

Università degli Studi di Padova  
Dipartimento di Scienze Statistiche  
Corso di Laurea Magistrale in  
Scienze Statistiche



**Quando il caldo è una minaccia per la salute.  
Analisi della mortalità in Veneto in relazione alle  
condizioni climatiche.**

Relatore: Prof.ssa Giovanna Boccuzzo  
Dipartimento di Scienze Statistiche

Correlatore: Dott. Ugo Fedeli  
Sistema Epidemiologico Regionale, Regione del Veneto

Laureanda: Giulia Capodaglio  
Matricola N. 1082576

Anno Accademico 2015/2016



*“Una meta trova il suo senso  
nel cammino percorso  
e nelle persone incontrate”*



# INDICE

---

INTRODUZIONE .....	1
1. IL CASO DI STUDIO	
1.1. Enunciazione del problema.....	5
1.2. Sviluppi in letteratura .....	6
2. DATI E FONTI	
2.1. Mortalità .....	9
2.2. Humidex, indicatore di disagio climatico.....	9
3. METODI	
3.1. Tassi di mortalità.....	13
3.2. Strumenti per l'analisi delle serie storiche .....	14
3.3. Modelli additivi generalizzati .....	17
4. RICERCA DEGLI ECCESSI DI MORTALITÀ	
4.1. Andamento della mortalità nel Veneto dal 2000 al 2013.....	24
4.2. La mortalità e la componente stagionale .....	30
4.3. Andamento della mortalità al netto della stagionalità.....	37
4.4. Identificazione degli eccessi di mortalità significativi.....	39
4.5. Mortalità e humidex: prime osservazioni .....	43
5. MODELLI PER LA RELAZIONE TRA MORTALITÀ E HUMIDEX	
5.1. Relazione tra mortalità e humidex per classi d'età .....	51
5.2. Effetto dell'humidex sulla mortalità prima e dopo i 75 anni .....	55
5.2.1. Popolazione di età minore di 75 anni .....	55
5.2.2. Popolazione di età maggiore di 75 anni.....	58
5.3. Discussione dei risultati .....	63
6. CONCLUSIONI .....	65
APPENDICE A.....	67
APPENDICE B .....	69
APPENDICE C .....	73
APPENDICE D.....	81
BIBLIOGRAFIA.....	85

SITOGRAFIA .....	88
------------------	----

# INTRODUZIONE

---

Nella presente tesi si ricerca la relazione tra la mortalità e le condizioni climatiche in estate, nel Veneto. L'interesse è individuare quando, e in presenza di quali condizioni climatiche, lo stato di salute e il benessere possono essere a rischio, e verificare inoltre se esistono differenze tra le diverse aree climatiche del Veneto: Litorale, Pianura, Pedemontana e Montana. Quindi gli obiettivi principali che si intende perseguire sono:

1. Individuare quando si sono verificati eccessi di mortalità significativi dal 2000 al 2013;
2. Identificare un valore soglia di humidex (indicatore di disagio fisico, basato su temperatura e umidità relativa) oltre il quale la mortalità aumenta evidentemente.

Si tratta di un tema che ha acquisito una notevole rilevanza soprattutto negli ultimi decenni, a causa dei cambiamenti climatici che si sono avvertiti in diverse zone del mondo. Uno degli effetti di tali cambiamenti è stato l'aumento delle ondate di calore, ossia condizioni meteorologiche che si verificano durante la stagione estiva, caratterizzate da temperature al di sopra dei valori usuali. Dagli studi epidemiologici emerge che i gruppi a più alto rischio sono costituiti da persone anziane (>75 anni) e da persone con alcune particolari patologie (Schifano et al. 2009). Le malattie cardiovascolari sono tra le principali cause di morte durante le ondate di calore (Lubczyńska et al. 2015), assieme alle malattie respiratorie (Oudin Astrom et al. 2015). Nel presente lavoro sono stati analizzati i decessi giornalieri, per tutte le cause, nella regione del Veneto dal 2000 al 2013, resi disponibili dal Sistema Epidemiologico Regionale (SER), in relazione ai valori di humidex (indicatore di disagio climatico basato su temperatura e umidità relativa) forniti dall'ARPAV (Agenzia Regionale per la Prevenzione e Protezione Ambientale del Veneto). Per l'analisi della serie storica della mortalità è stata applicata una modellazione ARIMA, con lo scopo di identificare gli eccessi di mortalità nei quattordici anni. Successivamente sono stati impiegati i Modelli Additivi Generalizzati per lo studio della relazione tra la mortalità e le condizioni climatiche.

La tesi si articola in sei capitoli: il primo descrive la problematica e gli sviluppi in letteratura, il secondo capitolo presenta i dati utilizzati e le relative fonti, il terzo capitolo descrive i metodi utilizzati. Il quarto e il quinto capitolo costituiscono il cuore delle analisi. Il quarto capitolo si focalizza sull'andamento della mortalità nel Veneto nei quattordici anni e sull'identificazione degli eccessi eventualmente verificatesi, e inoltre descrive le condizioni climatiche che si sono registrate nel Veneto negli stessi anni, 2000–2013. Il quinto capitolo entra nel vivo della questione d'interesse, ossia la relazione tra la mortalità e i valori di humidex, ricercando la soglia oltre la quale il rischio per la salute aumenta sensibilmente. Infine il capitolo 6 riporta i risultati salienti e le conclusioni generali del lavoro svolto.

Il lavoro è stato possibile grazie alla collaborazione del SER, in cui ho svolto lo stage, e dell'ARPAV.

Il SER è il Sistema Epidemiologico Regionale, istituito nel 1999, che supporta la programmazione socio-sanitaria regionale realizzando attività di: gestione diretta del Registro di Mortalità; epidemiologia descrittiva delle patologie, dei problemi di salute e dell'accesso ai servizi socio-sanitari della Regione Veneto; monitoraggio epidemiologico e valutazione dei percorsi e degli esiti di cura; studio dei fattori di rischio per l'incidenza delle patologie e del loro impatto sulla popolazione e delle opportunità di prevenzione primaria; stesura della relazione Socio-Sanitaria Regionale; integrazione informativa tra le basi di dati dei Registri di Mortalità e di Patologia afferenti ed i flussi informativi correnti; produzione e diffusione di informazioni di epidemiologia e sanità pubblica, in collaborazione con i registri di Patologia afferenti.

L'ARPAV è l'Agenzia Regionale per la Prevenzione e Protezione Ambientale del Veneto, istituita nel 1997, la quale realizza attività di prevenzione e controllo ambientale attraverso le strutture tecniche presenti nel territorio: prevenzione e controllo ambientale; previsione, informazione ed elaborazione meteo-climatica e radarmeteorologica; organizzazione e gestione del sistema informativo regionale per il monitoraggio ambientale ed epidemiologico in relazione ai fattori ambientali; promozione di attività di educazione ed informazione ambientale; fornitura di supporto tecnico-scientifico per la valutazione di impatto ambientale e per la determinazione del



danno ambientale; promozione di iniziative di ricerca di base e applicata sulle forme di tutela ambientale.



# 1. IL CASO DI STUDIO

---

## 1.1. Enunciazione del problema

Negli ultimi decenni, nelle varie zone del mondo si sono avvertiti dei cambiamenti delle condizioni climatiche. Uno degli effetti è stato l'aumento delle ondate di calore, che qualche tempo fa erano considerate solo eventi eccezionali. In molti Paesi, tra cui l'Italia, si ricorda in modo particolare l'estate del 2003 che è stata caratterizzata da temperature elevate e da vere e proprie ondate di calore.

In seguito a questo episodio, il Ministero della Salute ha disposto un'indagine epidemiologica sulla mortalità estiva, affidandola all'Istituto Superiore di Sanità. L'indagine basata sul confronto della mortalità nei mesi di luglio ed agosto fra il 2002 ed il 2003, ha stimato un eccesso di mortalità negli anziani, con più di 75 anni, di quasi 8.000 persone (Conti et al, 2004). Molte regioni a seguito di ciò hanno avviato attività di pianificazione, di educazione e di comunicazione, oltre alla elaborazione di progetti, raccomandazioni e linee guida locali. Inoltre a partire dal 2004, il Dipartimento della Protezione Civile, ha attivato il progetto "Sistema Nazionale di Sorveglianza, previsione e di allarme per la prevenzione degli effetti delle ondate di calore sulla salute della popolazione". Si tratta di un programma nazionale che prevede l'attivazione, nelle città italiane, tra cui anche Venezia e Verona, di sistemi di allarme per la prevenzione degli effetti del caldo sulla salute (Heat Health Watch Warning Systems, HHWWS).

Anche la regione Veneto ha elaborato, a partire dal 2004, un "Protocollo Sanitario Operativo" per la prevenzione della mortalità causata da elevate temperature nella popolazione anziana al di sopra di 75 anni o 65 anni con patologie croniche invalidanti, e con l'attivazione di piani operativi sociali al fine di intervenire prontamente negli stati di rischio e fragilità. Questi piani prevedono anche il coinvolgimento di altri enti del territorio a diretto contatto con le persone (ULSS, Servizi Sociali, Amministrazioni comunali, ecc.) e enti che svolgono attività inerenti alla problematica, tra cui ARPAV, il Sistema regionale della Prevenzione, il Sistema Epidemiologico Regionale (BUR n.46 del 8 maggio 2005: "Protocollo per la prevenzione delle patologie da elevate temperature nella popolazione anziana della Regione del Veneto- Estate 2015").

Conoscere quindi i fattori climatici (temperatura, umidità, vento, ecc..) ed ambientali che possono influire sullo stato di salute e sul benessere delle persone rappresenta una tematica assolutamente attuale e importante, che ha lo scopo di indirizzare interventi per la prevenzione e il controllo di situazioni di disagio, quando possibile.

## 1.2. Sviluppi in letteratura

Diversi sono gli studi che negli ultimi decenni hanno trattato questa problematica, condotti non solo in Italia e in Europa ma anche in America e in altre parti del mondo. Ciò che si nota è una sensibilità comune nel monitorare gli effetti delle condizioni climatiche ed ambientali sulla salute umana, con l'obiettivo principale di prevenzione e riduzione dei rischi.

Uno degli studi a cui si è fatto riferimento in questo lavoro per ricercare l'andamento della mortalità e gli eventuali eccessi è quello condotto in Scozia da Wagner et al. nel 2009: l'obiettivo era il monitoraggio degli eccessi di mortalità inattesi, per indirizzare le politiche sulla salute. I punti di forza di questo sistema di monitoraggio erano: l'elaborazione automatica dei decessi giornalieri, le correzioni giornaliere, e l'utilizzo di due modelli statistici per il calcolo della mortalità attesa, un'estensione al modello di Serfling e i Modelli Additivi Generalizzati (GAM). I modelli GAM sono risultati essere, grazie alla loro flessibilità, i più adatti per rilevare i picchi di mortalità, sulla base dell'andamento degli anni precedenti.

La stagionalità della mortalità è un elemento molto importante perché il suo impatto molto spesso varia tra gruppi di individui. Ad esempio lo studio condotto da Burkart et al. (2011) su dati del Bangladesh ha mostrato che l'effetto della stagione, o delle condizioni climatiche che caratterizzano la stagione, dipende dalle precondizioni di una sottopopolazione o regione. In particolare, nelle aree rurali è stata trovata una mortalità infantile maggiore, mentre le aree urbane sono risultate strettamente legate ad eccessi di mortalità estiva, soprattutto mortalità cardiovascolare.

Un altro studio di riferimento riguarda l'effetto delle ondate di calore sulla mortalità in gruppi suscettibili in due città diverse dell'Europa, una mediterranea (Roma) ed una del nord (Stoccolma), dal 2000 al 2008 (Oudin Astrom et al. 2015). Dai risultati è emerso che in entrambe le città la mortalità durante le ondate di calore è cresciuta per la

popolazione over 50, e soprattutto per alcuni sottogruppi di popolazione più fragili: individui con problemi cardiovascolari, respiratori o psichiatrici e diabetici. Gli indicatori di temperatura che sono stati utilizzati sono: minimo, massimo, media e temperatura massima apparente (MAT). L'effetto è stato stimato al netto del giorno della settimana, dei giorni di vacanza, e della stagionalità nel periodo. Per quanto riguarda le malattie cardiovascolari, il rischio più alto è stato trovato per l'ischemia (Lubczyńska et al. 2015). Il fatto che l'effetto delle temperature estreme colpisca in modo particolare gli anziani è stato ritrovato in molti studi (Schifano et al. 2009; Oudin Astrom et al. 2011; Xuan et al. 2014; Seposo et al. 2015) Inoltre la relazione sembra essere più forte nel giorno stesso dell'evento, ma il rischio rimane significativamente elevato per circa un giorno dopo l'evento. L'aumento nel rischio è comunque più evidente nei giorni con temperature più alte (Lubczyńska et al. 2015).

Uno studio condotto negli Stati Uniti nel 2014 ha ricercato la relazione tra la temperatura e la morte non accidentale. È stata esaminata l'associazione tra la temperatura e la mortalità in 7 anni, in 211 città degli Stati Uniti, con una particolare attenzione ai cambiamenti di tempo. La forma della relazione, non lineare, è risultata essere diversa tra località con differenti condizioni climatiche. Inoltre è stato trovato che il rischio dovuto all'effetto del caldo è influenzato significativamente dalla temperatura media estiva e dall'uso dell'aria condizionata. In modo particolare è emerso che per le persone che vivono nelle città degli Stati Uniti con temperature estive più basse, il rischio di mortalità, nei periodi caratterizzati da temperature elevate, è più alto (Nordio et al., 2015).

In Italia il Dipartimento di Epidemiologia del Lazio pubblica ogni anno dal 2003 un rapporto sull'influenza delle ondate di calore sulla salute, che presenta i risultati relativi ai sistemi di allarme HHWW (Heat/Health Watch/Warning Systems), all'attività di sorveglianza e ai piani operativi per la prevenzione. In questo progetto sono coinvolte diverse città italiane (34 nel 2014) tra cui anche Venezia (dal 2005) e Verona (dal 2006). Dal rapporto dell'estate del 2013 si può avere ad esempio un confronto tra le diverse città italiane, per ciò che è avvenuto nel periodo 15 Maggio – 15 Settembre del 2013: è stato osservato un eccesso di mortalità significativo a Milano (+14%), Verona (+13%), Venezia (+13%), Genova (+13%), Roma (+8%), Viterbo (+22%), Frosinone (+28%) e Latina

(+23%). Tra le città che hanno solo il sistema di rilevazione della mortalità giornaliera è stato osservato un eccesso a Padova (+12%), Trento (+19%) e Potenza (+24%). Nella maggior parte delle città del sud non sono state osservate condizioni a rischio per la salute nel corso dell'estate 2013 e la mortalità è stata uguale o inferiore all'atteso.

Uno studio condotto da Schifano et al. nel 2012 contribuisce ad avvalorare l'importanza di continuare a ricercare l'impatto sulla salute delle elevate temperature per poter in qualche modo indirizzare le politiche di prevenzione. Infatti lo studio appena specificato ha ricercato i cambiamenti nell'effetto del caldo nella mortalità tra gli anziani prima e dopo le ondate di calore del 2003 e successivamente l'introduzione delle attività di prevenzione del caldo, mettendo a confronto 34 maggiori città Italiane. I risultati hanno portato evidenza di un cambiamento: l'effetto sulla salute sembra essere stato inferiore dopo le attività di prevenzione attuate; questo è stato però attribuito sia alla diversa distribuzione della temperatura delle estati successive al 2003, sia all'introduzione delle misure di adattamento. Quindi la riduzione dell'effetto delle alte temperature suggerisce che i programmi di prevenzione possono mitigare l'impatto. Inoltre dai risultati è stato osservato anche un effetto per temperature un po' più basse, all'inizio dell'estate, quando la popolazione non si è ancora adattata e le attività di intervento non sono ancora completamente attivate.

## 2. DATI E FONTI

---

### 2.1. Mortalità

I dati relativi alla mortalità utilizzati in questo lavoro riguardano il numero di decessi (per tutte le cause) nel Veneto dall'1 gennaio 2000 al 31 dicembre 2013, per comune di residenza, età e sesso. I suddetti dati sono stati forniti dal Sistema Epidemiologico Regionale (SER) e sono contenuti nell'archivio generale delle schede di morte che fa parte del Datawarehouse Sanità della Regione del Veneto. Dall'anno 2006 la gestione del flusso informativo della mortalità è affidata al SER che si occupa sia della codifica delle cause di morte che della registrazione dei dati. Fino al 2005 il flusso regionale veniva invece generato dai dati inviati periodicamente dalle Aziende ULSS secondo specifiche tecniche regionali. La gestione centralizzata da parte del SER favorisce una migliore qualità dei dati: più omogeneità nella codifica delle cause di morte, controlli più agevoli e la riduzione degli errori legati all'informatizzazione dei dati.

### 2.2. Humidex, indicatore di disagio climatico

I dati sulle condizioni climatiche sono stati forniti dall'Agenzia Regionale per la Prevenzione e Protezione Ambientale del Veneto (ARPAV), la quale realizza attività di prevenzione e controllo ambientale attraverso le strutture tecniche presenti nel territorio.

Per monitorare l'intensità del caldo e il disagio fisico che esso può causare, esistono diversi indicatori chiamati *Indicatori di disagio climatico*. Uno di questi, molto utilizzato, è l'humidex. La ricerca e lo studio di questo indice è iniziata in Canada nel 1965<sup>1</sup>, con lo scopo di trovare una scala di valori che, a partire dalla temperatura e all'umidità relativa dell'aria, fosse in grado di descrivere il disagio che si prova nei classici giorni estivi caldo-umidi. Tale indice si basa su una semplice relazione empirica che prende in considerazione la temperatura dell'aria e la tensione di vapore. In realtà, anziché l'umidità relativa, questo indice usa un parametro correlato, ovvero la pressione di vapore, secondo la formula:

---

<sup>1</sup> Informazioni tratte da [www.centrometeo.com](http://www.centrometeo.com).

$$H = T_a + (0.5555 (e - 10)),$$

dove  $T_a$  è la temperatura in °C, mentre  $e$  è la pressione di vapore dell'aria (hPa). Però, dal momento che è più agevole conoscere l'umidità relativa (UR)<sup>2</sup>, anziché la pressione di vapore, conoscendo il legame tra queste ultime si può ottenere un'altra versione dell'indice Humidex:

$$H = T_a + (0.5555(0.06 * UR * 10^{0.03T_a} - 10))^3.$$

L'humidex non ha un'unità di misura propria, pertanto si può associare alla stessa unità della temperatura (°C) anche se si tratta di un indice e non di una grandezza fisica.

L'indice è applicabile tra le temperature di 20°C e 55°C. Esternamente a questo intervallo si ha sempre "benessere" sotto i 20°C ed "elevato pericolo" oltre i 55°C, qualunque sia il valore dell'umidità. I risultati che si ottengono applicando la formula sono mostrati in tabella 1.1 e le rispettive categorie di disagio fisico in tabella 1.2.

Per le analisi viene quindi utilizzato questo indicatore, e non la temperatura, perché esso misura la temperatura percepita, tenendo conto sia dell'umidità relativa che della temperatura, essendo che l'umidità dell'aria può aumentare la sensazione di caldo<sup>4</sup>.

I valori dell'humidex utilizzati in questo lavoro sono stati calcolati dall'ARPAV, a partire dai valori di temperatura e umidità rilevati dalle centraline disposte sul territorio. Sono state utilizzate le medie giornaliere di humidex per 33 centraline del Veneto, che rappresentano tutte le provincie e tutte le fasce climatiche (allegato A). Le fasce climatiche sono le aree di allerta climatica definite dall'ARPAV sulla base della somiglianza delle condizioni climatiche dei comuni. In figura 1.1 si riporta la mappa del Veneto comprendente la specificazione delle aree climatiche: Litorale, Pianura, Pedemontana, Val Belluna e Montana. A partire da questi dati, per gli scopi sopra specificati, sono state calcolate le medie giornaliere di humidex per fascia climatica

---

<sup>2</sup> Umidità relativa:  $UR = \frac{\text{umidità assoluta}}{\text{umidità di saturazione}} \times 100$ .

L'umidità assoluta è la quantità, espressa in grammi, di vapore acqueo contenuta in un metro cubo di aria ad una determinata temperatura. L'umidità di saturazione è la quantità massima, in grammi, di vapore acqueo che può contenere un metro cubo di aria a quella determinata temperatura

<sup>3</sup> Informazioni tratte da [www.centrometeo.com](http://www.centrometeo.com).

<sup>4</sup> L'umidità dell'aria può aumentare la sensazione di caldo perché un contenuto maggiore di vapore acqueo nell'aria rende meno agevole l'evaporazione di acqua contenuta nel sudore, processo fondamentale per il corpo umano per liberare calore in eccesso.



(aggregando la Val Belluna e la fascia Montana in un'unica area climatica, che sarà chiamata Montana, perché geograficamente vicine e anche perché non ci sono i dati delle centraline). Dalla tabella 1.1 si vede che l'humidex può variare tra i 13°C e i 93°C. Nella nostra regione l'intervallo è naturalmente più ristretto perché non si raggiungono temperature oltre i 40°C. In tabella 1.3 si mostrano alcune statistiche descrittive del valore medio giornaliero di humidex calcolato nel periodo 21/05 – 14/09, dal 2000 al 2013, separatamente per fascia climatica. Il valore minimo risulta inferiore a 13 perché qualche giorno ha registrato una temperatura inferiore ai 20°C.

**Tabella 1.1.** Valori di humidex per temperatura (°C, in riga) e umidità relativa (% , in colonna). I colori si riferiscono al grado di disagio fisico attribuito, le cui categorie sono riportate in tabella 1.2.

Fonte: <http://www.centrometeo.com>

	25%	30%	35%	40%	45%	50%	55%	60%	65%	70%	75%	80%	85%	90%	95%	100%
42°C	47	50	53	56	59	62	65	68	71	74	77	81	84	87	90	93
41°C	45	48	51	54	56	59	62	65	68	71	73	76	79	82	85	88
40°C	43	46	48	51	54	56	59	62	64	67	70	72	75	78	80	83
39°C	41	44	46	49	51	54	56	59	61	64	66	68	71	73	76	78
38°C	40	42	44	46	49	51	53	56	58	60	63	65	67	69	72	74
37°C	38	40	42	44	46	48	51	53	55	57	59	61	63	66	68	70
36°C	36	38	40	42	44	46	48	50	52	54	56	58	60	62	64	66
35°C	34	36	38	40	42	44	46	47	49	51	53	55	57	59	61	62
34°C	33	34	36	38	40	41	43	45	47	48	50	52	54	55	57	59
33°C	31	33	34	36	38	39	41	43	44	46	47	49	51	52	54	56
32°C	30	31	33	34	36	37	39	40	42	43	45	46	48	49	51	52
31°C	28	30	31	32	34	35	37	38	39	41	42	44	45	47	48	49
30°C	27	28	29	31	32	33	35	36	37	39	40	41	43	44	45	46
29°C	25	26	28	29	30	31	33	34	35	36	38	39	40	41	42	44
28°C	24	25	26	27	28	30	31	32	33	34	35	36	38	39	40	41
27°C	22	23	25	26	27	28	29	30	31	32	33	34	35	36	37	39
26°C	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36
25°C	20	21	22	22	23	24	25	26	27	28	29	30	31	32	33	34
24°C	18	19	20	21	22	23	24	24	25	26	27	28	29	30	31	31
23°C	17	18	19	20	20	21	22	23	24	24	25	26	27	28	29	29
22°C	16	17	17	18	19	20	20	21	22	23	23	24	25	26	26	27
21°C	15	15	16	17	17	18	19	20	20	21	22	22	23	24	25	25
20°C	13	14	15	15	16	17	17	18	19	19	20	21	21	22	23	23

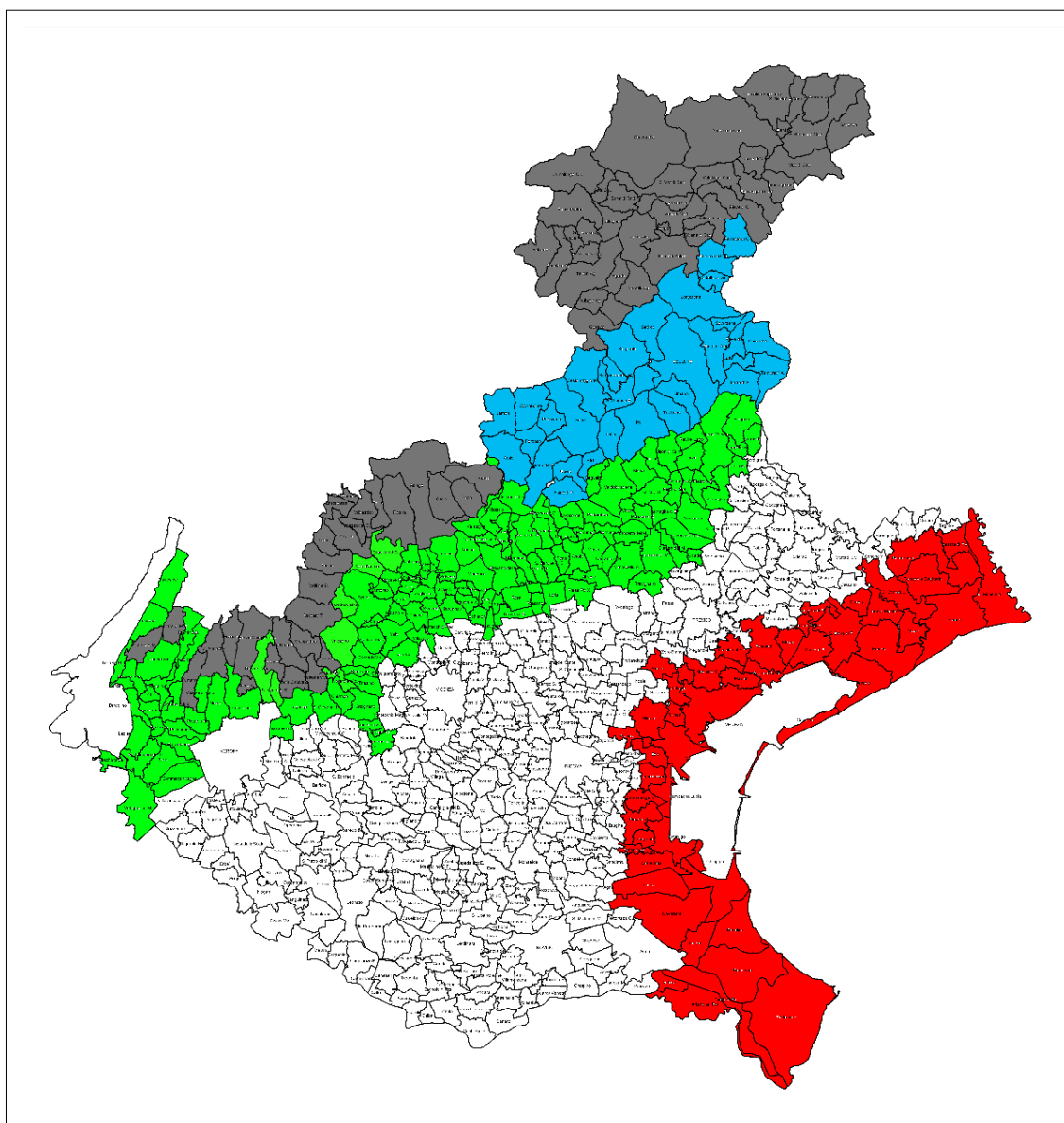
**Tabella 1.2.** Categorie di disagio fisico sulla base dei valori di humidex. I colori corrispondono a quelli della tabella 1.1.

CATEGORIE	Benessere	Disagio debole	Disagio moderato	Disagio intenso	Elevato pericolo
HUMIDEX (°C)	H < 27	27 ≤ H < 30	30 ≤ H < 40	40 ≤ H < 55	H ≥ 55

**Tabella 1.3** Statistiche descrittive dell'humidex, valore medio giornaliero calcolato nel periodo 21/05–14/09, dal 2000 al 2013, per fascia climatica. Fonte dei dati: ARPAV.

Fascia	Media	Dev std	Minimo	Massimo
<b>Litorale</b>	28.63	5.05	11.75	41.30
<b>Montana</b>	23.85	4.81	7.38	36.47
<b>Pedemontana</b>	27.54	5.07	9.44	38.95
<b>Pianura</b>	28.32	5.08	10.85	40.32

**Figura 1.1.** Suddivisione del Veneto per fasce climatiche: Litorale (rosso), Pianura (bianco), Pedemontana (verde), Val Belluna (azzurro), Montana (grigio). Fonte: ARPAV.



## 3. METODI

---

In questo capitolo si descrivono i metodi utilizzati per le analisi e le scelte fatte.

Sia la costruzione del dataset, che le analisi e la stima dei modelli sono stati condotti tramite il software statistico SAS 9.3 e 9.4.

### 3.1. Tassi di mortalità

Per vedere quale sia stato l'andamento della mortalità dal 2000 al 2013, vengono impiegati tassi di mortalità grezzi, specifici e standardizzati.

Il tasso **grezzo** di mortalità è il rapporto tra il numero di decessi nella popolazione, e la popolazione stessa in una determinata unità di tempo (generalmente l'anno):

$$\frac{\text{Numero di decessi}_t}{\text{Popolazione totale}_t} \times k$$

dove t è l'unità di tempo e k è la costante moltiplicativa (generalmente 1.000 o 100.000).

La popolazione totale rappresenta la massa a rischio al tempo t e viene qui calcolata come la popolazione mediamente presente nell'anno t.

I tassi **specifici** per età sono costruiti come i tassi grezzi ma per specifici gruppi di età in modo da confrontare la mortalità nelle differenti età o in periodi diversi di tempo:

$$\frac{\text{Numero di decessi}_{x,t}}{\text{Popolazione totale}_{x,t}} \times k$$

dove x rappresenta l'età o la classe d'età e k la costante moltiplicativa.

Il tasso **standardizzato (diretto)** di mortalità è il tasso che assumerebbe la popolazione in studio se avesse la stessa struttura per età di una popolazione standard. Questo tasso viene calcolato per confrontare la mortalità in popolazioni con struttura per età diversa, o in periodi diversi.

La popolazione standard scelta nel caso in studio è la popolazione del Veneto al 1° gennaio 2007, ed è stata applicata sia per i maschi che per le femmine volendo mettere a confronto anche i due sessi.

Il tasso standardizzato è così formulato:

$$T_{st} = \frac{\sum_i M_i * std_i}{\sum_i std_i} \times k$$

dove:

- $M_i$  è il tasso di mortalità nella popolazione in osservazione nella  $i$ -esima classe d'età,
- $std_i$  è la popolazione standard nella  $i$ -esima classe d'età
- $k = 100.000$

La standardizzazione è stata effettuata utilizzando le seguenti classi d'età: 0, 1-4, 5-14, 15-24, 25-44, 45-64, 65-74, 75-84, 85+.

Gli **intervalli di confidenza** per i tassi sono stati calcolati sulla base dell'approssimazione gaussiana, essendo il numero di casi maggiore di 20 (Fleiss, 1981).

Sia  $X$  la variabile casuale di riferimento che rappresenta il numero di casi e  $\mu$  la sua media, allora  $X \sim Poisson(\mu)$ . Secondo l'approssimazione gaussiana:

$$\frac{X-\mu}{\sqrt{\mu}} \rightarrow N(0,1)$$

e l'intervallo di confidenza si ricava risolvendo l'equazione di secondo grado in  $\mu$ :

$$z = \frac{|X-\mu|}{\sqrt{\mu}} < z_{\alpha/2} \text{ le cui soluzioni sono } \left( x + \frac{z_{\alpha/2}^2}{2} \right) \pm \left( \frac{z_{\alpha/2}^2}{2} \sqrt{z_{\alpha/2}^2 + 4x} \right).$$

### 3.2. Strumenti per l'analisi delle serie storiche

La serie giornaliera dei decessi in esame rappresenta la serie storica della mortalità in Veneto dal 2000 al 2013. A questa serie sono state applicate alcune trasformazioni con l'obiettivo principale di studiare il fenomeno della mortalità in questi anni, conoscere le sue componenti (trend e stagionalità) e verificare se ci siano stati degli eventi inattesi che abbiano modificato in alcuni periodi l'andamento consueto della mortalità.

Come primo strumento è stata utilizzata una media mobile semplice, centrata, la cui formulazione generale è la seguente:

$$y_t^* = \frac{1}{2m+1} \sum_{i=-m}^m y_{t+i}, \quad t = m+1, \dots, n-m$$

con  $m \in \mathbb{N}$ . Il numero di termini  $(2m+1)$  implicati nella trasformazione viene detto ordine della media mobile.

Successivamente, per depurare la serie dalla componente stagionale, è stata applicata una destagionalizzazione della serie mediante una procedura basata sull'uso di medie mobili semplici, che ipotizza una componente stagionale stabile. È stato ritenuto opportuno considerare la serie mensile anziché quella giornaliera, e quindi sono stati aggregati i decessi per mese.

La prima tappa della procedura consiste nella stima dei valori di trend-ciclo ( $y_t^{**}$ ) con media mobile centrata di ordine 13 (essendo la serie mensile) e nella successiva individuazione dei valori stimati per l'insieme della componente stagionale e aleatoria,  $\widehat{S\varepsilon}_t$ . Assumendo un modello additivo del tipo  $Y_t = T + S + I$  (T per la componente tendenziale ciclica, S per la stagionalità e I per la componente irregolare),  $\widehat{S\varepsilon}_t$  è ottenuta come differenza tra il valore della serie e la stima del valore di trend:  $\widehat{S\varepsilon}_t = y_t - y_t^{**}$ . Il valore ottenuto viene detto *indice specifico di stagionalità* ( $IS_T$ ), ed è caratterizzato dalla presenza della componente accidentale. Successivamente vengono calcolati i coefficienti grezzi di stagionalità:

$$\hat{S}_j^* = \frac{1}{N} \sum_{T=1}^N IS_{t,j} \quad \text{per } j = 1, \dots, 12,$$

dove N è il numero di anni in cui la serie storica è stata osservata e  $IS_{t,j}$  è l'indice specifico corrispondente al periodo j dell'anno T. Prima però di utilizzare i risultati ottenuti occorre verificare che a livello annuale la componente stagionale scompaia. Se ciò non avviene è necessario modificare i coefficienti stagionali affinché la condizione sia verificata, i coefficienti risultanti sono chiamati coefficienti *ideali* di stagionalità (Di Fonzo & Lisi, 2005).

Al fine di individuare i valori anomali della serie, è stato costruito un modello (S)ARIMA che approssimi adeguatamente il processo generatore dei dati in esame, mediante la procedura di Box e Jenkins. Questa procedura consta di tre fasi fondamentali:

- a) Identificazione dei parametri del modello;
- b) Stima dei parametri;
- c) Controllo diagnostico.

Un modello SARIMA è un'estensione dei processi ARIMA per trattare un andamento periodico (stazionario o non stazionario). La sua importante funzionalità è che descrive la correlazione tra valori consecutivi e la correlazione tra osservazioni che distano tra loro di un multiplo del periodo. Box e Jenkins (1976) hanno proposto di descrivere la struttura di dipendenza tra i periodi e tra le stagioni con il modello (S)ARIMA  $(p, d, q) \times (P, D, Q)_S$  così definito

$$\phi(B)\Phi(B^S)(1-B)^d(1-B^S)^DY_t = \phi_0 + \theta(B)\Theta(B^S)\varepsilon_t,$$

dove:

- $S$  è il periodo stagionale;
- $\phi(B)$  è l'operatore autoregressivo non stagionale di ordine  $p$ ;
- $\Phi(B^S)$  è l'operatore autoregressivo stagionale di ordine  $P$ ;
- $\theta(B)$  è l'operatore a media mobile non stagionale di ordine  $q$ ;
- $\Theta(B^S)$  è l'operatore a media mobile stagionale di ordine  $Q$ ;
- $(1-B)^d$  è l'operatore differenza non stagionale di ordine  $d$ ;
- $(1-B^S)^D$  è l'operatore differenza stagionale di ordine  $D$ .

Gli strumenti utilizzati per l'identificazione dei parametri del modello, sono la funzione di autocorrelazione e la funzione di autocorrelazione parziale (stimate entrambe sul campione). Sulla base della struttura della funzione di autocorrelazione empirica si è cercato di riconoscere la struttura di una funzione di autocorrelazione teorica. Per evitare il rischio di sovrapparametrizzazione si è ricorsi al criterio di informazione di Akaike (AIC) e al criterio bayesiano di Schwarz (SBC o BIC).

Per la stima dei parametri è stato utilizzato il metodo dei minimi quadrati non lineari. Infine è stata effettuata la diagnostica del modello per verificare se il modello è stato correttamente identificato e stimato. È stata quindi analizzata l'autocorrelazione dei residui, che non dovrebbe condurre al rifiuto dell'ipotesi *white noise* dei residui, e la normalità della loro distribuzione.

Individuato il modello migliore, si è proceduto con l'identificazione degli *outliers*, ossia quei valori che non sono imputabili, o almeno non totalmente, alla normale dinamica del fenomeno. Quindi i valori anomali vengono pensati come deterministici e stimati con opportuni modelli di regressione.

Sia  $\eta_t$  una variabile esplicativa che descrive alcuni tipi di cambiamenti nella media della risposta. Un outlier additivo in certi punti  $s$  corrisponde ad un effetto di shock  $\eta_t$  tale che  $\eta_s = 1$ , e  $\eta_t = 0$  negli altri punti.

Se si sta stimando un modello ARIMA,

$$D(B)Y_t = \mu_t + \frac{\theta(B)}{\phi(B)}a_t,$$

il processo di selezione comincia considerando gli effetti di shock che si generano per ciascuna misurazione non mancante. Questo significa testare

$H_0: \beta = 0$  versus  $H_1: \beta \neq 0$  nel modello  $D(B)(Y_t - \beta\eta_s) = \mu_t + \frac{\theta(B)}{\phi(B)}a_t$ , per ciascuno di questi effetti. Sono inclusi nel modello gli effetti di shock che soddisfano il livello di significatività prefissato.

Il precedente problema di verifica è equivalente a testare  $H_0: \beta = 0$  versus  $H_1: \beta \neq 0$  nel seguente modello di regressione con errori ARMA

$$N_t = \beta\zeta_t + \frac{\theta(B)}{\phi(B)}a_t$$

dove  $N_t$  è il processo "noise",  $\zeta_t$  l'effettivo segno di shock e  $a_t$  è il processo white noise. La stima ai minimi quadrati generalizzati di  $\beta$  e la sua varianza sono date da:

$$\hat{\beta} = \delta/k, \text{Var}(\hat{\beta}) = \sigma^2/k$$

dove  $\sigma^2$  è la varianza del processo white noise,  $\delta = \zeta^T \Omega^{-1} N$  e  $k = \zeta^T \Omega^{-1} \zeta$ , con  $\Omega$  è la matrice di varianza e covarianza associata al modello ARMA.

La statistica test  $\tau^2 = \delta^2/(\sigma^2 k)$  è utilizzata per testare la significatività di  $\beta$ , che sotto  $H_0$  ha una distribuzione approssimata chi quadro con 1 grado di libertà.

### 3.3. Modelli additivi generalizzati

Per studiare la relazione tra la mortalità e i valori di humidex, al netto della stagionalità e di altre variabili socio-demografiche, vengono utilizzati i modelli additivi generalizzati perché sono metodi statistici flessibili che consentono di identificare e caratterizzare anche effetti di regressione non lineari. Questi modelli vengono qui utilizzati per stimare e tener conto della stagionalità della mortalità e poi per stimare la relazione tra la mortalità e l'humidex, con lo scopo di catturare l'effetto non lineare.

In ambito di regressione, un modello additivo ha la seguente forma:

$$E(Y|X_1, X_2, \dots, X_p) = \alpha + f_1(X_1) + f_2(X_2) + \dots + f_p(X_p)$$

dove  $Y$  è l'outcome,  $X_1, X_2, \dots, X_p$  sono i predittori e  $f_1, f_2, \dots, f_p$  sono funzioni in una variabile, ciascuna con un proprio andamento, ed  $\alpha$  è la costante.

Per far fronte a quello che è essenzialmente un problema di identificabilità del modello, è necessario che le varie  $f_i$  siano centrate sullo 0, ossia

$$\sum_{i=1}^n f_j(x_{ij}) = 0, \quad (j = 1, \dots, p),$$

dove  $x_{ij}$  è la  $j$ -esima variabile per l'unità  $i$ .

Per dati di conteggio, in questo caso specifico il numero di decessi giornalieri, la distribuzione appropriata è quella di Poisson e la media condizionata della variabile risposta è quindi messa in relazione ai predittori mediante la funzione legame  $g(\cdot) = \log(\cdot)$ .

Dunque il modello acquisisce la seguente forma:

$$Y_t \sim \text{Poisson}(\mu_t)$$

$$\mu_Y(X) = E(Y|X_1, X_2, \dots, X_p)$$

$$\log[\mu_Y(X)] = \alpha + f_1(X_1) + f_2(X_2) + \dots + f_p(X_p).$$

Le funzioni  $f_i$  sono stimate in modo flessibile, utilizzando un algoritmo la cui struttura di base è uno scatterplot smoother (ad esempio le spline cubiche di lisciamento). Per i modelli additivi tale algoritmo è quello di *Backfitting*, invece quello generico per i modelli additivi generalizzati è l'algoritmo *Local Scoring* (Hastie, et al. 2009; Azzalini & Scarpa, 2012).

Un modello additivo ha la forma  $Y = \alpha + \sum_{j=1}^p f_j(X_j) + \varepsilon$ , dove il termine di errore  $\varepsilon$  ha media zero. Date le osservazioni  $x_i, y_i$ , un criterio analogo a quello della somma dei quadrati penalizzati (PRSS, *Penalized Residual Sum of Squares*) può essere indicato per questo problema,

$$PRSS(\alpha, f_1, f_2, \dots, f_p) = \sum_{i=1}^N \left( y_i - \alpha - \sum_{j=1}^p f_j(x_{ij}) \right)^2 + \sum_{j=1}^p \lambda_j \int f_j''(t_j)^2 dt_j$$



dove il primo termine misura la vicinanza ai dati e il secondo regola la curvatura della funzione. Il valore  $\lambda/(1 + \lambda)$  è il parametro di smoothing. Quando  $\lambda \rightarrow \infty$ , il parametro di liscio è vicino ad 1 e forma una curva più liscia, invece quando  $\lambda \rightarrow 0$ , il parametro di liscio è vicino a 0 e forma una curva più ruvida, ma che interpola meglio i dati.

È stato dimostrato matematicamente che la soluzione al problema di minimizzazione è rappresentata da una spline cubica naturale, i cui nodi sono i valori distinti di  $x$ :

$$\hat{f}(x) = \sum_{j=1}^{n_0} \theta_j N_j(x)$$

dove  $n_0$  è il numero degli  $x_i$  distinti (senza ripetizioni), e  $N_j(x)$  sono le funzioni spline cubiche naturali di base.<sup>5</sup>

L'algoritmo di backfitting è un processo iterativo basato su un metodo di stima non parametrico di funzioni in una variabile, che consente di stimare le funzioni  $f_j$ , e si compone dei seguenti passaggi:

1. Inizializzazione:  $\hat{\alpha} = \frac{1}{N} \sum_1^N y_i, \hat{f}_j \equiv 0, \forall i, j.$
2. Ciclo:  $j = 1, 2, \dots, p, \dots, 1, 2, \dots, p, \dots,$

$$\hat{f}_j \leftarrow S_j \left[ \{y_i - \hat{\alpha} - \sum_{k \neq j} \hat{f}_k(x_{ik})\}_1^N \right],$$

$$\hat{f}_j \leftarrow \hat{f}_j - \frac{1}{N} \sum_{i=1}^N \hat{f}_j(x_{ij}).$$

Finché le stime  $\hat{f}_j$  non si stabilizzano.

Si pone  $\hat{\alpha} = \frac{1}{N} \sum_1^N y_i$ , si applica una spline cubica  $S_j$  ad  $\{y_i - \hat{\alpha} - \sum_{k \neq j} \hat{f}_k(x_{ik})\}_1^N$ , come funzione di  $x_{ij}$ , per ottenere una nuova stima  $\hat{f}_j$ . Questo viene fatto a turno per ogni predittore, utilizzando la stima corrente delle altre funzioni  $\hat{f}_k$  nel calcolo  $y_i - \hat{\alpha} - \sum_{k \neq j} \hat{f}_k(x_{ik})$ . Il processo continua finché le stime non si stabilizzano.

---

<sup>5</sup> Il termine spline viene utilizzato in matematica per la costruzione di funzioni polinomiali a tratti, con lo scopo di approssimare nel modo più opportuno la funzione di interesse. Il principio di fondo è che a partire da una serie di punti scelti sull'asse  $x$  (chiamati nodi) si può costruire una funzione che passi esattamente attraverso i nodi ma sia libera negli altri punti

Per i modelli additivi generalizzati, il criterio appropriato è una log-verosimiglianza penalizzata. Per massimizzarla viene utilizzata la procedura di backfitting congiuntamente ad un massimizzatore della verosimiglianza. L'algoritmo generico per i modelli additivi generalizzati è chiamato Local Scoring; di seguito viene definita la sua struttura per modelli di Poisson (Hastie T, Tibshirani R. J, 1990):

1. Definizione valori di partenza:  $\hat{\alpha} = \log(\bar{y})$ , dove  $\bar{y} = \text{media}(y_i)$ ,  $\hat{f}_j \equiv 0 \forall j$ .
2. Definizione del predittore  $\hat{\eta}_i = \hat{\alpha} + \sum_j \hat{f}_j(x_{ij})$ .

Iterazione:

- a) Costruzione della variabile dipendente corretta

$$z_i = \hat{\eta}_i + \frac{(y_i - \mu)}{\mu}.$$

- b) Costruzione dei pesi  $w_i = \mu_i$ .
- c) Adattamento di un modello additivo alle  $z_i$  mediante i pesi  $w_i$ , utilizzando l'algoritmo di backfitting pesato. Questo da nuove stime  $\hat{\alpha}_j, \hat{f}_j \forall j$ .

3. Si continua il passo 2 fino a raggiungere la stabilità della stima.

Il criterio di convergenza utilizzato di default dalla procedura SAS per questi modelli

$$\text{è: } \frac{\sum_{i=1}^N w_i \sum_{j=1}^p (\hat{f}_j^{m-1}(x_{ij}) - \hat{f}_j^m(x_{ij}))^2}{\sum_{i=1}^N w_i (1 + \sum_{j=1}^p (\hat{f}_j^{m-1}(x_{ij}))^2)} \leq 10^{-8}$$

Per la selezione del parametro di lisciamento viene utilizzato il metodo della convalida incrociata generalizzata.

I modelli che saranno stimati, sono modelli semi-parametrici e hanno la seguente forma:

$$\log\{E[Y|X = x]\} = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + s(x_2).$$

Delle variabili per cui è richiesta una stima non parametrica, viene anche stimato l'effetto lineare che sarà quindi riportato nella parte parametrica del modello. L'effetto non lineare viene rappresentato graficamente, e vengono specificati il valore del parametro di lisciamento e i gradi di libertà utilizzati. È possibile inoltre calcolare l'effetto complessivo della variabile: indicata con P la predizione parziale della variabile (ossia la componente non lineare), l'effetto complessivo è dato da:  $P_{x_2} + \beta_2 * x_2$ .

Le scelte specifiche in merito ai modelli utilizzati saranno fornite in seguito con i risultati.

Come già stato detto al paragrafo precedente, essendo la variabile risposta il numero dei decessi, il modello scelto è quello di Poisson con funzione legame il logaritmo. La forma generale del modello di Poisson è la seguente:

$$\log\{E[Y|X = x]\} = \beta_0 + \beta_1 x_1 + \beta_2 x_2.$$

Il coefficiente di regressione  $\beta_k$  misura la variazione del logaritmo del valore atteso per una variazione unitaria della covariata k-sima, costanti i restanti regressori.

Per variabili di tipo dicotomico si confrontano i valori attesi condizionati per  $x_{i1} = 1$  ed  $x_{i1} = 0$  con riferimento ad un modello che contiene anche  $x_2$ , quindi si ottiene:

$$\frac{E(y_i|x_{i1} = 1, x_{i2})}{E(y_i|x_{i1} = 0, x_{i2})} = \frac{\exp(\beta_0 + \beta_1(1) + \beta_2 x_{i2})}{\exp(\beta_0 + \beta_1(0) + \beta_2 x_{i2})} = \exp(\beta_1).$$

L'insieme delle unità per cui  $x_{i1} = 1$  sono caratterizzate da un valore atteso condizionale che è  $\exp(\beta_1)$  volte quello delle unità per cui  $x_{i1} = 0$ , al netto delle altre variabili esplicative.  $\exp(\beta_1)$  può essere interpretato come rischio relativo, ossia il rischio di sperimentare l'evento in studio Y per la sottopopolazione definita da  $x_{i1} = 1$ , rispetto al rischio della sottopopolazione di riferimento per cui  $x_{i1} = 0$  (Rothman & Greenland, 1998).

Lo stesso ragionamento vale anche per una variabile categoriale, prendendo come riferimento una delle categorie. Le unità che appartengono alla categoria di riferimento saranno quelle per cui  $x_{i1} = 0$  e quindi rispetto a questa sarà calcolato l'effetto di ognuna delle altre categorie.



## 4. RICERCA DEGLI ECCESSI DI MORTALITÀ

---

In questo capitolo l'obiettivo è descrivere l'andamento della mortalità dal 2000 al 2013 in Veneto e verificare se ci sono stati degli eccessi di mortalità.

Nel primo paragrafo si esamina l'andamento della mortalità nel Veneto nei quattordici anni mediante il calcolo di tassi di mortalità annuali. Inoltre vengono anche calcolati i tassi di mortalità relativamente al periodo 21/05 – 14/09 di ogni anno per verificare se il trend è in linea con quello annuale e se si osservano dei picchi interessanti. Questo focus sul periodo estivo è di interesse perché successivamente l'obiettivo sarà capire se c'è una qualche relazione con le condizioni climatiche verificatesi nello stesso periodo.

Il secondo paragrafo si addentra nella serie storica della mortalità con l'obiettivo di studiare la componente stagionale che la caratterizza e rilevare se esistono differenze per classe d'età. La stagionalità è una componente importante che deve essere considerata per qualsiasi analisi sulla mortalità. Segue poi la destagionalizzazione della serie dei decessi con lo scopo di individuare se e quando si siano verificati degli eccessi di mortalità nei quattordici anni in esame, valutando la significatività statistica mediante modelli e test statistici opportuni.

Nel terzo paragrafo, viene depurata la serie dalla componente stagionale per evidenziare più chiaramente i possibili picchi anomali di mortalità. Nel quarto paragrafo si applica una modellazione ARIMA per ricercare il modello che meglio spiega la serie della mortalità in esame, con l'obiettivo finale di identificare i picchi anomali della serie, ossia gli eccessi di mortalità non spiegati dal modello.

L'ultimo paragrafo del capitolo è rivolto alla relazione tra la mortalità e l'humidex specificamente per estate e area climatica; si mettono a confronto le due serie trasformate mediante media mobile permettendo di fare alcune prime osservazioni per vedere se i picchi di mortalità riscontrati al paragrafo precedente sono avvenuti in periodi particolari per condizioni climatiche.

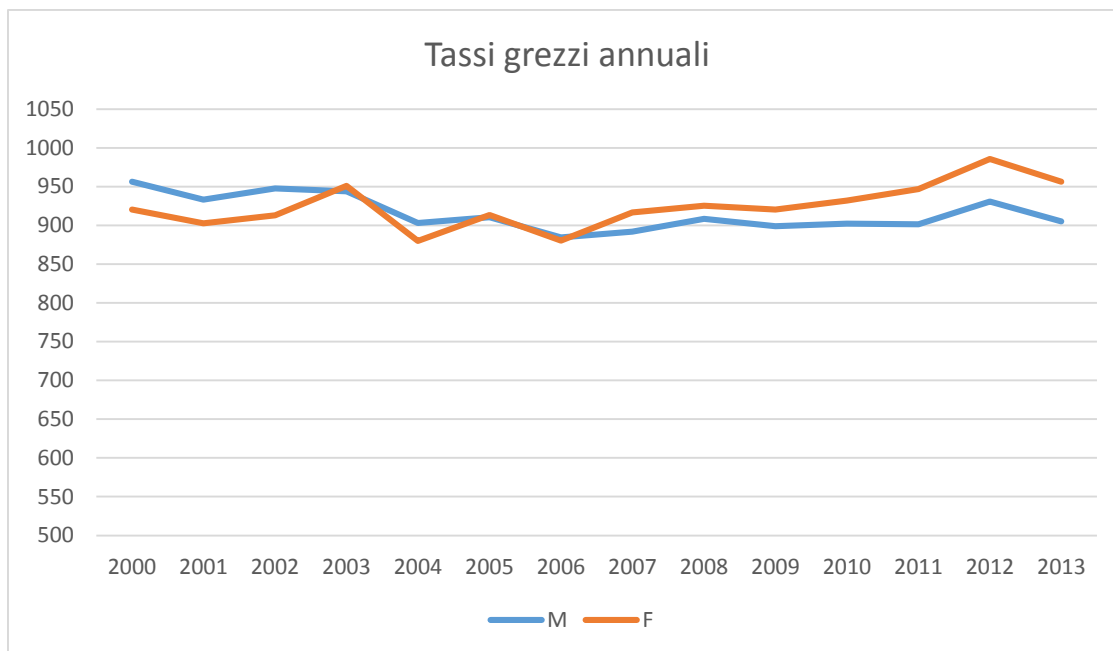
#### 4.1 Andamento della mortalità nel Veneto dal 2000 al 2013

Le figure 4.1 e 4.2 rappresentano l'andamento dei tassi di mortalità dal 2000 al 2013 in Veneto e la tabella 4.1 ne riporta i valori corrispondenti. Si nota, dai tassi standardizzati, un calo della mortalità evidente fino al 2006, soprattutto per i maschi. I tassi grezzi danno invece indicazione della mortalità reale, osservata, e mostrano una mortalità più elevata per le femmine soprattutto negli ultimi anni. Questo è dovuto alla diversa struttura per età, infatti la popolazione femminile è tendenzialmente più vecchia rispetto a quella maschile.

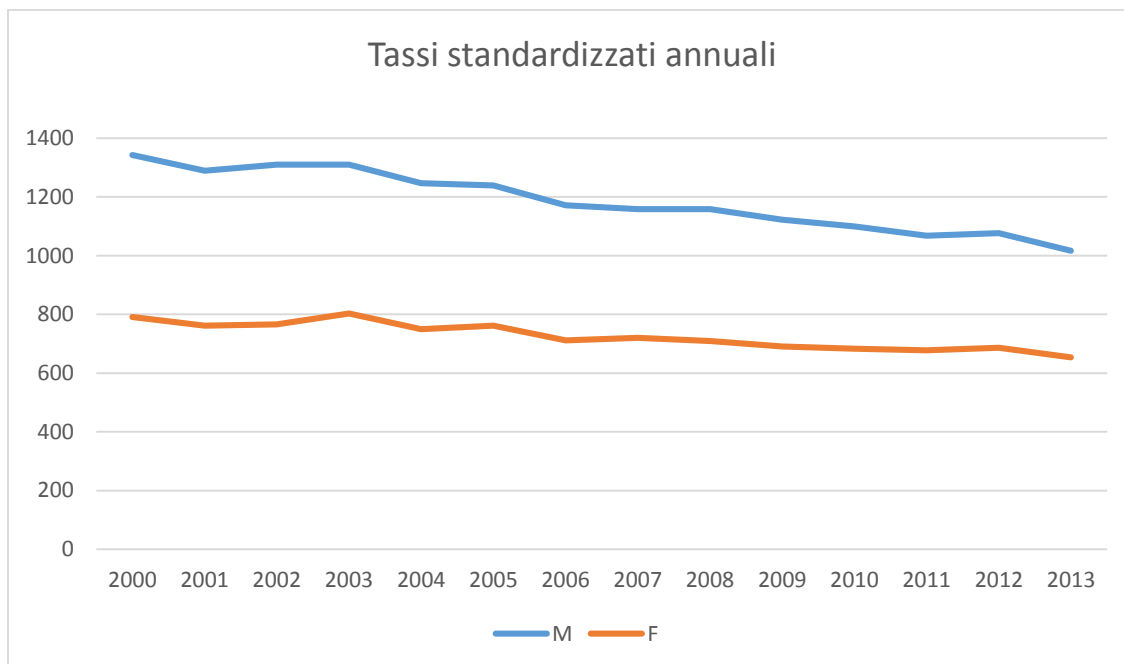
**Tabella 4.1.** Mortalità nel Veneto dal 2000 al 2013: numero di decessi (dei residenti) per sesso, tasso grezzo di mortalità e tasso standardizzato, per 100000 (standardizzazione diretta, pop. Standard: Veneto 1° gennaio 2007). Fonte dei dati: SER.

anno	MASCHI			FEMMINE		
	N	T. grezzo	T. std	N	T. grezzo	T. std
2000	20917	956,3	1342,5	21253	920,2	791,0
2001	20523	933,0	1288,8	20933	902,6	761,2
2002	20990	947,7	1310,2	21277	912,9	765,8
2003	21128	943,7	1309,7	22368	951,1	803,2
2004	20477	902,9	1247,2	20907	879,8	749,1
2005	20849	910,3	1238,7	21881	913,5	761,7
2006	20383	884,4	1171,1	21222	880,4	711,8
2007	20733	892,0	1158,7	22290	916,7	720,6
2008	21329	908,4	1158,7	22740	925,4	708,9
2009	21219	898,9	1122,0	22771	920,4	690,2
2010	21322	902,2	1099,3	23150	932,1	682,7
2011	21306	901,2	1067,9	23580	946,8	677,9
2012	22068	930,6	1076,2	24619	985,5	685,9
2013	21634	905,2	1016,8	24044	956,3	653,4

**Figura 4.1.** Tassi grezzi di mortalità (per tutte le cause) per sesso (per 100.000). Veneto, anni 2000-2013.



**Figura 4.2.** Tassi standardizzati di mortalità (dei residenti) per sesso (per 100000). Veneto, anni 2000-2013. Standardizzazione diretta, pop. standard: Veneto 1° gennaio 2007.

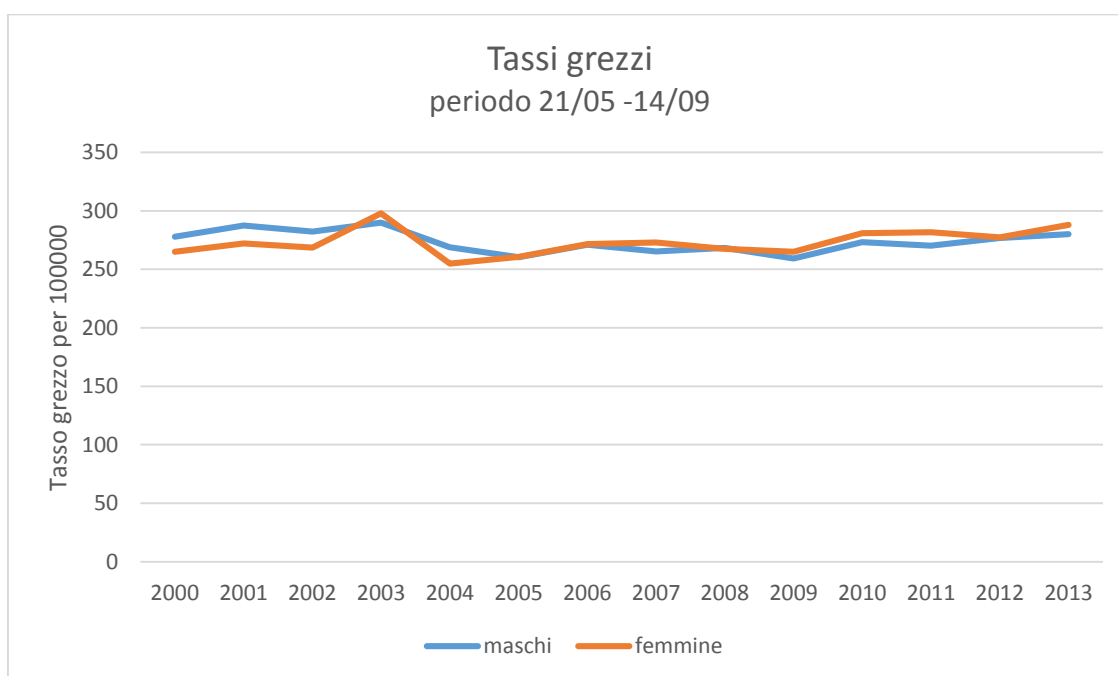


Le figure 4.3 e 4.4 mostrano invece l'andamento dei tassi di mortalità, grezzi e standardizzati, per il periodo estivo su cui si focalizzano le elaborazioni successive, che va dal 21 maggio al 14 settembre di ogni anno dal 2000 al 2013. Ciò che si nota è un andamento di trend simile a quello annuale, ed inoltre si nota una mortalità più elevata

nell'estate del 2003. Sempre in riferimento a questo periodo dell'anno, si mostrano anche i tassi specifici di mortalità per classe d'età (0-39, 40-64, 65-79, 80 e oltre) e relativi intervalli di confidenza (figure 4.5 – 4.8). Da questi si può notare che il calo della mortalità è più marcato per le classi 0-39 e 40-64 e, come osservato prima, soprattutto per i maschi. Inoltre si può vedere che il picco di mortalità del 2003 si distingue chiaramente solo per la classe degli over 80 e si potrebbe quindi pensare che durante l'estate del 2003 sia accaduto qualcosa che ha colpito maggiormente gli anziani.

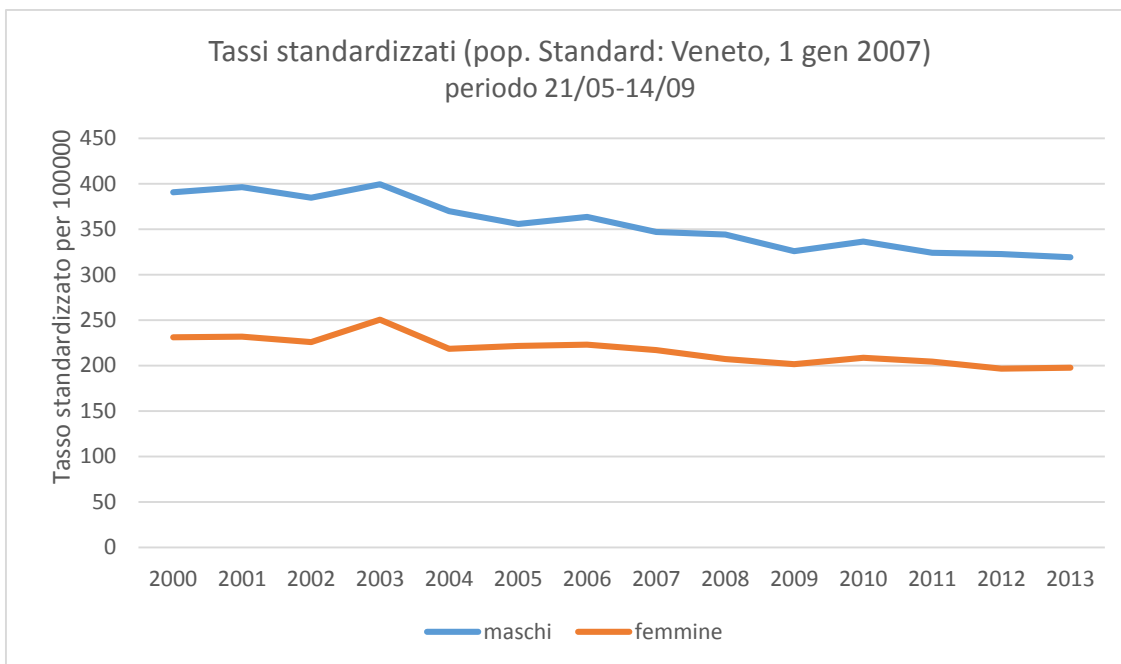
Per i valori dei tassi delle figure 4.5 – 4.8 si veda l'appendice B.

**Figura 4.3.** Tassi grezzi di mortalità (per tutte le cause) per sesso (per 100.000). Veneto, periodo 21/05-14/09, anni 2000-2013.

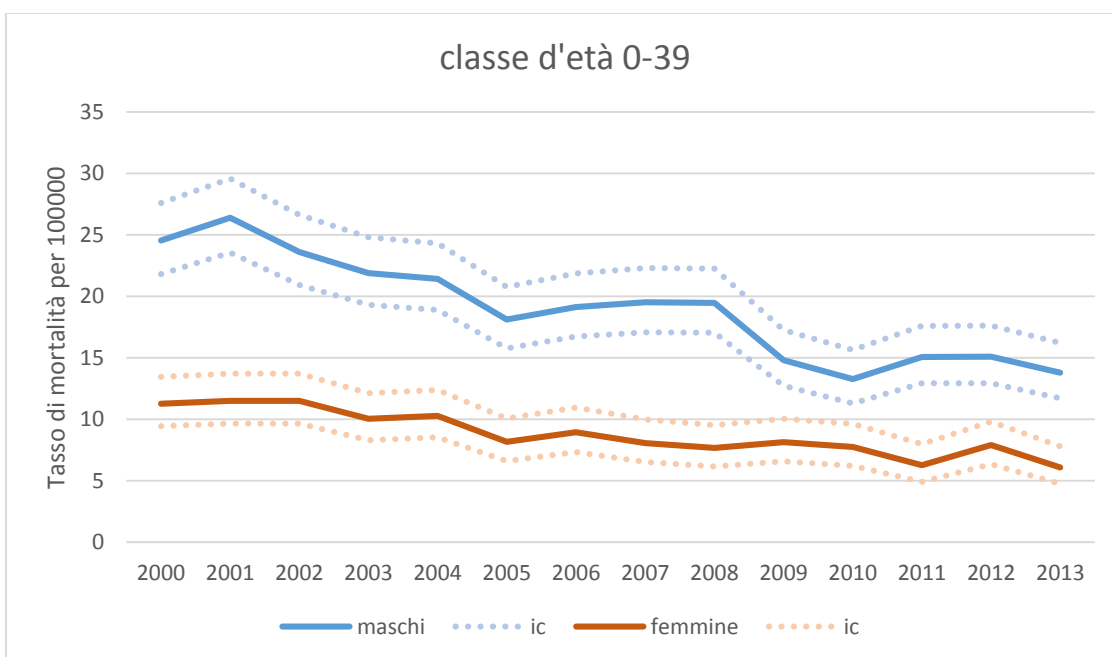




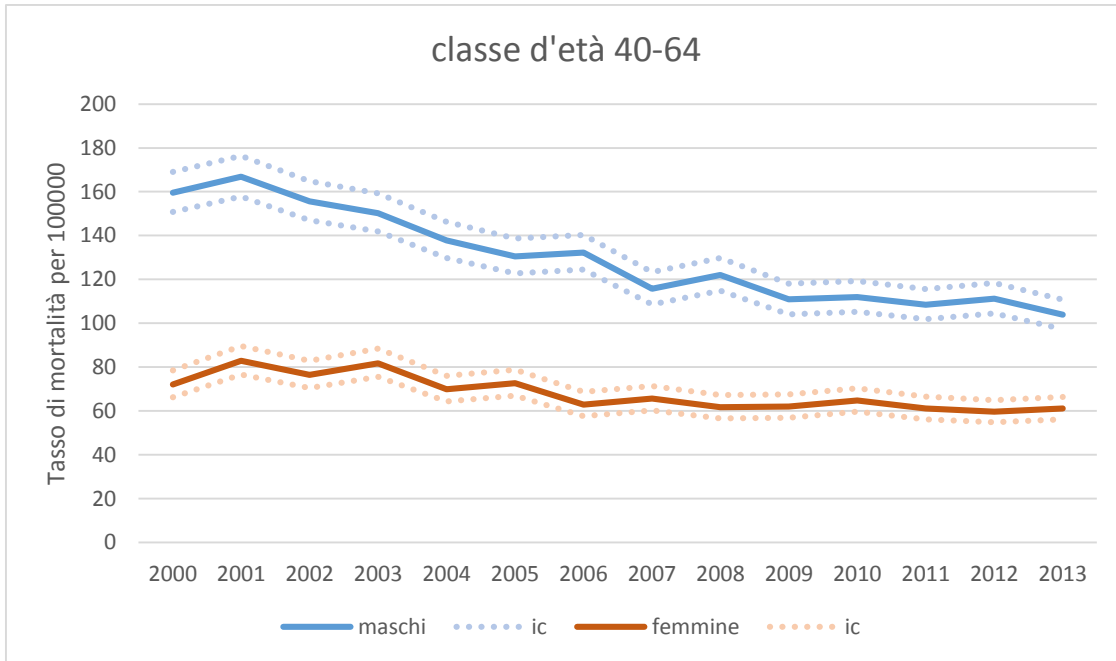
**Figura 4.4.** Tassi standardizzati di mortalità per sesso (per 100000). Veneto, periodo 21/05-14/09, anni 2000-2013. Standardizzazione diretta, pop. standard: Veneto 1° gennaio 2007.



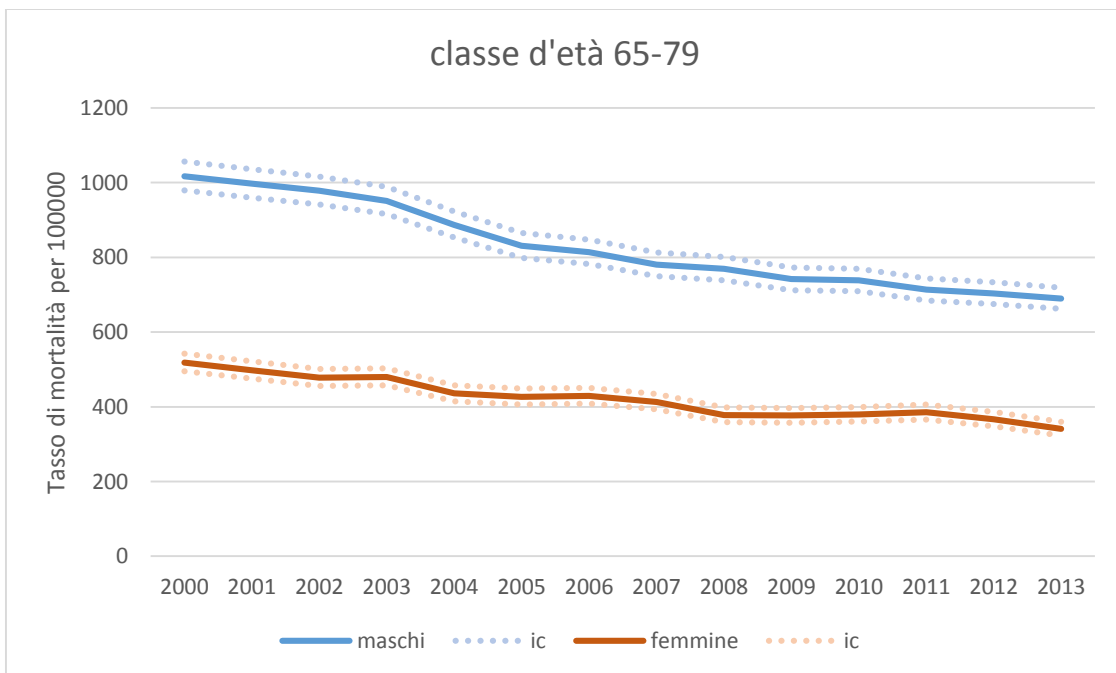
**Figura 4.5.** Tassi specifici di mortalità (con relativi intervalli di confidenza, in tratteggio), classe d'età 0-39, per sesso (per 100000). Veneto, periodo 21/05-14/09, anni 2000-2013.



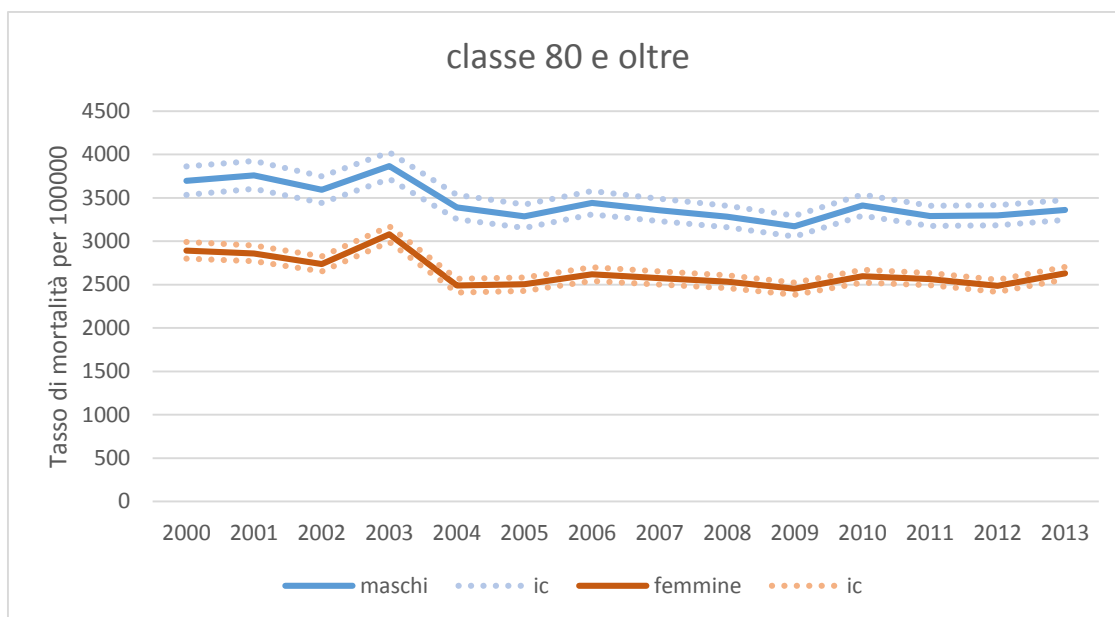
**Figura 4.6.** Tassi specifici di mortalità (con relativi intervalli di confidenza, in tratteggio), classe d'età 40-64, per sesso (per 100000). Veneto, periodo 21/05-14/09, anni 2000-2013



**Figura 4.7.** Tassi specifici di mortalità (con relativi intervalli di confidenza, in tratteggio), classe d'età 65-79, per sesso (per 100000). Veneto, periodo 21/05-14/09, anni 2000-2013.



**Figura 4.8.** Tassi specifici di mortalità (con relativi intervalli di confidenza, in tratteggio) per maschi e femmine di 80 o più anni (per 100000). Veneto, periodo 21/05-14/09, anni 2000-2013.



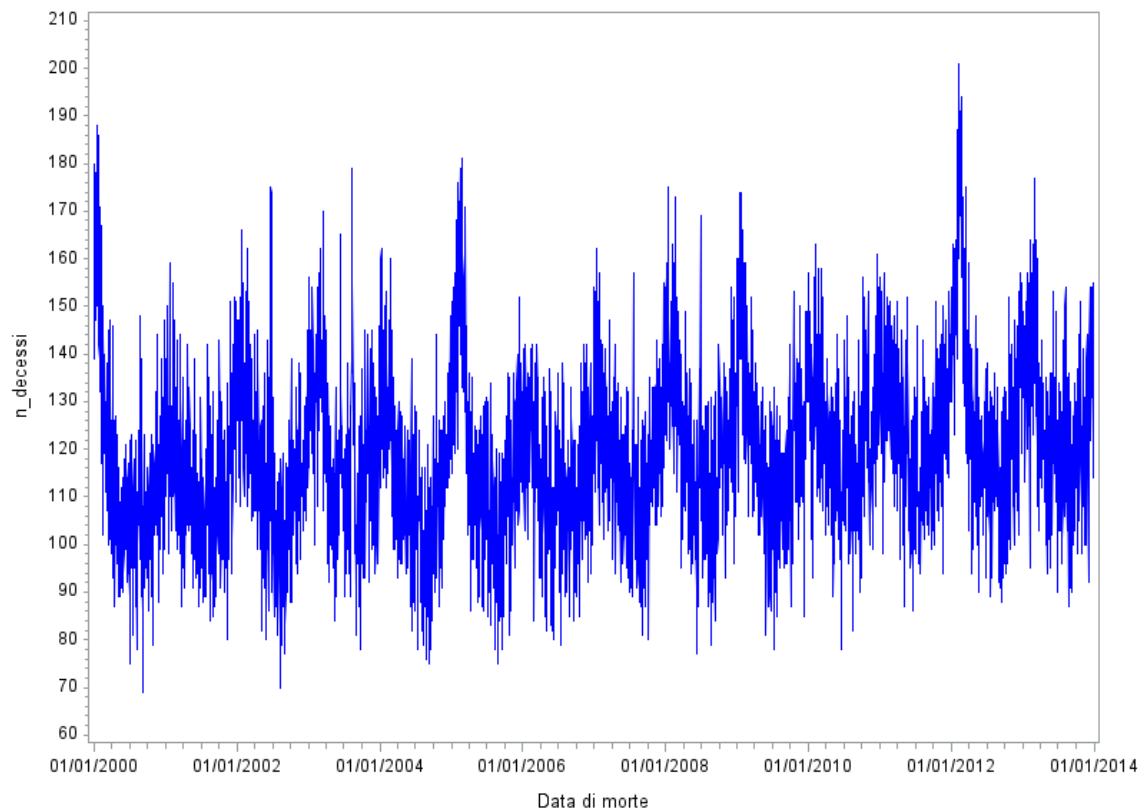
## 4.2 La mortalità e la componente stagionale

Si volge qui l'attenzione sulla serie dei decessi giornalieri, per valutarne l'andamento dal 2000 al 2013 e identificare le componenti che la caratterizzano. In modo particolare viene approfondita la stagionalità della serie, ricercando le eventuali differenze nell'andamento tra le diverse classi d'età.

Le figure 4.9 e 4.10 mostrano rispettivamente la serie originale dei decessi giornalieri, e la serie trasformata mediante media mobile di ordine 7. Come primo strumento di liscio della serie, è stata applicata una media mobile centrata di ordine 7 in quanto si tratta di una serie giornaliera.

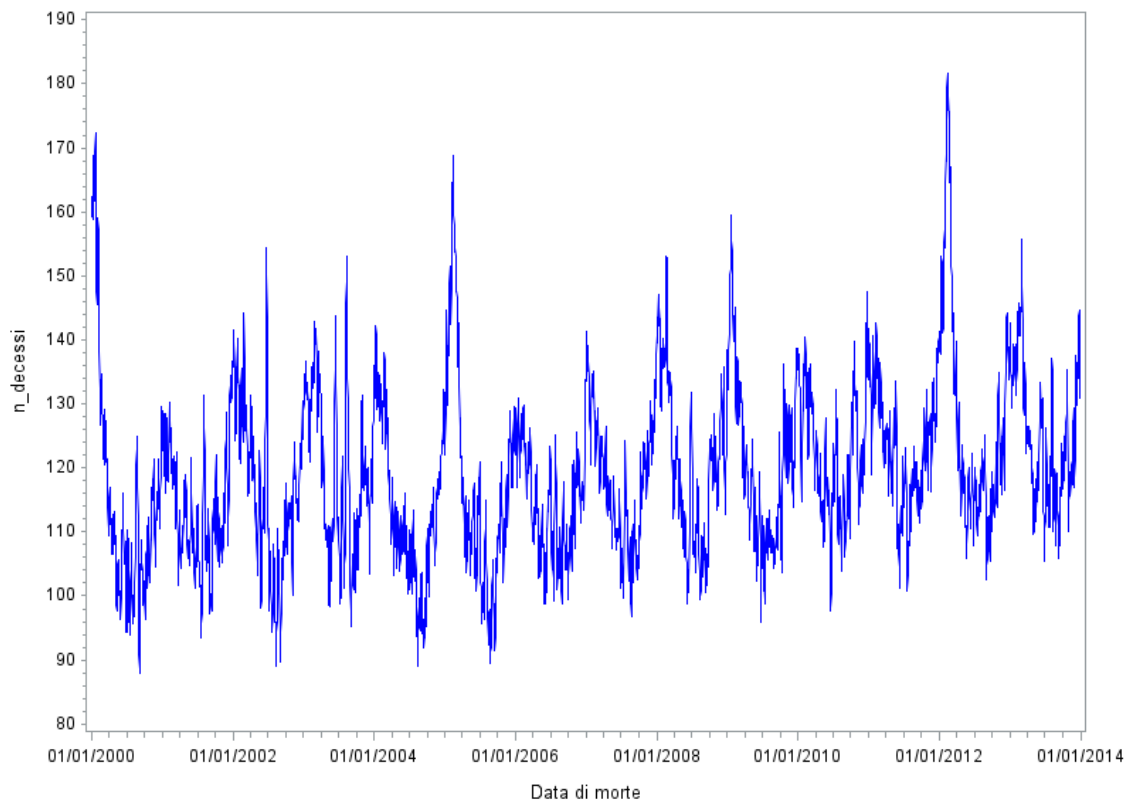
**Figura 4.9.** Serie grezza dei decessi giornalieri (per tutte le cause) del Veneto dal 2000 al 2013.

### Decessi giornalieri dal 2000 al 2013



**Figura 4.10.** Serie dei decessi del Veneto (per tutte le cause), media mobile 7 giorni

**Decessi giornalieri dal 2000 al 2013 (media mobile 7gg)**

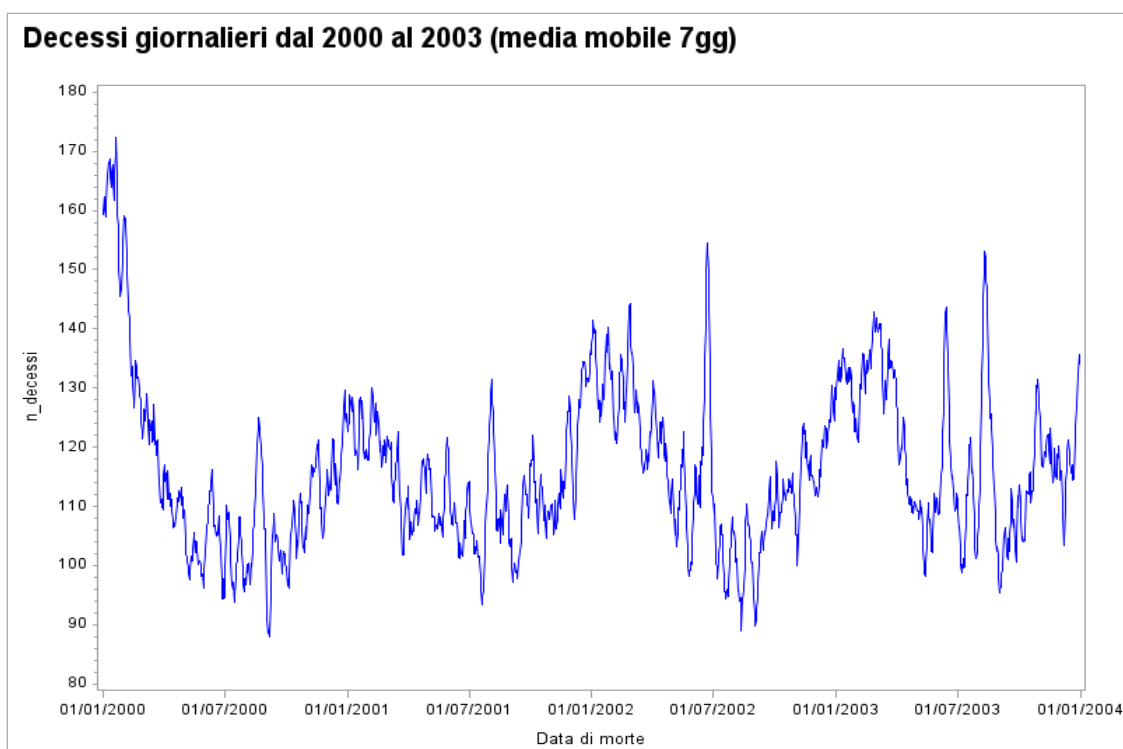


Da entrambi i grafici si riconosce chiaramente la componente stagionale della mortalità che raggiunge il suo massimo durante l'inverno e il minimo durante l'estate, assumendo così una forma ad U. Nelle località che presentano un clima caratterizzato da periodi stagionali, ricorre questa forma della mortalità, significativamente più elevata in inverno che in estate (Davis et al, 2004; Saez et al. 1995; Martin-Soler et al. 2014; Kalkstein A. J, 2013). In uno studio in cui sono stati analizzati dati di diverse nazioni (Australia, Brasile, Canada, Cina, Italia, Giappone, Corea del Sud, Spagna, Svezia, Taiwan, Thailandia, Uk e USA) è emerso che la maggior parte della mortalità collegata alla temperatura è attribuibile agli effetti del freddo (Gasparrini et al., 2015).

Alcuni studi epidemiologici hanno constatato che un fattore dell'eccesso di mortalità in inverno è l'influenza, la cui durata, e severità, possono variare di anno in anno (Davis et al. 2012; Nielsen et al, 2011; Klot et al 2012; Analitis et al. 2008). La ricerca dei possibili fattori di rischio della mortalità in inverno è comunque ancora aperta, con l'obiettivo di riuscire a prevenire le cause e quindi a ridurla.

Dalla figura 4.10 si possono riconoscere dei possibili picchi anomali, ad esempio nell'estate del 2002 e del 2003 (dettaglio in figura 4.11), e nell'inverno del 2005 e del 2012. Tuttavia queste sono solo alcune prime osservazioni, non ancora confermate da alcuna statistica. In figura 4.10 sembra esserci anche un leggero trend in salita negli ultimi anni, ma, visti i tassi di mortalità standardizzati, si può dire che questo trend sia dovuto all'invecchiamento della popolazione (i tassi standardizzati infatti mostravano un calo della mortalità non un incremento).

**Figura 4. 11.** Decessi giornalieri, media mobile 7 giorni. Veneto, anni 2000-2003.



La stagionalità è quindi parte integrante della serie e per raggiungere gli obiettivi prefissati il suo effetto dovrà essere considerato. Per questo motivo si vuole ora studiarla più a fondo per capire se il suo effetto è lo stesso per tutte le classi d'età, oppure no, e se esiste una stagionalità della mortalità anche solo nel periodo estivo.

Sono stati quindi utilizzati i modelli additivi generalizzati con l'utilizzo di spline di lisciamento per l'effetto della stagionalità ed è stato stimato un modello per ciascuna classe d'età per vedere se sussistono delle differenze (Wagner et al. 2013).

Ciascun modello comprende una variabile esplicativa per la stagionalità e una per il trend:

$$Y_t (= \text{decessi giorno } t) \sim \text{Poisson} (\mu_t)$$

$\log(\mu_{ts}^c) = \text{sexo} + \text{trend} + f(pt)$ , dove:

- $t$  è il giorno;
- $s$  è il sesso;
- $c$  è la classe d'età (0-14, 15-44, 45-64, 65-74, 75-84, 85+);
- trend è la variabile tempo che assume valore 1 il primo giorno della serie, 2 il secondo e così via. È la variabile che rappresenta il trend;
- $pt$  è la variabile che rappresenta il tempo intra-annuale, quindi serve per la stagionalità.  $Pt$  assume valore 0 l'1 gennaio e incrementa il suo valore di  $1/365$  (o  $1/366$ ) di giorno in giorno;
- $f$  è la funzione, stimata mediante spline di lisciamento, per la variabile  $pt$ ;

Per tener conto delle variazioni che ci possono essere nella numerosità delle coorti nel tempo è stato incluso nel modello il numero di individui della coorte (la popolazione residente a metà anno, di classe d'età  $c$  e sesso  $s$ ) come offset (in logaritmo).

Dal modello viene stimato sia l'effetto lineare della variabile stagionalità sia l'effetto non lineare (stimato mediante l'algoritmo Local Scoring, sulla base di spline di lisciamento). Nei grafici riportati in figura 4.12 si mostra l'effetto complessivo della variabile, ossia l'andamento della stagionalità annuale della mortalità, che è ottenuto dalla somma del termine lineare stimato con la predizione parziale (la componente non lineare).

Il parametro di lisciamento, e i gradi di libertà, sono stati scelti mediante convalida incrociata generalizzata. Anche se i gradi di libertà selezionati potrebbero essere tanti, sono stati comunque mantenuti perché qui l'obiettivo è individuare la forma che più approssima l'andamento della stagionalità annuale.

In tabella 4.2 si riportano le stime dei parametri ottenute da ciascun modello.

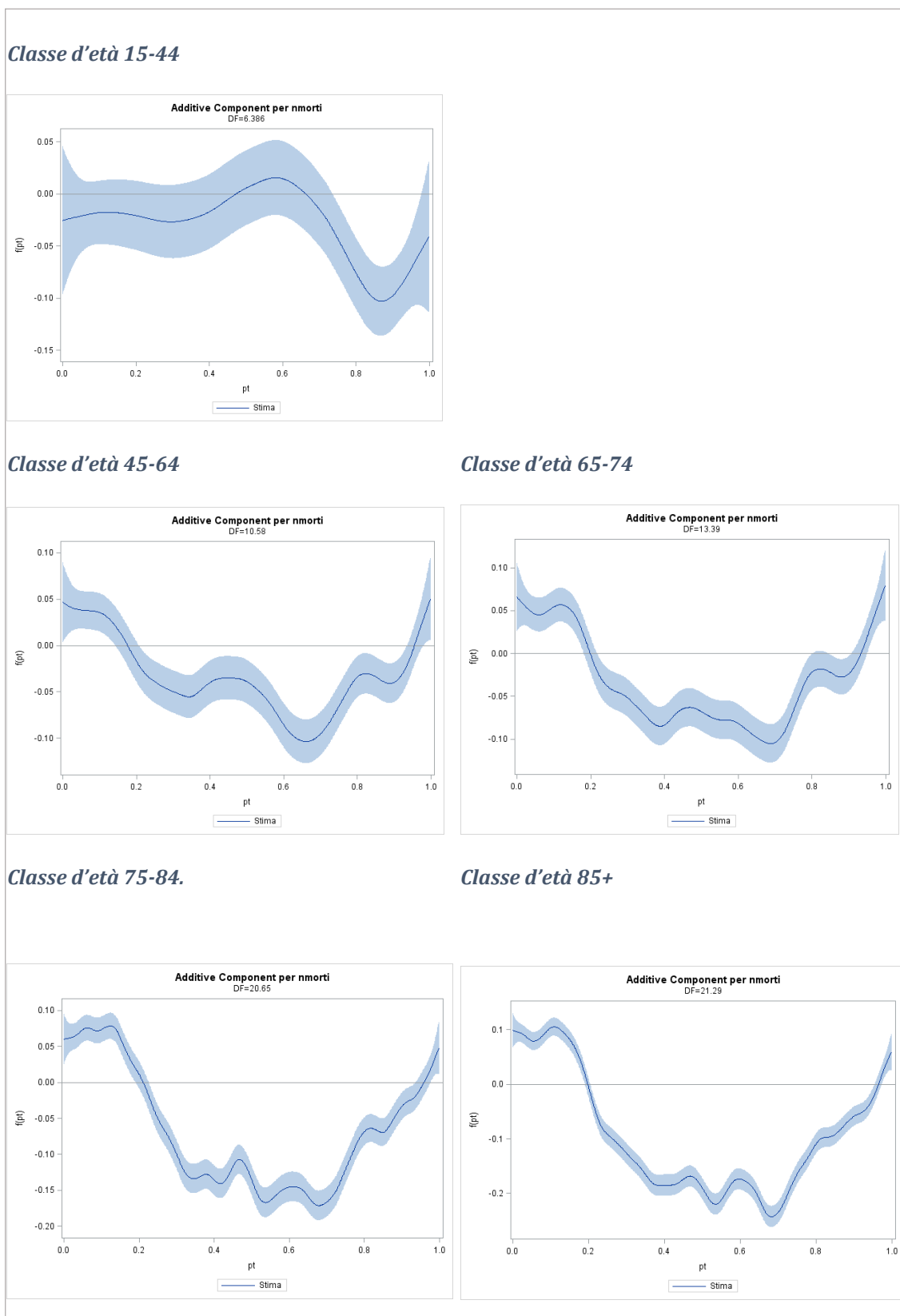
Emerge dunque una stagionalità significativa per tutte le classi d'età, a meno della classe 0-14, ma con alcune differenze nella forma della relazione. Per la classe 15-44 si nota un minimo in novembre, mentre prima non sembra esserci alcun effetto significativo. Per le classi 45-64, 65-74, 75-84, 85+ si nota una forma ad U simile con massimo assoluto in inverno e minimo in estate, e massimi locali in estate, via via più netti all'aumentare dell'età.

Concentrando l'attenzione sul periodo estivo (pt compreso tra 0.4 e 0.7, ossia da fine maggio a metà settembre): per la classe 45-64 si osserva il massimo locale all'inizio di giugno e poi scende; per la classe 65-74 si vede un aumento lieve da giugno che si mantiene fino a metà agosto quando comincia a diminuire; per la classe 75-84 sembra esserci un picco a fine giugno e uno più moderato ad agosto; infine per gli over 85 si notano gli stessi picchi della classe d'età precedente, ma più marcati.

Quindi, dopo quest'analisi sulla stagionalità, si conclude che c'è un effetto stagionale della mortalità soprattutto per le persone più anziane. In riferimento al periodo estivo sembrano distinguersi due picchi, uno a fine giugno e uno ad agosto.



**Figura 4. 12.** Effetto complessivo della variabile *pt* sulla mortalità, ossia andamento della mortalità nell'anno, per le classi d'età 15-44, 45-64, 65-74, 75-84 e 85+ (per la classe 0-14 non è risultato esserci una stagionalità significativa). In ascissa la variabile *pt* (0 l'1 gen, 1 il 31/12); in ordinata l'effetto.



**Tabella 4.2.** Stime dei parametri, modello GAM:  $\log(\mu_{ts}^c) = \text{sexo} + \text{trend} + f(\text{pt})$ .  
 \*\*\* significativo all'1%, \*\* significativo al 5%, \*significativo al 10%.

	<b>0-14</b>	<b>15-44</b>	<b>45-64</b>	<b>65-74</b>	<b>75-84</b>	<b>85+</b>
<b>Intercetta</b>	-12,40***	-13,09***	-11,58***	-10,28***	-9,12***	-7,78***
<b>trend</b>	-0,00006***	-0,00007***	-0,00009***	-0,00009***	-0,00005***	-0,00002***
<b>Maschi vs Femmine</b>	-0,02	0,40***	0,64***	0,76***	0,58***	0,24***
<b>Linear(pt)</b>	-0,001	-0,056**	-0,063***	-0,059***	-0,123***	-0,169***

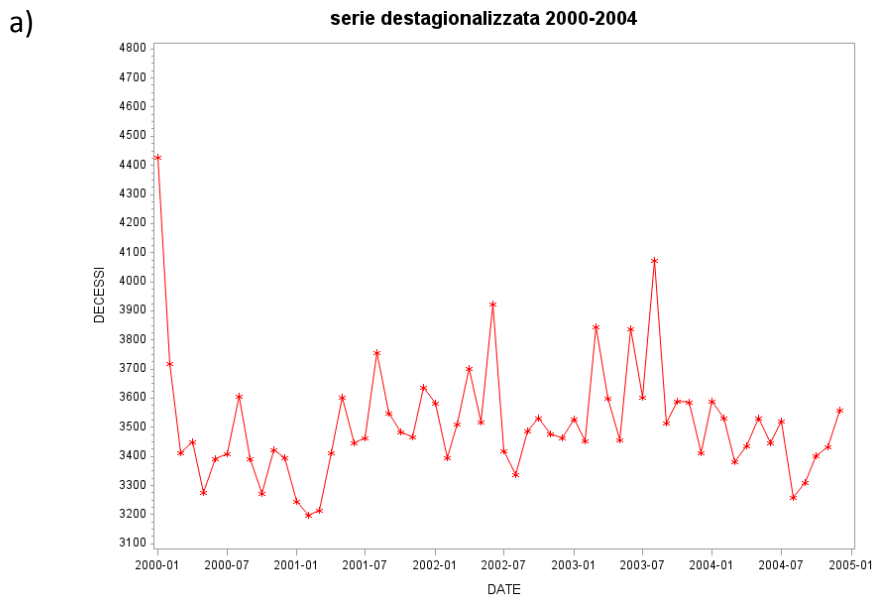
### 4.3 Andamento della mortalità al netto della stagionalità

Studiata la componente stagionale della mortalità, è di interesse depurare la serie da essa per evidenziare più chiaramente gli eventuali picchi anomali di mortalità.

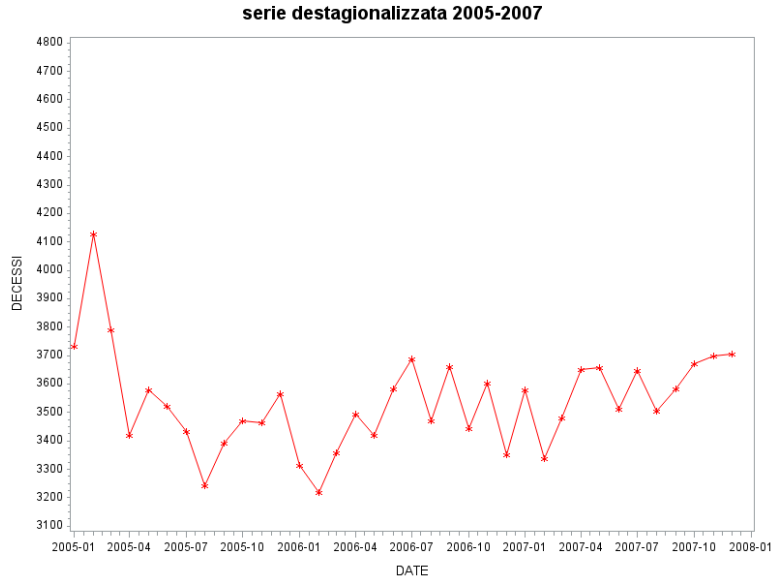
È stata quindi operata una destagionalizzazione, assumendo un modello additivo del tipo  $Y_t = T + S + I$ , dove T rappresenta il trend, S la componente stagionale e I quella irregolare. Si è ritenuto conveniente utilizzare per la destagionalizzazione la serie mensile anziché giornaliera, e quindi sono stati aggregati i decessi per mese. La serie destagionalizzata è riportata nelle figure 4.13 (a-d).

Si può notare che, come atteso, la componente stagionale non compare più, ed emergono in modo particolare alcuni possibili eccessi di mortalità in alcuni mesi: gennaio 2000, luglio 2002, marzo, giugno e agosto 2003, febbraio 2005, febbraio 2012. Nel paragrafo successivo l'obiettivo sarà quello di capire quali picchi anomali si possono ritenere statisticamente significativi.

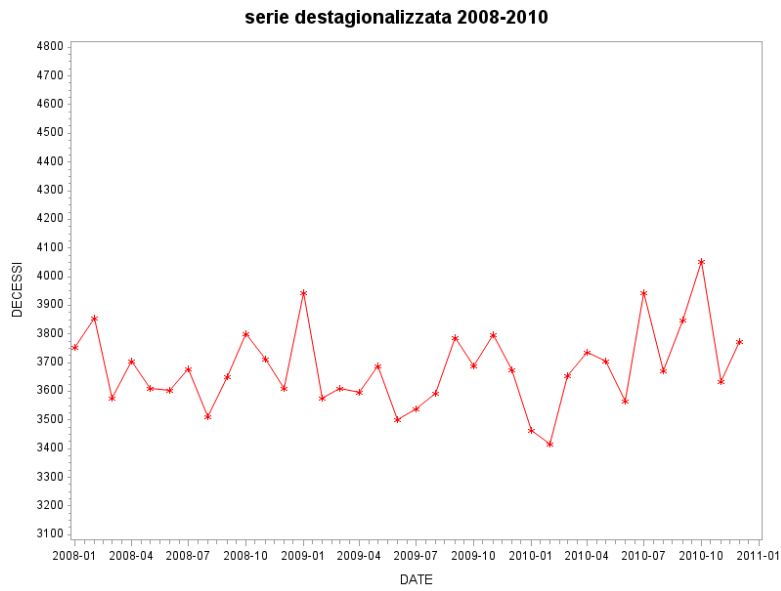
**Figure 4.13.** Serie mensile della mortalità destagionalizzata, a) anni 2000 - 2004, b) anni 2005 - 2007, c) anni 2008 - 2010, d) 2011 - 2013.



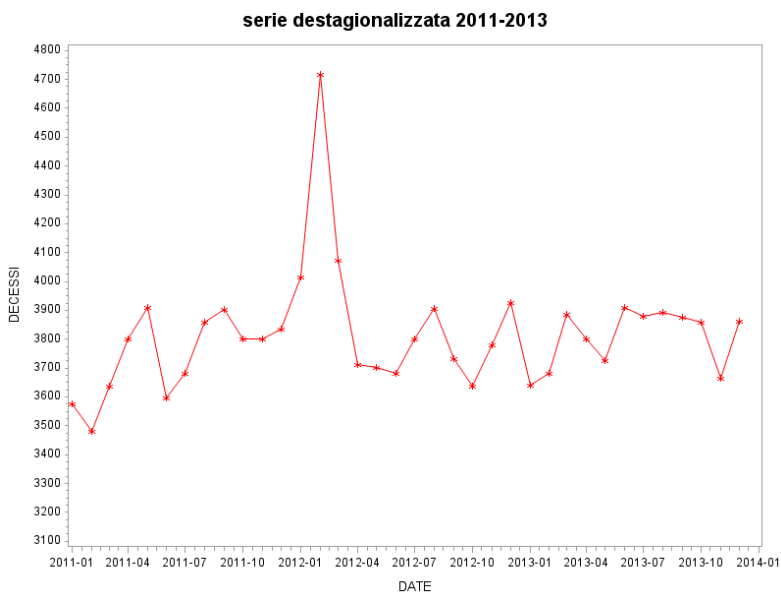
b)



c)



d)



#### 4.4 Identificazione degli eccessi di mortalità significativi

Dopo aver indagato sulle componenti della serie storica della mortalità, si vuole a questo punto applicare una modellazione ARIMA con scopo ultimo l'identificazione degli outliers della serie. Una volta trovato il modello ARIMA che meglio spiega il processo generatore della mortalità nel Veneto, si ricercano gli eventuali outliers/shock della serie non spiegati da tale modello. Quest'ultimi sono di particolare interesse perché indicano che probabilmente qualche altro fenomeno interagisce con la mortalità. La domanda successiva sarà se questo fenomeno può essere identificato nelle condizioni climatiche.

Il modello SARIMA assunto, dopo alcuni passaggi di identificazione e confronto con altri modelli (riportati nel dettaglio in Appendice C) è il seguente:

$$\text{SARIMA}(1, 0, 0) \times (0, 1, 1)_{12},$$

ossia un modello Stagionale Autoregressivo Integrato a Media Mobile, con un parametro di differenziazione stagionale, uno in media mobile per la componente stagionale (di ordine 12), più un parametro autoregressivo di ordine 1. In tabella 4.3 sono riportate le stime dei parametri, con rispettivo errore standard e statistica test per l'ipotesi nulla  $\theta = 0$  o  $\phi = 0$ .

**Tabella 4.3.** Stima dei parametri del modello SARIMA (1, 0, 0) x (0, 1, 1)<sub>12</sub>.

Stima dei minimi quadrati condizionati					
Parametro	Stima	Errore standard	Valore t	p-value	Ritardo
$\mu$	23.78072	11.39222	2.09	0.0385	0
$\theta_1$	0.62365	0.06464	9.65	<.0001	12
$\phi_1$	0.35067	0.07573	4.63	<.0001	1

Quindi l'espressione del modello stimato è la seguente:

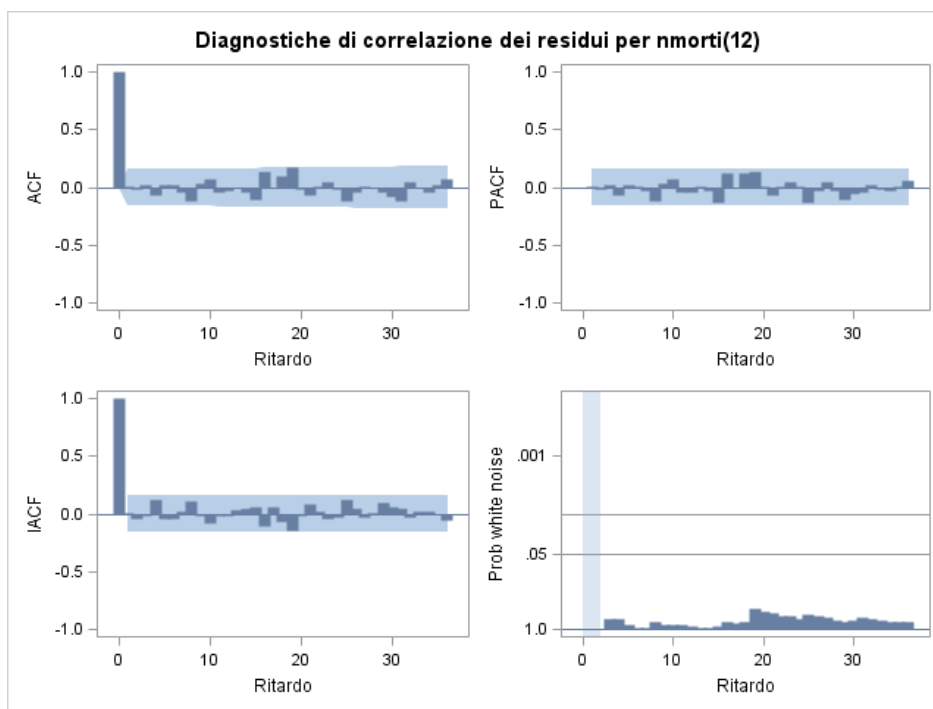
$$(1 - 0.35067 * B)(1 - B^{12})(Y_t - 23,78072) = (1 - 0.62365 * B^{12})\xi_t$$

Di seguito sono invece riportati gli strumenti per la diagnostica del modello: l'autocorrelazione dei residui per verificare l'ipotesi di processo white noise (tabella 4.4), i grafici di autocorrelazione globale e parziale (figura 4.14), e istogramma e q-q plot per la normalità dei residui (figura 4.15).

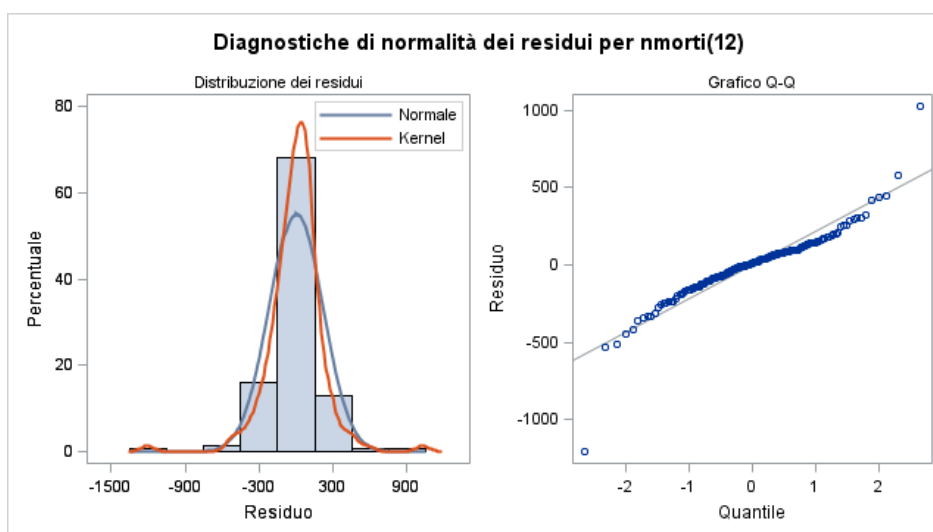
**Tabella 4. 4.** Controllo autocorrelazione dei residui, statistica di Ljung-Box.

Controllo autocorrelazione dei residui										
Fino al ritardo	Stat. Ljung Box	DF	p-value	Autocorrelazioni						
6	1.02	4	0.9072	0.008	-0.024	0.025	-0.066	0.022	0.015	
12	5.01	10	0.8904	-0.039	-0.115	0.031	0.073	-0.045	-0.027	
18	12.23	16	0.7281	-0.002	-0.041	-0.110	0.131	-0.011	0.098	
24	18.83	22	0.6556	0.169	-0.014	-0.067	0.013	0.052	-0.015	
30	23.53	28	0.7061	-0.123	-0.042	0.007	-0.002	-0.043	-0.076	
36	28.07	34	0.7530	-0.114	0.051	-0.005	-0.038	0.026	0.070	

**Figura 4.14.** Diagnostiche di correlazione dei residui: autocorrelazione globale (ACF), autocorrelazione parziale (PACF) e autocorrelazione parziale inversa (IPACF).



**Figura 4.15.** Diagnostiche di mortalità dei residui: istogramma e grafico q-q plot.



I residui appaiono incorrelati, l'ipotesi di residui *white noise* non è rigettata e anche la normalità dei residui sembra essere avvalorata. Quindi il modello si può ritenere buono per spiegare il fenomeno in esame.

Quindi, stimato il modello, sono stati identificati gli *outliers* (tabella 4.5).

**Tabella 4.5.** *Outliers identificati, con significatività 5%, del modello stimato SARIMA(1, 0, 0)x(0, 1, 1)<sub>12</sub>. In azzurro i mesi invernali e in rosso quelli estivi.*

ID temporale	Tipo	Stima	Chi-quadrato	p-value
2012-02	Additivo	899.54814	52.03	<.0001
2000-01	Additivo	898.91026	41.28	<.0001
2003-08	Additivo	590.05958	21.89	<.0001
2005-02	Additivo	531.10932	17.88	<.0001
2002-06	Additivo	411.50620	10.47	0.0012
2003-03	Additivo	367.91053	9.26	0.0023
2008-02	Additivo	354.62946	8.81	0.0030
2003-06	Additivo	343.86462	9.21	0.0024
2010-10	Additivo	338.44360	8.99	0.0027
2000-02	Additivo	365.33966	9.00	0.0027
2005-03	Additivo	321.64835	8.71	0.0032
2009-01	Additivo	309.32305	8.15	0.0043
2012-01	Additivo	299.05188	7.38	0.0066
2010-07	Additivo	283.43514	7.78	0.0053
2001-08	Additivo	288.35279	7.62	0.0058
2012-03	Additivo	285.03962	7.47	0.0063
2006-12	Additivo	-269.24547	7.18	0.0074
2000-08	Additivo	279.03984	6.66	0.0099
2001-05	Additivo	227.03465	5.13	0.0235
2006-01	Additivo	-214.97428	4.84	0.0278
2011-04	Shift	88.24840	4.73	0.0296
2001-01	Additivo	-200.33729	4.09	0.0431
2013-06	Additivo	208.95231	4.04	0.0444
2006-10	Additivo	-184.44314	3.87	0.0491

La procedura ha identificato 24 *outliers*, tutti di tipo additivo tranne 1. L'outlier additivo si manifesta con una brusca variazione della serie in un dato istante temporale, il cui effetto è però immediatamente riassorbito nell'istante successivo. Esso può dipendere da eventi eccezionali. L'outlier di tipo shift rappresenta un cambiamento di livello, ossia è un cambiamento brusco in un dato istante temporale che è destinato però a perdurare nel tempo. Esso corrisponde ad una rottura nell'andamento dei valori medi della serie, ove il cambiamento intervenuto rimane un dato permanente.

Concludendo, dalla ricerca sono emersi essere valori anomali della serie della mortalità i seguenti mesi:

- gennaio, febbraio, agosto 2000;
- gennaio, maggio, agosto 2001;
- giugno 2002;
- marzo, giugno, agosto 2003;
- febbraio, marzo 2005;
- gennaio, ottobre, dicembre 2006;
- febbraio 2008;
- gennaio 2009;
- luglio, ottobre 2010;
- gennaio, febbraio, marzo 2012;
- giugno 2013.

Alcuni di questi valori già erano stati identificati dai grafici della serie destagionalizzata, ma solo ora è possibile affermare che effettivamente sono dei picchi di mortalità significativi, al netto della stagionalità della serie.



#### 4.5 Mortalità e humidex: prime osservazioni

La relazione che sussiste tra la mortalità giornaliera e i valori di humidex non è una relazione lineare semplice. La correlazione di Pearson infatti risulta bassa: 0.18 per il Litorale, 0.28 per la Pianura, 0.25 per la Pedemontana, 0.12 per la Montana.

Per introdurre allo studio di questa relazione, che sarà approfondito nel capitolo successivo, si riportano una serie di statistiche descrittive delle due serie nei 14 anni di riferimento, e anche una raffigurazione dell'andamento osservato delle due serie storiche insieme, separatamente per fascia climatica.

Le tabelle 4.6 e 4.7 mostrano, per fascia climatica, il numero medio di decessi al giorno e il numero totale di decessi nel periodo 21/05 – 14/09, il minimo, il massimo, la media e la deviazione standard dell'humidex e anche la percentuale di giorni di disagio moderato (humidex compreso tra 30°C e 35°C) e intenso (humidex maggiore di 35°C) sempre nello stesso periodo.

Per la fascia Litorale si nota che il 2003, seguito poi dal 2009, è stato l'anno con il più alto numero di giorni con humidex superiore ai 30°C, ma i valori massimi di humidex (41°C) sono stati raggiunti negli anni 2010 e 2013.

Nella fascia Montana invece non si sono verificati giorni di disagio intenso, ma nel 2003 ci sono stati ben 23 giorni di disagio moderato, il numero superiore rispetto a tutti i quattordici anni considerati, seguono poi il 2012 e il 2013 con 17 e 16 giorni. I valori massimi di humidex raggiunti in questa fascia non superano i 36°C.

Nella Pedemontana, nel 2003 ci sono stati parecchi giorni di disagio intenso e moderato, rispettivamente 51 e 20; il secondo anno più caldo sembra essere stato il 2012 con un totale di 50 giorni di disagio.

Nella Pianura il 2003, il 2009 e il 2012 sono stati gli anni con il maggior numero di giorni di disagio: nel 2003 ci sono stati 46 giorni di disagio moderato e 30 di disagio intenso, nel 2009 ci sono stati 48 i giorni di disagio moderato e 10 di disagio intenso, nel 2012 i giorni di disagio moderato sono stati 40 e di intenso 17. Il valori di humidex nella Pianura non hanno mai superato i 40°C negli anni in esame.

In tutte le fasce è evidente che l'estate che ha registrato il più alto numero di giorni di disagio moderato e intensa è stata quella del 2003.

**Tabella 4.6.** Statistiche descrittive dei decessi e dei valori di humidex giornalieri, nel periodo 21/05 - 14/09 di ogni anno, per le fasce climatiche Litorale e Montana. In rosso alcuni valori rilevanti.

Anno periodo 21/05 - 14/09	Media di decessi al giorno nel periodo 21/05 – 14/09	Numero assoluto di decessi nel periodo 21/05 – 14/09	HUMIDEX				Disagio moderato	Disagio intenso
			min	media	max	Std	(hum. tra 30 e 35) % giorni	(hum. > 35) % giorni
<b>LITORALE</b>								
2000	22	2531	18.62	27.33	37.27	4.42	25.64	4.27
2001	22	2536	15.79	27.75	37.72	5.12	35.90	4.27
2002	21	2512	17.28	27.94	37.71	4.02	25.64	5.13
2003	23	2655	16.18	30.67	38.33	5.29	38.46	23.93
2004	21	2499	13.76	27.25	38.83	4.91	26.50	3.42
2005	21	2443	15.93	28.09	39.65	4.57	23.93	6.84
2006	22	2612	13.29	28.01	38.15	5.98	24.79	14.53
2007	22	2562	16.10	28.30	37.48	4.57	33.33	5.98
2008	22	2626	16.61	29.99	39.47	4.73	35.90	15.38
2009	22	2611	18.32	29.51	37.92	4.85	39.32	12.82
2010	24	2753	19.63	28.54	41.30	5.05	29.91	10.26
2011	23	2672	19.37	29.26	38.75	4.30	32.48	11.11
2012	23	2697	17.37	30.00	39.42	4.88	31.62	17.09
2013	24	2809	11.75	28.21	41.12	6.27	29.91	13.68
<b>MONTANA</b>								
2000	7	858	14.29	23.26	30.31	4.04	2.56	0
2001	8	902	11.72	23.76	32.43	5.21	7.69	0
2002	8	892	14.41	23.71	33.90	3.94	5.98	0
2003	7	869	13.15	25.79	33.44	5.02	20.51	0
2004	7	837	10.34	23.00	32.43	4.56	4.27	0
2005	8	895	11.66	23.31	34.04	4.54	11.11	0
2006	8	890	8.12	22.99	32.35	5.76	10.26	0
2007	8	887	10.62	22.72	32.03	4.45	4.27	0
2008	8	895	13.39	24.19	33.08	4.17	8.55	0
2009	7	835	13.08	24.13	33.41	4.48	6.84	0
2010	8	908	13.44	23.85	35.06	5.16	11.11	0.85
2011	8	941	15.62	24.24	33.25	4.10	11.11	0
2012	8	908	13.07	25.21	33.99	4.62	15.38	0
2013	9	1023	7.38	23.79	36.47	6.04	13.68	0.85

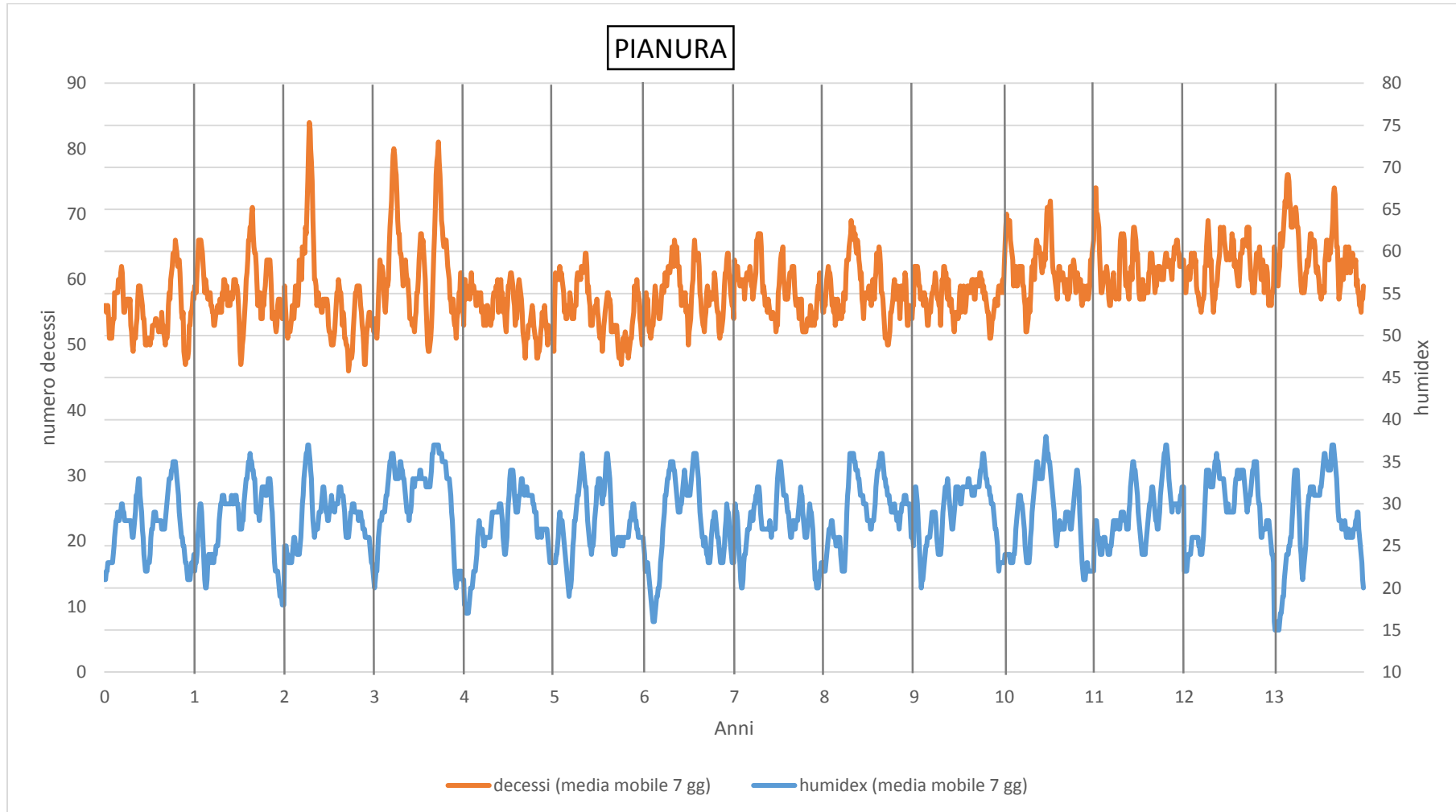
**Tabella 4. 7.** Statistiche descrittive dei decessi e dei valori di humidex giornalieri, nel periodo 21/05 - 14/09 di ogni anno, per le fasce climatiche Pedemontana e Pianura. In rosso alcuni valori rilevanti.

Anno periodo 21/05 - 14/09	Media di decessi al giorno nel periodo 21/05 – 14/09	Numero assoluto di decessi nel periodo 21/05 – 14/09	HUMIDEX**				Disagio moderato  (hum. tra 30 e 35)	Disagio intenso  (hum. > 35)
			min	media	max	Std		
<b>PEDEMONTANA</b>								
2000	20	2342	17.40	26.80	35.61	4.34	26.50	0.85
2001	20	2370	14.40	27.44	36.70	5.39	37.61	2.56
2002	21	2435	17.53	27.74	38.52	4.21	19.66	5.98
2003	23	2720	16.20	<b>30.31</b>	38.54	5.39	<b>44.44</b>	<b>17.95</b>
2004	20	2314	13.37	26.12	36.37	4.75	21.37	1.71
2005	20	2363	15.20	27.40	38.94	4.84	23.93	6.84
2006	21	2500	11.39	26.73	36.86	5.99	27.35	6.84
2007	21	2486	14.37	26.68	35.87	4.64	21.37	1.71
2008	21	2461	15.96	28.28	36.64	4.61	35.04	5.13
2009	21	2510	15.46	27.87	36.77	4.80	36.75	3.42
2010	22	2565	16.08	27.12	38.95	5.21	23.08	8.55
2011	22	2570	18.85	27.81	36.88	4.19	27.35	5.13
2012	23	2636	14.60	<b>28.60</b>	36.49	4.72	<b>36.75</b>	<b>5.98</b>
2013	22	2600	9.44	26.71	38.69	6.25	23.93	8.55
<b>PIANURA</b>								
2000	55	6442	18.25	27.17	35.83	4.36	25.64	3.42
2001	58	6802	15.18	27.81	37.48	5.37	39.32	4.27
2002	57	6650	18.03	28.25	38.82	4.16	24.79	6.84
2003	61	7174	16.67	<b>30.89</b>	38.86	5.42	<b>39.32</b>	<b>25.64</b>
2004	55	6445	14.22	27.08	37.80	4.86	28.21	2.56
2005	55	6468	15.82	27.76	38.81	4.71	23.93	6.84
2006	58	6758	12.21	27.33	37.40	5.89	29.06	8.55
2007	58	6796	15.24	27.50	36.50	4.56	28.21	3.42
2008	58	6833	16.12	29.37	38.08	4.74	35.90	11.11
2009	57	6700	16.90	29.02	37.66	4.86	<b>41.03</b>	<b>9.40</b>
2010	62	7196	17.75	28.08	40.32	5.23	27.35	11.11
2011	62	7221	19.70	28.91	37.46	4.17	30.77	9.40
2012	62	7220	15.93	<b>29.64</b>	38.16	4.73	<b>34.19</b>	<b>14.53</b>
2013	64	7438	10.85	27.71	39.52	6.27	28.21	11.97

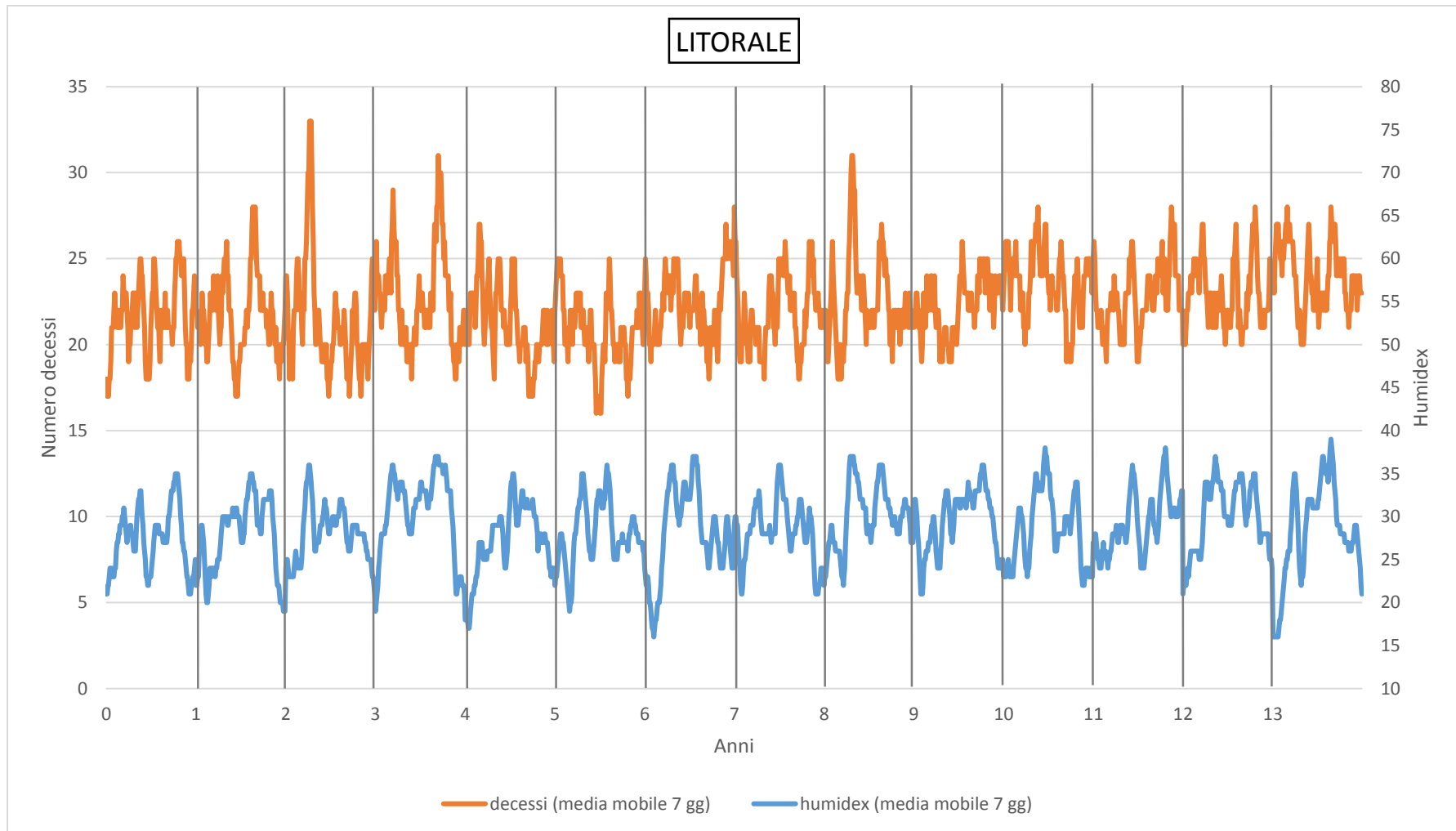
Le figure 4.16, 4.17, 4.18 e 4.19 illustrano invece la distribuzione dell'humidex e dei decessi per giorno, dal 2000 al 2013, trasformata mediante media mobile di 7 giorni (perché si tratta di una serie giornaliera), distintamente per le quattro aree climatiche. Dai grafici si possono distinguere curve diverse di humidex da un anno all'altro. Ad esempio si è visto dalle tabelle precedenti che oltre al 2003 ci sono stati altre estati calde in cui si sono raggiunti valori alti di humidex, ma da queste figure è possibile vedere graficamente che nel 2003, a differenza degli altri anni, i valori si sono mantenuti alti anche fuori dalle due ondate di calore.

Tutti i risultati confermano che l'estate del 2003 è stata caratterizzata da condizioni climatiche ben diverse dagli altri anni: è l'estate in cui i valori di humidex si sono mantenuti alti per più giorni, a differenza di altri anni in cui si sono verificati dei picchi di humidex ma meno intensi e più brevi.

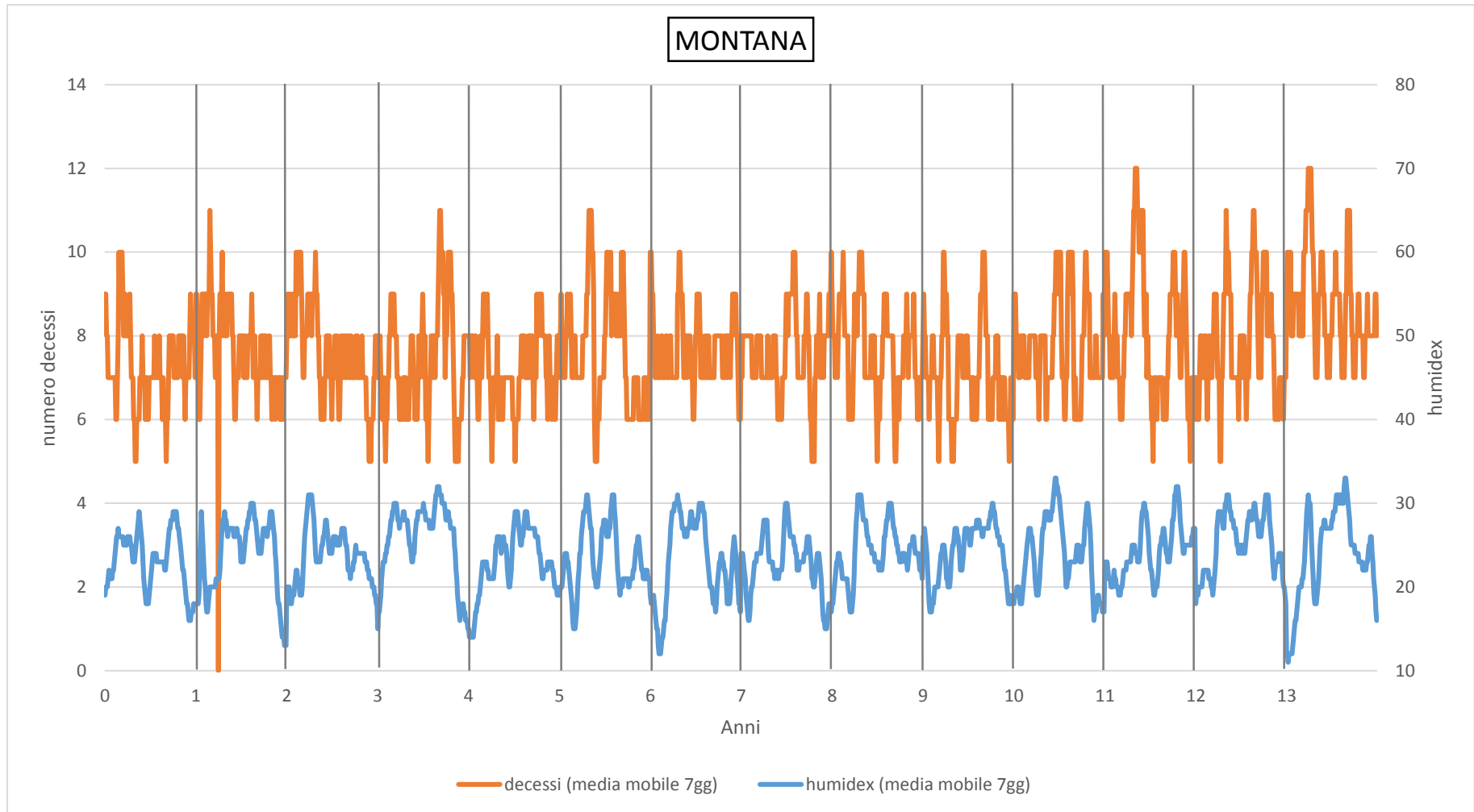
**Figura 4.16.** Numero decessi giornalieri e valore di humidex medio dello stesso giorno, media mobile 7gg, periodo 21/05 – 14/09, anni dal 2000 al 2013, PIANURA. Le linee verticali dividono i periodi da un anno all'altro. Le medie mobili sono state calcolate separatamente per anno, essendo la serie interrotta.



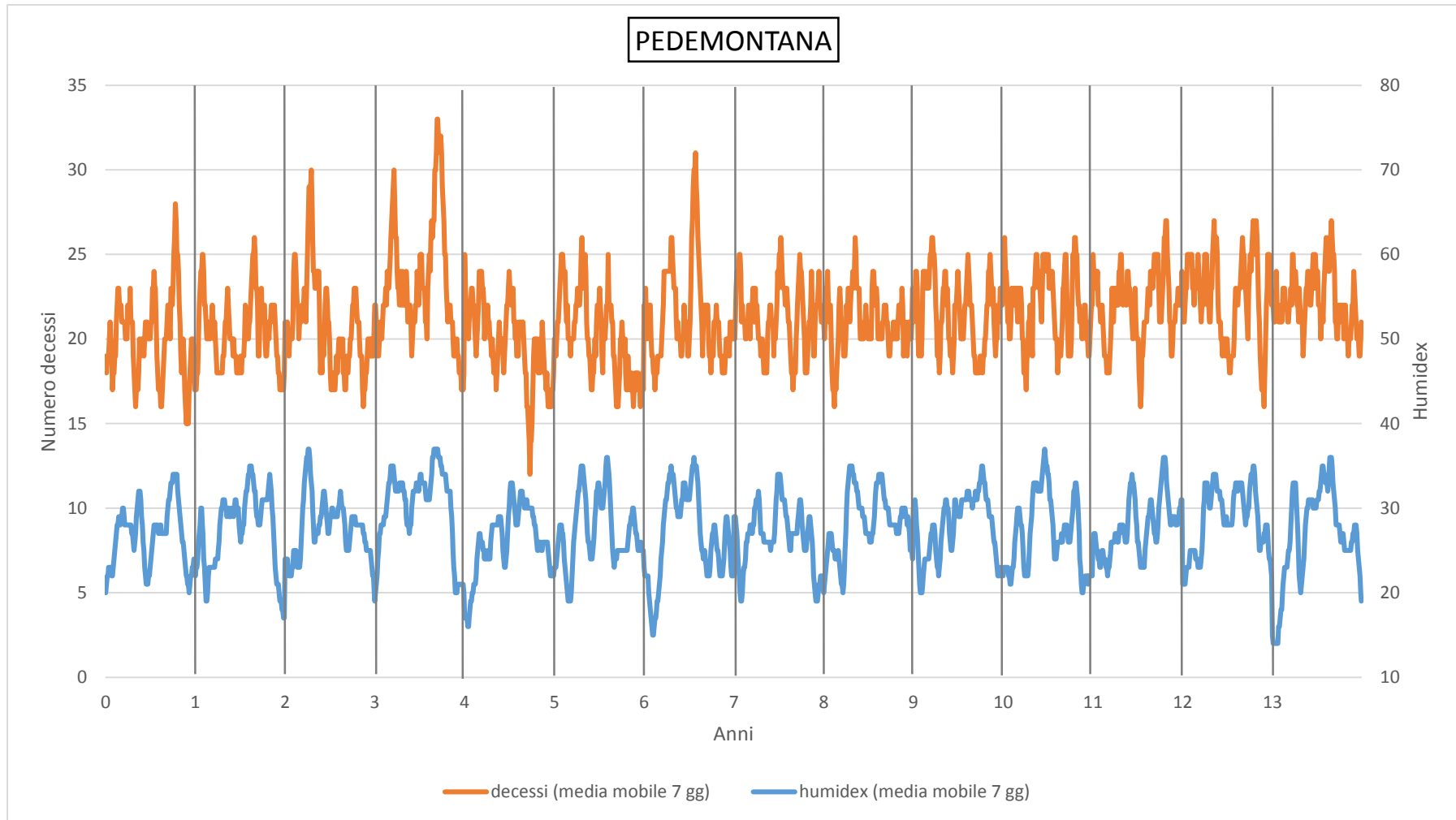
**Figura 4.17.** Numero decessi giornalieri e valore di humidex medio dello stesso giorno, media mobile 7gg, anni dal 2000 al 2013, periodo 21/05 – 14/09, LITORALE.  
Le linee verticali dividono i periodi da un anno all'altro. Le medie mobili sono state calcolate separatamente per anno, essendo la serie interrotta.



**Figura 4.18.** Numero decessi giornalieri e valore di humidex medio dello stesso giorno, media mobile 7gg, anni dal 2000 al 2013, periodo 21/05 – 14/09, MONTANA. Le linee verticali dividono i periodi da un anno all'altro. Le medie mobili sono state calcolate separatamente per anno, essendo la serie interrotta.



**Figura 4. 19.** Numero decessi giornalieri e valore di humidex medio dello stesso giorno, media mobile 7gg, anni dal 2000 al 2013, periodo 21/05 – 14/09, PEDEMONTANA. Le linee verticali dividono i periodi da un anno all'altro. Le medie mobili sono state calcolate separatamente per anno, essendo la serie interrotta.





## 5. MODELLI PER LA RELAZIONE TRA MORTALITÀ E HUMIDEX

---

Dalle elaborazioni precedenti è emerso che, dal 2000 al 2013, nel Veneto si sono verificati degli eccessi di mortalità in alcuni periodi dell'anno, e in particolare anche in alcune estati. Obiettivo di questo capitolo è capire se questi picchi di mortalità estiva possono essere stati causati dalle condizioni climatiche e in particolar modo dall'effetto simultaneo di temperatura e umidità relativa. A questo scopo si ricerca la relazione tra la mortalità (giornaliera) e l'humidex con l'attesa di individuare quali sono le condizioni climatiche maggiormente correlate alla mortalità nel periodo estivo; in particolare identificare il valore critico di humidex oltre il quale la mortalità aumenta evidentemente.

Le analisi vengono effettuate separatamente per area climatica dal momento che, come precedentemente detto, presentano caratteristiche climatiche differenti.

### 5.1 Relazione tra mortalità e humidex per classi d'età

Il primo modello gam, semi-parametrico, stima la relazione tra mortalità e humidex al netto del giorno della settimana, del sesso, dell'anno e della stagionalità nel periodo, per ogni classe d'età e fascia climatica. In tutto quindi sono stati stimati 24 modelli (le 4 fasce climatiche per le 6 classi d'età). È stato scelto di stimare i modelli separatamente per classe d'età per verificare se esistono differenze significative nella relazione mortalità - humidex tra le classi d'età, ed indirizzare di conseguenza le analisi successive.

Di seguito la formulazione del modello sopra descritto:

$$Y_t = \text{Mortalità}_t \sim \text{Poisson}(\mu_t)$$

$$\begin{aligned} \log[E(Y|X)_t^{c,f}] \\ = \text{intercetta} + \text{dow} + \text{sesso} + \text{anno} + s(\text{doy}) + s(\text{humidex}_{\text{ma5gg}}) \\ + \text{offset} \end{aligned}$$

dove:

- $t$  è il giorno,

- $c$  è la classe d'età (0-14, 15-44, 45-64, 65-74, 75-84, 85+),
- $f$  è la fascia climatica (Pianura, Pedemontana, Montana, Litorale),
- $dow$  è la variabile categoriale "giorno della settimana",
- $doy$  è la variabile che identifica il giorno del periodo e assume valore 1 il primo giorno (21 maggio) e incrementa il suo valore di 1 giorno dopo giorno fino a 117 (ultimo giorno del periodo, 14 settembre),
- $humidex_{ma5gg}$  è la media mobile dei valori di humidex dal giorno  $t-4$  al giorno  $t$ .

È stato scelto di utilizzare come esplicativa la media mobile dell'humidex fino ai quattro giorni precedenti per tener conto non solo delle condizioni climatiche del giorno della morte, ma anche dei giorni precedenti. Si pensa infatti che non siano solo le condizioni climatiche del giorno stesso ad influire sullo stato di salute, ma anche quelle dei giorni immediatamente precedenti. Per verifica era stato stimato anche il modello con il valore dell'humidex del giorno stesso, e in termini di AIC non era risultato essere migliore di questo.

Per tener conto delle variazioni che vi possono essere nella numerosità delle coorti è stato incluso nel modello il numero di individui della coorte (la popolazione residente a metà anno, di classe d'età  $c$  e sesso  $s$ ) come offset (in logaritmo).

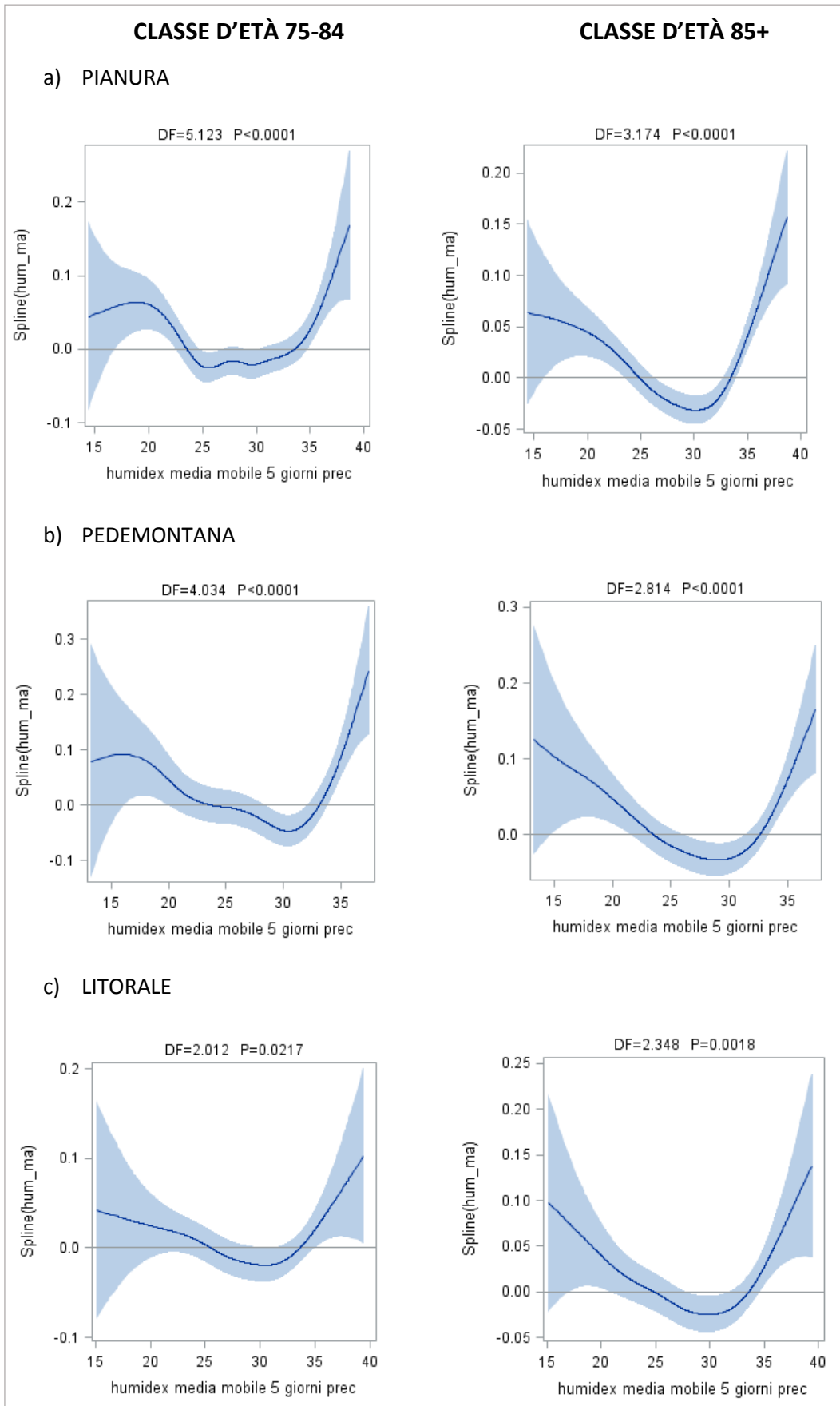
Le figure 5.1 riportano i grafici dell'effetto non lineare della variabile humidex sulla mortalità, per le classi d'età 75-84 anni e over 85 nella Pianura, nella Pedemontana e nel Litorale. Per le altre classi d'età quest'effetto non è risultato significativo. Nella fascia Montana invece non è risultato significativo per nessuna classe d'età.

All'interno di ogni fascia si nota che l'andamento è simile nelle 2 classi d'età (75-84 e 85+), con la differenza che sembra più marcato per la classe 85+. L'andamento suggerisce che intorno ai 34/35°C di humidex il disagio aumenta velocemente, intensificandosi dopo gli 85 anni, in tutte e tre le fasce con qualche differenza: nella fascia Pedemontana sembra iniziare un po' prima, mentre nella fascia Litorale rispetto alla Pianura l'effetto cresce meno velocemente.

Le stime relative ai parametri lineari del modello sono riportate in Appendice D. Per l'analisi e la discussione di esse si rimanda ai modelli presentati al par. 5.2.

Da questo modello quindi si deduce che esiste un effetto non lineare significativo dell'humidex sulla mortalità, per gli over 75 anni e nelle fasce climatiche Pianura, Pedemontana e Litorale. Ottenuto dunque questo risultato, il modello successivamente stimato, e descritto al seguente paragrafo, separa i minori di 75 anni dagli over 75, rilevando però all'interno di queste fasce le eventuali differenze di livello. Inoltre, vista la significatività dell'effetto non lineare, si mantiene nel modello questa componente e se ne valuta l'effetto combinato con la componente lineare, ossia l'effetto complessivo dell'humidex sulla mortalità.

**Figure 5. 1.** Effetto non lineare della variabile humidex sulla mortalità, con limiti di confidenza al 95%, per le classi d'età 75-84 (a sinistra) e 85+(a destra). a) PIANURA, b) PEDEMONTANA, c) LITORALE.



## 5.2 Effetto dell'humidex sulla mortalità prima e dopo i 75 anni

Considerati i risultati del modello precedente, il modello 2 viene stimato utilizzando 75 anni come cut-off e inserendo la variabile età nel modello per catturare le differenze di livello tra le diverse classi d'età. Dunque sono stati stimati otto modelli (le quattro fasce climatiche per le due classi d'età) così formulati:

$$\begin{aligned}\log[E(Y|X)_t^{c,f}] \\ &= \textit{intercetta} + \textit{dow} + \textit{sezzo} + \textit{età} + \textit{anno} + s(\textit{doy}) \\ &+ s(\textit{humidex}_{ma5gg}) + \textit{offset}\end{aligned}$$

Dove:

- t è il giorno;
- C è la variabile che separa i minori di 75 anni da coloro che hanno 75 anni o più;
- f è la fascia climatica (Pianura, Pedemontana, Montana, Litorale);

e le variabili esplicative sono:

- Dow, il giorno della settimana (riferimento: mercoledì);
- Sesso (riferimento: femmine);
- Età rappresenta la classe d'età (riferimento: 65-74 per il modello <75; 85+ per il modello ≥75);
- Anno (riferimento: 2003)
- Doy, la variabile che identifica il giorno del periodo, assume valore 1 il primo giorno (21 maggio) e incrementa di 1 il suo valore giorno dopo giorno fino a 117 (ultimo giorno del periodo, 14 settembre)
- $\textit{humidex}_{ma5gg}$  è la media mobile dei valori di humidex dal giorno t-4 al giorno t

### 5.2.1 POPOLAZIONE DI ETÀ MINORE DI 75 ANNI

La tabella 5.1 contiene i risultati delle stime dei parametri lineari del modello, suddivisi per fascia climatica. Le figure 5.2 e 5.3 rappresentano, rispettivamente, l'effetto non lineare e l'effetto complessivo della variabile humidex (media mobile 5 giorni precedenti) sulla mortalità, sempre dei minori di 75 anni, solo per la Pianura, essendo questa l'unica area climatica in cui l'effetto è risultato significativo. L'effetto complessivo

è stato calcolato come somma dei due effetti: indicata con P la predizione parziale della variabile (ossia la componente non lineare), l'effetto complessivo è dato da:  $P_{humidex_i} + \beta * humidex_i$ .

Dai due grafici si nota che la media dell'humidex nei 5 giorni precedenti ha un effetto significativo sulla mortalità dai 33-35°C, nella Pianura.

I risultati confermano che l'entità del rischio di mortalità aumenta all'aumentare dell'età. In tutti i modelli stimati emerge però che la classe 0-14 ha un rischio superiore rispetto alla classe 15-44. Questo è dovuto al fatto che la classe 0-14 contiene la mortalità infantile, cioè i morti durante il primo anno di vita. Infatti è riconosciuto che l'andamento della mortalità al variare dell'età ha il suo primo valore massimo in corrispondenza del primo anno di vita, e poi questo numero si contrae fino a raggiungere i valori più bassi intorno ai dieci anni.

Anche tra maschi e femmine risulta una differenza significativa: il rischio di mortalità per i maschi è superiore a quello delle femmine dell'86% ( $e^{0.624}$ ). Per quanto riguarda il giorno della settimana e l'anno, emerge che al sabato, rispetto al mercoledì, sembra esserci una diminuzione della mortalità (ma significativa solo al 10%), e solo gli anni 2000 e 2002 non risultano significativamente diversi dal 2003 e dopo il 2003 sembra esserci un trend in diminuzione.

Per la fascia Pedemontana, risulta significativo l'effetto lineare dell'humidex ( $e^{0.008} = 1.008$ ) ma non l'effetto non lineare. Anche per questa fascia climatica la classe 65-74 risulta essere la più a rischio rispetto alle altre classi, infatti per le classi d'età 0-14 e 45-64 il rischio è inferiore del 70% ( $e^{-1.249}, e^{-1.23}$ ), invece per la classe 15-44 il rischio è inferiore del 90%.

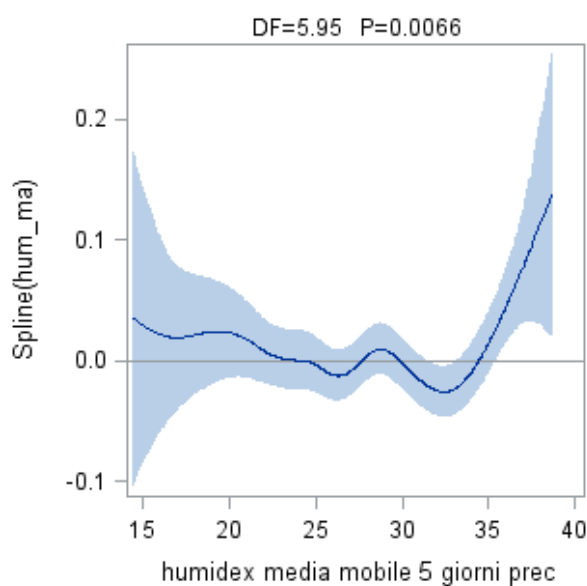
Per i maschi il rischio è superiore del 47% ( $e^{0.386}$ ) rispetto alle femmine.

Per la fascia Montana risultano differenze significative solo per il sesso e la classe d'età. Anche qui il rischio è maggiore per i maschi, ma in misura minore rispetto alle altre fasce, infatti è del 27%. Tra le classi d'età si leggono i seguenti risultati: rispetto alla classe 65-74, la classe 0-14 ha un rischio minore del 40%, la classe 15-44 minore dell'80% e la classe 45-64 minore del 35%. Si conferma così il trend delle altre fasce climatiche.

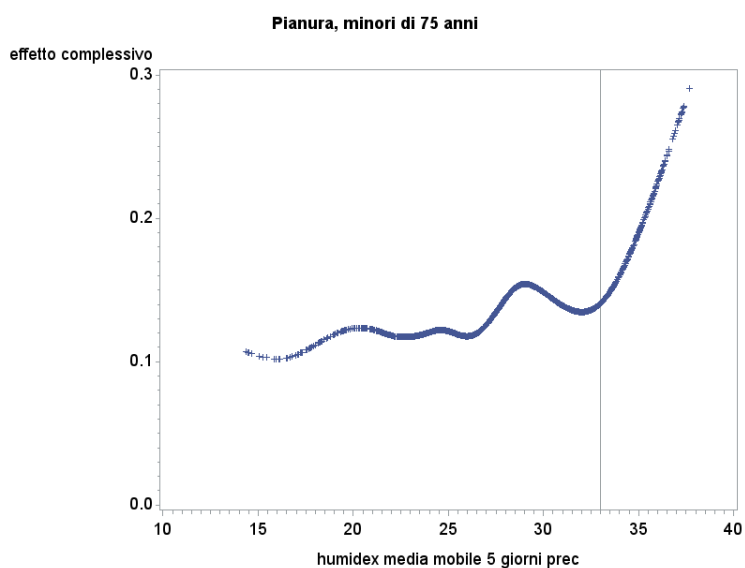
Nemmeno per la fascia Litorale sembra esserci un effetto non lineare significativo dell'humidex; invece appare un effetto lineare lieve ( $e^{0.004} = 1.004$ ) significativo al 5%. In questa fascia climatica i maschi hanno un rischio superiore del 60% rispetto alle femmine.

Le estati 2004 – 2008 non sono risultate statisticamente diverse da quella del 2003, mentre le altre estati sì.

**Figura 5.2.** Effetto non lineare della variabile humidex (media mobile backward 5 gg), sulla mortalità, e relativo intervallo di confidenza al 95%. Minori di 75 anni, PIANURA.



**Figura 5.3.** Effetto complessivo della variabile humidex sulla mortalità.



**Tabella 5.1.** *Modello 2, stime dei parametri lineari, minori di 75 anni, per ciascuna fascia climatica.*

\*\*\* significativo all'1%, \*\* significativo al 5%, \*significativo al 10%.

Minori di 75 anni	Pianura	Pedemontana	Montana	Litorale
Intercetta	-10,485***	-10,228***	-9,416***	-10,252***
dow Friday vs Wednesday	-0,016	-0,034	0,002	-0,010
dow Monday vs Wednesday	-0,008	0,017	-0,023	0,004
dow Sunday vs Wednesday	0,019	-0,001	0,012	0,021
dow Tuesday vs Wednesday	-0,009	-0,009	0,008	0,024
dow Saturday vs Wednesday	-0,037*	-0,018	-0,033	0,004
dow Thursday vs Wednesday	-0,001	-0,016	0,051	-0,023
secco Maschi vs Femmine	0,624***	0,386***	0,243***	0,459***
eta 0-14 vs 65-74	-1,930***	-1,249***	-0,514***	-1,024***
eta 15-44 vs 65-74	-2,648***	-2,163***	-1,520***	-2,071***
eta 45-64 vs 65-74	-1,326***	-1,230***	-1,039***	-1,187***
anno 2000 vs 2003	0,057*	0,058	-0,015	0,132***
anno 2001 vs 2003	0,087***	0,014	0,018	0,106**
anno 2002 vs 2003	0,010	0,019	0,041	0,107**
anno 2004 vs 2003	-0,068**	-0,043	-0,018	0,011
anno 2005 vs 2003	-0,115***	-0,127**	0,001	-0,061
anno 2006 vs 2003	-0,138***	-0,094*	-0,084	-0,054
anno 2007 vs 2003	-0,179***	-0,188***	-0,038	-0,026
anno 2008 vs 2003	-0,180***	-0,172***	-0,045	-0,089*
anno 2009 vs 2003	-0,234***	-0,168***	-0,082	-0,119**
anno 2010 vs 2003	-0,239***	-0,214***	-0,083	-0,086*
anno 2011 vs 2003	-0,234***	-0,244***	-0,074	-0,176***
anno 2012 vs 2003	-0,260***	-0,254***	-0,079	-0,162***
anno 2013 vs 2003	-0,311***	-0,257***	-0,112	-0,140***
Linear(doy)	-0,001***	-0,001**	0,000	0,000
Linear(hum_ma)	0,005***	0,008***	0,003	0,004**

## 5.2.2 POPOLAZIONE DI ETÀ MAGGIORE DI 75 ANNI

Le figure 5.4 rappresentano l'effetto non lineare delle variabili doy e humidex, invece le figure 5.5 rappresentano l'effetto complessivo della variabile humidex, per la Pianura, la Pedemontana e il Litorale (per la fascia Montana non sono significativi). La tabella 5.2 riporta i coefficienti stimati dal modello per i parametri lineari, per coloro che hanno più 75 anni.

Per questa classe d'età la componente non lineare della variabile humidex è significativa per tutte le fasce climatiche a meno della Montana. Si nota un effetto significativo sulla mortalità per tutte e tre le fasce.



Guardando l'andamento dell'effetto complessivo si nota, in generale, che esso aumenta al crescere dei valori di humidex, e in particolare aumenta più sensibilmente dopo i 30°C. È stato ricercato in modo empirico la posizione del punto di flesso a partire dal quale la mortalità aumenta repentinamente, calcolando l'incremento dell'effetto sulla mortalità ogni 2 unità di humidex e individuando quindi il valore di humidex per cui l'incremento risultava maggiore. In riferimento alla Pianura, l'effetto sembra aumentare dai 15°C ai 20°C, poi si stabilizza fino ai 25°C, e dopo i 25°C continua ad aumentare, ma l'incremento maggiore è risultato dai 32°C, punto in cui la curva aumenta di nuovo la sua pendenza. L'effetto nella Pedemontana è molto simile, qui il punto di flesso è leggermente più basso. L'incremento dell'effetto dai 31 ai 33°C è circa del 14%, e dai 33 ai 35°C aumenta ancora del 23%.

Per la fascia Litorale invece l'effetto sembra essere minore rispetto alla Pianura e alla Pedemontana, ma si nota anche qui un incremento sensibile dopo i 33°C, che risulta essere del 25%.

Per quanto riguarda le esplicative inserite nel modello, il giorno della settimana risulta significativo e in particolare nella fascia della Pianura e del Litorale sembra che il venerdì e il sabato la mortalità sia inferiore (di circa il 5%) rispetto al mercoledì. Per verifica, è stato stimato un modello utilizzando un'altra fonte dei decessi, indipendente dalle schede di morte, ed è emerso lo stesso risultato. Questa variabile è stata inserita nel modello in quanto in letteratura è considerata un confondente.

Si nota inoltre che la mortalità nel 2003 risulta significativamente superiore rispetto alle estati successive, nella Pianura, nella Pedemontana e nel Litorale. Nella Pianura la differenza più bassa si nota con l'estate del 2010. Invece nella fascia Pedemontana e Litorale si nota un trend in calo negli anni.

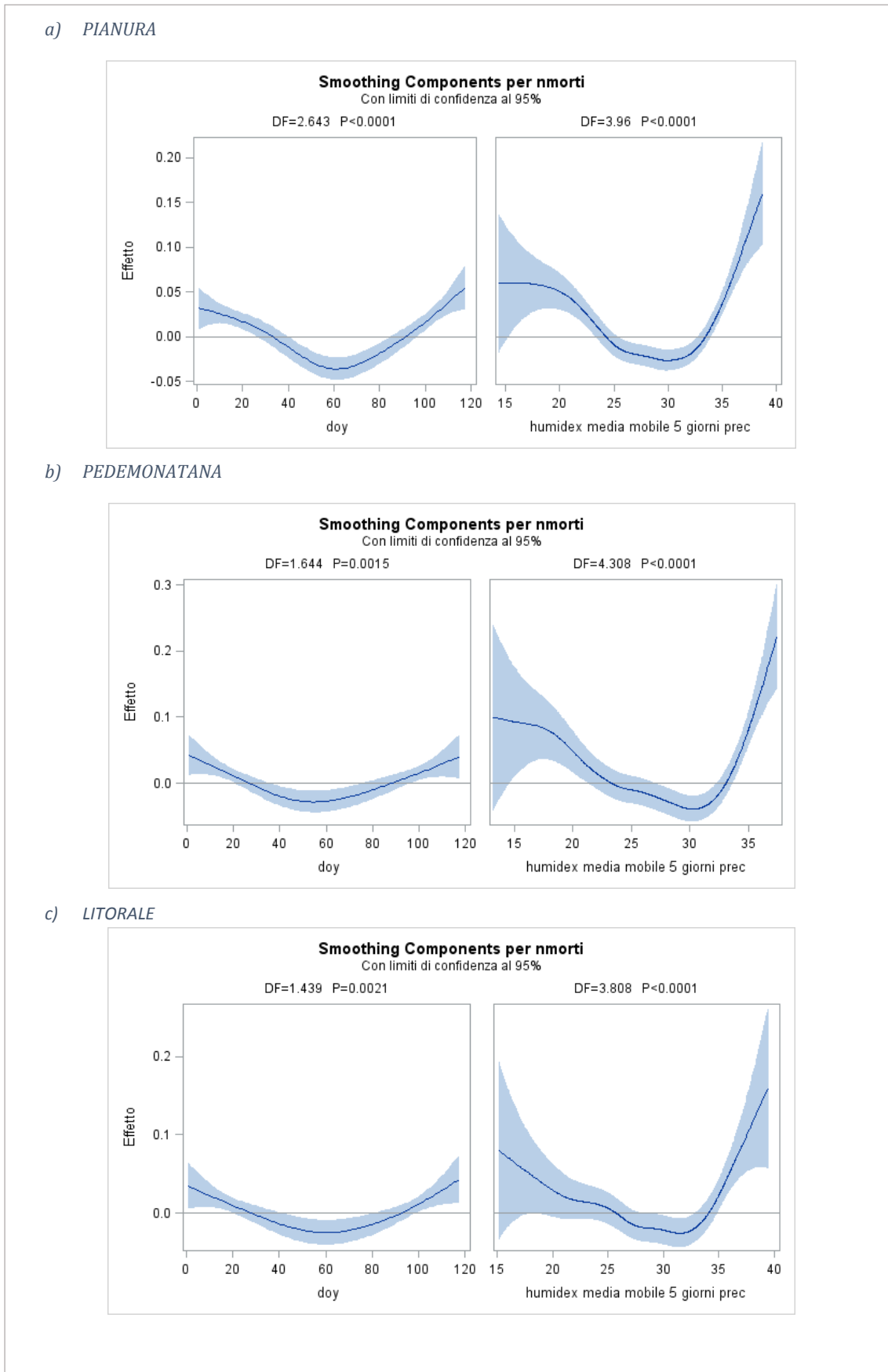
Come per la popolazione fino ai 75 anni, anche quella più anziana, la differenza tra maschi e femmine è significativa, a sfavore dei maschi i quali hanno un rischio di morire più alto, rispetto alle femmine, del 50% ( $e^{0.402}$ ) nella Pianura, del 60% ( $e^{0.45}$ ) nella Pedemontana, dell'80% ( $e^{0.621}$ ) nella Montana e del 55% ( $e^{0.435}$ ) nel Litorale.

**Tabella 5.2.** Modello 2, stime dei parametri lineari, **maggiori di 75 anni**, per ciascuna fascia climatica.

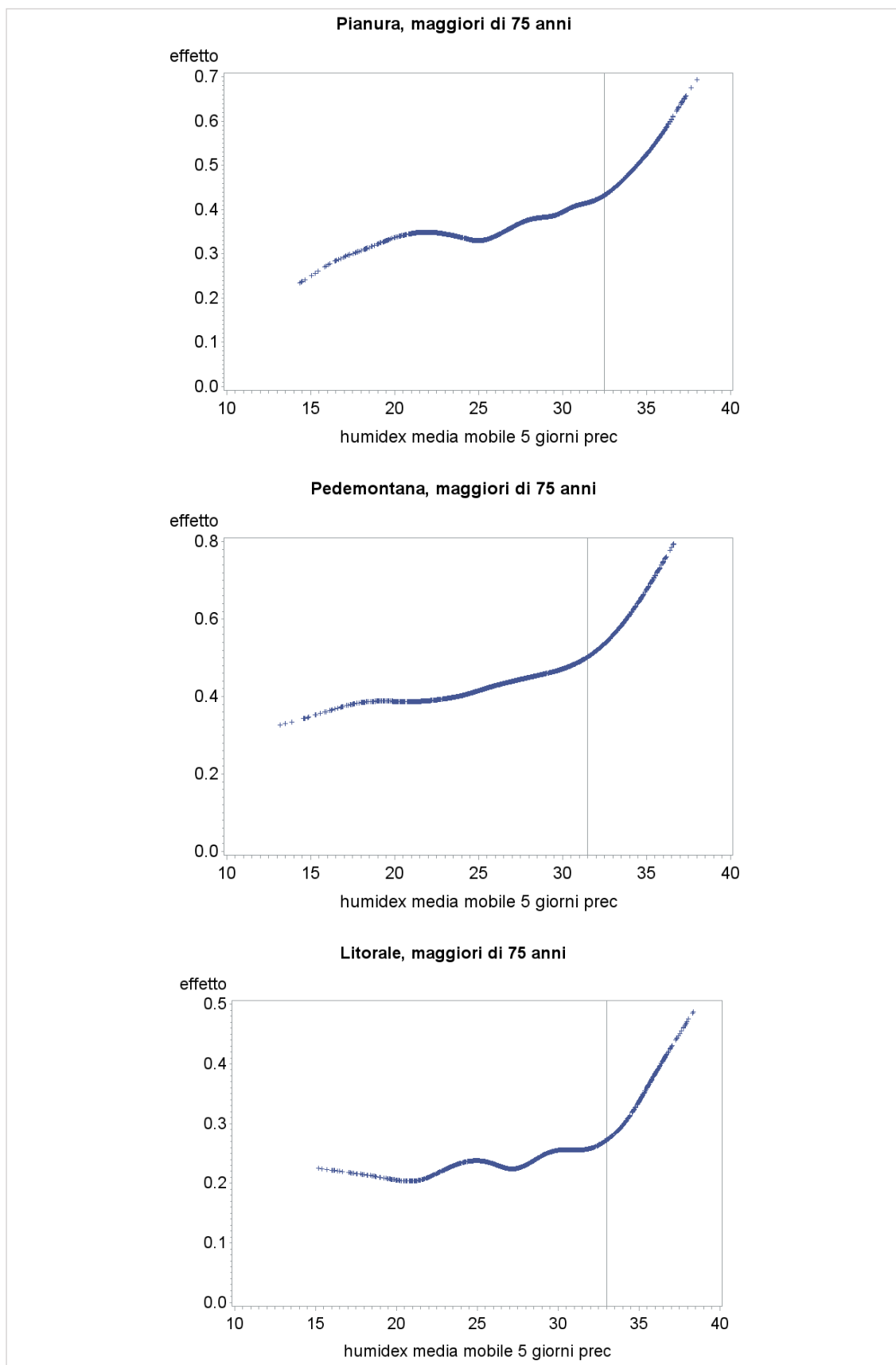
\*\*\* significativo all'1%, \*\* significativo al 5%, \*significativo al 10%.

<b>Maggiori di 75 anni</b>	<b>Pianura</b>	<b>Pedemontana</b>	<b>Montana</b>	<b>Litorale</b>
<b>Intercetta</b>	-8,308***	-8,345***	-7,992***	-8,159***
<b>dow Friday vs Wednesday</b>	-0,053***	0,024	0,004	-0,041*
<b>dow Monday vs Wednesday</b>	0,000	0,043*	0,001	0,007
<b>dow Sunday vs Wednesday</b>	-0,008	0,064***	-0,005	-0,007
<b>dow Tuesday vs Wednesday</b>	-0,018	0,038	-0,002	0,010
<b>dow Saturday vs Wednesday</b>	-0,045***	0,018	-0,048	-0,049**
<b>dow Thursday vs Wednesday</b>	-0,004	0,046*	-0,008	0,016
<b>sesto Maschi vs Femmine</b>	0,402***	0,450***	0,621***	0,435***
<b>eta 75-84 vs 85+</b>	-1,252***	-1,242***	-1,148***	-1,226***
<b>anno 2000 vs 2003</b>	-0,032	-0,082**	0,031	-0,053
<b>anno 2001 vs 2003</b>	-0,024	-0,094***	-0,014	-0,082**
<b>anno 2002 vs 2003</b>	-0,049**	-0,107***	-0,021	-0,062*
<b>anno 2004 vs 2003</b>	-0,042**	-0,097***	-0,019	-0,068*
<b>anno 2005 vs 2003</b>	-0,054**	-0,099***	0,066	-0,065*
<b>anno 2006 vs 2003</b>	-0,038*	-0,062*	0,037	-0,062*
<b>anno 2007 vs 2003</b>	-0,050**	-0,091***	-0,043	-0,098***
<b>anno 2008 vs 2003</b>	-0,119***	-0,168***	-0,050	-0,130***
<b>anno 2009 vs 2003</b>	-0,151***	-0,158***	-0,123**	-0,165***
<b>anno 2010 vs 2003</b>	-0,099***	-0,168***	-0,054	-0,124***
<b>anno 2011 vs 2003</b>	-0,123***	-0,200***	-0,087	-0,174***
<b>anno 2012 vs 2003</b>	-0,169***	-0,232***	-0,081	-0,207***
<b>anno 2013 vs 2003</b>	-0,135***	-0,253***	-0,048	-0,201***
<b>Linear(doy)</b>	-0,001***	-0,001***	-0,001***	0,000
<b>Linear(hum_ma)</b>	0,014***	0,017***	0,011***	0,009***

**Figure 5.4.** Modello 2. Effetto non lineare della variabile humidex (media mobile backward 5 gg) e doy, sulla mortalità, e relativo intervallo di confidenza al 95%. **Maggiori di 75 anni**, a) PIANURA, b) PEDEMONTANA, c) LITORALE.



**Figure 5.5** Modello 2. Effetto complessivo (lineare più non lineare) della variabile humidex sulla mortalità, per i **maggiori di 75 anni**, Pianura, Pedemontana e Litorale. La linea verticale è posta in prossimità del valore di humidex per cui l'effetto sale notevolmente.



### 5.3 Discussione dei risultati

I risultati dei modelli stimati conducono quindi ad affermare che il perdurare per qualche giorno di condizioni climatiche di disagio influisce sullo stato di salute e sul benessere della persona, soprattutto se anziana. In particolare, facendo riferimento ai valori di humidex (indicatore di disagio fisico, basato su temperatura e umidità relativa dell'aria) è emerso che il rischio di morire aumenta sensibilmente se la media dell'humidex in cinque giorni consecutivi supera i 32 – 33°C. L'effetto è stato stimato al netto della stagionalità, del sesso, dell'età, del giorno della settimana e dell'anno.

Questo effetto si nota nella fascia climatica della Pianura, della Pedemontana e del Litorale, con qualche leggera differenza per valori precedenti ai 30°C, ma dopo questa soglia l'effetto è chiaramente simile. Nella fascia Montana invece, dove nei quattordici anni considerati pochissimi giorni hanno registrato valori di humidex superiori ai 35°C (tabella 4.6), è emerso solo un effetto lineare lieve.

In generale è stato inoltre confermato un rischio maggiore per i maschi e all'aumentare dell'età, concordemente ai risultati dei tassi di mortalità (capitolo 4). Per le differenze riscontrate sui giorni della settimana saranno fatti degli opportuni approfondimenti a cura del SER.



## 6. CONCLUSIONI

---

Questa tesi ha trattato la relazione tra la mortalità e le condizioni climatiche nel Veneto, prendendo in considerazione gli anni dal 2000 al 2013. Gli obiettivi posti all'inizio erano: individuare quando si sono verificati eccessi di mortalità significativi nei quattordici anni in esame, e identificare un valore soglia di humidex (indicatore di disagio fisico, basato su temperatura e umidità relativa) oltre il quale la mortalità aumenta evidentemente.

La sensibilità verso il problema dell'impatto sulla salute delle variazioni climatiche è cresciuta negli ultimi decenni a causa dei cambiamenti climatici che hanno interessato il pianeta. In modo particolare un campanello d'allarme è stata l'estate del 2003 che fu caratterizzata, in parecchie zone del mondo, da temperature elevate, più alte rispetto alle temperature usuali estive, che hanno causato situazioni di disagio, anche molto intenso e pericoloso. Tali situazioni di disagio hanno destato preoccupazioni e quindi hanno portato all'attivazione di una serie di programmi d'azione per far fronte alla problematica, ma soprattutto per prevenirla quando possibile. In letteratura sono diversi gli studi che hanno trattato questo problema e negli ultimi anni gli sviluppi sono stati notevoli. Dai principali risultati è emerso che durante le ondate di calore si è verificato un evidente aumento della mortalità generale. Inoltre ad esserne interessati sono stati soprattutto gli anziani e gli individui che presentano particolari patologie collegate con la temperatura.

Il lavoro descritto in questa tesi si è focalizzato sulla mortalità nel Veneto dal 2000 al 2013, e ne ha studiato l'andamento negli anni, ha verificato se ci sono stati degli eccessi di mortalità significativi ed ha ricercato la relazione di tali eccessi estivi con le condizioni climatiche.

Attraverso la modellazione ARIMA sono stati trovati eccessi di mortalità in alcuni mesi invernali (gennaio 2000, febbraio 2005, gennaio e febbraio 2012) e anche in alcuni mesi estivi (giugno 2002, marzo, giugno e agosto 2003, luglio 2010 e giugno 2013). Dallo studio della stagionalità della serie storica la mortalità invernale si è confermata essere più ingente di quella estiva; questa è infatti una caratteristica propria dei Paesi caratterizzati da stagioni; il risultato è in linea con quelli presenti in letteratura.

Per quanto riguarda l'estate, la relazione ricercata tra la mortalità e i valori di humidex è risultata essere una relazione non lineare, e per questo motivo è stato opportuno utilizzare i Modelli Additivi Generalizzati. Dallo studio di questa relazione è emerso che l'impatto dell'humidex sulla mortalità è evidente soprattutto per le persone con più di settantacinque anni e che il rischio di morire aumenta sensibilmente se la media dell'humidex in cinque giorni consecutivi supera i 32 – 33°C. Questo effetto è stato riscontrato per la fascia climatica della Pianura, della Pedemontana e del Litorale, con qualche leggera differenza per valori precedenti ai 30°C, ma dopo questa soglia l'effetto si può dire simile. Nella fascia Montana invece, dove nei quattordici anni considerati pochissimi giorni hanno registrato valori di humidex superiori ai 35°C, è emerso solo un effetto lineare lieve.

Come già era stato anticipato dall'andamento dei tassi di mortalità, per le femmine il rischio sembra essere inferiore rispetto ai maschi.

Quindi quando sono previste tali situazioni di disagio climatico, e considerati come riferimento i valori di humidex risultati critici, è necessario attivare ed intensificare attività di prevenzione, di comunicazione, di educazione soprattutto verso gli anziani e i gruppi di individui più esposti al rischio perché essi non sottovalutino le conseguenze a cui potrebbe portare un loro comportamento azzardato o ingenuo.

Lo studio potrebbe essere proseguito in seguito considerando anche le cause di mortalità, per poter studiare in modo più specifico la relazione all'interno di gruppi omogenei per causa.

Un altro possibile sviluppo è lo studio della relazione tra la mortalità e le condizioni climatiche considerando anche l'effetto dell'inquinamento dell'aria, che potrebbe interagire con le condizioni climatiche.



# APPENDICE A

---

## CENTRALINE ARPAV

**Tabella A.1.** Centraline del Veneto da cui sono stati prelevati i dati di temperatura e umidità relativa, per provincia e fascia climatica, per calcolare l'humidex medio giornaliero dal 21/05 al 14/09 per gli anni 2000-2013.

FASCIA	PROVINCIA	STAZIONE
Val Belluna	BL	Belluno
	BL	Feltre
Pedemontana	VR	Castelnuovo Del Garda
	VR	San Pietro In Cariano
	VR	Montecchia Di Crosara
	VI	Brendola
	VI	Malo
	VI	Bassano
	TV	Maser
	TV	Conegliano
	TV	Vittorio Veneto
Pianura	VR	Villafranca Di Verona
	VR	Salizzole
	VR	Legnago
	VI	Montegalda
	VI	Vicenza
	VI	Noventa Vicentina
	VI	Rosa'
	PD	Cittadella
	PD	Legnaro
	PD	Padova
	PD	Este
	RO	Adria
	RO	Rovigo
	RO	Trecenta
	TV	Breda di Piave
	TV	Oderzo
	TV	Treviso
	TV	Castelfranco Veneto
Litorale	RO	Porto Tolle
	RO	Rosolina
	VE	Cavallino Treporti
	VE	Portogruaro



# APPENDICE B

## TASSI GREZZI E STANDARDIZZATI DI MORTALITÀ

**Tabella B. 1.** Tassi di mortalità annuali, grezzi e standardizzati, del Veneto, anni 2000-2013, per sesso. Standardizzazione diretta, popolazione standard: 1° gennaio 2007.

anno	MASCHI			FEMMINE		
	N	T. grezzo	T. std	N	T. grezzo	T. std
2000	20917	956,3	1342,5	21253	920,2	791,0
2001	20523	933,0	1288,8	20933	902,6	761,2
2002	20990	947,7	1310,2	21277	912,9	765,8
2003	21128	943,7	1309,7	22368	951,1	803,2
2004	20477	902,9	1247,2	20907	879,8	749,1
2005	20849	910,3	1238,7	21881	913,5	761,7
2006	20383	884,4	1171,1	21222	880,4	711,8
2007	20733	892,0	1158,7	22290	916,7	720,6
2008	21329	908,4	1158,7	22740	925,4	708,9
2009	21219	898,9	1122,0	22771	920,4	690,2
2010	21322	902,2	1099,3	23150	932,1	682,7
2011	21306	901,2	1067,9	23580	946,8	677,9
2012	22068	930,6	1076,2	24619	985,5	685,9
2013	21634	905,2	1016,8	24044	956,3	653,4

**Tabella B. 2.** Tassi di mortalità grezzi e standardizzati, periodo 21/05 - 14/09 di ogni anno dal 2000 al 2013, del Veneto, per sesso. Standardizzazione diretta, popolazione standard: Veneto, 1° gennaio 2007.

anno	MASCHI			FEMMINE		
	N	T. grezzo	T. std	N	T. grezzo	T. std
<b>2000</b>	6058	277,7	390,6	6109	265,1	231,3
<b>2001</b>	6304	287,3	396,4	6301	272,2	232,0
<b>2002</b>	6227	282,3	384,8	6238	268,4	225,8
<b>2003</b>	6448	289,9	399,4	6968	298,0	250,4
<b>2004</b>	6059	268,8	369,8	6031	254,9	218,6
<b>2005</b>	5942	260,3	356,0	6221	260,6	221,8
<b>2006</b>	6229	270,9	363,5	6531	271,7	223,1
<b>2007</b>	6130	265,2	347,1	6601	273,0	217,2
<b>2008</b>	6276	268,4	344,2	6539	267,3	207,4
<b>2009</b>	6114	259,2	325,9	6542	264,9	201,7
<b>2010</b>	6460	273,3	336,5	6962	280,8	208,5
<b>2011</b>	6390	270,3	324,0	7014	281,8	204,4
<b>2012</b>	6549	276,9	322,8	6912	277,3	196,5
<b>2013</b>	6660	280,0	319,0	7210	287,9	197,7

**Tabella B. 3. Tassi specifici di mortalità, e relativi intervalli di confidenza, periodo 21/05 – 14/09 di ogni anno dal 2000 al 2013, del Veneto, per sesso. Classe d'età 0-39.**

	MASCHI 0-39			FEMMINE 0-39		
	Tasso specifico	IC		Tasso specifico	IC	
2000	24,5	21,8	27,6	11,3	9,4	13,4
2001	26,4	23,6	29,6	11,5	9,6	13,7
2002	23,6	20,9	26,6	11,5	9,6	13,7
2003	21,9	19,3	24,8	10,0	8,3	12,1
2004	21,4	18,9	24,3	10,3	8,5	12,4
2005	18,1	15,8	20,8	8,2	6,7	10,1
2006	19,1	16,7	21,9	9,0	7,3	10,9
2007	19,5	17,1	22,3	8,1	6,5	9,9
2008	19,5	17,1	22,2	7,7	6,2	9,5
2009	14,8	12,7	17,3	8,1	6,6	10,0
2010	13,3	11,3	15,6	7,7	6,2	9,6
2011	15,1	12,9	17,6	6,3	4,9	8,0
2012	15,1	12,9	17,6	7,9	6,4	9,8
2013	13,8	11,7	16,2	6,1	4,7	7,8

**Tabella B. 4. Tassi specifici di mortalità, e relativi intervalli di confidenza, periodo 21/05 – 14/09 di ogni anno dal 2000 al 2013, del Veneto, per sesso. Classe d'età 40-64.**

	MASCHI 40-64			FEMMINE 40-64		
	Tasso specifico	IC		Tasso specifico	IC	
2000	159,6	150,7	169,0	72,1	66,2	78,5
2001	166,8	157,8	176,4	82,8	76,6	89,6
2002	155,7	147,0	164,8	76,5	70,5	83,0
2003	150,3	141,8	159,2	81,7	75,6	88,4
2004	137,7	129,7	146,2	69,9	64,3	76,0
2005	130,4	122,7	138,6	72,7	67,0	78,8
2006	132,2	124,5	140,3	62,9	57,7	68,7
2007	115,7	108,6	123,3	65,6	60,3	71,4
2008	122,0	114,8	129,7	61,7	56,6	67,3
2009	110,9	104,0	118,1	62,0	56,9	67,5
2010	112,0	105,2	119,3	64,8	59,6	70,4
2011	108,5	101,8	115,6	61,1	56,2	66,5
2012	111,2	104,5	118,4	59,6	54,8	65,0
2013	103,8	97,4	110,7	61,1	56,2	66,4

**Tabella B.5. Tassi specifici di mortalità, e relativi intervalli di confidenza, periodo 21/05 – 14/09 di ogni anno dal 2000 al 2013, del Veneto, per sesso. Classe d'età 65-79.**

	MASCHI 65-79			FEMMINE 65-79		
	Tasso specifico	IC		Tasso specifico	IC	
2000	1017,1	979,2	1056,4	518,5	495,4	542,6
2001	997,0	959,8	1035,6	498,1	475,6	521,7
2002	977,9	941,3	1015,9	477,9	455,9	501,1
2003	951,3	915,6	988,5	479,7	457,6	502,8
2004	887,0	852,9	922,5	435,8	414,9	457,8
2005	831,3	798,8	865,2	427,0	406,4	448,6
2006	814,1	782,3	847,3	429,1	408,6	450,7
2007	780,7	749,8	812,9	412,9	392,8	434,0
2008	769,0	738,5	800,6	378,1	358,9	398,3
2009	741,8	712,2	772,7	376,6	357,6	396,7
2010	738,6	709,2	769,1	379,4	360,3	399,5
2011	713,3	684,4	743,3	385,5	366,2	405,8
2012	703,5	675,3	732,9	366,8	348,2	386,5
2013	689,9	662,3	718,8	340,9	323,0	359,7

**Tabella B. 6. Tassi specifici di mortalità, e relativi intervalli di confidenza, periodo 21/05 – 14/09 di ogni anno dal 2000 al 2013, del Veneto, per sesso. Classe d'età 80+.**

	MASCHI 80+			FEMMINE 80+		
	Tasso specifico	IC		Tasso specifico	IC	
2000	3695,2	3534,4	3863,3	2893,1	2800,1	2989,2
2001	3760,2	3603,1	3924,2	2858,6	2768,8	2951,4
2002	3590,7	3442,0	3745,8	2736,9	2651,5	2825,1
2003	3866,9	3716,3	4023,6	3080,9	2992,2	3172,2
2004	3389,8	3252,4	3533,0	2487,4	2409,5	2567,9
2005	3286,1	3154,4	3423,4	2503,7	2427,2	2582,6
2006	3441,5	3310,0	3578,2	2619,2	2542,3	2698,3
2007	3358,1	3231,5	3489,6	2573,0	2498,4	2649,9
2008	3282,0	3160,0	3408,7	2532,0	2459,3	2606,9
2009	3170,4	3052,8	3292,5	2451,4	2380,9	2524,0
2010	3412,8	3293,0	3537,0	2595,3	2523,5	2669,1
2011	3289,9	3174,5	3409,4	2562,4	2492,1	2634,7
2012	3295,8	3182,1	3413,5	2484,2	2415,6	2554,8
2013	3359,1	3246,3	3475,8	2629,5	2559,5	2701,4



# APPENDICE C

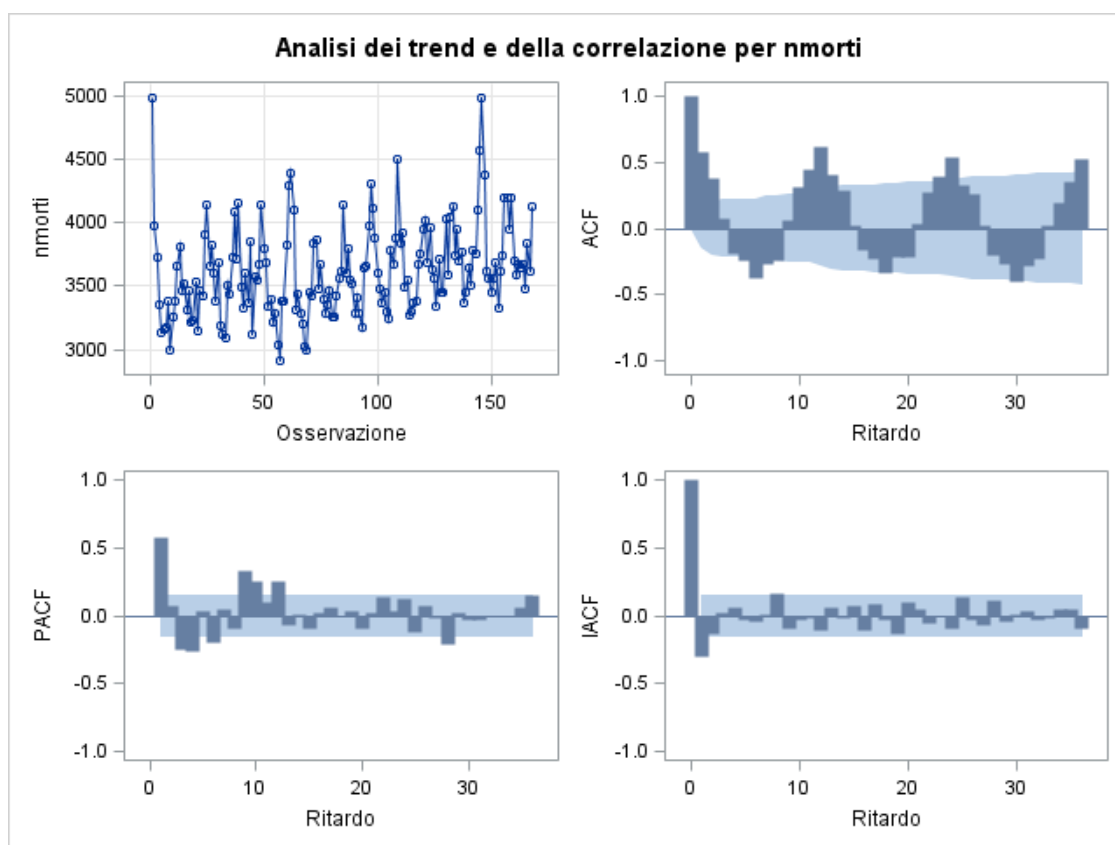
## MODELLAZIONE ARIMA DELLA SERIE STORICA DEI DECESSI DEL VENETO DAL 2000 AL 2013 - PROCEDURA DI BOX E JENKINS

### Identificazione dei parametri del modello e stima

La serie sottoposta ad analisi è costituita dai decessi totali mensili nel Veneto dal 2000 al 2013.

Come primo passo si analizza l'autocorrelazione globale e parziale della serie (figura C.1) per l'identificazione dei parametri del modello.

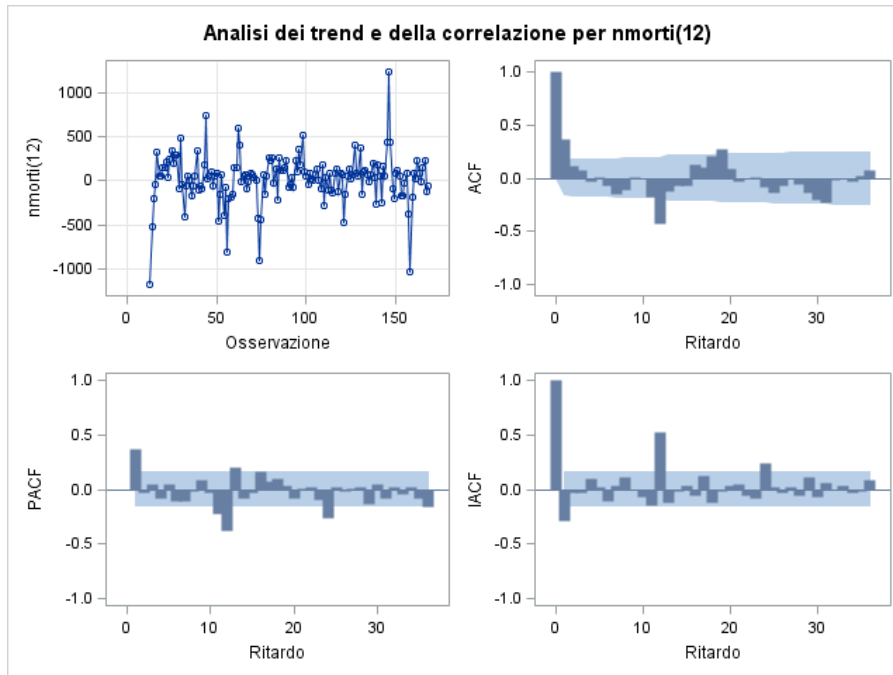
**Figura C.1.** Grafici di autocorrelazione globale e parziale della serie dei decessi mensili del Veneto, dal 2000 al 2013.



Si nota che le autocorrelazioni ai ritardi stagionali decrescono a zero molto lentamente, quindi sottopongo la serie ad una differenziazione stagionale del tipo  $(1-B^S)^D y_t$  con  $S=12$  e  $D=1$ .

Quindi le autocorrelazioni, della serie differenziata, sono quelle riportate in figura C.2.

**Figura C. 2.** Autocorrelazione globale (ACF) e parziale (PACF) della serie sottoposta a differenziazione stagionale del tipo  $(1-B^S)^D y_t$  con  $S=12$  e  $D=1$ .



L'effetto stagionale si mostra ridotto, ma sembra ancora visibile. Per questo motivo provo a stimare un modello autoregressivo di ordine 12, ossia la variabile del numero di decessi viene modellata come valore medio più una frazione della sua devianza dal suo valore medio nell'anno precedente, più un errore casuale. I risultati sono riportati di seguito:

1. Primo modello SARIMA  $(0,0,0)x(1,1,0)_{12}$

**Tabella C.1.** Risultati del primo modello stimato: SARIMA  $(0,0,0)x(1,1,0)_{12}$

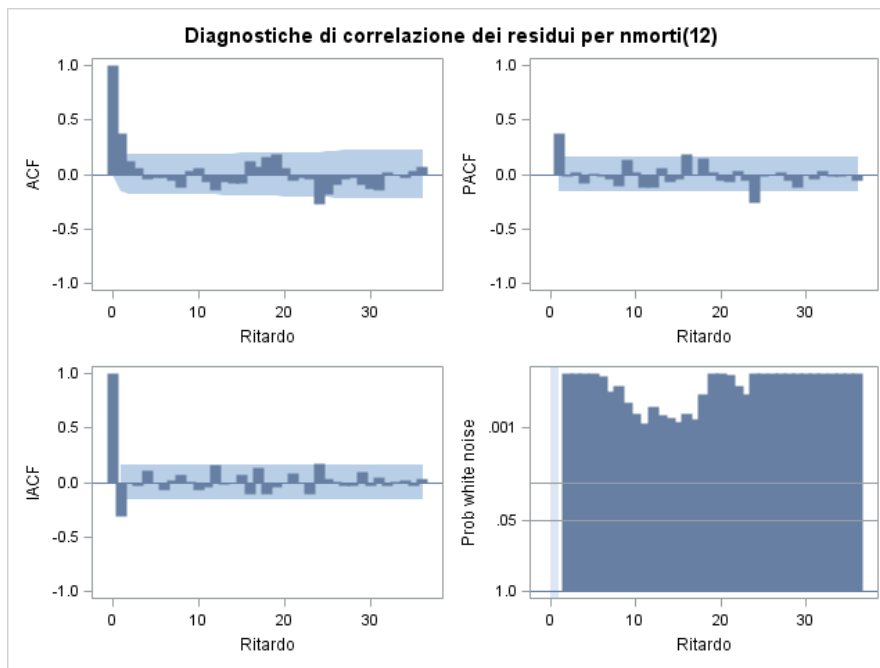
Modello per la variabile nmorti					
Media stimata	26.67615				
Periodi di differenziazione	12				
Stima dei minimi quadrati condizionati					
Parametro	Stima	Errore standard	Valore t	Pr apross >  t	Ritardo
$\mu$	26.67615	13.59940	1.96	0.0516	0
$\phi_1$	-0.49033	0.07640	-6.42	<.0001	12
Stima costante	39.75633				
Stima varianza	61354.11				
Stima errore std	247.6976				
AIC	2164.505				
SBC	2170.605				
Numero residui	156				



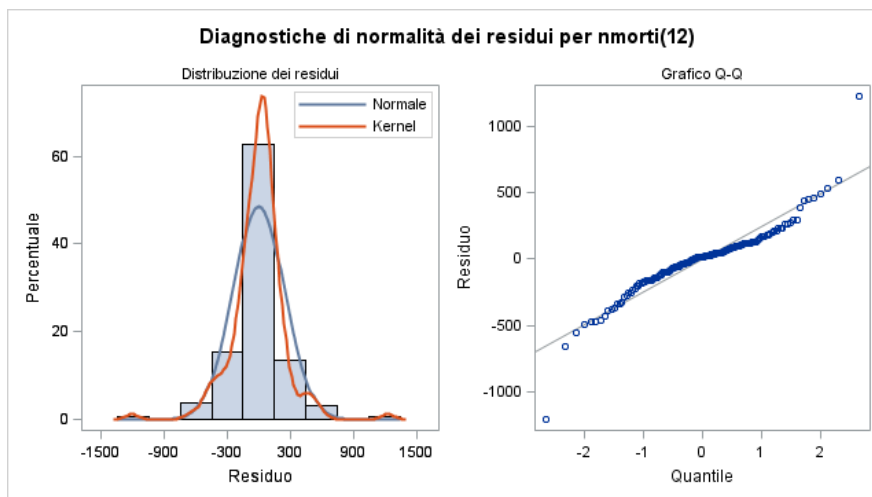
**Tabella C. 2.** Autocorrelazione dei residui, modello SARIMA (0,0,0)x(1,1,0)<sub>12</sub>.

Controllo autocorrelazione dei residui									
Fino al ritardo	Ljung-Box	DF	p-value	Autocorrelazioni					
6	25.48	5	0.0001	0.372	0.121	0.059	-0.038	-0.027	-0.031
12	33.60	11	0.0004	-0.062	-0.124	0.027	0.060	-0.066	-0.143
18	45.00	17	0.0002	-0.066	-0.087	-0.083	0.123	0.069	0.161
24	65.90	23	<.0001	0.182	0.062	-0.052	-0.032	-0.043	-0.266
30	80.13	29	<.0001	-0.188	-0.100	-0.044	-0.027	-0.091	-0.135
36	85.89	35	<.0001	-0.148	0.024	-0.003	-0.033	0.030	0.067

**Figura C.3.** Autocorrelazione dei residui, globale (ACF) e parziale (PACF), modello SARIMA(0,0,0)x(1,1,0)<sub>12</sub>.



**Figura C.4.** Diagnostiche di normalità dei residui, istogramma e qq-plot, modello SARIMA(0,0,0)x(1,1,0)<sub>12</sub>.



L'autocorrelazione dei residui sembra ancora presente, l'ipotesi white noise è rifiutata. Guardando ai grafici delle autocorrelazioni dei residui, suppongo che probabilmente occorre aggiungere un parametro autoregressivo di ordine 1.

2. Secondo modello stimato: SARIMA (1,0,0)x(1,1,0)<sub>12</sub>

**Tabella C.3.** Risultati del secondo modello stimato: SARIMA (1,0,0)x(1,1,0)<sub>12</sub>

<b>Modello per la variabile nmorti</b>	
<b>Media stimata</b>	25.08157
<b>Periodi di differenziazione</b>	12

<b>Stima dei minimi quadrati condizionati</b>					
<b>Parametro</b>	<b>Stima</b>	<b>Errore standard</b>	<b>Valore t</b>	<b>Pr apross &gt;  t </b>	<b>Ritardo</b>
$\mu$	25.08157	16.94880	1.48	0.1410	0
$\phi_1$	0.29036	0.06944	4.18	<.0001	1
$\Phi_2$	-0.42854	0.07409	-5.78	<.0001	12

<b>Stima costante</b>	28.54733
<b>Stima varianza</b>	55426.32
<b>Stima errore std</b>	235.4279
<b>AIC</b>	2149.638
<b>SBC</b>	2158.788
<b>Numero residui</b>	156

**Tabella C.4.** Autocorrelazione dei residui, modello SARIMA (1,0,0)x(1,1,0)<sub>12</sub>

<b>Controllo autocorrelazione dei residui</b>									
<b>Fino al ritardo</b>	<b>Ljung- Box</b>	<b>DF</b>	<b>p-value</b>	<b>Autocorrelazioni</b>					
<b>6</b>	2.25	4	0.6906	0.083	-0.009	0.074	-0.037	0.011	-0.006
<b>12</b>	14.96	10	0.1334	-0.060	-0.146	0.039	0.080	-0.089	-0.185
<b>18</b>	24.39	16	0.0813	0.103	-0.011	-0.098	0.146	-0.022	0.107
<b>24</b>	41.14	22	0.0079	0.179	0.029	-0.084	-0.004	0.049	-0.221
<b>30</b>	48.79	28	0.0088	-0.157	-0.061	-0.024	-0.006	-0.057	-0.089
<b>36</b>	55.98	34	0.0102	-0.140	0.098	0.010	-0.051	0.028	0.059

Il modello non è ancora buono. Ipotizzo quindi un modello SARIMA (1,0,0)x(1,1,1).

3. Terzo modello SARIMA (1,0,0)x(1,1,1)

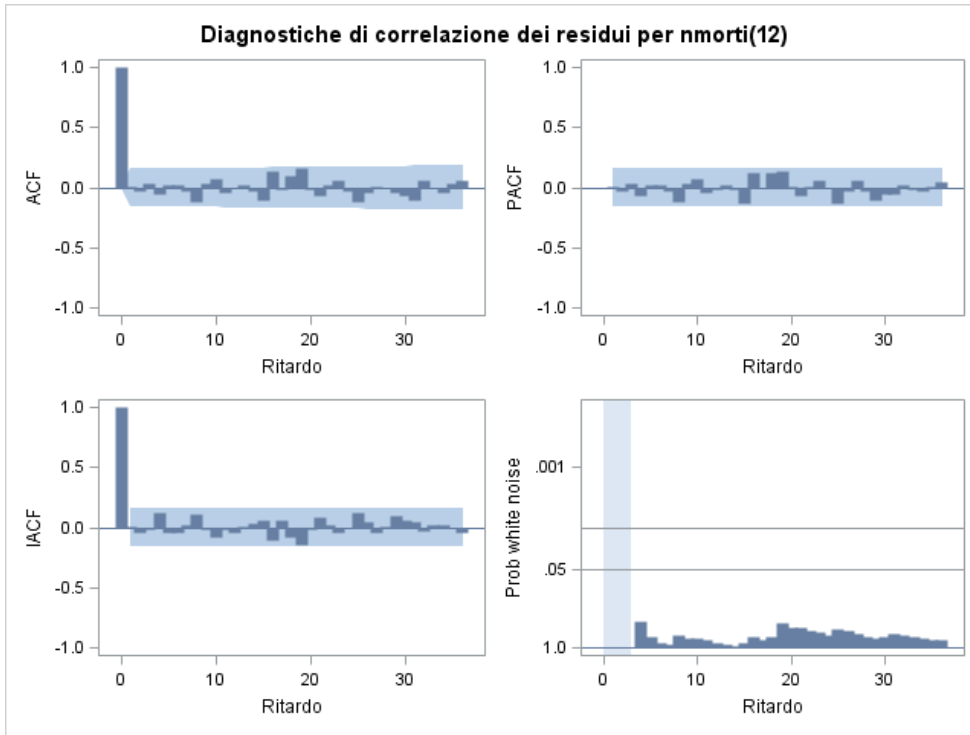
*Tabella C.5. Risultati del terzo modello stimato: SARIMA (1,0,0)x(1,1,1)<sub>12</sub>*

<b>Modello per la variabile nmorti</b>					
<b>Media stimata</b>		24.12259			
<b>Periodi di differenziazione</b>		12			
<b>Stima dei minimi quadrati condizionati</b>					
Parametro	Stima	Errore standard	Valore t	Pr apross >  t	Ritardo
MU	24.12259	11.30097	2.13	0.0344	0
MA1,1	0.58816	0.10338	5.69	<.0001	12
AR1,1	0.35072	0.07599	4.62	<.0001	1
AR1,2	-0.06236	0.11877	-0.53	0.6003	12
<b>Stima costante</b>		17.1666			
<b>Stima varianza</b>		48196.15			
<b>Stima errore std</b>		219.5362			
<b>AIC</b>		2128.81			
<b>SBC</b>		2141.009			
<b>Numero residui</b>		156			

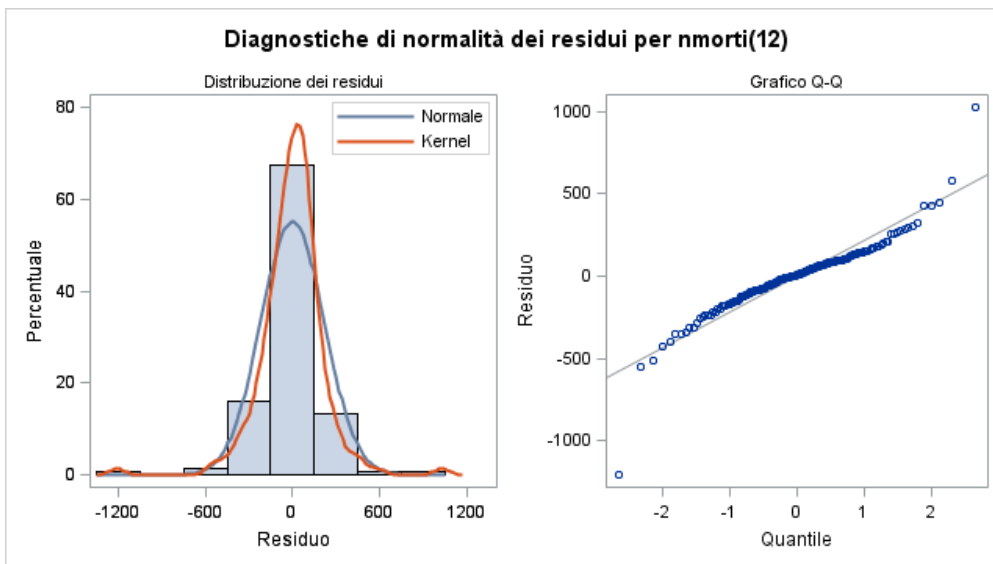
*Tabella C. 6. Autocorrelazione dei residui, modello SARIMA (1,0,0)x(1,1,1)<sub>12</sub>*

<b>Controllo autocorrelazione dei residui</b>									
Fino al ritardo	Ljung-Box	DF	p-value	Autocorrelazioni					
<b>6</b>	0.98	3	0.8067	0.007	-0.024	0.028	-0.062	0.018	0.022
<b>12</b>	5.02	9	0.8321	-0.034	-0.120	0.031	0.076	-0.044	-0.007
<b>18</b>	12.17	15	0.6659	0.014	-0.037	-0.109	0.133	-0.019	0.096
<b>24</b>	18.70	21	0.6042	0.163	-0.013	-0.069	0.014	0.057	-0.033
<b>30</b>	23.23	27	0.6728	-0.123	-0.042	0.003	0.001	-0.043	-0.071
<b>36</b>	27.49	33	0.7379	-0.109	0.058	-0.003	-0.039	0.030	0.061

**Figura C.5.** Autocorrelazione dei residui, globale (ACF) e parziale (PACF), modello SARIMA(0,0,0)x(1,1,1)<sub>12</sub>.



**Figura C.6.** Diagnostiche di normalità dei residui, istogramma e qq-plot, modello SARIMA(0,0,0)x(1,1,0)<sub>12</sub>.



Si osserva che aggiungendo il parametro a media mobile di ordine 12, il parametro autoregressivo al ritardo 12 non è più significativo. Dal controllo dell'autocorrelazione dei residui si vede che l'autocorrelazione sembra essere stata annullata, quindi i residui sembrano essere white noise.

4. Elimino quindi dal modello precedentemente stimato il parametro autoregressivo al ritardo 12, quindi stimo un modello con una differenziazione per la componente stagionale, un parametro autoregressivo e un parametro a media mobile per la componente stagionale:

$$\text{SARIMA}(1,0,0)\times(0,1,1)$$

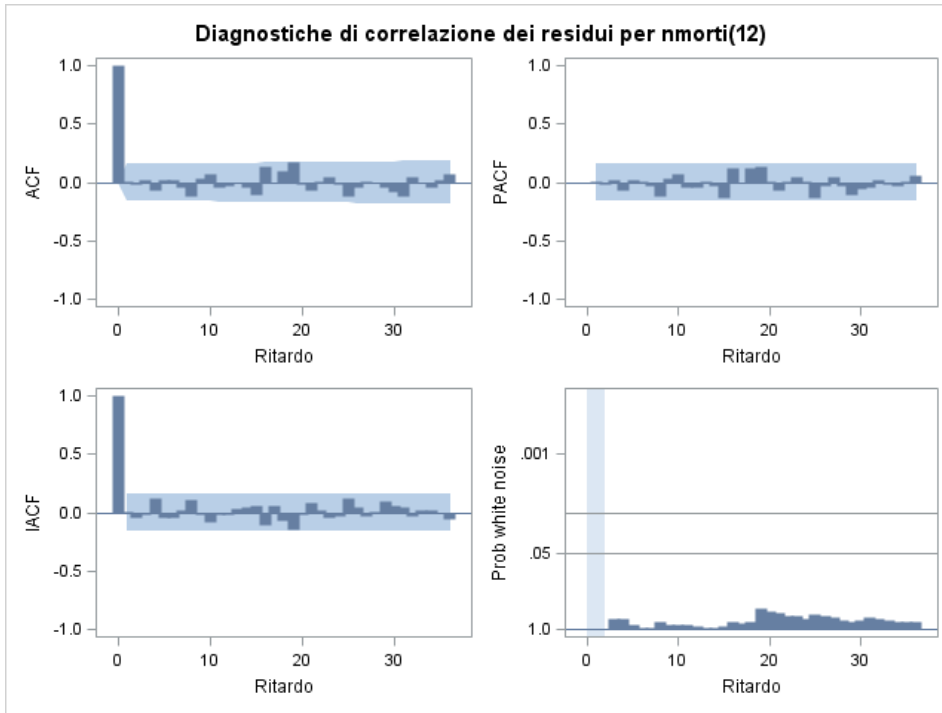
**Tabella C. 7.** Risultati del quarto modello stimato SARIMA(1,0,0)x(0,1,1).

<b>Modello per la variabile nmorti</b>					
<b>Media stimata</b>		23.78072			
<b>Periodi di differenziazione</b>		12			
<b>Stima dei minimi quadrati condizionati</b>					
<b>Parametro</b>	<b>Stima</b>	<b>Errore standard</b>	<b>Valore t</b>	<b>Pr apross &gt;  t </b>	<b>Ritardo</b>
<b>MU</b>	23.78072	11.39222	2.09	0.0385	0
<b>MA1,1</b>	0.62365	0.06464	9.65	<.0001	12
<b>AR1,1</b>	0.35067	0.07573	4.63	<.0001	1
<b>Stima costante</b>		15.44157			
<b>Stima varianza</b>		47986.1			
<b>Stima errore std</b>		219.0573			
<b>AIC</b>		2127.152			
<b>SBC</b>		2136.301			
<b>Numero residui</b>		156			

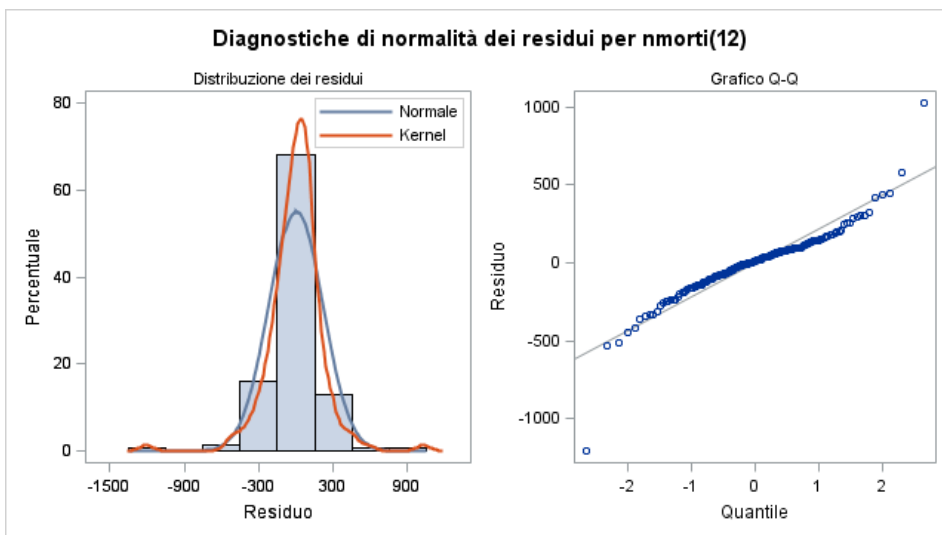
**Tabella C.8.** Autocorrelazione dei residui, modello SARIMA(1,0,0)x(0,1,1).

<b>Controllo autocorrelazione dei residui</b>									
<b>Fino al ritardo</b>	<b>Ljung-Box</b>	<b>DF</b>	<b>p-value</b>	<b>Autocorrelazioni</b>					
<b>6</b>	1.02	4	0.9072	0.008	-0.024	0.025	-0.066	0.022	0.015
<b>12</b>	5.01	10	0.8904	-0.039	-0.115	0.031	0.073	-0.045	-0.027
<b>18</b>	12.23	16	0.7281	-0.002	-0.041	-0.110	0.131	-0.011	0.098
<b>24</b>	18.83	22	0.6556	0.169	-0.014	-0.067	0.013	0.052	-0.015
<b>30</b>	23.53	28	0.7061	-0.123	-0.042	0.007	-0.002	-0.043	-0.076
<b>36</b>	28.07	34	0.7530	-0.114	0.051	-0.005	-0.038	0.026	0.070

**Figura C.7.** Autocorrelazione dei residui, globale (ACF) e parziale (PACF), modello SARIMA(1,0,0)x(0,1,1).



**Figura C.8.** Diagnostiche di normalità dei residui, istogramma e qq-plot, modello SARIMA(1,0,0)x(0,1,1).



Confrontando i valori di AIC e BIC quest'ultimo modello sembra essere il migliore; anche l'autocorrelazione dei residui lo conferma, in quanto l'ipotesi di residui white noise non può essere rigettata. Inoltre le diagnostiche di normalità dei residui non sembrano condurre al suo rifiuto.

# APPENDICE D

## RISULTATI MODELLO 1

**Tabella D.1. Modello 1. Stime dei parametri lineari per classe d'età, PIANURA.**

\*\*\* significativo all'1%, \*\* significativo al 5%, \*significativo al 10%.

parametro	classe d'età					
	0-14	15-44	45-64	65-74	75-84	85+
Intercetta	-12,047 ***	-13,152***	-11,695***	-10,554***	-9,553***	-8,334***
Friday vs Wednesday	0,011	0,142**	0,004	-0,054*	-0,061***	-0,046**
Monday vs Wednesday	0,006	0,062	-0,026	-0,007	-0,015	0,013
Sunday vs Wednesday	0,028	0,081	0,055	-0,016	-0,008	-0,008
Tuesday vs Wednesday	-0,008	0,051	-0,027	-0,007	-0,019	-0,017
Saturday vs Wednesday	0,051	0,214***	-0,066***	-0,067**	-0,068***	-0,026
Thursday vs Wednesday	0,013	0,041	-0,008	-0,003	0,018	-0,024
secco Maschio vs Femmina	-0,100	0,249***	0,563***	0,741***	0,591***	0,224***
anno 2000 vs 2003	0,138	0,063	-0,037	0,117***	0,000	-0,055*
anno 2001 vs 2003	0,082	0,088	0,075	0,087**	-0,006	-0,036
anno 2002 vs 2003	0,093	0,004	-0,011	0,022	-0,009	-0,082***
anno 2004 vs 2003	0,036	0,006	-0,127**	-0,054	-0,06**	-0,025
anno 2005 vs 2003	-0,036	-0,085	-0,100**	-0,134***	-0,077***	-0,034
anno 2006 vs 2003	-0,079	-0,011	-0,119**	-0,179***	-0,081***	0,001
anno 2007 vs 2003	-0,007	-0,052	-0,193***	-0,196***	-0,112***	0,005
anno 2008 vs 2003	-0,059	-0,094	-0,206***	-0,176***	-0,180***	-0,067**
anno 2009 vs 2003	-0,056	-0,164	-0,260***	-0,246***	-0,185***	-0,124***
anno 2010 vs 2003	-0,138	-0,145	-0,273***	-0,235***	-0,130***	-0,073***
anno 2011 vs 2003	-0,177	-0,090	-0,272***	-0,232***	-0,184***	-0,076***
anno 2012 vs 2003	-0,051	-0,171	-0,265***	-0,289***	-0,219***	-0,130***
anno 2013 vs 2003	-0,135	-0,152	-0,331***	-0,328***	-0,233***	-0,063**
Linear(doy)	0	0	-0,001***	-0,001***	-0,001***	-0,001***
Linear(hum_ma)	0,003	0,009*	0,003	0,005***	0,012***	0,016***

**Tabella D.2. Modello 1. Stime dei parametri lineari per classe d'età, PEDEMONTANA.**

\*\*\* significativo all'1%, \*\* significativo al 5%, \*significativo al 10%.

parametro	classe d'età					
	0-14	15-44	45-64	65-74	75-84	85+
Intercetta	-11,125***	-12,271***	-11,467***	-10,312***	-9,524***	-8,387***
Friday vs Wednesday	-0,036	-0,036	-0,016	-0,050	-0,035	0,073**
Monday vs Wednesday	0,005	-0,083	0,021	0,028	0,013	0,071**
Sunday vs Wednesday	0,053	0,006	0,009	-0,013	0,018	0,102***
Tuesday vs Wednesday	0,056	-0,041	-0,003	-0,007	-0,005	0,076**
Saturday vs Wednesday	0,021	-0,017	0,033	-0,056	-0,022	0,051
Thursday vs Wednesday	0,012	-0,112	0,035	-0,034	-0,012	0,096***
sexo Maschio vs Femmina	-0,060	0,056	0,306***	0,501***	0,562***	0,343***
anno 2000 vs 2003	0,144	0,008	0,067	0,069	-0,008	-0,145***
anno 2001 vs 2003	0,027	0,004	0,046	0,000	-0,024	-0,151***
anno 2002 vs 2003	0,313	0,084	0,017	0,007	-0,062	-0,146***
anno 2004 vs 2003	-0,032	-0,018	0,017	-0,082	-0,067	-0,140***
anno 2005 vs 2003	-0,042	-0,005	-0,071	-0,186***	-0,137***	-0,066
anno 2006 vs 2003	-0,071	0,007	-0,026	-0,149**	-0,133***	-0,002
anno 2007 vs 2003	-0,001	-0,045	-0,117	-0,262***	-0,089***	-0,101**
anno 2008 vs 2003	-0,120	-0,079	-0,123	-0,219***	-0,236***	-0,114**
anno 2009 vs 2003	-0,128	-0,096	-0,110	-0,220***	-0,180***	-0,146***
anno 2010 vs 2003	-0,123	-0,086	-0,125	-0,303	-0,204***	-0,145***
anno 2011 vs 2003	-0,116	-0,040	-0,269***	-0,268***	-0,217***	-0,188***
anno 2012 vs 2003	-0,111	0,045	-0,233***	-0,317***	-0,304***	-0,184***
anno 2013 vs 2003	-0,112	-0,035	-0,195**	-0,329***	-0,288***	-0,231***
Linear(doy)	0,000	-0,001	-0,001	-0,001**	0,000	-0,001***
Linear(hum_ma)	0,000	0,010	0,008**	0,011***	0,013***	0,018***



**Tabella D.3. Modello 1. Stime dei parametri lineari per classe d'età, MONTANA.**

\*\*\* significativo all'1%, \*\* significativo al 5%, \*significativo al 10%.

parametro	classe d'età					
	0-14	15-44	45-64	65-74	75-84	85+
Intercetta		-11,063***	-10,361***	-9,479***	-9,067***	-8,021***
Friday vs Wednesday		0,058	-0,027	0,013	0,024	-0,007
Monday vs Wednesday		-0,057	-0,061	0,004	0,025	-0,013
Sunday vs Wednesday		0,000	0,013	0,022	-0,027	0,013
Tuesday vs Wednesday		0,100	0,008	0,012	0,046	-0,034
Saturday vs Wednesday		0,085	-0,057	-0,024	-0,014	-0,067
Thursday vs Wednesday		0,056	0,029	0,075	-0,009	-0,002
secco Maschio vs Femmina		0,104	0,109*	0,358***	0,598***	0,643***
anno 2000 vs 2003		0,193	-0,043	0,003	0,120	-0,046
anno 2001 vs 2003		0,152	0,066	-0,026	0,073	-0,094
anno 2002 vs 2003	numero di oss. inferiore al numero dei parametri	0,037	-0,018	0,088	0,040	-0,080
anno 2004 vs 2003		0,044	0,010	-0,041	-0,010	-0,035
anno 2005 vs 2003		0,102	0,014	-0,014	-0,001	0,130
anno 2006 vs 2003		0,029	-0,142	-0,070	-0,017	0,085
anno 2007 vs 2003		0,117	-0,089	0,001	-0,043	-0,052
anno 2008 vs 2003		0,195	-0,112	-0,025	-0,013	-0,088
anno 2009 vs 2003		0,160	-0,045	-0,130	-0,054	-0,186**
anno 2010 vs 2003		0,107	-0,178	-0,050	-0,071	-0,047
anno 2011 vs 2003		0,370	-0,073	-0,122	-0,044	-0,128*
anno 2012 vs 2003		0,392	-0,124	-0,095	-0,097	-0,081
anno 2013 vs 2003	0,229	-0,093	-0,175	-0,059	-0,050	
Linear(doy)		0,000	0,000	-0,001	-0,001	-0,001***
Linear(hum_ma)		0,005	0,003	0,003	0,006*	0,014***

**Tabella D. 4. Modello 1. Stime dei parametri lineari per classe d'età, PIANURA.**

\*\*\* significativo all'1%, \*\* significativo al 5%, \*significativo al 10%.

parametro	classe d'età					
	0-14	15-44	45-64	65-74	75-84	85+
Intercetta	-10,707***	-12,192***	-11,359***	-10,328***	-9,434***	-8,107***
Friday vs Wednesday	-0,010	-0,077	0,009	-0,006	-0,025	-0,057*
Monday vs Wednesday	-0,001	-0,067	-0,017	0,027	-0,005	0,019
Sunday vs Wednesday	0,065	-0,070	0,066	0,001	-0,008	-0,006
Tuesday vs Wednesday	0,051	-0,081	0,060	0,011	0,010	0,008
Saturday vs Wednesday	-0,029	0,026	0,041	-0,023	-0,059*	-0,042
Thursday vs Wednesday	0,054	-0,075	-0,007	-0,029	0,034	0,000
secco Maschio vs Femmina	-0,077	0,054	0,406***	0,554***	0,536***	0,330***
anno 2000 vs 2003	0,011	0,117	0,116	0,155**	0,023	-0,126**
anno 2001 vs 2003	0,115	0,112	0,040	0,159**	-0,048	-0,112**
anno 2002 vs 2003	-0,072	0,026	0,098	0,135**	-0,008	-0,115**
anno 2004 vs 2003	-0,102	-0,014	-0,004	0,033	-0,065	-0,073
anno 2005 vs 2003	-0,117	-0,031	0,014	-0,107	-0,050	-0,088*
anno 2006 vs 2003	-0,117	0,048	-0,043	-0,073	-0,132***	0,013
anno 2007 vs 2003	-0,141	0,027	-0,062	-0,013	-0,116**	-0,081*
anno 2008 vs 2003	-0,155	0,086	-0,003	-0,181***	-0,141***	-0,118**
anno 2009 vs 2003	-0,176	0,058	-0,108	-0,145**	-0,198***	-0,138***
anno 2010 vs 2003	-0,208	-0,001	-0,105	-0,086	-0,157***	-0,095**
anno 2011 vs 2003	-0,182	-0,004	-0,173**	-0,192***	-0,203***	-0,146***
anno 2012 vs 2003	-0,208	0,087	-0,171**	-0,181***	-0,265***	-0,160***
anno 2013 vs 2003	-0,181	0,137	-0,193**	-0,136**	-0,294***	-0,126***
Linear(doy)	0,000	0,000	-0,001**	0,000	0,000	0,000
Linear(hum_ma)	-0,002	0,008	0,004	0,004	0,010***	0,008***

# BIBLIOGRAFIA

---

Analitis A, Katsouyanni K, Biggeri A, Baccini M, Forsberg B, Bisanti L, Kirchmayer U, Ballester F, Cadum E, Goodman P. G, Hojs A, Sunyer J, Tiittanen P, Michelozzi P (2008). "Effect of cold Weather on Mortality: Results From 15 European Cities Within the PHEWE Project". *American Journal of Epidemiology*, vol 168, No. 12.

Azzalini A, Scarpa B (2012). "Data Analysis and Data Mining: an introduction". New York: Oxford University Press.

Burkart K, Khan M. H, Krämer A, Breitner S, Schneider A, Endlicher W. R (2011). "Seasonal variation of all-cause and cause-specific mortality by age, gender, and socioeconomic condition in urban and rural areas of Bangladesh". *International Journal for Equity in Health*, 10:32.

Conti S, Meli P, Minelli G, Solimini R, Toccaceli V, Vichi M, Beltrano C, Perini L (2004). "Epidemiologic Study of mortality during the Summer 2003 heat wave in Italy". *Environmental research* 98: 390-399.

Davis R. E, Knappenberger PC, Michaels PJ, Novicoff WM (2004). "Seasonality of climate-human mortality relationship in US cities and impacts of climate change". *Clim Res* 26: 61-76.

Davis R. E, Rossier C. E, Enfield K. B (2012). "The Impact of Weather on Influenza and Pneumonia Mortality in New York City, 1975–2002: A Retrospective Study". *PLoS ONE* 7(3): e34091.

Di Fonzo T, Lisi F (2005). "Serie Storiche Economiche". Roma: Carocci.

Fleiss Joseph L (1981). "Statistical methods for rates and proportions". New York: Wiley & Sons.

Gasparri A, et al. (2015). "Mortality risk attributable to high and low ambient temperature: a multicountry observational study". *Lancet* 386: 369-75.

Hastie T, Tibshirani R, Friedman J (2009). "The Elements of Statistical Learning, Data Mining, Inference, and Prediction". New York: Springer.

Hastie T, Tibshirani R.J. (1990). "Generalized additive models". London: Chapman and Hall.

Kalkstein A. J (2013). "Regional Similarities in Seasonal Mortality across the United States: An Examination of 28 Metropolitan Statistical Areas". *PLOS ONE*, 8(5): e63971.

Lubczyńska M. J, Christophi A. C, Lelieveld J (2015). "Heat-related cardiovascular mortality risk in Cyprus: a case-crossover study using a distributed lag non-linear model". *Environmental Health*, 14:39.

Martin-Soler H, Gonseth S, Gubelmann C, Stringhini S, Bovet P, Chen P-C, Wojtyniak B, Paccaud F, Tsai D-H, Zdrojewski T, Marques-Vidal P (2014). "Seasonal Variation of Overall and Cardiovascular Mortality: A Study in 19 Countries from Different Geographic Locations". *PLOS ONE*, 9(11): e113500.

Nielsen J, Mazick A, Steffen G, Molbak (2011). "Excess mortality related to seasonal influenza and extreme temperatures in Denmark, 1994-2010". *BMC Infectious Diseases*, 11:350.

Nordio F, Zanobetti A, Colicino E, Kloog I, Schwartz J. (2015) "Changing patterns of the temperature-mortality association by time and location in the US, and implications for climate change". *Environmental International* 81: 80-86.

Oudin Astrom D, Bertil F, Joacim R (2011). "Heat wave impact on morbidity and mortality in elderly population: a review of recent studies". *Maturitas* 69: 99-105.

Oudin Astrom D, Schifano P, Asta F, Lallo A, Michelozzi P, Rocklöv J, Forsberg B (2015). "The effect of heat waves on mortality in susceptible groups: a cohort study of a mediterranean and a northern European City". *Environmental health*, 14:30.

Rothman K. J, Greenland S (1998). "Modern Epidemiology" (2nd ed.). Philadelphia: Lippincott-Raven.

Saez M, Sunyer J, Castellsagué J, Murillo C, Mantò J (1995). "Relationship between Weather Temperature and Mortality: a Time Series Analysis Approach in Barcelona". *International Journal of Epidemiology*. Vol. 24, No. 3.

Schifano P, Cappai G, De Sario M, Michelozzi P, Marino C, Bargagli A.M, Perucci C (2009). "Susceptibility to heat wave-related mortality: a follow-up study of a cohort of elderly in Rome". *Environmental Health*, 8:50.

Schifano P, Leone M, De Sario M, de'Donato F, Bargagli A.M, D'Ippoliti D, Marino C, Michelozzi P (2012). "Changes in the effects of heat on mortality among the elderly from 1998–2010: results from a multicenter time series study in Italy". *Environmental Health*, 11:58.

Seposo X. T, Dang T. N, Honda Y (2015). "Evaluating the Effects of Temperature on Mortality in Manila City (Philippines) from 2006-2010 Using a Distributed Lag Nonlinear Model". *Int. J. Environ. Res. Public Health*, 12: 6842-6857.

Von Klot S, Zanobetti A, Schwartz J (2012). Influenza epidemics, seasonality, and effects of cold weather on cardiac mortality. *Environmental Health*, 11:74.

Wagner A. P, McKenzie E, Robertson C, McMenemy J, Reynolds A, Murdoch H (2013). "Automated mortality monitoring in Scotland from 2009". *Euro Surveill*, 18 (15).

Xuan, Egondi, Ngoan, Toan, Huong (2013). "Seasonality in mortality and its relationship to temperature among the older population in Hanoi, Vietnam". *Global Health Action*, 7:23115.

# SITOGRAFIA

---

“Bur n. 46 del 08 maggio 2015” in

<http://bur.regione.veneto.it/BurvServices/pubblica/DettaglioDgr.aspx?id=297456>.

“Rapporto ondate di calore estate 2013” in

[http://www.deplazio.net/it/rapporti/cat\\_view/45-rapporti-2013](http://www.deplazio.net/it/rapporti/cat_view/45-rapporti-2013).

“L’indice Humidex (H)” in <http://www.centrometeo.com/articoli-reportage-appfondimenti/climatologia/5008-indici-humidex-temperatura-equivalente>.

## RINGRAZIAMENTI

Desidero ringraziare i miei genitori che mi hanno sostenuta fin qui e mi hanno permesso di raggiungere questo traguardo. Insieme a loro ringrazio Andrea, le mie sorelle, i miei amici e tutta la mia famiglia della quale sono molto fiera ed orgogliosa. Un grazie a tutti i professori del Dipartimento di Scienze Statistiche che in questi anni ho conosciuto, in particolare ringrazio la professoressa Boccuzzo che mi ha guidata in questi mesi di tesi e la professoressa Bisaglia per il gentile contributo.

Infine un grazie speciale a tutto lo staff del Sistema Epidemiologico Regionale per avermi ben accolta nei quattro mesi di stage; grazie per la disponibilità e per l'esperienza che mi avete dato modo di fare. Ringrazio in particolare il dott. Fedeli e il dott. Avossa, che hanno seguito con attenzione le elaborazioni, e tutti gli statistici del SER per i preziosi consigli. Da ultimo, ma non per importanza, ringrazio i miei colleghi d'ufficio per il supporto che non è mai mancato.

Infine un grazie al dott. Veronese dell'ARPAV per la disponibilità e la collaborazione nella realizzazione di questo lavoro.

GRAZIE di cuore a tutti.