marazzina e Edit Rroji (2023). «An investigation of the Volatility Adjustment». In: *Decisions in Economics and Finance*, pp. 1–31.

Bianchetti, Marco e Mattia Carlicchi (2013). «Markets evolution after the credit crunch». In: *arXiv preprint arXiv:1301.7078*.

Commissione Europea (gen. 2015). *Regolamento delegato (UE) 2015/35 della Commissione,* Gazzetta ufficiale dell'Unione europea, L 12/1 del 17.1.2015.

Dozat, Timothy (2016). «Incorporating nesterov momentum into adam». In.

EIOPA (2024). *Technical documentation of the methodology to derive EIOPA's risk-free interest rate term structures*.

Ferrario, Andrea, Alexander Noll e Mario V Wuthrich (2020). «Insights from inside neural networks». In: *Available at SSRN 3226852*.

Hornik, Kurt, Maxwell Stinchcombe e Halbert White (1989). «Multilayer feedforward networks are universal approximators». In: *Neural networks* 2.5, pp. 359–366.

Karlin, Samuel e Yosef Rinott (1983). «M-matrices as covariance matrices of multinormal distributions». In: *Linear algebra and its applications* 52, pp. 419–438.

Kort, Jan de e MH Vellekoop (2016). «Term structure extrapolation and asymptotic forward rates». In: *Insurance: Mathematics and Economics* 67, pp. 107–119.

Lagerås, Andreas e Mathias Lindholm (2016). «Issues with the Smith–Wilson method». In: *Insurance: Mathematics and Economics* 71, pp. 93–102.

Nelson, Charles R e Andrew F Siegel (1987). «Parsimonious modeling of yield curves». In: *Journal of business*, pp. 473–489.

Parlamento europeo e Consiglio dell'Unione europea (apr. 2014). *Direttiva 2014/51/UE del Parlamento europeo e del Consiglio*. Gazzetta ufficiale dell'Unione europea, L 153 del 22.5.2014, pag. 1.

Pelagatti, Matteo M (2015). *Time series modelling with unobserved components*. CRC Press.

Smith e Wilson (2015). *Smith–Wilson Risk-Free Interest Rate Extrapolation Tool*. Ver. 1.1. Excel file. URL: https://register.eiopa.europa.eu/Publications/Standards/Smith-Wilson%20Risk-Free%20Interest%20Rate%20Extrapolation%20Tool%20v1.1_7122015.xlsb.

Svensson, Lars EO (1994). *Estimating and interpreting forward interest rates: Sweden 1992-1994*.

Unione Europea (2025). *Direttiva 2009/138/CE del Parlamento Europeo e del Consiglio*. Versione consolidata al 17 gennaio 2025.

Wuthrich, Mario V e Christoph Buser (2023). «Data analytics for non-life insurance pricing». In: *Swiss Finance Institute Research Paper* 16-68.

Wüthrich, Mario V e Michael Merz (2023). *Statistical foundations of actuarial learning and its applications*. Springer Nature.