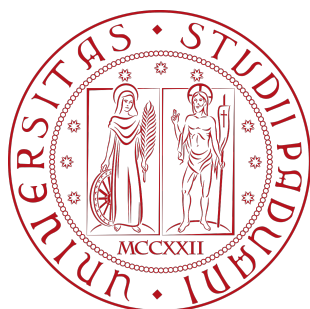


Università degli Studi di Padova
Dipartimento di Scienze Statistiche
Corso di Laurea Triennale in
Statistica per le Tecnologie e le Scienze



RELAZIONE FINALE
**UN APPROCCIO NON PARAMETRICO PER
L'INDIVIDUAZIONE DEL PLATEAU DI MORTALITÀ**

Relatore: Prof. Stefano Mazzuco
Dipartimento di Scienze Statistiche

Laureanda: Aurora Valtulina
Matricola N. 2079527

Anno Accademico 2024/2025

Indice

Introduzione	7
Obiettivo	8
Late-life mortality deceleration	9
Metodi	11
P-splines	11
Strumenti per lo studio del plateau di mortalità	15
Dati	17
Risultati	19
Adattamento delle P-splines	19
Confronto con il modello di Gompertz	22
Analisi della derivata prima	26
Conclusioni	31
Bibliografia	32
Appendice A	34
Appendice B	43

Elenco delle figure

1	Confronto tra i log-tassi stimati tramite P-splines unidimensionali e la crescita lineare del modello di Gompertz, per Stato, nell'anno 2019.	23
2	Confronto tra i log-tassi stimati tramite P-splines unidimensionali, con $\lambda=100$, e la crescita lineare del modello di Gompertz, per Stato, nell'anno 2019.	24
3	Confronto tra i log-tassi stimati tramite P-splines bidimensionali e la crescita lineare del modello di Gompertz, per Stato, nell'anno 2019.	25
4	Confronto tra i log-tassi stimati tramite P-splines bidimensionali, con $\lambda_{age}=10$ e $\lambda_{year}=3.162278$, e la crescita lineare del modello di Gompertz, per Stato, nell'anno 2019.	25
5	Derivate prime dei log-tassi stimati tramite P-splines unidimensionali, nei vari anni, per Stato.	27
6	Derivate prime dei log-tassi stimati tramite P-splines unidimensionali, con $\lambda=100$, nei vari anni, per Stato.	28
7	Derivate prime dei log-tassi stimati tramite P-splines bidimensionali, nei vari anni, per Stato.	29
8	Derivate prime dei log-tassi stimati tramite P-splines bidimensionali, con $\lambda_{age}=10$ e $\lambda_{year}=3.162278$, nei vari anni, per Stato.	29

9	Andamento delle derivate prime dei log-tassi stimati per ciascuno Stato nel 2019. In alto a sinistra la derivata fa riferimento ai tassi stimati con P-splines unidimensionali, in alto a destra con P-splines bidimensionali, in basso a sinistra con P-splines unidimensionali con $\lambda=100$, in basso a destra con P-splines bidimensionali con $\lambda_{age}=10$ e $\lambda_{year}=3.162278$.	30
10	Confronto tra i tassi di mortalità reali e quelli stimati tramite P-splines unidimensionali, per Stato, nell'anno 1960.	34
11	Confronto tra i tassi di mortalità reali e quelli stimati tramite P-splines unidimensionali, per Stato, nell'anno 2019.	35
12	Confronto tra i tassi di mortalità reali e quelli stimati tramite P-splines unidimensionali, con $\lambda=100$, per Stato, nell'anno 1960.	36
13	Confronto tra i tassi di mortalità reali e quelli stimati tramite P-splines unidimensionali, con $\lambda=100$, per Stato, nell'anno 2019.	36
14	Confronto tra i tassi di mortalità reali e quelli stimati tramite P-splines bidimensionali, per Stato, nell'anno 1960. .	37
15	Confronto tra i tassi di mortalità reali e quelli stimati tramite P-splines bidimensionali, per Stato, nell'anno 2019. .	37
16	Confronto tra i tassi di mortalità reali e quelli stimati tramite P-splines biidimensionali, con $\lambda_{age}=10$ e $\lambda_{year}=3.162278$, per Stato, nell'anno 1960.	38
17	Confronto tra i tassi di mortalità reali e quelli stimati tramite P-splines bidimensionali, con $\lambda_{age}=10$ e $\lambda_{year}=3.162278$, per Stato, nell'anno 2019.	39
18	Residui di devianza standardizzati di tutti gli anni, per Stato, con stime ottenute tramite P-splines unidimensionali.	40
19	Residui di devianza standardizzati di tutti gli anni, per Stato, con stime ottenute tramite P-splines unidimensionali, con $\lambda=100$	40

20	Residui di devianza standardizzati di tutti gli anni, per Stato, con stime ottenute tramite P-splines bidimensionali.	41
21	Residui di devianza standardizzati di tutti gli anni, per Stato, con stime ottenute tramite P-splines bidimensionali, con $\lambda_{age}=10$ e $\lambda_{year}=3.162278$.	41

Elenco delle tabelle

1	Gradi di libertà (<i>gdl</i>) e valori di AIC complessivi per Stato, risultanti dall'adattamento di P-splines unidimensionali.	20
2	Gradi di libertà (<i>gdl</i>) e valori di AIC complessivi per Stato, risultanti dall'adattamento di P-splines unidimensionali, con $\lambda=100$	20
3	Valori dei parametri di penalizzazione delle P-splines bidimensionali, per Stato.	21
4	Gradi di libertà (<i>gdl</i>) e valori di AIC complessivi per Stato, risultanti dall'adattamento di P-splines bidimensionali.	21
5	Gradi di libertà (<i>gdl</i>) e valori di AIC complessivi per Stato, risultanti dall'adattamento di P-splines bidimensionali, con $\lambda_{age}=10$ e $\lambda_{year}=3.162278$	22
6	Modellazione tramite P-splines unidimensionali per l'Italia: anno, λ , gradi di libertà (<i>gdl</i>), AIC e RMSE.	43
7	Modellazione tramite P-splines unidimensionali per il Giappone: anno, λ , gradi di libertà (<i>gdl</i>), AIC e RMSE.	45
8	Modellazione tramite P-splines unidimensionali per la Svezia: anno, λ , gradi di libertà (<i>gdl</i>), AIC e RMSE.	47
9	Modellazione tramite P-splines unidimensionali per gli USA: anno, λ , gradi di libertà (<i>gdl</i>), AIC e RMSE.	49
10	Modellazione tramite P-splines unidimensionali, con $\lambda=100$, per l'Italia: anno, gradi di libertà (<i>gdl</i>), AIC e RMSE.	51
11	Modellazione tramite P-splines unidimensionali, con $\lambda=100$, per il Giappone: anno, gradi di libertà (<i>gdl</i>), AIC e RMSE.	54

12	Modellazione tramite P-splines unidimensionali, con $\lambda=100$, per la Svezia: anno, gradi di libertà (<i>gdl</i>), AIC e RMSE. .	56
13	Modellazione tramite P-splines unidimensionali, con $\lambda=100$, per gli USA: anno, gradi di libertà (<i>gdl</i>), AIC e RMSE. .	58
14	Valori dell'errore quadratico medio per le stime ottenute con P-splines bidimensionali, per Stato.	60
15	Valori dell'errore quadratico medio per le stime ottenute con P-splines bidimensionali, con $\lambda_{age}=10$ e $\lambda_{year}=3.162278$, per Stato.	63

Introduzione

Lo studio dei tassi di mortalità costituisce un pilastro fondamentale nell'analisi della dinamica demografica, della longevità e del processo di invecchiamento. Nel corso del tempo, in numerosi Paesi, si è verificato un aumento dell'aspettativa di vita e dell'età modale alla morte, come conseguenza di un declino dei tassi di mortalità. Infatti, secondo Rau et al. (2008), l'estensione della longevità è stata inizialmente dovuta a una riduzione della mortalità infantile e giovanile, ma a partire dalla metà del XX secolo è stata principalmente guidata da un calo dei tassi di mortalità della tarda età. Insieme al parallelo declino della fecondità, l'aumento dell'aspettativa di vita, che, da uno studio di Oppen e Vaupel (2002), per le donne sembra apprestarsi attorno ai 2.5 anni per decade a partire dal 1840, ha portato ad avere una fetta di popolazione sempre più ampia composta da persone in età anziana, che quindi costituiscono un peso importante nelle decisioni riguardanti il sistema sanitario, pensionistico e di previdenza sociale, influenzando dunque la politica di ogni Stato. Alla luce di questa dinamica l'analisi della mortalità in età avanzate ha acquisito crescente rilevanza, portando numerosi studiosi a concentrarsi sulla modellazione dei relativi tassi. Tra i fenomeni più discussi figura la possibile esistenza di un plateau di mortalità a età estreme, in cui i tassi di mortalità anziché continuare a crescere esponenzialmente secondo la legge di Gompertz (1825), tendono a decelerare o stabilizzarsi, suggerendo quindi la possibile assenza di un limite biologico. Sono stati proposti diversi metodi per studiare tale fenomeno, che fanno principalmente riferimento a un approccio parametrico. I modelli parametrici esprimono il tasso di mortalità come una funzione matematicamente definita da un numero limitato di parametri. Essi solitamente presentano un'elevata interpretabilità e semplicità, tuttavia mostrano dei limiti non trascurabili, specialmente quando si tratta dello studio della tarda età. Impongono infatti una forma funzionale rigida

che potrebbe non adattarsi correttamente ai dati osservati, oppure non cogliere variazioni locali. Dunque, negli ultimi decenni si sono diffusi anche approcci non parametrici e semi-parametrici, che permettono una maggiore adattabilità ai dati, consentendo quindi di esplorare fenomeni come la late-life mortality deceleration in modo flessibile e robusto, mantenendo una buona interpretabilità.

Obiettivo

La presente tesi vuole proporre uno studio della mortalità a età avanzate sfruttando un approccio non parametrico, considerando i tassi istantanei di mortalità provenienti da Paesi pionieri della longevità, quali Giappone, Italia e Svezia, o con particolari caratteristiche demografiche, come gli USA, che invece presentano una mortalità in età giovanile e adulta maggiore rispetto agli altri Stati e quindi una speranza di vita inferiore. In particolare vengono applicati due tipi di P-splines, Penalized splines, proposte da Eilers e Marx (1996), che permettono di ottenere curve lisce dei tassi di mortalità, sfruttando una penalizzazione che consente di limitare il rumore presente nei dati osservati e quindi evitare fenomeni di overfitting. Vengono utilizzate P-splines unidimensionali sui tassi di mortalità di ogni anno, applicando una penalizzazione sull'età, e P-spline bidimensionali, che modellano invece le superfici di mortalità, imponendo una lisciatura sia sull'asse dell'età che su quello degli anni, garantendo coerenza tra i profili di mortalità nel tempo. Per entrambe le modellazioni è stata valutata la scelta di parametri di penalizzazione univoci, per garantire un confronto più robusto tra le curve di mortalità. Ogni curva stimata, che rappresenta l'andamento dei tassi di mortalità rispetto all'età, viene comparata con un modello di Gompertz, in particolare viene osservata la progressione dei log-tassi stimati rispetto all'andamento lineare che dovrebbero avere in caso di crescita esponenziale della mortalità. Inoltre, viene calcolata analiticamente la derivata

prima del logaritmo dei tassi stimati, in questo modo si possono trarre indicazioni sul tasso di crescita relativo della mortalità, infatti una decrescita della derivata è indice di un rallentamento della mortalità, mentre se tende a zero, di una possibile stabilizzazione e quindi di un plateau. Questo permette di avere informazioni sul fenomeno della late-life mortality deceleration, valutando se c'è un effettivo rallentamento dei tassi di mortalità nelle età avanzate.

Late-life mortality deceleration

Per gran parte dell'età adulta i tassi di mortalità aumentano esponenzialmente con l'età, secondo la legge di Gompertz, tuttavia alcuni studi evidenziano un rallentamento della crescita nelle età molto avanzate, fino ad una possibile convergenza ad un valore finito, il cosiddetto plateau di mortalità. Questa teoria è nota come late-life mortality deceleration, abbreviata in LLMD. Sono state ipotizzate varie cause per questo fenomeno, come l'eterogeneità selettiva, trattata da Vaupel et al. (1979), secondo cui i soggetti più fragili tendono a morire a età più giovani, per cui i sopravvissuti a età avanzate tendono ad avere caratteristiche fisiche favorevoli, come sistemi immunitari più robusti o stili di vita sani, il che porta a una decelerazione apparente della mortalità. Un'altra interpretazione è legata al rischio individuale, ipotizzata da Custinger et al. (1992), che suggerisce che la decelerazione della mortalità si verifichi a un livello individuale, poiché alcune cause di morte diventano meno probabili, ad esempio perché gli individui anziani tendono a essere meno attivi o i meccanismi di deterioramento dell'organismo possono rallentare o stabilizzarsi in età molto avanzate. Negli ultimi anni, grazie a una maggior qualità dei dati e all'aumento della longevità, a cui è conseguito un incremento consistente nel numero di centenari e supercentenari, sono stati condotti numerosi studi demografici, che hanno portato a risultati diversi. Sebbene una parte di queste analisi neghi l'evidenza di

un plateau di mortalità, come quella condotta da Gavrilova e Gavrilov (2020), che ha mostrato che la mortalità dopo i 110 anni non è costante, specialmente se si considerano i dati più recenti, oppure lo studio di Dang et al. (2023) sui semi-supercentenari francesi, che ha concluso che un livello di mortalità costante oltre i 105 anni può essere statisticamente escluso, sono altrettanto presenti studi a favore della teoria della late-life mortality deceleration, come quello di Barbi et al. (2018), in cui viene confermata la presenza di un plateau di mortalità oltre i 105 anni sui dati italiani, oppure quello di Gampe (2021) che verifica che una forza di mortalità costante oltre i 108 anni si adatta bene ai dati italiani e francesi, mentre Einmahl et al. (2019) lo confermano per i dati olandesi. Questi risultati portano ad ipotizzare che se il plateau di mortalità non viene osservato, probabilmente è dovuto al fatto che non si hanno ancora abbastanza dati per età sufficientemente elevate. Tuttavia, uno studio condotto da Newman (2018) asserisce che gli errori statistici siano la principale causa di un'apparente decelerazione della mortalità umana. Questi errori fanno sia riferimento al campionamento demografico, ai registri di nascita e morte e alla segnalazione dell'età, ma anche alla crescente incertezza delle stime dei tassi con l'aumentare dell'età, che non permette di comprendere se il rallentamento della mortalità sia frutto di errori di stima oppure sia reale. L'assenza di un risultato univoco e robusto è proprio legata alla presenza di un numero ancora ridotto di centenari e supercentenari, oltre alla qualità non ottimale di alcuni dati. Dunque, studi futuri che potrebbero usufruire di una miglior qualità dei dati, potrebbero dare conclusioni affidabili sulla presenza o assenza del plateau di mortalità umana.

Metodi

P-splines

I metodi splines rivestono un ruolo importante nell'ambito dell'analisi non parametrica. Essi consentono di ridurre o controllare oscillazioni eccessive nell'interpolazione dei dati, suddividendo il dominio in sottointervalli e adattando polinomi di basso grado su ciascuno di essi, garantendo continuità e regolarità nel passaggio tra un tratto e l'altro. Nel corso degli anni sono stati presentati diversi tipi di splines, come le smoothing splines, le regression splines o le B-splines, proposte da de Boor (1978). Nella presente tesi viene usato un approccio basato sulle P-splines, un'estensione delle B-splines che applica una penalizzazione sulle differenze finite dei coefficienti di regressione. Le B-splines sono funzioni composte da tratti polinomiali connessi in corrispondenza di specifici valori x del dominio, chiamati nodi. Una B-spline di grado q è formata da $q+1$ polinomi di grado q , che si uniscono in q nodi interni, dove si ha continuità delle derivate fino all'ordine $q-1$. Ogni B-spline è positiva solo su un dominio coperto da $q+2$ nodi, mentre altrove è nulla. Per un dato x , $q+1$ B-splines sono non nulle e tranne ai confini del dominio una B-spline si sovrappone con $2q$ tratti polinomiali delle funzioni B-splines vicine. Quando si applicano le B-splines il dominio viene suddiviso in n' sottointervalli, da $n'+1$ nodi. Ogni intervallo viene coperto da $q+1$ B-splines di grado q , dunque il numero totale di nodi per la costruzione delle B-splines è $n'+2q+1$, mentre il numero di B-splines usate è $n = n'+q$. I nodi possono essere o meno equidistanti, ma indipendentemente dal loro tipo di posizionamento, un numero elevato può portare all'overfitting. Sia $\mathbf{B}_j(x; q)$ il valore della j -esima B-spline di grado q nel punto x . L'adattamento di una curva spline ai dati (x, y) si ottiene con una combinazione lineare, con coefficienti α_j , delle B-splines,

secondo:

$$\hat{y}(x) = \sum_{j=1}^n \hat{a}_j \mathbf{B}_j(x; q).$$

Per stimare tale curva è necessario minimizzare la funzione dei minimi quadrati:

$$S = \sum_{i=1}^m \left\{ y_i - \sum_{j=1}^n a_j \mathbf{B}_j(x_i) \right\}^2$$

con m che indica il numero totale di osservazioni. Al fine di evitare un eccessivo adattamento ai dati, Eilers e Marx suggerirono una penalizzazione sulle differenze finite dei coefficienti di B-splines adiacenti, proponendo così le P-splines:

$$S = \sum_{i=1}^m \left\{ y_i - \sum_{j=1}^n a_j \mathbf{B}_j(x_i) \right\}^2 + \lambda \sum_{j=d+1}^n (\Delta^d a_j)^2$$

con d che indica l'ordine della differenza e λ parametro di penalizzazione. Diversamente da altri metodi non parametrici, le P-splines conservano i momenti dei dati e la curva adattata non tende verso zero per valori fuori dal dominio.

Quando si usano le P-splines per la modellazione dei tassi di mortalità, si hanno a disposizione due dataset in input: il numero di decessi e degli esposti al rischio, entrambi organizzati in matrici $m \times n$, in cui le righe indicano le età e le colonne gli anni. Si hanno quindi $\mathbf{Y} = (y_{ij})$ matrice dei decessi ed $\mathbf{E} = (e_{ij})$ matrice degli esposti al rischio.

$$\mathbf{Y} = \begin{bmatrix} y_{11} & y_{12} & \cdots & y_{1n} \\ y_{21} & y_{22} & \cdots & y_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ y_{m1} & y_{m2} & \cdots & y_{mn} \end{bmatrix} \quad \mathbf{E} = \begin{bmatrix} e_{11} & e_{12} & \cdots & e_{1n} \\ e_{21} & e_{22} & \cdots & e_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ e_{m1} & e_{m2} & \cdots & e_{mn} \end{bmatrix}.$$

Si assume che il numero di morti all'età i nell'anno j , quindi y_{ij} , sia distribuito secondo una Poisson di media $e_{ij} \mu_{ij}$, dove μ_{ij} rappresenta

la forza di mortalità, approssimabile con i tassi di mortalità. Si ha dunque: $y_{ij} \sim \mathcal{P}(e_{ij} \mu_{ij})$. Se si lavora in un contesto unidimensionale, considerando fisso un anno o un'età, si estrae da ciascuna matrice una colonna o una riga e, adottando un modello lineare generalizzato con distribuzione di Poisson, si ottiene che il logaritmo dei valori attesi dei decessi è rappresentato da:

$$\ln[\mathbf{E}(y)] = \ln(e) + \ln(\mu) = \ln(e) + \eta.$$

Si descrive quindi la log-mortalità, ossia il predittore η , come una combinazione lineare di B-splines:

$$\eta = \mathbf{B} \boldsymbol{\alpha}$$

con \mathbf{B} matrice di k basi spline equispaziate e $\boldsymbol{\alpha}$ i rispettivi coefficienti. In un modello lineare generalizzato, il termine di penalità viene aggiunto alla log-verosimiglianza, ottenendo così:

$$l^* = l(\boldsymbol{\alpha}; \mathbf{B}, \mathbf{y}) - \frac{1}{2} \lambda \|D_d \boldsymbol{\alpha}\|^2 = l(\boldsymbol{\alpha}; \mathbf{B}, \mathbf{y}) - \frac{1}{2} \boldsymbol{\alpha}^\top P \boldsymbol{\alpha}$$

dove $l(\boldsymbol{\alpha}; \mathbf{B}, \mathbf{y}) \propto \sum_{x=i}^m \{y_x \ln(\mu_x) - e_x \mu_x\}$ è la log-verosimiglianza nel caso di una Poisson, $P = \lambda D_d^\top D_d$ è la matrice di penalizzazione, con λ rispettivo parametro e D_d la matrice delle differenze finite di ordine d . Nell'analisi della mortalità vengono predilette B-splines di grado 3 e differenze di ordine 2. Per la stima dei parametri è necessario massimizzare l^* . Le equazioni di verosimiglianza penalizzata sono date da $\mathbf{B}^\top(\mathbf{y} - \boldsymbol{\mu}) = P \boldsymbol{\alpha}$ e possono essere risolte tramite una versione penalizzata dell'algoritmo dei minimi quadrati pesati iterati, proposto da Nelder e Wedderburn (1972):

$$(\mathbf{B}^\top \tilde{\mathbf{W}} \mathbf{B} + P) \tilde{\boldsymbol{\alpha}} = \mathbf{B}^\top \tilde{\mathbf{W}} \tilde{\mathbf{z}}, \quad \text{con} \quad \tilde{\mathbf{z}} = \frac{\mathbf{y} - \tilde{\boldsymbol{\mu}}}{\tilde{\boldsymbol{\mu}}} + \mathbf{B} \tilde{\boldsymbol{\alpha}}$$

e passo di aggiornamento: $\hat{\boldsymbol{\alpha}}_{t+1} = \left(\mathbf{B}^\top \hat{\mathbf{W}}_t \mathbf{B} + P \right)^{-1} \mathbf{B}^\top \hat{\mathbf{W}}_t \hat{\mathbf{z}}_t$

$\tilde{\boldsymbol{\mu}}$ e $\tilde{\boldsymbol{\alpha}}$ sono le approssimazioni correnti del sistema e $\tilde{\mathbf{W}}$ è la matrice dei pesi, che in questo caso è data da $\tilde{\mathbf{W}} = \text{diag}(\tilde{\boldsymbol{\mu}})$. Se si vuole modellare la superficie di mortalità, considerando quindi età, anno e tasso di mortalità, è possibile applicare P-splines bidimensionali, con due parametri di penalizzazione, uno che fa riferimento all'età e l'altro all'anno. In questo modo è possibile ridurre il rumore, in quanto una penalizzazione sull'asse degli anni forza i tassi a cambiare lentamente nel tempo, garantendo coerenza delle curve di mortalità per età nel tempo. Inoltre permette di avere una maggiore flessibilità, poiché possono essere scelti parametri di lisciatura diversi per età e per anni. Siano \mathbf{B}_x , di dimensione $m \times k_x$, e \mathbf{B}_t , di dimensione $n \times k_t$, le matrici di regressione B-spline rispettivamente di età e di anni, con k_x e k_t numero delle B-splines nei rispettivi domini. La matrice di regressione per un modello bidimensionale è data dal prodotto di Kronecker $\mathbf{B} = \mathbf{B}_t \otimes \mathbf{B}_x$. I tassi di mortalità vengono modellati in modo analogo al caso unidimensionale, considerando però il termine di penalizzazione bidimensionale:

$$P = \lambda_x (I_{k_t} \otimes D_x^\top k_x) + \lambda_t (D_t^\top D_t \otimes I_{k_x})$$

dove D_x è la matrice delle differenze finite che agisce sulle colonne di \mathbf{A} e D_t quella che agisce sulle righe, con \mathbf{A} matrice dei coefficienti di regressione di dimensione $k_x \times k_t$, λ_x e λ_t i rispettivi parametri di penalizzazione e I_{k_t} e I_{k_x} sono le matrici identità di ordine k_t e k_x . La stima dei parametri può essere ottenuta come nel caso unidimensionale, tuttavia per una maggiore facilità e stabilità nei calcoli, è possibile sfruttare la proprietà delle P-splines bidimensionali di appartenere alla classe dei GLAM, Generalized Linear Array Model, e applicare il rispettivo algoritmo di stima. Per entrambi i metodi è possibile calcolare la matrice

cappello H_λ e il numero di gradi di libertà $gdl(\alpha, \lambda)$.

$$H_\lambda = \mathbf{B} \left(\mathbf{B}^\top \tilde{\mathbf{W}} \mathbf{B} + P \right)^{-1} \mathbf{B}^\top \tilde{\mathbf{W}}$$

$$gdl(\alpha, \lambda) = \text{tr}(H_\lambda)$$

Per la scelta del λ ottimale, cioè quello che bilancia bias e varianza nella costruzione del modello, è possibile svolgere una selezione sulla base del Criterio di Informazione di Akaike, AIC, definito come:

$$\text{AIC}(\lambda) = \text{Dev}(\mathbf{y}; \alpha, \lambda) + 2 \cdot gdl(\alpha, \lambda)$$

con $\text{Dev}(\mathbf{y}; \alpha, \lambda) = 2 \sum_{i=1}^n \left\{ y_i \log \left(\frac{y_i}{\hat{\mu}_i} \right) - (y_i - \hat{\mu}_i) \right\}$ e n numero di osservazioni. Alternativamente è possibile seguire l'indicazione del Criterio di Informazione Bayesiano, BIC:

$$\text{BIC}(\lambda) = \text{Dev}(\mathbf{y}; \alpha, \lambda) + \ln(n) \cdot gdl(\alpha, \lambda).$$

L'AIC tende a favorire penalizzazioni, λ , più basse e quindi una curva più flessibile, mentre il criterio BIC seleziona penalizzazioni più alte e quindi una curva più liscia.

Strumenti per lo studio del plateau di mortalità

Per lo studio del comportamento dei tassi di mortalità in età avanzate, sono state svolte due tipologie di analisi, la prima confronta la crescita esponenziale prevista dal modello di Gompertz con quella risultante dalle stime ottenute tramite P-splines, la seconda analizza la derivata prima rispetto all'età, calcolata analiticamente sui log-tassi stimati. Entrambe danno indicazioni su un possibile rallentamento della mortalità nella tarda età.

Il modello di Gompertz prevede una crescita esponenziale della mortalità

per tutta l'età adulta, dai 30 anni fino agli 80 anni. Il tasso di mortalità istantaneo è quindi modellato come $\mu(x) = ae^{bx}$, con x che indica l'età, a il livello iniziale di mortalità e b il tasso di crescita esponenziale della mortalità con l'età. Il rispettivo logaritmo ha forma:

$$\log \mu(x) = \log a + bx$$

dunque, il modello prevede che il logaritmo del tasso di mortalità in funzione dell'età abbia un andamento lineare. Non è quindi atteso un rallentamento della mortalità con l'età, per questo è utile confrontare il modello di Gompertz con i tassi di mortalità stimati tramite P-splines, lavorando in termini di logaritmo. Se la crescita lineare del logaritmo dei tassi di mortalità prevista da Gompertz non viene rispettata dalle stime ottenute con le P-splines, in particolare se queste, per età avanzate, presentano una curvatura sistematicamente sotto la retta, si ha l'indicazione di un possibile rallentamento nella crescita dei tassi di mortalità e, se tende ad un valore costante, di un plateau di mortalità.

Il secondo metodo usato per indagare un possibile rallentamento nella mortalità, consiste nell'analisi dell'andamento della derivata prima rispetto all'età della curva stimata. Quando si modellano i tassi di mortalità tramite P-splines unidimensionali, il log-tasso stimato, ossia il predittore η , è ottenuto come combinazione lineare di B-splines, dunque la derivata rispetto all'età è espressa come:

$$\frac{d}{dx}\eta(x) = \frac{d}{dx}(\mathbf{B}_q(x)) \boldsymbol{\alpha} = \mathbf{B}'_q(x) \boldsymbol{\alpha}$$

con x che indica l'età, \mathbf{B} matrice di regressione B-spline, $\boldsymbol{\alpha}$ coefficienti di regressione e q grado della B-spline. De Boor propose la seguente formula per il calcolo della derivata prima rispetto a x di una B-spline di ordine q :

$$\frac{d\mathbf{B}_{j,q}(x)}{dx} = q \left(\frac{\mathbf{B}_{j,q-1}(x)}{k_{j+q} - k_j} - \frac{\mathbf{B}_{j+1,q-1}(x)}{k_{j+q+1} - k_{j+1}} \right)$$

con $\{k_i\}$ vettore dei nodi e j indice della B-spline. Nello studio della mortalità solitamente $q=3$, per cui si ha:

$$\frac{d\mathbf{B}_{j,3}(x)}{dx} = 3 \left(\frac{\mathbf{B}_{j,2}(x)}{k_{j+3} - k_j} - \frac{\mathbf{B}_{j+1,2}(x)}{k_{j+3+1} - k_{j+1}} \right).$$

Quando invece la modellazione avviene tramite P-splines bidimensionali, con \mathbf{B}_x matrice di regressione B-spline per l'età e \mathbf{B}_t la rispettiva per gli anni, la derivata prima del log-tasso di mortalità rispetto all'età segue la proprietà del prodotto tensoriale del calcolo differenziale e ha forma:

$$\frac{\partial}{\partial x} \eta(x, t) = \mathbf{B}'_x(x) A \mathbf{B}_t(t)^\top$$

con $\mathbf{B}'_x(x)$ calcolata come nel caso unidimensionale. La derivata prima del log-tasso di mortalità misura la variazione relativa del tasso, se essa ha un valore costante, allora si ha una crescita esponenziale come previsto dal modello di Gompertz, mentre se per età avanzate si osserva una decrescita, allora si ha un rallentamento nella curva di mortalità e se tende a zero, il tasso si stabilizza indicando un possibile plateau.

Dati

Per le analisi e le illustrazioni presenti in questa tesi, sono stati impiegati i dati di mortalità di quattro Stati disponibili nello Human Mortality Database (HMD): Italia, Giappone, Svezia e USA. I Paesi scelti sono caratterizzati da un'elevata longevità, che permette uno studio più robusto della mortalità ad età avanzate. Fanno eccezione gli USA, che sono stati inclusi per un confronto demografico. Per ogni Paese si è fatto uso di M_x , tasso di mortalità per l'età x , E_x , numero di esposti al rischio in età x ed è stato ricavato D_x , numero di decessi all'età x , tramite la formula inversa $D_x = M_x E_x$. Gli anni presi in considerazione vanno dal 1960 al 2019, in modo da escludere particolari oscillazioni dei dati dovute a guerre o epidemie che porterebbero ad una distorsione dei

risultati. Il range di età considerato è invece 30-110+, escludendo quindi la mortalità infantile e giovanile per ottenere una migliore modellazione della mortalità per le età avanzate. I dati utilizzati fanno riferimento all'intera popolazione e sono stati preferiti i tassi di mortalità reali M_x a quelli della tavola di mortalità m_x , in quanto questi ultimi presentano una lisciatura che andrebbe a distorcere l'analisi.

Risultati

Adattamento delle P-splines

Per la modellazione dei tassi di mortalità tramite P-splines sono state utilizzate le funzioni `Mort1DSmooth` e `Mort2DSmooth` del pacchetto `MortSmooth` di Camarda (2012) e tutte le analisi sono state svolte in R. La scelta del parametro di penalizzazione è stata effettuata mediante AIC, per garantire una maggior flessibilità delle curve e quindi un miglior adattamento ai dati. Nel caso delle P-splines unidimensionali sono stati modellati i tassi di mortalità di ciascuno Stato, considerando singolarmente ogni anno. In quelle bidimensionali, invece, per ciascun Paese sono stati presi in considerazione contemporaneamente i tassi di mortalità di tutti gli anni, per garantire coerenza tra i profili di mortalità nel tempo. In generale, le curve di mortalità stimate sembrano cogliere bene l'andamento dei tassi reali. La bontà di adattamento viene trattata in Appendice A.

Nell'utilizzo di P-splines unidimensionali il parametro di penalizzazione selezionato assume valori differenti in base all'anno e allo Stato di riferimento, come riportato nelle tabelle in Appendice B. Tuttavia, il numero di λ diversi è limitato a: 0.001, 0.003, 0.010, 0.032, 0.316, 1.000, 3.162, 10.000, 31.623, 100.000, 316.228, 1000.000. In Tabella 1 sono riportati, per ogni Paese, i valori dei gradi di libertà effettivi e i rispettivi AIC calcolati complessivamente. Il numero di gradi di libertà indica la complessità effettiva del modello. In questo caso, a fronte di un numero di basi splines complessive che varia tra 1113 e 1140 a seconda dello Stato, con un'alternanza negli anni tra le 18 e le 19 basi, il confronto con i gradi di libertà suggerisce che la curva possiede una certa flessibilità, filtrando il rumore, ma garantendo regolarità.

Tabella 1: Gradi di libertà (*gdl*) e valori di AIC complessivi per Stato, risultanti dall'adattamento di P-splines unidimensionali.

	ITA	JPN	SWE	USA
<i>gdl</i>	789.9943	788.7122	617.35	818.5377
AIC	5583.951	5573.368	5365.828	5618.496

È stata considerata la scelta, tra i diversi valori di λ selezionati, di un parametro di penalizzazione uguale per tutti gli Stati e gli anni in analisi, valutandolo in termini di minimo AIC complessivo per il periodo compreso tra il 1990 e il 2019, al fine di garantire maggior aderenza ai dati più recenti e più qualitativi. È stato selezionato un valore di λ pari a 100. In Tabella 2 vengono riportati i gradi di libertà effettivi e i rispettivi AIC complessivi di questo adattamento, per i diversi Paesi. Si osserva che la scelta di un parametro univoco non porta a grandi variazioni in termini di gradi di libertà e quindi di complessità del modello, ma provoca un generale innalzamento dell'AIC, specialmente per gli USA, dove si ha un aumento su tutto il periodo considerato, ma particolarmente rilevante per i primi anni. Dunque, per la modellazione delle curve di mortalità sui singoli anni, risulta preferibile considerare il miglior λ del caso specifico, per garantire un miglior bilanciamento tra bontà di adattamento e complessità del modello.

Tabella 2: Gradi di libertà (*gdl*) e valori di AIC complessivi per Stato, risultanti dall'adattamento di P-splines unidimensionali, con $\lambda=100$.

	ITA	JPN	SWE	USA
<i>gdl</i>	796.8287	837.1095	622.8636	928.703
AIC	7860.838	9518.035	5502.163	32280.39

Per le P-splines bidimensionali, invece, la stima dei parametri di penalizzazione per ogni Stato è univoca per struttura e viene riportata in Tabella 3, mentre in Tabella 4 sono indicati i gradi di libertà e l'AIC complessivi dei modelli. Anche in questo caso si osservano dei valori dei

λ diversi in base al Paese, ma si ha sempre una penalizzazione sull'asse degli anni minore di quella applicata sull'asse dell'età, stando a sottolineare una variazione della longevità e della qualità dei dati nel tempo, che rende necessario adattare soluzioni più flessibili nella dimensione temporale. Inoltre, il numero di gradi di libertà effettivi, se confrontato con il numero di basi splines, 285, indica che la curva è abbastanza flessibile per seguire l'andamento dei dati, ignorando però le fluttuazioni casuali.

Tabella 3: Valori dei parametri di penalizzazione delle P-splines bidimensionali, per Stato.

	ITA	JPN	SWE	USA
λ_{age}	31.62278	31.62278	100	10
λ_{year}	1	10	3.162278	3.162278

Tabella 4: Gradi di libertà (gdl) e valori di AIC complessivi per Stato, risultanti dall'adattamento di P-splines bidimensionali.

	ITA	JPN	SWE	USA
gdl	175.0598	175.8624	138.1631	206.1972
AIC	4975.031	4975.901	4938.183	5006.149

Anche in questo caso è stata valutata, in modo analogo alle P-splines unidimensionali, la scelta di parametri di penalizzazione uguali per tutti gli Stati. In particolare, sulla base del minor AIC, è stato selezionato per l'età $\lambda_{age}=10$, mentre per gli anni $\lambda_{year}=3.162278$. In Tabella 5 sono riportati i gradi di libertà effettivi e l'AIC per i diversi Paesi. La scelta di una penalizzazione univoca non causa grandi variazioni in termini di complessità del modello, valutata tramite i gradi di libertà, ma porta a un innalzamento dell'AIC, che indica la preferenza di λ diversi per i vari Stati, se si vuole garantire maggior coerenza con i dati e limitare la complessità del modello.

Tabella 5: Gradi di libertà (gdl) e valori di AIC complessivi per Stato, risultanti dall'adattamento di P-splines bidimensionali, con $\lambda_{age}=10$ e $\lambda_{year}=3.162278$.

	ITA	JPN	SWE	USA
gdl	212.144	219.4837	181.0169	206.1972
AIC	21731.97	20994.321	6040.108	5006.149

Confronto con il modello di Gompertz

Per verificare la presenza di una decelerazione dei tassi di mortalità nelle età più avanzate, dunque la teoria della late-life mortality deceleration, è stata condotta una prima valutazione sul tipo di crescita dei log-tassi stimati. Su questi è stato adattato un modello lineare, in quanto la legge di Gompertz prevede un andamento esponenziale dei tassi di mortalità, quindi lineare dei relativi logaritmi. In particolare, per stimare il modello, sono stati considerati i log-tassi previsti dalle P-splines, compresi nel range d'età 70-100 anni, poiché si è osservato che, nel tempo, la crescita esponenziale si è progressivamente spostata verso destra sull'asse dell'età, in conseguenza a una maggiore longevità. In questo modo è possibile cogliere in maniera più accurata l'effettivo trend esponenziale e verificare se, alle età più avanzate, i tassi stimati continuano a seguirlo o se invece mostrino un rallentamento. Per tutti gli Stati e per tutti gli anni analizzati, si osserva una decelerazione dei log-tassi di mortalità stimati rispetto alla crescita lineare prevista dal modello di Gompertz. Tuttavia, questo rallentamento può essere considerato effettivo solo a partire dagli anni 2000 circa, poiché per il periodo precedente i dati presentano un'instabilità non trascurabile per le età più elevate. In Figura 1 viene riportato il confronto tra i log-tassi stimati tramite P-splines unidimensionali e la crescita lineare prevista da Gompertz, per i singoli Paesi analizzati nell'anno 2019. Per una maggior interpretabilità grafica, sono state rappresentate le età comprese tra 80 e 110 anni. Si

osserva una curvatura dei log-tassi stimati a partire da 95-100 anni, che porta a un progressivo discostamento al di sotto della retta del modello di Gompertz, indicando un rallentamento della crescita esponenziale della mortalità per le età estreme. In nessun caso analizzato, tuttavia, i log-tassi sembrano stabilizzarsi ad un valore costante e quindi ad un plateau. Anche nel caso di un parametro di penalizzazione univoco, $\lambda=100$, i risultati, osservabili in Figura 2, sembrano mantenere la stessa tendenza, con un rallentamento dei log-tassi stimati rispetto alla crescita lineare di Gompertz, senza però avere una convergenza ad un valore finito. Inoltre, l'entità del discostamento è equivalente a quella precedente.

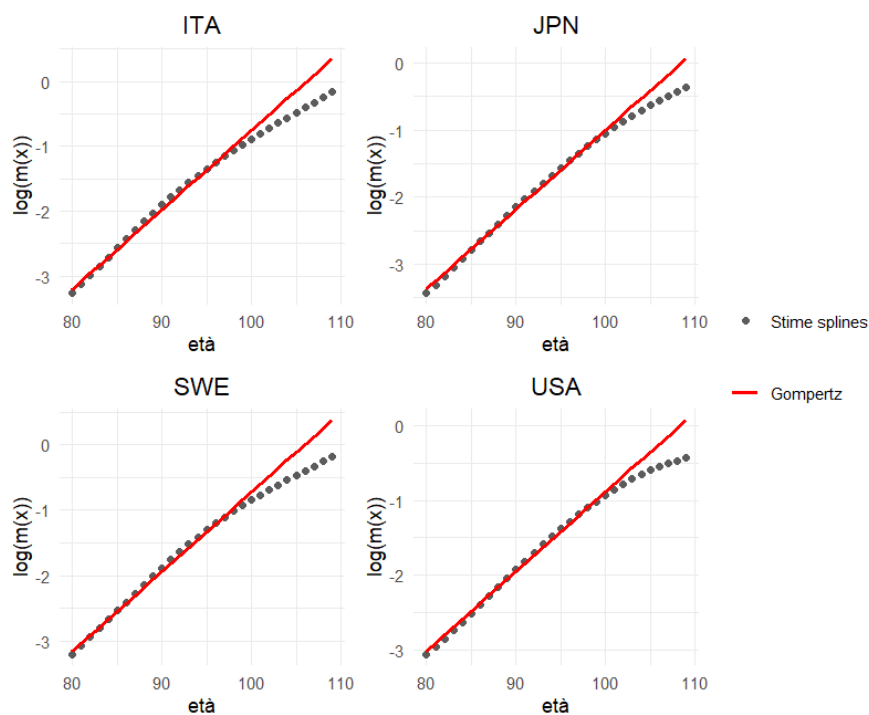


Figura 1: Confronto tra i log-tassi stimati tramite P-splines unidimensionali e la crescita lineare del modello di Gompertz, per Stato, nell'anno 2019.

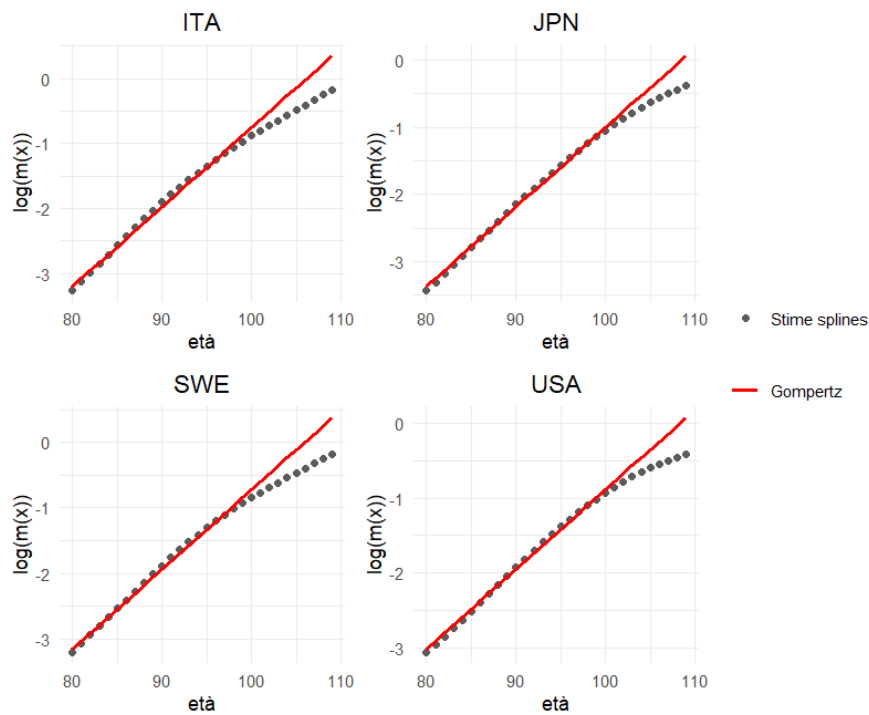


Figura 2: Confronto tra i log-tassi stimati tramite P-splines unidimensionali, con $\lambda=100$, e la crescita lineare del modello di Gompertz, per Stato, nell'anno 2019.

In Figura 3 viene riportato il confronto tra i log-tassi stimati tramite P-splines bidimensionali e la crescita lineare prevista da Gompertz, per tutti i Paesi analizzati nell'anno 2019. Anche in questo caso si osserva un rallentamento generale, con una curvatura dei log-tassi stimati a partire sempre dai 95-100 anni, che li porta ad essere sistematicamente sotto la retta, evidenziando un rallentamento della mortalità per le età estreme, ma non la presenza di un plateau, poiché non c'è una stabilizzazione ad un valore costante. Considerazioni analoghe possono essere tratte anche utilizzando parametri di penalizzazioni uguali per tutti gli Stati, $\lambda_{age}=10$ e $\lambda_{year}=3.162278$, osservando anche un discostamento dalla retta della stessa entità. In Figura 4 vengono riportati i confronti per i singoli Paesi nell'anno 2019.

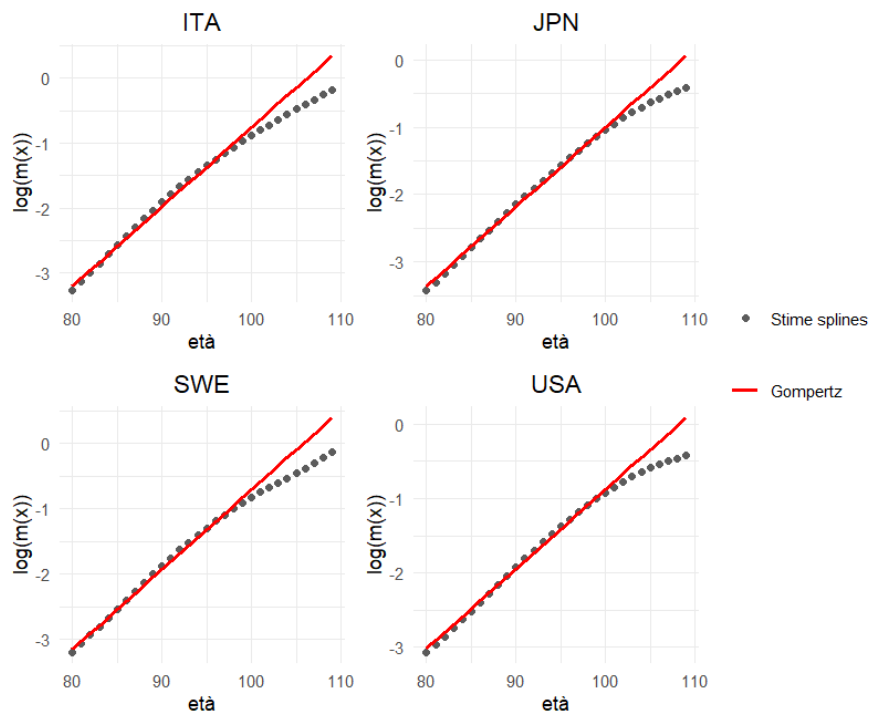


Figura 3: Confronto tra i log-tassi stimati tramite P-splines bidimensionali e la crescita lineare del modello di Gompertz, per Stato, nell'anno 2019.

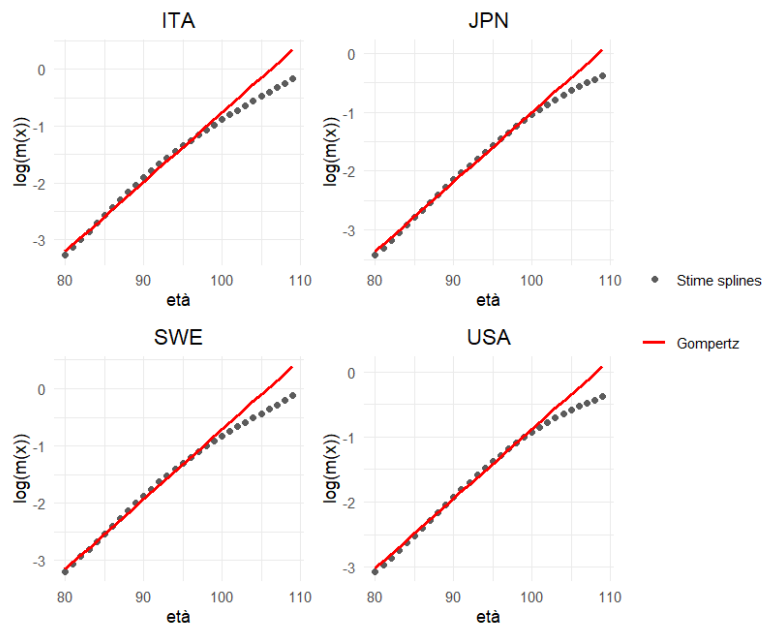


Figura 4: Confronto tra i log-tassi stimati tramite P-splines bidimensionali, con $\lambda_{age}=10$ e $\lambda_{year}=3.162278$, e la crescita lineare del modello di Gompertz, per Stato, nell'anno 2019.

Si può quindi concludere che, da un confronto con il modello di Gompertz, emerge un rallentamento dei tassi di mortalità per le età estreme, sia nel caso di stime ottenute utilizzando P-splines unidimensionali che bidimensionali, con parametri di penalizzazione variabili in base all'anno e allo Stato, oppure univoci. Non si ha però indicazione di un plateau di mortalità, in quanto non è stata osservata la stabilizzazione verso un valore finito in nessuno dei casi analizzati.

Analisi della derivata prima

Il secondo metodo utilizzato per lo studio della mortalità nella tarda età consiste nell'analisi dell'andamento della derivata prima rispetto all'età, della curva dei log-tassi stimati. Infatti, se i tassi seguissero la legge di Gompertz, la derivata dovrebbe assumere un valore costante, al contrario, una sua decrescita indicherebbe un rallentamento della mortalità, mentre una stabilizzazione attorno allo zero, suggerirebbe la presenza di un plateau. In Figura 5 è riportato l'andamento delle derivate dei log-tassi di mortalità stimati tramite P-splines unidimensionali, considerando parametri di penalizzazione diversi in base all'anno, per ciascuno Stato in analisi. Al fine di garantire una maggiore interpretabilità grafica, sono stati riportati unicamente i valori per le età superiori ai 40 anni. Per i primi anni considerati si osserva un valore costante della derivata fino agli 85 anni circa, indice di una crescita esponenziale della mortalità, seguito da un netto declino alle età avanzate, imputabile a un'eccessiva instabilità dei tassi reali. Le derivate che fanno riferimento al periodo più recente, invece, appaiono costanti fino ai 70 anni, per poi crescere nuovamente con un picco intorno ai 90 anni, e successivamente diminuire senza tuttavia attestarsi a zero, stando a indicare un possibile rallentamento della mortalità, ma non la presenza di un plateau. Si osserva dunque come, con l'avanzare degli anni, l'inizio della crescita esponenziale, quindi della derivata, si sia progressivamente spostato ver-

so destra sull'asse dell'età, sottolineando un aumento della longevità. Le medesime considerazioni possono essere tratte anche applicando un parametro di penalizzazione univoco, $\lambda=100$, come è osservabile in Figura 6. Inoltre, la scelta di un λ fissato non comporta variazioni nei profili delle derivate, se non un aumento nel range d'età 85-95 anni, che porta a stimare crescite esponenziali più ripide in quella fascia d'età. In Figura 7 è riportato l'andamento delle derivate dei log-tassi di mortalità stimati tramite P-splines bidimensionali, nei vari anni per tutti i Paesi in analisi, utilizzando parametri di penalizzazione diversi in base allo Stato, mentre in Figura 8 vengono considerati fissi $\lambda_{age}=10$ e $\lambda_{year}=3.162278$. Anche in questi casi, per i primi anni in analisi, si osserva un forte declino della derivata alle età avanzate, a fronte di un valore costante fino a circa gli 85 anni, mentre, per gli anni più recenti, la derivata tende ad assumere un valore costante fino ai 70 anni, per poi aumentare, sottolineando nuovamente uno spostamento della crescita esponenziale verso una maggiore età.

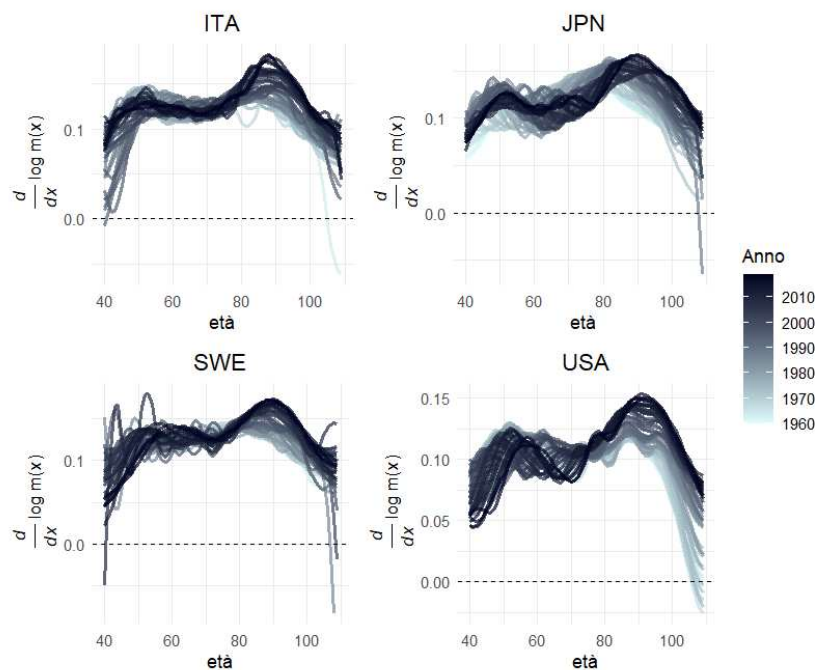


Figura 5: Derivate prime dei log-tassi stimati tramite P-splines unidimensionali, nei vari anni, per Stato.

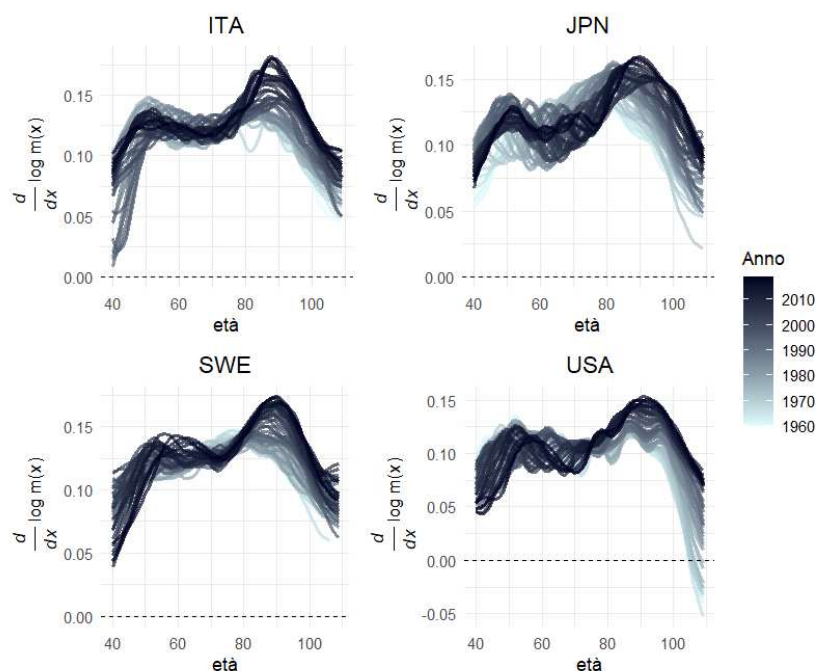


Figura 6: Derivate prime dei log-tassi stimati tramite P-splines unidimensionali, con $\lambda=100$, nei vari anni, per Stato.

Viene poi raggiunto il picco intorno ai 90 anni, seguito da una decrescita e quindi da un rallentamento della mortalità, senza però una convergenza verso zero, dunque senza l'evidenza di un plateau. Il confronto tra le derivate dei diversi Stati nel 2019, riportato in Figura 9, sottolinea ancora una volta la tendenza dei tassi di mortalità a presentare l'inizio della crescita esponenziale intorno ai 70 anni. Per Italia e Giappone si osserva un aumento sostenuto della derivata con l'età fino a circa i 90 anni, indicando un regolare incremento della mortalità nelle età anziane a fronte di bassi livelli di mortalità infantile e giovanile. Successivamente la derivata tende a diminuire, segnalando un rallentamento nella crescita dei tassi. La Svezia presenta un andamento simile, ma con valori più elevati tra i 55 e i 70 anni, che suggeriscono una mortalità relativamente maggiore in quella fascia di età, per poi riallinearsi con il profilo di Italia e Giappone. Gli USA, invece, presentano valori della derivata sistematicamente inferiori rispetto agli altri Stati, evidenziando una crescita dei tassi con l'età meno ripida. Ciò riflette livelli di mortalità più elevati

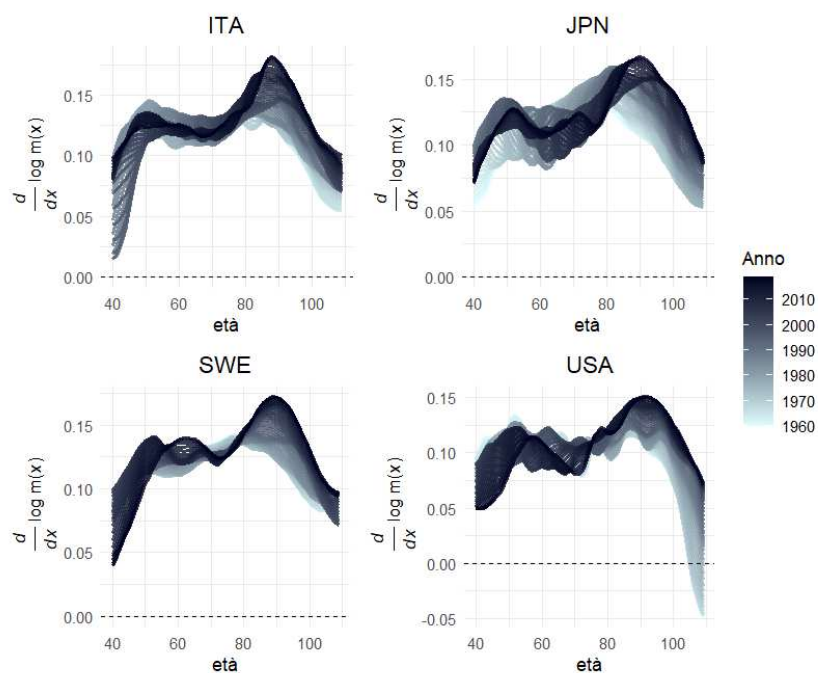


Figura 7: Derivate prime dei log-tassi stimati tramite P-splines bidimensionali, nei vari anni, per Stato.

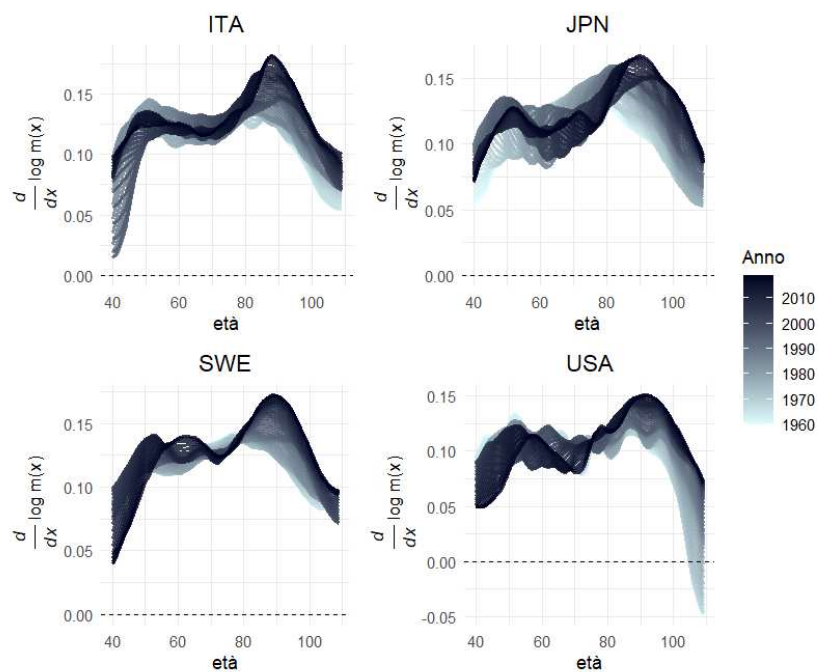


Figura 8: Derivate prime dei log-tassi stimati tramite P-splines bidimensionali, con $\lambda_{age}=10$ e $\lambda_{year}=3.162278$, nei vari anni, per Stato.

nelle età giovani e adulte rispetto agli altri Paesi, tuttavia anche in questo caso si osserva un rallentamento a partire dai 100 anni circa.

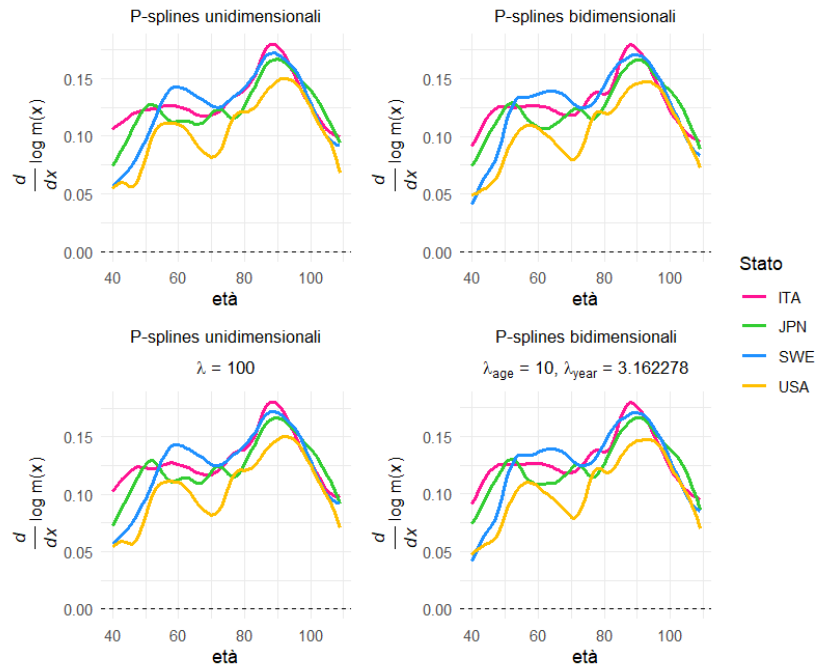


Figura 9: Andamento delle derivate prime dei log-tassi stimati per ciascuno Stato nel 2019. In alto a sinistra la derivata fa riferimento ai tassi stimati con P-splines unidimensionali, in alto a destra con P-splines bidimensionali, in basso a sinistra con P-splines unidimensionali con $\lambda=100$, in basso a destra con P-splines bidimensionali con $\lambda_{age}=10$ e $\lambda_{year}=3.162278$.

Si può dunque concludere che, anche dall'analisi della derivata dei log-tassi di mortalità rispetto all'età, si ha indicazione di un rallentamento della mortalità alle età estreme, usando sia una modellazione con P-splines unidimensionali che bidimensionali, con parametri di penalizzazione univoci o meno. Tuttavia, non emerge la presenza di un plateau, poiché non si osserva una convergenza a zero per nessuno dei casi in analisi.

Conclusioni

L'obiettivo principale di questo elaborato è stato analizzare l'andamento della mortalità alle età estreme, adottando un approccio non parametrico in grado di garantire una maggiore flessibilità nella modellazione dei tassi di mortalità, senza imporre una forma funzionale rigida tipica degli approcci parametrici. La modellazione della curva di mortalità tramite P-splines sia unidimensionali che bidimensionali, ha mostrato un buon adattamento per il periodo più recente, mentre per gli anni iniziali le stime non sembrano cogliere adeguatamente l'andamento oscillante dei tassi reali alle età più avanzate, verosimilmente dovuto alla scarsa qualità dei dati disponibili. Dallo studio dell'andamento della curva di mortalità è emerso un suo rallentamento alle età più estreme, ma non una stabilizzazione ad un plateau. Il confronto con la legge di Gompertz, condotto per ciascuno degli Stati in analisi, ha infatti mostrato una decelerazione della mortalità a partire dai 95 anni, rispetto alla crescita esponenziale prevista, senza tuttavia segnalare la stabilizzazione verso un valore finito. Lo studio della derivata prima dei log-tassi di mortalità calcolata rispetto all'età, ha confermato queste evidenze: alle età estreme si osserva una decrescita della derivata, ma non un avvicinamento a zero che indicherebbe la presenza di un plateau. I risultati ottenuti suggeriscono pertanto la presenza di un rallentamento della mortalità alle età più elevate, per ciascuno degli Stati in analisi, ma non si hanno indicazioni di una stabilizzazione verso un plateau. Ciò non consente, allo stato attuale, di confermare pienamente la teoria della late-life mortality deceleration. Tuttavia, un possibile aumento della longevità e la conseguente disponibilità futura di un maggior numero di dati relativi a centenari e supercentenari, potranno consentire analisi più approfondite e fornire indicazioni più solide in merito alla verifica di tale teoria e in particolare della possibile evidenza di un plateau di mortalità umana.

Bibliografia

- Barbi, E., Lagona, F., Marsili, M., Vaupel, J. W., & Wachter, K. W. (2018). The plateau of human mortality: Demography of longevity pioneers. *Science*, *360*(6396), 1459–1461.
- Camarda, C. G. (2012). MortalitySmooth: An R Package for Smoothing Poisson Counts with P-Splines. *Journal of Statistical Software*, *50*(1), 1–24.
- Currie, I. D., Durban, M., & Eilers, P. H. C. (2004). Smoothing and forecasting mortality rates. *Statistical Modelling*, *4*(4), 279–298.
- Curtsinger, J. W., Fukui, H. H., Townsend, D. R., & Vaupel, J. W. (1992). Demography of genotypes: Failure of the limited life-span paradigm in *Drosophila melanogaster*. *Science*, *258*(5081), 461–463.
- Dang, L. H. K., Camarda, C. G., Meslé, F., Ouellette, N., Robine, J., & Vallin, J. (2023). The question of the human mortality plateau: Contrasting insights by longevity pioneers. *Demographic Research*, *48*, 321–338.
- De Boor, C. (1978). *A practical guide to Splines*. Springer.
- Eilers, P. H. C., & Marx, B. D. (1996). Flexible smoothing with B-splines and penalties. *Statistical Science*, *11*(2), 89–121.
- Einmahl, J. J., Einmahl, J. H. J., & De Haan, L. (2019). Limits to human life span through extreme value theory. *Journal of the American Statistical Association*, *114*(527), 1075–1080.
- Gampe, J. (2020). Mortality of Supercentenarians: Estimates from the Updated IDL. In H. Maier, J. Gampe, B. Jeune, J.-M. Robine & J. W. Vaupel (Cur.), *Supercentenarians* (pp. 29–35). Springer.
- Gavrilov, L. A., & Gavrilova, N. S. (2011). Mortality measurement at advanced ages. *North American Actuarial Journal*, *15*(3), 432–447.

- Gompertz, B. (1825). On the Nature of the Function Expressive of the Law of Human Mortality and on a New Mode of Determining the Value of Life Contingencies. *Philosophical Transactions of the Royal Society London*, 115, 513–585.
- Horiuchi, S., & Wilmoth, J. R. (1998). Deceleration in the age pattern of mortality at older ages. *Demography*, 35(4), 391–412.
- Nelder, J. A., & Wedderburn, R. W. M. (1972). Generalized Linear Models. *Journal of the Royal Statistical Society*, 135, 370–384.
- Newman, S. J. (2018). Errors as a primary cause of late-life mortality deceleration and plateaus. *PLoS Biology*, 16(12), e2006776.
- Oeppen, J., & Vaupel, J. W. (2002). Broken limits to life expectancy. *Science*, 296(5570), 1029–1031.
- Rau, R., Soroko, E., Jasilionis, D., & Vaupel, J. W. (2008). Continued reductions in mortality at advanced ages. *Population and Development Review*, 34(4), 747–768.
- Vaupel, J. W., Manton, K. G., & Stallard, E. (1979). The impact of heterogeneity in individual frailty on the dynamics of mortality. *Demography*, 16(3), 439–454.

Appendice A

La bontà di adattamento delle curve splines ai dati, viene inizialmente valutata mediante un confronto grafico tra i tassi di mortalità osservati e quelli stimati. Per le P-splines unidimensionali, in Figura 10 viene riportata tale comparazione per ciascuno Stato analizzato nel 1960: l'adattamento è buono fino a circa i 95 anni, successivamente la curva stimata non sembra cogliere adeguatamente le oscillazioni dei tassi reali, che risultano molto instabili, probabilmente a causa della scarsità di dati disponibili per le età estreme. In Figura 11 è invece riportato il confronto relativo all'anno 2019, per ciascun Paese. Si osserva un netto miglioramento nell'adattamento, anche per età molto avanzate, ciò è attribuibile a una minor variabilità dei tassi osservati, dovuta a una maggior longevità e quindi a un numero superiore di centenari e supercentenari.

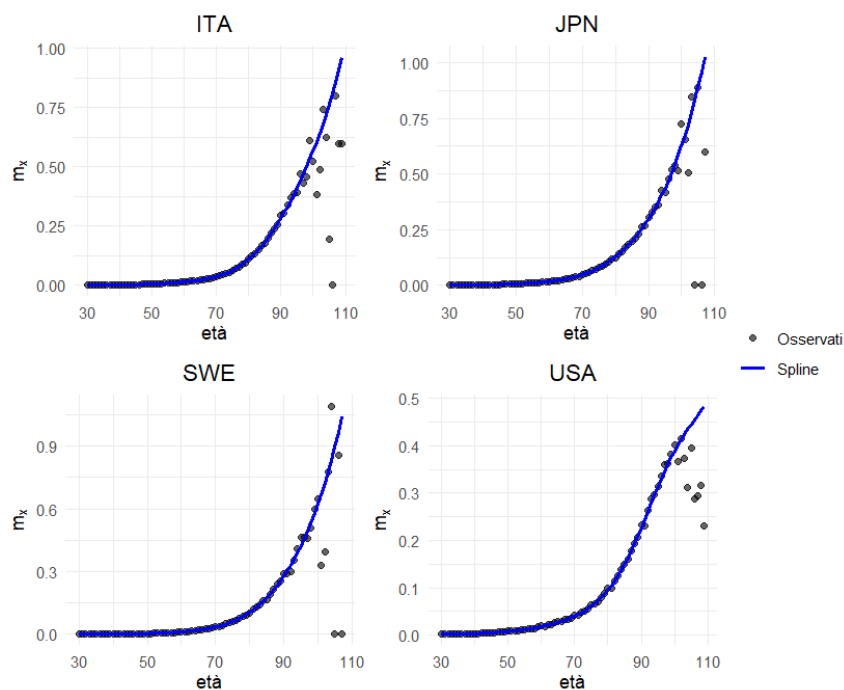


Figura 10: Confronto tra i tassi di mortalità reali e quelli stimati tramite P-splines unidimensionali, per Stato, nell'anno 1960.

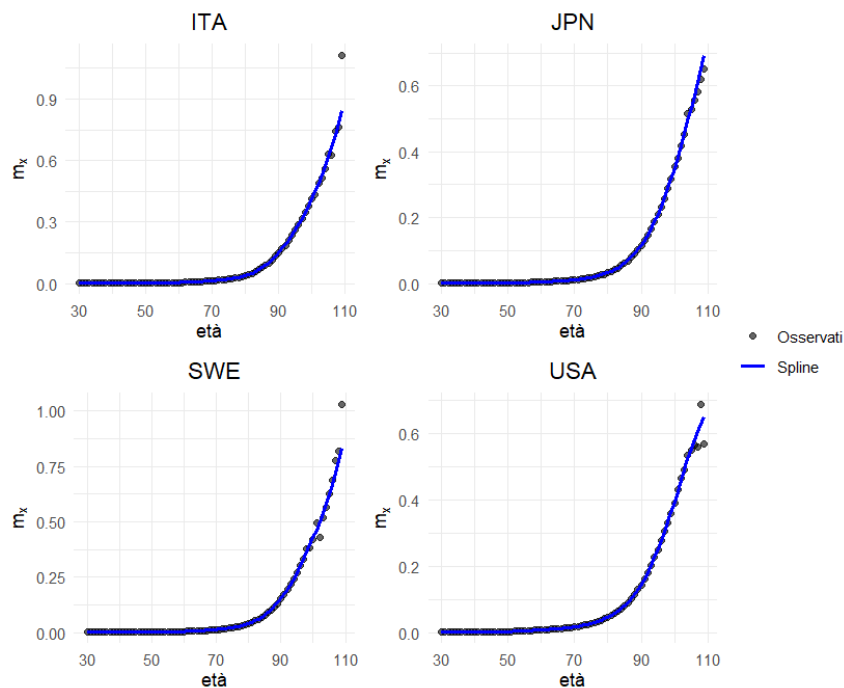


Figura 11: Confronto tra i tassi di mortalità reali e quelli stimati tramite P-splines unidimensionali, per Stato, nell'anno 2019.

L'utilizzo di un parametro di penalizzazione univoco, $\lambda=100$, non sembra comportare un peggioramento nell'adattamento, specialmente per i dati più recenti, come è osservabile in Figura 13. Tuttavia, per i primi anni analizzati, emerge una maggior flessibilità della curva stimata nelle età più avanzate, come riportato in Figura 12, che porta a cogliere fluttuazioni casuali dovute all'incertezza dei dati come vero segnale e quindi rende il risultato meno robusto. Anche l'applicazione di un parametro di penalizzazione bidimensionale sembra garantire un buon adattamento ai dati. In particolare, in Figura 14 si nota una maggior flessibilità della curva per le età estreme, che riesce quindi, rispetto alle P-splines unidimensionali, a cogliere meglio l'andamento dei tassi reali del primo periodo, talvolta però adattandosi eccessivamente alle fluttuazioni casuali dovute all'incertezza dei dati. In Figura 15, invece, è osservabile il miglioramento nell'adattamento per gli anni più recenti, dovuto anche alla maggior stabilità dei tassi.

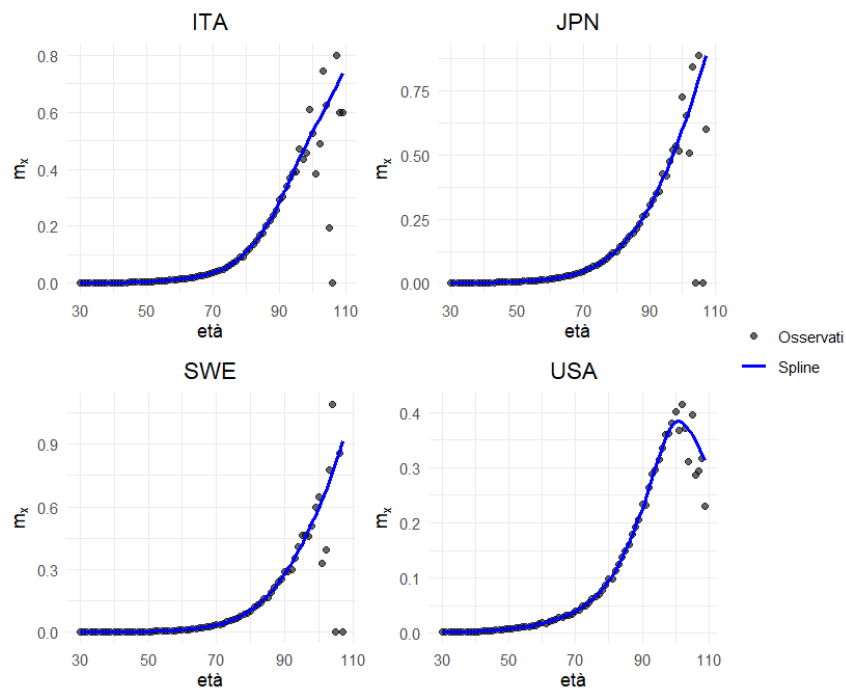


Figura 12: Confronto tra i tassi di mortalità reali e quelli stimati tramite P-splines unidimensionali, con $\lambda=100$, per Stato, nell'anno 1960.

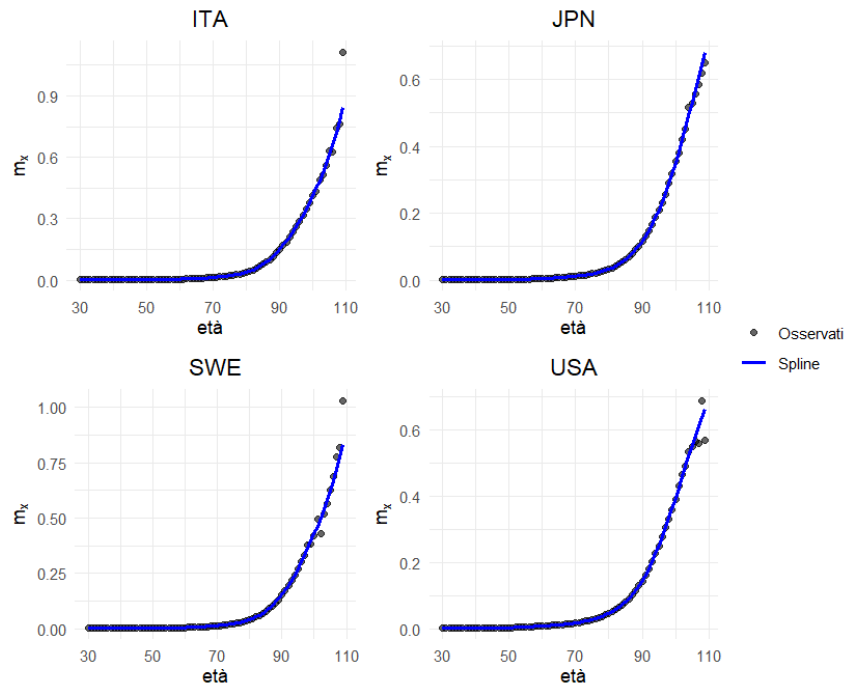


Figura 13: Confronto tra i tassi di mortalità reali e quelli stimati tramite P-splines unidimensionali, con $\lambda=100$, per Stato, nell'anno 2019.

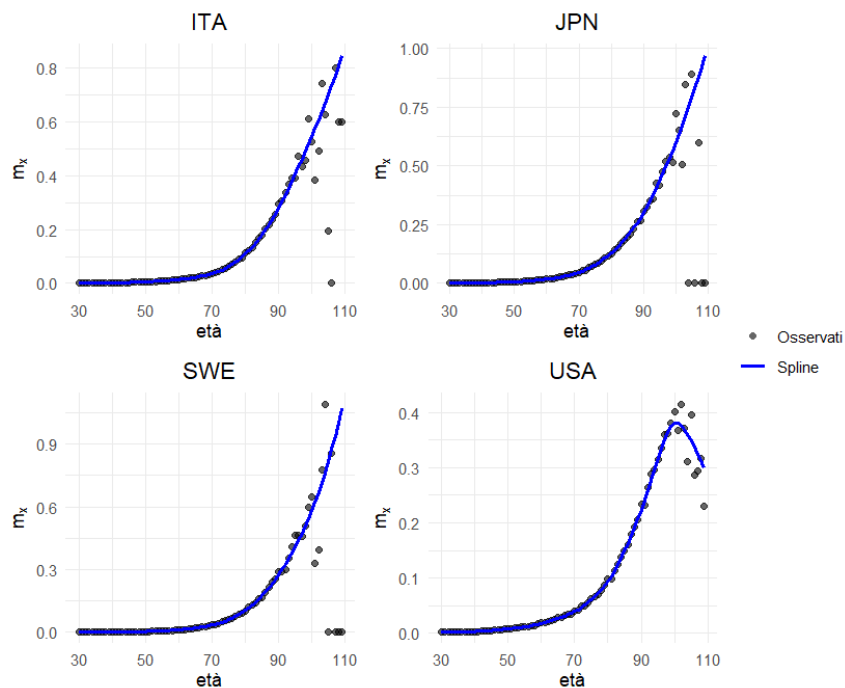


Figura 14: Confronto tra i tassi di mortalità reali e quelli stimati tramite P-splines bidimensionali, per Stato, nell'anno 1960.

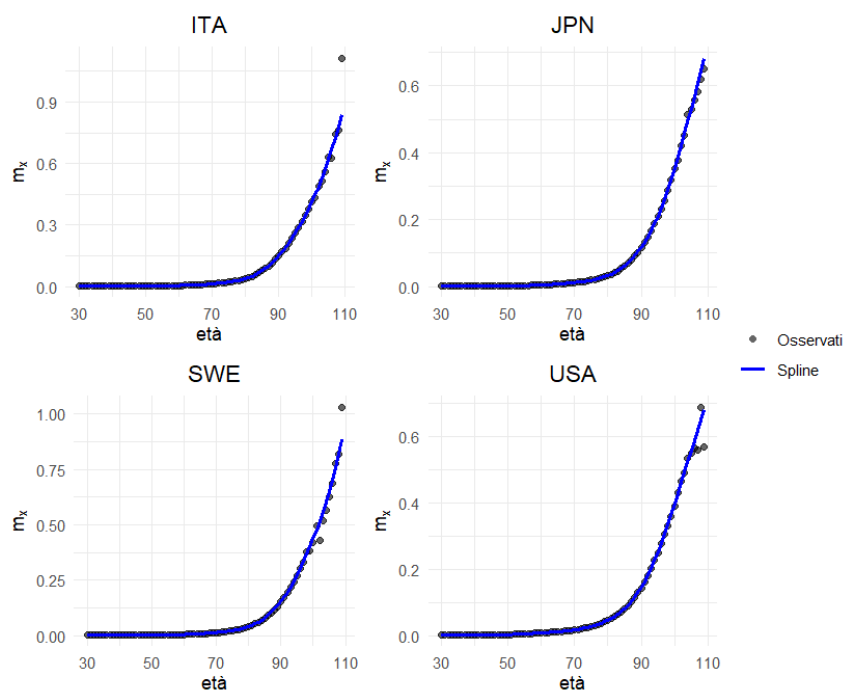


Figura 15: Confronto tra i tassi di mortalità reali e quelli stimati tramite P-splines bidimensionali, per Stato, nell'anno 2019.

L'utilizzo di parametri di penalizzazione uguali per tutti gli Stati, $\lambda_{age}=10$

e $\lambda_{year}=3.162278$, sembra mantenere un buon adattamento ai dati, come è osservabile in Figura 16 e Figura 17. Anche in questo caso si osserva un'eccessiva flessibilità delle curve per i primi anni, che porta a cogliere fluttuazioni casuali, tipiche dei tassi della tarda età del primo periodo analizzato, come vero segnale. Complessivamente quindi, si osserva un buon adattamento fino ai 95 anni, mentre per le età avanzate l'adeguatezza delle stime migliora con il tempo, come emerge anche dalle tabelle in Appendice B, che mostrano una decrescita dell'errore quadratico medio, RMSE, negli anni. Questa tendenza è verosimilmente dovuta a un aumento della qualità e della quantità dei dati per la tarda età, che ha portato a minori oscillazioni dei tassi. Inoltre, si osserva la tendenza generale negli anni, a uno spostamento dell'inizio della crescita esponenziale verso destra sull'asse delle età, indice di un aumento della longevità.

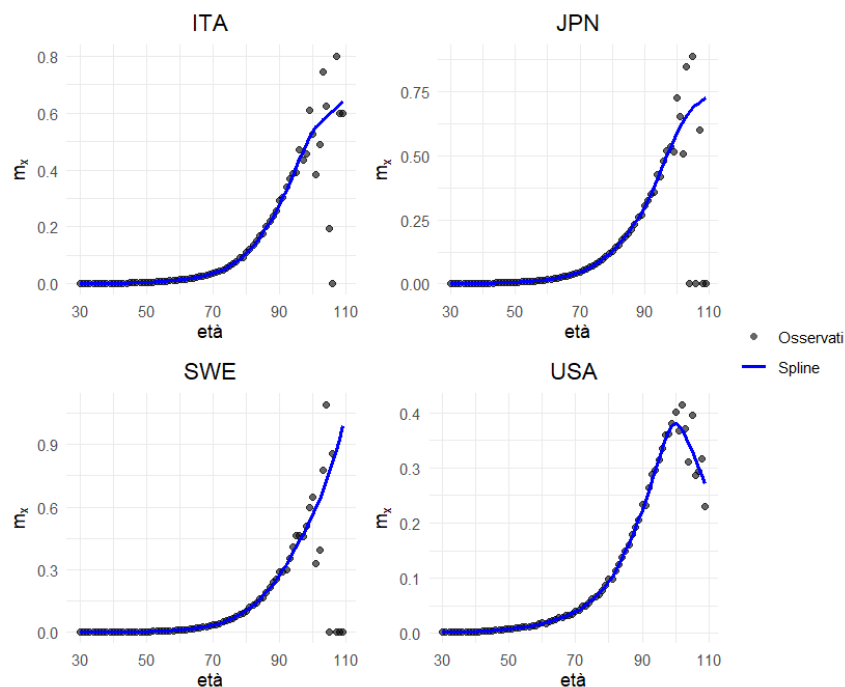


Figura 16: Confronto tra i tassi di mortalità reali e quelli stimati tramite P-splines biidimensionali, con $\lambda_{age}=10$ e $\lambda_{year}=3.162278$, per Stato, nell'anno 1960.

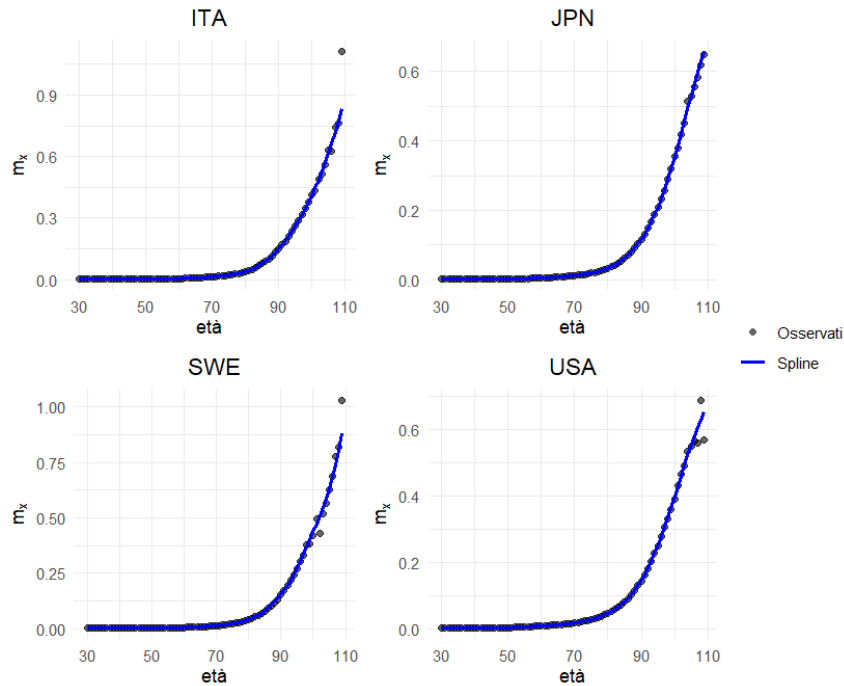


Figura 17: Confronto tra i tassi di mortalità reali e quelli stimati tramite P-splines bidimensionali, con $\lambda_{age}=10$ e $\lambda_{year}=3.162278$, per Stato, nell'anno 2019.

La bontà di adattamento ai dati reali della curva di mortalità stimata, è stata valutata anche mediante l'analisi dei residui di devianza standardizzati. In Figura 18 vengono riportati i residui per tutti gli anni per ciascuno degli Stati considerati, ottenuti utilizzando P-splines unidimensionali, mentre in Figura 19 i rispettivi ricavati utilizzando un parametro di penalizzazione univoco, $\lambda=100$. Si osserva un buon adattamento complessivo per ogni Paese, con i residui distribuiti in modo casuale e compresi tra -2 e 2, con eventuali valori estremi che fanno riferimento ai primi anni analizzati. Le medesime considerazioni sono ricavabili anche dall'analisi dei residui di devianza standardizzati ottenuti dall'utilizzo di P-splines bidimensionali, come si può notare in Figura 20 e Figura 21, in cui emerge anche un leggero miglioramento, in termini di valori fuori dal range, rispetto al caso unidimensionale.

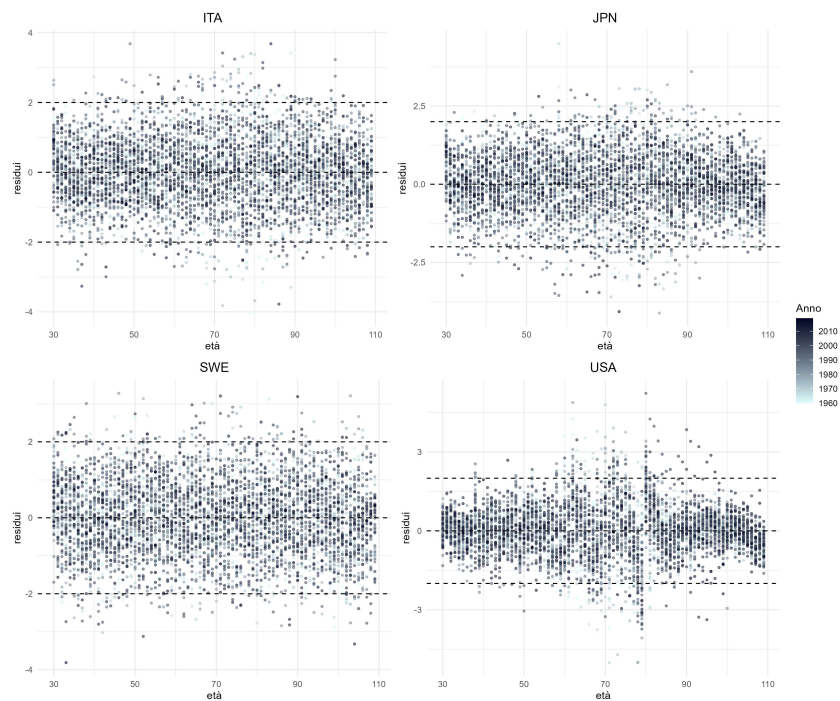


Figura 18: Residui di devianza standardizzati di tutti gli anni, per Stato, con stime ottenute tramite P-splines unidimensionali.

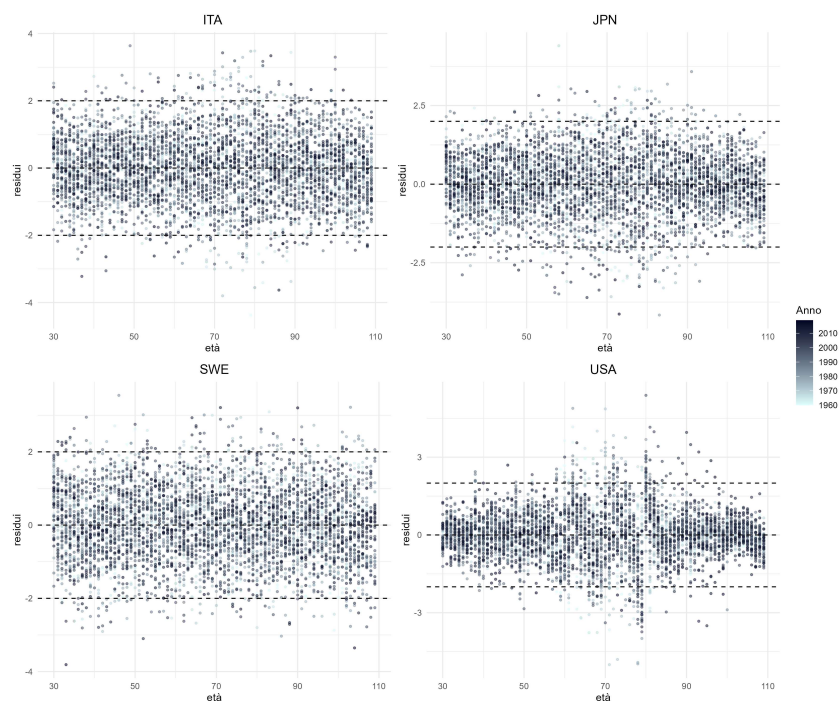


Figura 19: Residui di devianza standardizzati di tutti gli anni, per Stato, con stime ottenute tramite P-splines unidimensionali, con $\lambda=100$.

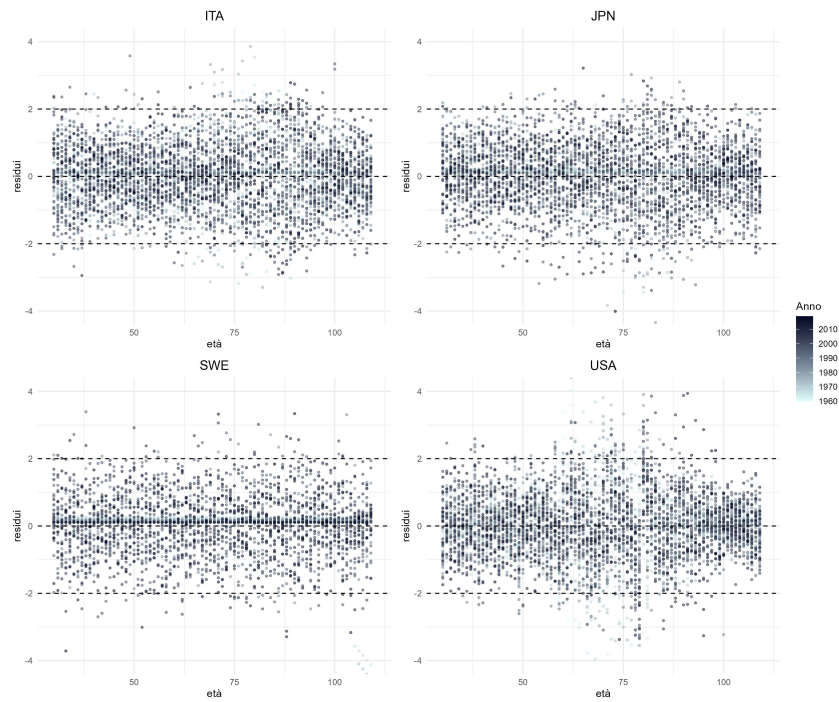


Figura 20: Residui di devianza standardizzati di tutti gli anni, per Stato, con stime ottenute tramite P-splines bidimensionali.

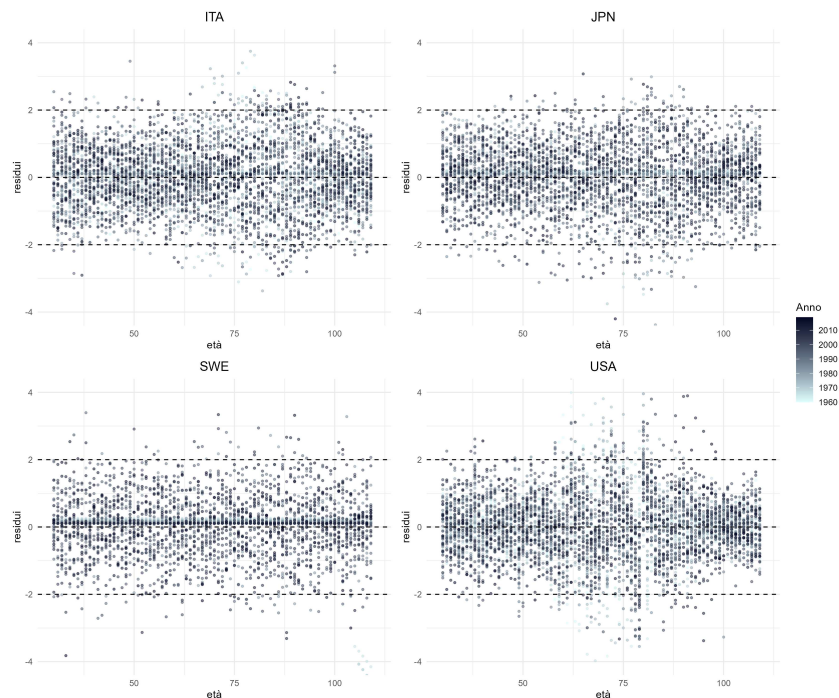


Figura 21: Residui di devianza standardizzati di tutti gli anni, per Stato, con stime ottenute tramite P-splines bidimensionali, con $\lambda_{age}=10$ e $\lambda_{year}=3.162278$.

Si può quindi concludere che sia le P-splines unidimensionali che quelle bidimensionali forniscono una buona stima delle curve di mortalità, anche con l'utilizzo di parametri di penalizzazione univoci. Dunque, i risultati ottenuti sono considerabili attendibili. Inoltre, l'utilizzo di P-splines bidimensionali sembra portare a un miglior adattamento rispetto alle P-splines unidimensionali, fornendo un leggero miglioramento sia dal punto di vista grafico, che in termini di RMSE e residui di devianza standardizzati.

Appendice B

Tabella 6: Modellazione tramite P-splines unidimensionali per l'Italia: anno, λ , gradi di libertà (*gdl*), AIC e RMSE.

Anno	λ	<i>gdl</i>	AIC	RMSE
1960	316.228	9.884	89.930	0.1278
1961	316.228	10.531	90.538	0.1726
1962	316.228	10.299	90.298	0.1274
1963	316.228	9.703	88.700	0.1227
1964	3.162	15.757	95.759	0.0948
1965	100.000	11.558	89.585	0.1479
1966	100.000	11.804	90.821	0.0987
1967	316.228	10.182	90.180	0.1582
1968	316.228	9.795	89.804	0.0620
1969	31.623	13.741	93.744	0.1749
1970	31.623	13.870	93.885	0.2521
1971	316.228	9.828	88.850	0.0725
1972	31.623	13.697	93.713	0.1216
1973	100.000	12.199	92.212	0.2634
1974	31.623	12.979	91.982	0.1852
1975	100.000	12.355	92.364	0.1575
1976	31.623	13.883	93.884	0.3314
1977	100.000	11.905	91.897	0.1253
1978	100.000	12.348	92.392	0.1146
1979	100.000	12.176	92.179	0.2434
1980	316.228	11.124	91.137	0.1877
1981	100.000	12.855	92.859	0.2574
1982	316.228	11.220	91.224	0.1139
1983	31.623	14.475	94.496	0.0619
1984	100.000	12.956	92.982	0.1252

Tabella 6 (continua)

Anno	λ	<i>gdl</i>	AIC	RMSE
1985	316.228	11.001	91.060	0.0475
1986	31.623	13.967	94.016	0.4462
1987	0.003	17.750	96.752	0.0458
1988	31.623	14.514	94.515	0.0583
1989	100.000	12.575	92.592	0.0478
1990	100.000	13.194	93.197	0.0840
1991	31.623	14.449	94.484	0.0749
1992	100.000	12.974	92.996	0.0318
1993	100.000	13.099	93.106	0.0544
1994	0.316	17.687	97.728	0.0355
1995	100.000	12.037	92.082	0.0883
1996	10.000	15.408	95.417	0.0406
1997	10.000	15.506	95.510	0.0249
1998	100.000	12.789	92.791	0.0578
1999	100.000	12.712	92.723	0.0405
2000	100.000	13.098	93.108	0.0191
2001	316.228	11.210	91.224	0.0653
2002	100.000	12.991	93.019	0.0351
2003	316.228	11.844	91.873	0.0417
2004	31.623	14.471	94.491	0.0481
2005	31.623	14.975	94.974	0.0214
2006	100.000	13.378	93.419	0.0392
2007	100.000	13.058	93.070	0.0548
2008	1.000	17.071	97.087	0.0268
2009	100.000	13.356	93.368	0.0376
2010	100.000	13.071	93.103	0.0458
2011	100.000	13.200	93.202	0.0411
2012	316.228	11.622	91.649	0.0139
2013	100.000	13.309	93.313	0.0261

Tabella 6 (continua)

Anno	λ	<i>gdl</i>	AIC	RMSE
2014	0.316	17.755	97.754	0.0158
2015	0.032	18.652	98.674	0.0123
2016	316.228	10.879	90.924	0.0304
2017	10.000	16.146	96.129	0.0313
2018	1.000	17.307	97.306	0.0088
2019	316.228	11.817	91.880	0.0307

Tabella 7: Modellazione tramite P-splines unidimensionali per il Giappone: anno, λ , gradi di libertà (*gdl*), AIC e RMSE.

Anno	λ	<i>gdl</i>	AIC	RMSE
1960	316.228	7.947	85.9834	0.1673
1961	100.000	9.782	88.7406	0.1805
1962	316.228	7.975	86.0032	0.2149
1963	31.623	11.080	89.1210	0.6641
1964	316.228	8.097	85.1442	0.6142
1965	316.228	8.415	84.3899	0.5948
1966	1000.000	6.585	82.5874	0.1149
1967	316.228	8.226	85.2293	0.1049
1968	316.228	8.030	86.0699	0.1664
1969	100.000	10.019	87.9790	0.3623
1970	316.228	8.574	84.5693	0.1267
1971	1000.000	6.813	83.8446	0.1141
1972	316.228	8.309	86.3529	0.1241
1973	100.000	9.985	88.9774	0.3501
1974	316.228	8.243	87.2597	0.2047
1975	1000.000	6.748	84.7594	0.2530
1976	1000.000	7.403	85.3853	0.1479
1977	1000.000	6.682	85.7178	0.5898
1978	316.228	8.202	86.2059	0.1028

Tabella 7 (continua)

Anno	λ	<i>gdl</i>	AIC	RMSE
1979	316.228	8.836	87.8210	0.2295
1980	1000.000	7.023	87.0243	0.2160
1981	31.623	12.459	92.4566	0.1908
1982	1000.000	7.293	87.2865	0.2647
1983	316.228	8.231	88.2797	0.1473
1984	100.000	10.383	90.3863	0.2730
1985	100.000	10.004	89.0052	0.5825
1986	0.001	17.805	96.8241	0.1091
1987	100.000	10.241	88.2437	0.1682
1988	316.228	8.474	87.5021	0.1451
1989	316.228	8.513	87.5094	0.1079
1990	316.228	8.051	87.0933	0.1938
1991	100.000	11.147	91.1271	0.1244
1992	100.000	10.623	90.6596	0.1002
1993	316.228	9.098	89.0838	0.5664
1994	316.228	9.119	89.0905	0.0794
1995	0.010	18.349	98.3235	0.1304
1996	100.000	10.924	90.9173	0.1180
1997	100.000	10.299	90.3438	0.0808
1998	0.000316	18.688	98.6743	0.0351
1999	1000.000	7.551	87.5329	0.1060
2000	100.000	10.741	90.7084	0.1230
2001	316.228	8.551	88.5593	0.0813
2002	10.000	14.048	94.1033	0.1656
2003	0.003	18.699	98.6935	0.0820
2004	100.000	10.835	90.8271	0.1357
2005	31.623	12.371	92.4052	0.2726
2006	316.228	8.283	87.3510	0.1143
2007	31.623	12.751	92.7499	0.1065

Tabella 7 (continua)

Anno	λ	<i>gdl</i>	AIC	RMSE
2008	31.623	12.611	92.6070	0.1091
2009	100.000	10.509	90.5773	0.0675
2010	0.010	18.315	98.3157	0.1659
2011	100.000	10.733	90.7224	0.0659
2012	100.000	10.887	90.8837	0.0312
2013	31.623	12.407	92.4433	0.0612
2014	31.623	12.665	92.6619	0.0330
2015	100.000	10.571	90.5728	0.0250
2016	100.000	10.254	90.2652	0.0391
2017	100.000	10.108	90.1092	0.0398
2018	100.000	11.109	91.1052	0.0278
2019	100.000	10.676	90.6613	0.0255

Tabella 8: Modellazione tramite P-splines unidimensionali per la Svezia: anno, λ , gradi di libertà (*gdl*), AIC e RMSE.

Anno	λ	<i>gdl</i>	AIC	RMSE
1960	316.228	7.947	85.983	0.1673
1961	100.000	9.782	88.741	0.1805
1962	316.228	7.975	86.003	0.2149
1963	31.623	11.080	89.121	0.6641
1964	316.228	8.097	85.144	0.6142
1965	316.228	8.415	84.390	0.5948
1966	1000.000	6.585	82.587	0.1149
1967	316.228	8.226	85.229	0.1049
1968	316.228	8.030	86.070	0.1664
1969	100.000	10.019	87.979	0.3623
1970	316.228	8.574	84.569	0.1267
1971	1000.000	6.813	83.845	0.1141
1972	316.228	8.309	86.353	0.1241

Tabella 8 (continua)

Anno	λ	<i>gdl</i>	AIC	RMSE
1973	100.000	9.985	88.977	0.3501
1974	316.228	8.243	87.260	0.2047
1975	1000.000	6.748	84.759	0.2530
1976	1000.000	7.403	85.385	0.1479
1977	1000.000	6.682	85.718	0.5898
1978	316.228	8.202	86.206	0.1028
1979	316.228	8.836	87.821	0.2295
1980	1000.000	7.023	87.024	0.2160
1981	31.623	12.459	92.457	0.1908
1982	1000.000	7.293	87.286	0.2647
1983	316.228	8.231	88.280	0.1473
1984	100.000	10.383	90.386	0.2730
1985	100.000	10.004	89.005	0.5825
1986	0.001	17.805	96.824	0.1091
1987	100.000	10.241	88.244	0.1682
1988	316.228	8.474	87.502	0.1451
1989	316.228	8.513	87.509	0.1079
1990	316.228	8.051	87.093	0.1938
1991	100.000	11.147	91.127	0.1244
1992	100.000	10.623	90.660	0.1002
1993	316.228	9.098	89.084	0.5664
1994	316.228	9.119	89.091	0.0794
1995	0.010	18.349	98.324	0.1304
1996	100.000	10.924	90.917	0.1180
1997	100.000	10.299	90.344	0.0808
1998	0.000	18.688	98.674	0.0351
1999	1000.000	7.551	87.533	0.1060
2000	100.000	10.741	90.708	0.1230
2001	316.228	8.551	88.559	0.0813

Tabella 8 (continua)

Anno	λ	<i>gdl</i>	AIC	RMSE
2002	10.000	14.048	94.103	0.1656
2003	0.003	18.699	98.694	0.0820
2004	100.000	10.835	90.827	0.1357
2005	31.623	12.371	92.405	0.2726
2006	316.228	8.283	87.351	0.1143
2007	31.623	12.751	92.750	0.1065
2008	31.623	12.611	92.607	0.1091
2009	100.000	10.509	90.577	0.0675
2010	0.010	18.315	98.316	0.1659
2011	100.000	10.733	90.722	0.0659
2012	100.000	10.887	90.884	0.0312
2013	31.623	12.407	92.443	0.0612
2014	31.623	12.665	92.662	0.0330
2015	100.000	10.571	90.573	0.0250
2016	100.000	10.254	90.265	0.0391
2017	100.000	10.108	90.109	0.0398
2018	100.000	11.109	91.105	0.0278
2019	100.000	10.676	90.661	0.0255

Tabella 9: Modellazione tramite P-splines unidimensionali per gli USA: anno, λ , gradi di libertà (*gdl*), AIC e RMSE.

Anno	λ	<i>gdl</i>	AIC	RMSE
1960	100.000	10.586	90.626	0.0458
1961	31.623	13.068	93.075	0.0312
1962	31.623	12.944	92.882	0.0291
1963	31.623	12.587	92.576	0.0265
1964	31.623	13.273	93.250	0.0238
1965	31.623	12.844	92.834	0.0269
1966	31.623	13.285	93.258	0.0191

Tabella 9 (continua)

Anno	λ	<i>gdl</i>	AIC	RMSE
1967	31.623	13.719	93.687	0.0268
1968	100.000	11.150	91.148	0.0370
1969	100.000	11.893	91.898	0.0278
1970	31.623	13.808	93.802	0.0161
1971	31.623	14.343	94.339	0.0082
1972	31.623	14.193	94.176	0.0252
1973	31.623	13.734	93.723	0.0171
1974	31.623	14.809	94.816	0.0127
1975	31.623	13.687	93.681	0.0217
1976	31.623	13.662	93.691	0.0166
1977	31.623	13.396	93.362	0.0139
1978	31.623	13.135	93.110	0.0207
1979	31.623	14.392	94.386	0.0132
1980	100.000	12.230	92.169	0.0219
1981	100.000	12.548	92.527	0.0204
1982	31.623	14.465	94.454	0.0202
1983	100.000	12.480	92.490	0.0201
1984	100.000	13.353	93.361	0.0214
1985	100.000	13.012	92.981	0.0217
1986	100.000	12.535	92.515	0.0204
1987	100.000	12.614	92.598	0.0295
1988	100.000	12.610	92.587	0.0234
1989	100.000	13.228	93.207	0.0286
1990	316.228	10.956	90.909	0.0287
1991	100.000	13.357	93.355	0.0147
1992	100.000	13.380	93.412	0.0207
1993	31.623	14.889	94.900	0.0149
1994	100.000	13.758	93.810	0.0163
1995	100.000	13.687	93.679	0.0127

Tabella 9 (continua)

Anno	λ	<i>gdl</i>	AIC	RMSE
1996	100.000	13.226	93.224	0.0122
1997	316.228	11.655	91.598	0.0256
1998	316.228	11.278	91.312	0.0198
1999	100.000	13.799	93.797	0.0065
2000	316.228	12.409	92.439	0.0279
2001	1.000	17.641	97.668	0.0066
2002	31.623	15.502	95.461	0.0128
2003	31.623	15.668	95.671	0.0174
2004	100.000	14.300	94.296	0.0101
2005	100.000	13.787	93.847	0.0100
2006	100.000	13.976	93.970	0.0096
2007	100.000	14.183	94.241	0.0126
2008	100.000	14.027	94.036	0.0215
2009	100.000	14.562	94.579	0.0201
2010	100.000	14.224	94.223	0.0165
2011	316.228	13.191	93.254	0.0105
2012	100.000	14.931	94.936	0.0122
2013	100.000	14.997	94.997	0.0165
2014	100.000	14.574	94.573	0.0099
2015	316.228	13.069	93.116	0.0152
2016	100.000	15.155	95.155	0.0060
2017	100.000	15.022	95.073	0.0133
2018	3.162	17.456	97.459	0.0045
2019	31.623	16.295	96.297	0.0127

Tabella 10: Modellazione tramite P-splines unidimensionali, con $\lambda=100$, per l'Italia: anno, gradi di libertà (*gdl*), AIC e RMSE.

Anno	<i>gdl</i>	AIC	RMSE
1960	12.879	176.492	0.0974

Tabella 10 (continua)

Anno	<i>gdl</i>	AIC	RMSE
1961	12.884	126.887	0.1574
1962	12.980	153.785	0.1228
1963	12.428	168.316	0.1221
1964	12.910	123.028	0.1275
1965	12.600	163.539	0.1484
1966	12.424	129.940	0.0950
1967	13.011	163.811	0.1344
1968	13.081	211.726	0.0628
1969	13.071	151.819	0.1742
1970	13.065	138.099	0.2493
1971	12.541	163.768	0.0711
1972	13.076	157.193	0.1213
1973	13.132	154.641	0.2641
1974	12.572	176.596	0.1836
1975	13.179	146.016	0.1518
1976	13.197	150.383	0.3321
1977	13.172	186.801	0.1205
1978	13.164	144.218	0.1145
1979	13.164	159.620	0.2413
1980	13.222	113.235	0.1863
1981	13.169	110.154	0.2578
1982	13.166	104.905	0.1122
1983	13.237	104.141	0.0602
1984	13.182	104.977	0.1251
1985	13.217	118.071	0.0461
1986	13.226	145.635	0.4457
1987	12.692	97.795	0.0954
1988	13.293	106.786	0.0656
1989	13.310	137.126	0.0452

Tabella 10 (continua)

Anno	<i>gdl</i>	AIC	RMSE
1990	13.393	103.455	0.0827
1991	13.448	124.091	0.0827
1992	13.479	120.709	0.0319
1993	13.478	113.398	0.0534
1994	13.503	108.411	0.0293
1995	13.553	207.677	0.0856
1996	13.554	167.606	0.0542
1997	13.529	144.205	0.0311
1998	13.559	139.920	0.0566
1999	13.543	143.859	0.0405
2000	13.530	116.660	0.0185
2001	13.559	122.875	0.0639
2002	13.558	124.879	0.0344
2003	13.589	92.991	0.0442
2004	13.507	128.936	0.0501
2005	13.558	94.862	0.0216
2006	13.556	102.301	0.0384
2007	13.563	120.955	0.0553
2008	13.579	129.934	0.0275
2009	13.591	105.300	0.0372
2010	13.584	121.577	0.0447
2011	13.587	114.124	0.0407
2012	13.599	103.106	0.0153
2013	13.568	106.792	0.0258
2014	13.562	86.251	0.0220
2015	13.655	109.857	0.0318
2016	13.580	147.363	0.0297
2017	13.642	94.799	0.0315
2018	13.603	80.282	0.0110

Tabella 10 (continua)

Anno	<i>gdl</i>	AIC	RMSE
2019	13.578	94.161	0.0309

Tabella 11: Modellazione tramite P-splines unidimensionali, con $\lambda=100$, per il Giappone: anno, gradi di libertà (*gdl*), AIC e RMSE.

Anno	<i>gdl</i>	AIC	RMSE
1960	12.845	202.516	0.1358
1961	12.851	195.435	0.1338
1962	12.793	180.519	0.1346
1963	13.283	199.994	0.1074
1964	13.314	181.416	0.1699
1965	13.348	152.193	0.2262
1966	13.332	197.425	0.3947
1967	13.032	178.770	0.0700
1968	13.069	244.437	0.0827
1969	12.968	249.285	0.1274
1970	12.914	203.996	0.1266
1971	13.459	161.186	0.1473
1972	13.477	176.746	0.1192
1973	13.585	185.179	0.1879
1974	13.605	149.631	0.1364
1975	12.889	177.594	0.1323
1976	12.989	184.645	0.0889
1977	13.596	171.952	0.2355
1978	13.608	125.644	0.0989
1979	13.610	144.240	0.0955
1980	13.689	177.882	0.0785
1981	13.702	156.048	0.1592
1982	13.698	154.032	0.1175
1983	13.786	166.914	0.1184

Tabella 11 (continua)

Anno	<i>gdl</i>	AIC	RMSE
1984	13.793	135.485	0.0913
1985	13.833	164.140	0.0575
1986	13.838	129.315	0.0314
1987	13.839	147.362	0.0916
1988	13.928	131.122	0.0673
1989	13.946	134.259	0.0658
1990	14.003	149.526	0.0795
1991	14.008	133.573	0.0670
1992	14.038	139.492	0.0481
1993	14.081	112.747	0.0599
1994	14.121	131.008	0.0668
1995	14.209	149.420	0.0449
1996	14.166	116.147	0.0121
1997	14.230	129.673	0.0358
1998	14.285	123.377	0.0437
1999	14.362	149.168	0.0382
2000	14.363	158.157	0.0342
2001	14.410	123.810	0.0282
2002	14.431	139.773	0.0412
2003	14.499	136.933	0.0103
2004	14.546	171.609	0.0441
2005	14.607	148.560	0.0139
2006	14.624	148.288	0.0085
2007	14.647	145.431	0.0150
2008	14.692	146.902	0.0133
2009	14.698	159.677	0.0139
2010	14.736	177.370	0.0138
2011	14.804	159.949	0.0139
2012	14.750	149.309	0.0208

Tabella 11 (continua)

Anno	<i>gdl</i>	AIC	RMSE
2013	14.745	122.851	0.0081
2014	14.736	123.457	0.0095
2015	14.726	191.313	0.0119
2016	14.747	173.578	0.0068
2017	14.745	129.678	0.0068
2018	14.736	172.824	0.0157
2019	14.746	145.074	0.0059

Tabella 12: Modellazione tramite P-splines unidimensionali, con $\lambda=100$, per la Svezia: anno, gradi di libertà (*gdl*), AIC e RMSE.

Anno	<i>gdl</i>	AIC	RMSE
1960	9.853	98.284	0.1513
1961	9.713	85.601	0.1805
1962	9.658	89.547	0.1982
1963	9.648	101.970	0.6539
1964	9.932	94.343	0.6226
1965	10.097	83.761	0.5969
1966	10.109	99.139	0.1068
1967	10.030	90.770	0.0919
1968	10.008	101.075	0.1677
1969	9.981	86.282	0.3622
1970	10.151	78.494	0.1283
1971	10.108	90.440	0.1010
1972	10.079	89.749	0.1200
1973	9.979	88.670	0.3501
1974	9.925	89.442	0.1992
1975	9.967	91.223	0.2596
1976	10.202	68.138	0.1318
1977	10.088	98.705	0.6045

Tabella 12 (continua)

Anno	<i>gdl</i>	AIC	RMSE
1978	10.220	102.957	0.1022
1979	10.161	72.799	0.2297
1980	10.605	100.043	0.1881
1981	10.568	89.267	0.1855
1982	10.562	82.919	0.2689
1983	10.568	119.560	0.1376
1984	10.568	98.963	0.2728
1985	10.205	98.933	0.5826
1986	10.188	106.409	0.5993
1987	10.311	91.408	0.1683
1988	10.292	94.852	0.1440
1989	10.220	88.352	0.1060
1990	10.232	114.690	0.1997
1991	10.698	74.213	0.1255
1992	10.667	92.621	0.0999
1993	10.711	81.095	0.5768
1994	10.676	79.582	0.0699
1995	10.680	100.685	0.3339
1996	10.687	81.414	0.1190
1997	10.666	108.213	0.0759
1998	10.643	77.722	0.1305
1999	10.684	75.671	0.1062
2000	10.637	86.305	0.1237
2001	10.673	105.132	0.0688
2002	10.668	100.412	0.1672
2003	10.652	78.441	0.3385
2004	10.598	81.316	0.1359
2005	10.620	95.523	0.2694
2006	10.210	98.282	0.1015

Tabella 12 (continua)

Anno	<i>gdl</i>	AIC	RMSE
2007	10.642	81.419	0.1051
2008	10.653	88.426	0.1186
2009	10.640	96.490	0.0675
2010	10.611	92.845	0.1734
2011	10.602	85.386	0.0657
2012	10.660	81.708	0.0313
2013	10.643	95.614	0.0549
2014	10.655	86.313	0.0382
2015	10.656	94.349	0.0253
2016	10.683	111.391	0.0388
2017	10.686	120.697	0.0415
2018	10.694	75.138	0.0282
2019	10.637	88.978	0.0254

Tabella 13: Modellazione tramite P-splines unidimensionali, con $\lambda=100$, per gli USA: anno, gradi di libertà (*gdl*), AIC e RMSE.

Anno	<i>gdl</i>	AIC	RMSE
1960	14.932	1405.114	0.0157
1961	14.914	933.927	0.0257
1962	14.959	1050.346	0.0169
1963	14.993	1387.198	0.0118
1964	14.991	876.248	0.0144
1965	15.011	1198.138	0.0163
1966	15.041	919.502	0.0129
1967	15.005	661.989	0.0189
1968	15.083	1192.681	0.0189
1969	15.093	748.264	0.0197
1970	15.087	663.932	0.0129
1971	15.121	455.550	0.0061

Tabella 13 (continua)

Anno	<i>gdl</i>	AIC	RMSE
1972	15.162	528.257	0.0219
1973	15.197	760.496	0.0113
1974	15.181	325.217	0.0110
1975	15.180	756.499	0.0163
1976	15.204	777.797	0.0127
1977	15.224	952.151	0.0097
1968	15.254	1161.402	0.0133
1979	15.261	482.670	0.0104
1980	15.321	695.197	0.0101
1981	15.332	573.773	0.0128
1982	15.326	483.849	0.0180
1983	15.367	618.222	0.0113
1984	15.408	370.415	0.0163
1985	15.447	476.381	0.0191
1986	15.473	660.097	0.0141
1987	15.511	644.220	0.0192
1988	15.546	670.749	0.0119
1989	15.575	456.168	0.0217
1990	15.598	575.685	0.0179
1991	15.628	439.470	0.0083
1992	15.636	433.013	0.0126
1993	15.719	482.526	0.0123
1994	15.725	366.406	0.0117
1995	15.750	395.897	0.0078
1996	15.730	523.301	0.0081
1997	15.708	426.313	0.0166
1998	15.711	523.110	0.0126
1999	15.716	362.452	0.0065
2000	15.709	287.243	0.0205

Tabella 13 (continua)

Anno	<i>gdl</i>	AIC	RMSE
2001	15.731	249.835	0.0108
2002	15.737	293.792	0.0126
2003	15.737	250.984	0.0171
2004	15.705	264.238	0.0094
2005	15.720	366.737	0.0086
2006	15.711	329.751	0.0064
2007	15.711	286.887	0.0078
2008	15.731	321.315	0.0164
2009	15.726	222.921	0.0173
2010	15.741	286.293	0.0126
2011	15.756	195.991	0.0060
2012	15.771	179.264	0.0108
2013	15.792	175.139	0.0164
2014	15.811	240.712	0.0103
2015	15.850	227.137	0.0094
2016	15.889	169.055	0.0058
2017	15.909	190.157	0.0122
2018	15.915	170.100	0.0049
2019	15.929	158.227	0.0135

Tabella 14: Valori dell'errore quadratico medio per le stime ottenute con P-splines bidimensionali, per Stato.

Anno	ITA	JPN	SWE	USA
1960	0.0894	0.1167	0.1071	0.0124
1961	0.1305	0.1259	0.1558	0.0258
1962	0.1003	0.0961	0.0828	0.0140
1963	0.1124	0.0707	0.6552	0.0104
1964	0.0726	0.0846	0.6175	0.0120
1965	0.1513	0.1632	0.6050	0.0143

Tabella 14 (continua)

Anno	ITA	JPN	SWE	USA
1966	0.0460	0.4054	0.0381	0.0162
1967	0.0981	0.0746	0.0657	0.0224
1968	0.0876	0.0232	0.1417	0.0175
1969	0.1559	0.1009	0.3577	0.0214
1970	0.2548	0.0779	0.1271	0.0139
1971	0.0732	0.1290	0.0647	0.0064
1972	0.1230	0.1060	0.0519	0.0194
1973	0.2550	0.1479	0.3261	0.0098
1974	0.1897	0.1102	0.1022	0.0089
1975	0.1082	0.0791	0.2754	0.0139
1976	0.3577	0.0845	0.0274	0.0126
1977	0.0740	0.2296	0.5903	0.0107
1978	0.1256	0.0483	0.1033	0.0081
1979	0.2543	0.0889	0.2106	0.0072
1980	0.1628	0.1004	0.0928	0.0123
1981	0.2505	0.1351	0.1543	0.0104
1982	0.0572	0.1212	0.2522	0.0170
1983	0.0731	0.1243	0.0319	0.0115
1984	0.0853	0.0793	0.2588	0.0121
1985	0.0464	0.0665	0.5902	0.0149
1986	0.4574	0.0404	0.5970	0.0115
1987	0.0870	0.0949	0.1657	0.0130
1988	0.0462	0.0625	0.1160	0.0090
1989	0.0475	0.0599	0.1013	0.0170
1990	0.0748	0.1024	0.2040	0.0137
1991	0.0827	0.0599	0.0680	0.0069
1992	0.0313	0.0399	0.0483	0.0161
1993	0.0821	0.0655	0.5907	0.0108
1994	0.0290	0.0582	0.0594	0.0115

Tabella 14 (continua)

Anno	ITA	JPN	SWE	USA
1995	0.0749	0.0281	0.3482	0.0063
1996	0.0437	0.0246	0.0557	0.0121
1997	0.0364	0.0490	0.0687	0.0127
1998	0.0531	0.0403	0.0966	0.0111
1999	0.0368	0.0310	0.1280	0.0112
2000	0.0163	0.0302	0.1118	0.0193
2001	0.0636	0.0421	0.0558	0.0074
2002	0.0330	0.0433	0.1652	0.0116
2003	0.0609	0.0108	0.3727	0.0135
2004	0.0532	0.0471	0.1336	0.0100
2005	0.0273	0.0140	0.2715	0.0113
2006	0.0286	0.0092	0.0944	0.0068
2007	0.0600	0.0167	0.1112	0.0045
2008	0.0244	0.0104	0.0603	0.0120
2009	0.0318	0.0123	0.0785	0.0166
2010	0.0338	0.0166	0.1414	0.0087
2011	0.0350	0.0185	0.0659	0.0054
2012	0.0285	0.0155	0.0356	0.0095
2013	0.0227	0.0079	0.0535	0.0182
2014	0.0300	0.0083	0.0308	0.0149
2015	0.0455	0.0133	0.0239	0.0083
2016	0.0257	0.0051	0.0462	0.0075
2017	0.0311	0.0077	0.0463	0.0110
2018	0.0090	0.0160	0.0263	0.0051
2019	0.0322	0.0034	0.0204	0.0128

Tabella 15: Valori dell'errore quadratico medio per le stime ottenute con P-splines bidimensionali, con $\lambda_{age}=10$ e $\lambda_{year}=3.162278$, per Stato.

Anno	ITA	JPN	SWE	USA
1960	0.0894	0.1167	0.1071	0.0124
1961	0.1305	0.1259	0.1558	0.0258
1962	0.1003	0.0961	0.0828	0.0140
1963	0.1124	0.0707	0.6552	0.0104
1964	0.0726	0.0846	0.6175	0.0120
1965	0.1513	0.1632	0.6050	0.0143
1966	0.0460	0.4054	0.0381	0.0162
1967	0.0981	0.0746	0.0657	0.0224
1968	0.0876	0.0232	0.1417	0.0175
1969	0.1559	0.1009	0.3577	0.0214
1970	0.2548	0.0779	0.1271	0.0139
1971	0.0732	0.1290	0.0647	0.0064
1972	0.1230	0.1060	0.0519	0.0194
1973	0.2550	0.1479	0.3261	0.0098
1974	0.1897	0.1102	0.1022	0.0089
1975	0.1082	0.0791	0.2754	0.0139
1976	0.3577	0.0845	0.0274	0.0126
1977	0.0740	0.2296	0.5903	0.0107
1978	0.1256	0.0483	0.1033	0.0081
1979	0.2543	0.0889	0.2106	0.0072
1980	0.1628	0.1004	0.0928	0.0123
1981	0.2505	0.1351	0.1543	0.0104
1982	0.0572	0.1212	0.2522	0.0170
1983	0.0731	0.1243	0.0319	0.0115
1984	0.0853	0.0793	0.2588	0.0121
1985	0.0464	0.0665	0.5902	0.0149
1986	0.4574	0.0404	0.5970	0.0115
1987	0.0870	0.0949	0.1657	0.0130

Tabella 15 (continua)

Anno	ITA	JPN	SWE	USA
1988	0.0462	0.0625	0.1160	0.0090
1989	0.0475	0.0599	0.1013	0.0170
1990	0.0748	0.1024	0.2040	0.0137
1991	0.0827	0.0599	0.0680	0.0069
1992	0.0313	0.0399	0.0483	0.0161
1993	0.0821	0.0655	0.5907	0.0108
1994	0.0290	0.0582	0.0594	0.0115
1995	0.0749	0.0281	0.3482	0.0063
1996	0.0437	0.0246	0.0557	0.0121
1997	0.0364	0.0490	0.0687	0.0127
1998	0.0531	0.0403	0.0966	0.0111
1999	0.0368	0.0310	0.1280	0.0112
2000	0.0163	0.0302	0.1118	0.0193
2001	0.0636	0.0421	0.0558	0.0074
2002	0.0330	0.0433	0.1652	0.0116
2003	0.0609	0.0108	0.3727	0.0135
2004	0.0532	0.0471	0.1336	0.0100
2005	0.0273	0.0140	0.2715	0.0113
2006	0.0286	0.0092	0.0944	0.0068
2007	0.0600	0.0167	0.1112	0.0045
2008	0.0244	0.0104	0.0603	0.0120
2009	0.0318	0.0123	0.0785	0.0166
2010	0.0338	0.0166	0.1414	0.0087
2011	0.0350	0.0185	0.0659	0.0054
2012	0.0285	0.0155	0.0356	0.0095
2013	0.0227	0.0079	0.0535	0.0182
2014	0.0300	0.0083	0.0308	0.0149
2015	0.0455	0.0133	0.0239	0.0083
2016	0.0257	0.0051	0.0462	0.0075

Tabella 15 (continua)

Anno	ITA	JPN	SWE	USA
2017	0.0311	0.0077	0.0463	0.0110
2018	0.0090	0.0160	0.0263	0.0051
2019	0.0322	0.0034	0.0204	0.0128