

Indice

1	Introduzione	3
2	Panoramica sui modelli già esistenti	5
3	Presentazione modello	7
4	Applicazione al dominio a forma di C	9
5	Confronto con gli altri modelli	11
6	Produzione di rifiuti nella provincia di Venezia	13
6.1	L'inclusione dell'effetto del turismo	14
6.2	Trattamento del dominio	14
6.2.1	Regression splines	15
6.2.2	Scelte particolari tra i comuni	15
6.2.3	Triangolazione del dominio	16
6.3	Applicazione del modello senza covariate	17
6.3.1	Analisi preliminare e ricerca del miglior λ	17
6.3.2	Risultati	18
6.4	Applicazione del modello con covariate	20
6.4.1	Analisi preliminare e ricerca del miglior λ	20
6.4.2	Risultati	21
7	Conclusioni e sviluppi futuri	23

Capitolo 1

Introduzione

Il presente lavoro di tesi illustra un modello statistico spatio-temporal spline regression (STR-PDE) per l’analisi funzionale di dati distribuiti in spazio e tempo. Quanto fatto può essere considerato un’estensione dei modelli proposti in [6], che studiano la possibilità di costruire una stima funzionale per dati distribuiti su un dominio spaziale attraverso l’approssimazione in basi di elementi finiti. Il modello STR-PDE, invece, sviluppa una tecnica analoga permettendo la variazione temporale alla stima funzionale precedente. Di conseguenza, può essere considerato un buon strumento per lo studio di fenomeni varianti in spazio e in tempo. Dalla modellizzazione matematica è stato sviluppato un algoritmo e il codice R per il calcolo della soluzione numerica della stima.

Il lavoro è motivato dalla ricerca di un buon metodo di analisi di un dataset contenente le misurazioni della produzione dei rifiuti urbani pro capite nei comuni della provincia di Venezia tra il 1997 e il 2011. I dati sono stati raccolti ed elaborati dall’Agenzia Regionale per la Prevenzione e Protezione Ambientale del Veneto (Arpav) e sono disponibili sul sito di Open Data Veneto¹ per la consultazione e il trattamento. Sono disponibili le misurazioni per tutta la regione del Veneto, ma per semplicità computazionale e per l’elevato interesse della laguna veneta sarà analizzata solo la provincia di Venezia. Il modello STR-PDE permette di stimare l’andamento della produzione dei rifiuti su tutta la regione e ad ogni istante di tempo nell’intervallo considerato, garantendo una chiara visualizzazione del fenomeno.

Il lavoro di tesi sarà strutturato come segue. Nel Capitolo 2 è riportato un excursus sui metodi simili già esistenti in letteratura. Nel Capitolo 3 è presentata la costruzione del modello matematico STR-PDE. Nel Capitolo 4 si hanno i primi risultati, derivanti dall’applicazione del modello e del codice R al caso del dominio a forma di C descritto in [5] e [7], per il quale è possibile valutare la bontà delle stime ottenute grazie alla perfetta

¹<http://dati.veneto.it/dataset/produzione-annua-di-rifiuti-urbani-totale-e-pro-capite-1997-2011>

conoscenza del fenomeno reale in ogni punto e in ogni istante. Nel capitolo 5 il modello STR-PDE è paragonato ad altri metodi già esistenti per il confronto delle stime ottenute. Nel Capitolo 6 si ha l'applicazione allo studio della produzione dei rifiuti nella provincia di Venezia, e infine nel Capitolo 7 sono raccolte le conclusioni e i possibili sviluppi futuri.

Capitolo 2

Panoramica sui modelli già esistenti

Prima dello sviluppo del modello STR-PDE, è stato necessario uno studio dell'ambito in cui si inserisce il modello. Sono già presenti in letteratura tecniche per l'analisi funzionale di dati di questo tipo.

Capitolo 3

Presentazione modello

Capitolo 4

Applicazione al dominio a forma di C

Capitolo 5

Confronto con gli altri modelli

Il modello STSR rappresenta una generalizzazione del caso puramente spaziale proposto in [6] e, come già evidenziato nel Capitolo 2, non è l'unico modello disponibile per l'analisi di dati distribuiti sia in spazio che in tempo. Pertanto è necessario che sia valutata la bontà del modello attraverso il confronto con altri metodi esistenti e studiati in letteratura, al fine di poter dire se e quanto il modello STR-PDE possa rappresentare un miglioramento in questo campo.

L'articolo [1] propone l'analisi di dati di questo tipo attraverso modelli misti additivi generalizzati (GAMM) di interazione spazio-tempo. Questo metodo è generalizzato, quindi può essere usato per spiegare anche funzioni del valore atteso della risposta. Nel nostro caso, per avvicinarci al caso STR-PDE, si ipotizza che la risposta sia pari alla somma di una funzione e di un eventuale termine con covariata. Alla funzione è associato lo smoothing secondo il prodotto tensoriale dei termini marginali in spazio e tempo con le loro penalizzazioni. Quindi la costruzione dei GAMM è molto simile a quella analizzata in STR-PDE, e grazie al codice implementato nel pacchetto R *mgcv* è possibile scegliere tra più tipi di modelli. In particolare ne saranno studiati due, i più simili al modello STR-PDE:

- TPS, in cui sono poste marginalmente *cubic regression splines* in tempo e *thin plate splines* in spazio;
- SOAP, che considera *cubic regression splines* in tempo e *soap film smoothing* in spazio.

Un altro metodo da confrontare è sicuramente il kriging (KRIG) spazio-temporale. Le stime sono ottenute fissando un variogramma separabile e marginalmente esponenziale in spazio e tempo. I parametri del variogramma sono stimati dal variogramma empirico, e successivamente è possibile eseguire la stima grazie al codice implementato nel pacchetto R *spacetime*.

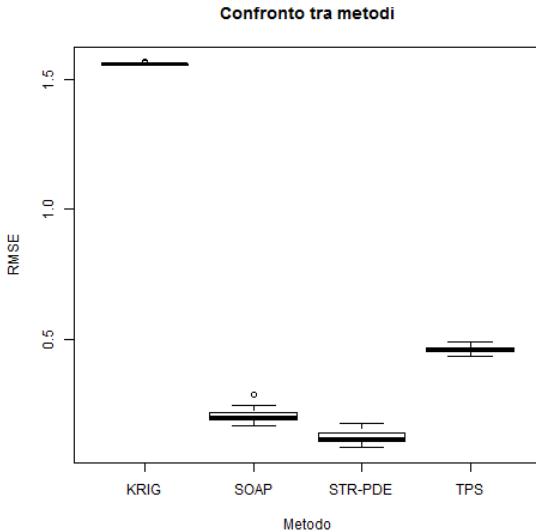


Figura 5.1: Confronto del RMSE, caso senza covariate

I quattro modelli sono confrontati sull'esempio del dominio a forma di C proposto precedentemente, poiché garantisce di poter conoscere in ogni punto spaziale e ad ogni istante temporale il valore esatto della funzione. La triangolazione e i dati sono ricavati esattamente allo stesso modo di quanto fatto nel Capitolo 4. In aggiunta è stata costruita una griglia spazio-temporale di punti per la validazione: sono stati presi 80 punti equispaziati in $(-1, +3.5)$ per l'ascissa, 40 punti in equispaziati $(-1, +1)$ per l'ordinata e 20 istanti in $(0, 2\pi)$ per il tempo. Ovviamente la validazione è stata studiata soltanto sui punti che ricadevano all'interno del dominio a forma di C.

I modelli sono stati confrontati attraverso il Root Mean Square Error (RMSE) prodotto sui punti di validazione. Quindi se V è l'insieme dei punti della griglia interni al dominio, e Mod rappresenta la stima ottenuta dal modello, si avrà:

$$\text{RMSE}_V(\text{Mod}) = \sqrt{\frac{\sum_{(\underline{p}_i, t_i) \in V} (\text{Mod}(\underline{p}_i, t_i) - g(\underline{p}_i) \cos(t_i))^2}{\text{card}(V)}}$$

Il procedimento è stato iterato 50 volte, per poter escludere possibili andamenti particolari dovuti alla generazione del rumore.

Nel caso senza covariate si hanno i risultati riportati in figura 5.1. Sono stati tracciati i boxplot dei valori di RMSE calcolati nelle 50 iterazioni per ogni metodo. Si hanno alti valori per KRIG e TPS (e questo fa presupporre una stima non buona), ma il modello che commette meno errore è STR-PDE.

AGGIUNTA DEI GRAFICI..

Capitolo 6

Produzione di rifiuti nella provincia di Venezia

Come è già stato brevemente accennato in precedenza, l'applicazione scelta per il modello STR-PDE riguarda i dati della produzione di rifiuti urbani nel periodo di anni dal 1997 al 2011 nella provincia di Venezia. Per rifiuti urbani si intendono rifiuti domestici, prodotti in locali, aree pubbliche, parchi o giardini, spiagge o provenienti dalla pulizia delle strade o di altri luoghi pubblici. Non sono conteggiati i rifiuti speciali (tra cui ad esempio gli industriali, agricoli o provenienti da attività commerciali o di costruzione) o pericolosi (per i quali esistono programmi di smaltimento particolari).

In realtà i dati raccolti dall'Agenzia Regionale per la Prevenzione e Protezione Ambientale del Veneto (Arpav) riguardano tutto il Veneto. Tuttavia è stata analizzata solo la provincia di Venezia per due motivi. Inanzitutto l'interesse particolare per la zona della laguna veneta, in cui è rilevante il ruolo del turismo (come si potrà notare in seguito). Inoltre considerare tutto il Veneto aumenta notevolmente le dimensioni delle matrici in gioco e causa una grossa spesa computazionale per la ricerca della soluzione. Quindi è stato scelto di concentrarsi su un dominio più piccolo ma nel quale è possibile notare più facilmente le particolarità dell'andamento del fenomeno della produzione dei rifiuti urbani.

Per ogni comune della provincia di Venezia e per ogni anno è disponibile il numero di rifiuti totali raccolti in tonnellate e la popolazione residente. La popolazione è certamente un valore influente per la produzione di rifiuti, perciò la quantità di riferimento non sarà il valore dei rifiuti totali raccolti in ogni anno per comune, ma il valore pro capite.

Le coordinate spaziali dei comuni sono la longitudine e la latitudine, disponibili on line¹. Nel caso dei comuni con dato replicato (sez 6.2.2), sono state scaricate da Google Maps.

¹<http://www.dossier.net/utilities/coordinate-geografiche/>

6.1 L'inclusione dell'effetto del turismo

L'inclusione della popolazione residente nella risposta tramite la scelta di usare i valori pro capite è necessaria, poichè permette di depurare la risposta da una variabile che per sua natura la influenzerebbe. Ma sarebbe un errore fermarsi solo alla popolazione residente, poichè anche i turisti rappresentano una componente non trascurabile di produzione di rifiuti urbani.

Nella provincia di Venezia sono presenti molte zone di elevata attrazione turistica. La più rilevante di queste è Venezia, ma si hanno anche zone balneari (come Lido di Venezia, Cavallino-Treporti, Jesolo, San Michele al Tagliamento, Bibione, ecc...). L'informazione scelta per sintetizzare l'attività turistica è il numero di posti letto totali del comune, valore disponibile grazie all'applicativo dell'Istat *Atlante Statistico dei Comuni*² per ogni anno. Il totale dei posti letto per comune è la somma di vari tipi di attività non solamente alberghiere (ad esempio sono conteggiati anche esercizi complementari, bed & breakfast, campeggi) e saranno considerati normalizzati per la popolazione residente per uniformità con la risposta. I valori ricavati saranno inseriti nel modello come possibile covariata.

6.2 Trattamento del dominio

Per poter studiare il problema a livello computazionale occorre avere una buona approssimazione della frontiera della regione. Questa è disponibile nel pacchetto R *raster* che descrive dati geografici di moltissime zone del mondo sia a livello nazionale che locale (nel caso italiano province e comuni) tramite poligoni molto precisi.

Una volta scaricata la provincia di Venezia si è riscontrato subito un problema: la regione è composta da un insieme di 101 poligoni distinti (a causa delle numerose isole di cui è composta la laguna di Venezia) e ogni poligono ha un alto numero di vertici (ad esempio, la prima delle due regioni corrispondenti all'entroterra aveva 10538 vertici). Non è possibile analizzare il problema su un territorio così descritto, perciò è stata necessaria una analisi iniziale della frontiera per ridurne la complessità.

Oltre all'entroterra (composto da due poligoni) stati scelti solo le più importanti isole della laguna veneta: Venezia, Murano, Lido di Venezia e Pellestrina (più rilevanti a livello di popolazione e turismo). I poligoni sono stati semplificati in tutti i casi e uniti tra loro con ponti dove era possibile. Tra le isole collegate solamente via mare con il resto del territorio sono state simulati ponti in corrispondenza delle trafficate linee di trasporto pubblico con traghetto.

²<http://www.istat.it/it/archivio/113712>

6.2.1 Regression splines

Per ridurre l'elevato numero di vertici di ognuno dei poligoni considerati si è scelto di ricorrere ad un'analisi di smoothing di dati funzionali. Ad ogni poligono è associata una coppia di funzioni: la latitudine e la longitudine rispetto all'ascissa curvilinea (disponibili per punti, corrispondenti ai vertici) che sono state rappresentate in basi e valutate in un numero molto inferiore di punti, con i quali si è costruita la nuova definizione della regione.

Per avere una rappresentazione tramite funzioni di base di queste funzioni sono state provate più tecniche, ma la scelta definitiva è ricaduta sulle *Regression Splines* cubiche senza penalizzazione della derivata seconda. Infatti i risultati non sono stati migliori negli altri casi a causa della zona interna alla laguna di Venezia, fortemente frastagliata: penalizzare la derivata seconda eliminava troppe asperità presenti sulle coste del territorio, mentre con *Kernel Smoothing* sono state ricavate regioni che, dopo la triangolazione, presentavano troppi triangoli composti solamente da punti di frontiera (e quindi senza dati) rispetto agli altri metodi.

Una volta fissato un ragionevole numero di punti con cui descrivere la regione sono stati eseguiti più tentativi per decidere il miglior numero di basi necessario per descrivere le due funzioni. Il criterio di scelta è stato complesso, poichè sono stati esclusi i valori che generavano intersezioni nella nuova descrizione della regione e comuni esterni alla frontiera. Ma la scelta del miglior numero di basi per *Regression Splines* è ricaduta sul valore che, una volta eseguito lo smoothing della regione, causava la minor distanza tra i nuovi punti della regione e il poligono iniziale. In figura 6.1 è riportato il risultato dello smoothing sul primo poligono che descrive l'entroterra della provincia di Venezia (descritta con 100 punti, molto meno dei 10538 iniziali).

Dopo aver ripetuto l'analisi per ognuna delle isole scelte inizialmente la descrizione finale è stata ricavata unendo tra loro tutti i nuovi poligoni. In seguito è stata eliminata una zona costiera dell'entroterra della laguna di Venezia che, sebbene sia presente sia in *raster* che nei grafici di Google Maps, corrisponde ad una parte fangosa e paludosa e quindi disabitata. Non essendo possibile che su di essa siano prodotti rifiuti è stata tagliata dalla regione. Per questo motivo si troverà sempre una zona non analizzata sui grafici con mappe da Google Maps. In figura 6.2 è riportata la descrizione finale del dominio con i punti spaziali considerati (si consulti anche la sezione 6.2.2).

6.2.2 Scelte particolari tra i comuni

L'uso di valori pro capite per rifiuti e posti letto consente di replicare del comune anche su altri punti in cui risulta necessario. Ad esempio, le isole di Murano, Lido di Venezia e Pellestrina non sono sedi di comune, ma si riferiscono a Venezia. Quindi il dato di Venezia è stato replicato in queste

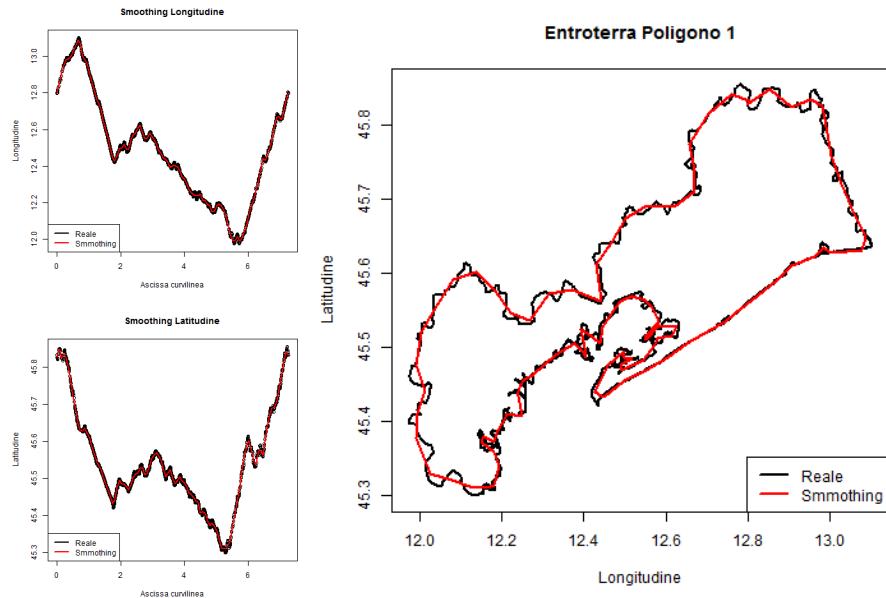


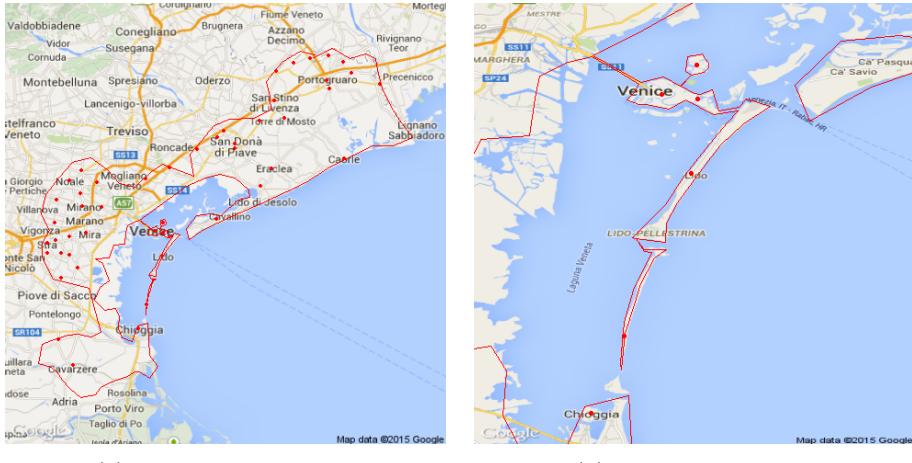
Figura 6.1: Smoothing con *Regression Splines* cubiche per il primo poligono dell'entroterra della provincia di Venezia

isole ad ogni anno, per avere un valore di riferimento in quanto zone distaccate. Come si può notare in figura 6.2(b) anche nell'isola di Venezia è stato duplicato il dato, per avere una triangolazione senza troppi triangoli composti solo da punti di frontiera in una zona della provincia di Venezia di particolare rilevanza.

Un caso particolare riguarda il comune di Cavallino-Treporti, che è stato istituito nel 1999 da una parte dei territori del comune di Venezia. La separazione all'interno dei dati, però, è presente dal 2002. Di conseguenza prima di questo anno il dato in Cavallino-Treporti è una replica del dato di Venezia.

6.2.3 Triangolazione del dominio

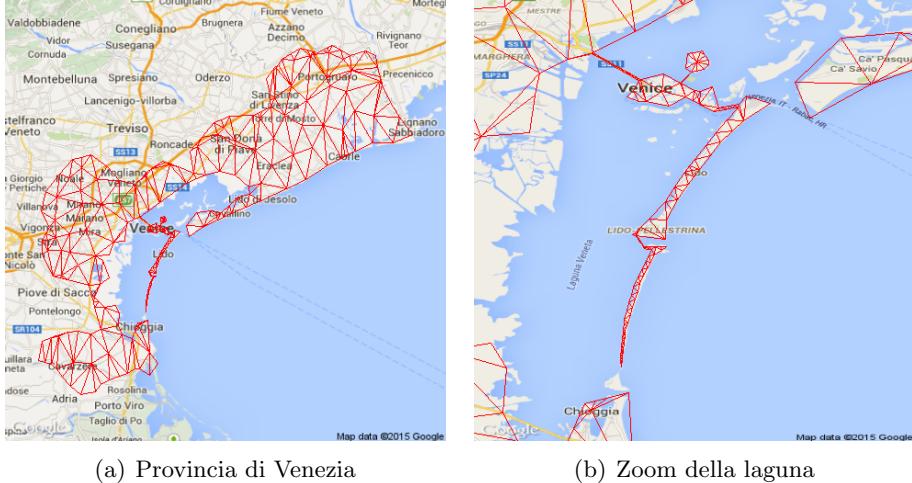
La triangolazione è stata prodotta tramite il pacchetto R *RTriangle*. Poiché nella zona ad est il numero di capoluoghi di comune (e quindi di nodi della triangolazione) è minore rispetto al resto della regione, è stata fissata un'area massima per i triangoli generati da *RTriangle*. Questo ha reso la triangolazione più fitta anche dove non lo sarebbe stata, e garantirà una stima della risposta più precisa nella zona balneare ad est, che come si potrà notare in seguito ha una grande importanza per la distribuzione dei rifiuti. Affinchè questo sia possibile sono stati aggiunti nuovi punti spaziali, che restano senza dato per tutta l'analisi. In fig. 6.3 si ha la triangolazione finale che sarà usata da ora in avanti.



(a) Provincia di Venezia

(b) Zoom della laguna

Figura 6.2: Frontiera e punti spaziali per la provincia di Venezia



(a) Provincia di Venezia

(b) Zoom della laguna

Figura 6.3: Triangolazione della provincia di Venezia

In conclusione, il dominio è descritto da 414 punti (41 con dati, 373 di frontiera o aggiunti dalla triangolazione) e da 475 triangoli.

6.3 Applicazione del modello senza covariate

6.3.1 Analisi preliminare e ricerca del miglior λ

Le basi in spazio scelte per l'applicazione del modello in questo caso sono gli elementi finiti lineari. Ad ognuno dei punti spaziali (interni o di frontiera) è associata una funzione di base, coerentemente con la triangolazione prodot-

ta. Di conseguenza, si ha $N = 414$ mentre il numero di punti con dati n è minore ed è pari a 41.

In tempo, esattamente come nel caso del dominio a forma di C, sono state scelte come funzioni di base le B-splines cubiche. L'intervallo temporale per la descrizione del dominio è [1997, 2011], e i dati sono disponibili con cadenza annuale. Anche in questo caso assumiamo che il numero di basi sia pari al numero di istanti temporali a disposizione, quindi $M = m = 15$.

Prima di calcolare i risultati dell'analisi occorre fissare i parametri λ_S e λ_T . Il procedimento è perfettamente analogo a quello ricavato nel caso del dominio a forma di C, minimizzando la quantità $GCV(\underline{\lambda})$ in (CITAZIONE NECESSARIA). In tabella sono disponibili i risultati.

Intervalli per $\log_{10} \lambda_S$ e $\log_{10} \lambda_T$	Miglior valore
$\log_{10} \lambda_S \in \{-5, -4, \dots, +1\}$	$\underline{\lambda} = (10^0, 10^{-3})$
$\log_{10} \lambda_T \in \{-5, -4, \dots, +1\}$	
$\log_{10} \lambda_S \in \{-1, -0.75, \dots, +1\}$	$\underline{\lambda} = (10^{-0.5}, 10^{-3.25})$
$\log_{10} \lambda_T \in \{-4, -3.75, \dots, -2\}$	
$\log_{10} \lambda_S \in \{-1, -0.875, \dots, +0\}$	$\underline{\lambda} = (10^{-0.375}, 10^{-3.25})$
$\log_{10} \lambda_T \in \{-3.75, -3.625, \dots, -2.75\}$	

Tabella 6.1: Analisi di $GCV(\underline{\lambda})$

6.3.2 Risultati

La produzione dei rifiuti è stata analizzata con $\underline{\lambda} = (10^{-0.375}, 10^{-3.25})$. In fig. 6.4 sono riportati i risultati ottenuti nei 15 anni a disposizione. Questi grafici (grazie anche alla visualizzazione sulle mappe di Google Maps) permettono di sottendere il profilo della funzione nei vari istanti temporali, e avere un'idea dell'evoluzione della produzione di rifiuti a livello geografico.

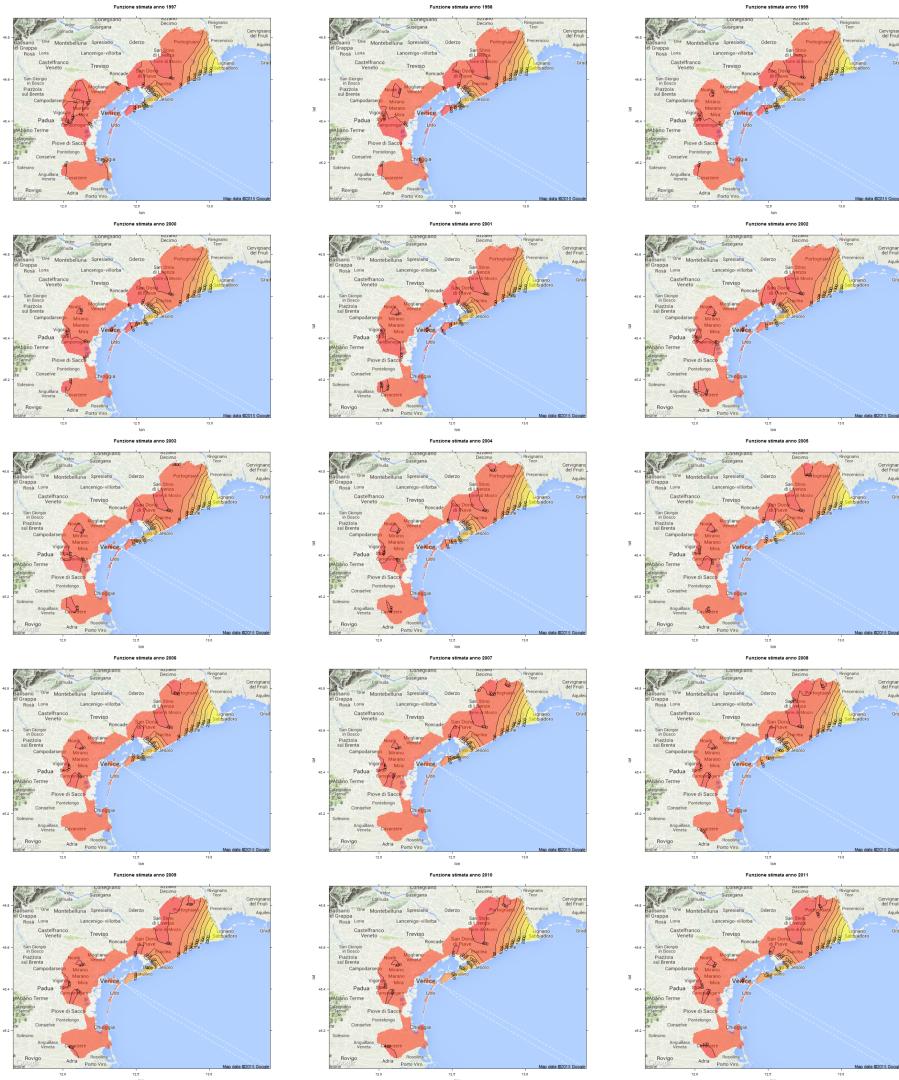


Figura 6.4: Estimated spatio-temporal field of the Venice waste data with the STSR method without covariates at fixed time instants: from year 1997 to year 2011.

Come si può notare dai grafici, ci sono due zone ad elevato valore di produzione dei rifiuti pro capite, corrispondenti a zone di elevato interesse turistico per le spiagge presenti. Nella parte ad est della provincia di Venezia, tra le altre, si hanno le località turistiche di Bibione (confinante con Lingano Sabbiadoro, che però è oltre il confine veneto) e Caorle. Scendendo verso sud-ovest si nota che anche Jesolo causa un innalzamento della funzione che descrive la risposta.

La produzione dei rifiuti è particolarmente alta in queste zone turistiche, e contrariamente a ciò che si poteva immaginare, molto di più di Venezia.

Quindi già adesso possiamo immaginare che il turismo sarà rilevante nell'analisi della produzione di rifiuti, e la causa non sarà Venezia (dove la produzione non è eccessivamente diversa dagli altri comuni) ma le località balneari.

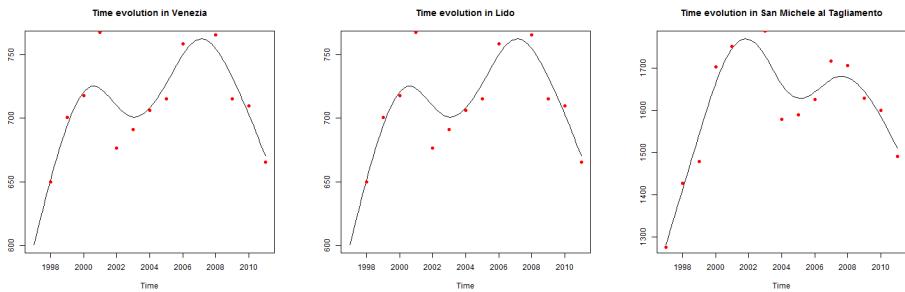


Figura 6.5: Estimated spatio-temporal field of the Venice waste data with the STSR method without covariates at three fixed spatial points: Venice, Lido and San Michele al Tagliamento.

In fig. ?? Sono riportati gli sviluppi temporali stimati in alcuni comuni dati. In particolare sono stati scelti Venezia, Jesolo e Caorle). Si può notare come la funzione stimata spieghi bene l'andamento tracciato dai dati iniziali senza cadere nell'eccessiva interpolazione.

6.4 Applicazione del modello con covariate

6.4.1 Analisi preliminare e ricerca del miglior λ

Le basi in spazio e in tempo sono esattamente le stesse del caso precedente. Prima di eseguire l'analisi, come al solito, occorre cercare buoni valori per $\underline{\lambda}$. Dalla tabella si possono seguire le iterazioni eseguite e il risultato finale.

Intervalli per $\log_{10} \lambda_S$ e $\log_{10} \lambda_T$	Miglior valore
$\log_{10} \lambda_S \in \{-5, -4, \dots, +1\}$	$\underline{\lambda} = (10^0, 10^{-3})$
$\log_{10} \lambda_T \in \{-5, -4, \dots, +1\}$	
$\log_{10} \lambda_S \in \{-1, -0.75, \dots, +1\}$	$\underline{\lambda} = (10^{-0.5}, 10^{-3.25})$
$\log_{10} \lambda_T \in \{-4, -3.75, \dots, -2\}$	
$\log_{10} \lambda_S \in \{-1, -0.875, \dots, +0\}$	$\underline{\lambda} = (10^{-0.375}, 10^{-3.25})$
$\log_{10} \lambda_T \in \{-3.75, -3.625, \dots, -2.75\}$	

Tabella 6.2: Analisi di GCV($\underline{\lambda}$)

6.4.2 Risultati

Le analisi sono state eseguite con $\lambda = (10^{-0.375}, 10^{-3.25})$. Come si può notare dai grafici in fig. 6.6, dove è riportata la stima della funzione $f(\underline{p}, t)$ in ogni punti senza l'aggiunta della parte spiegata dalla covariata, nelle zone dove la produzione di rifiuti è massima si hanno ora i valori più bassi.

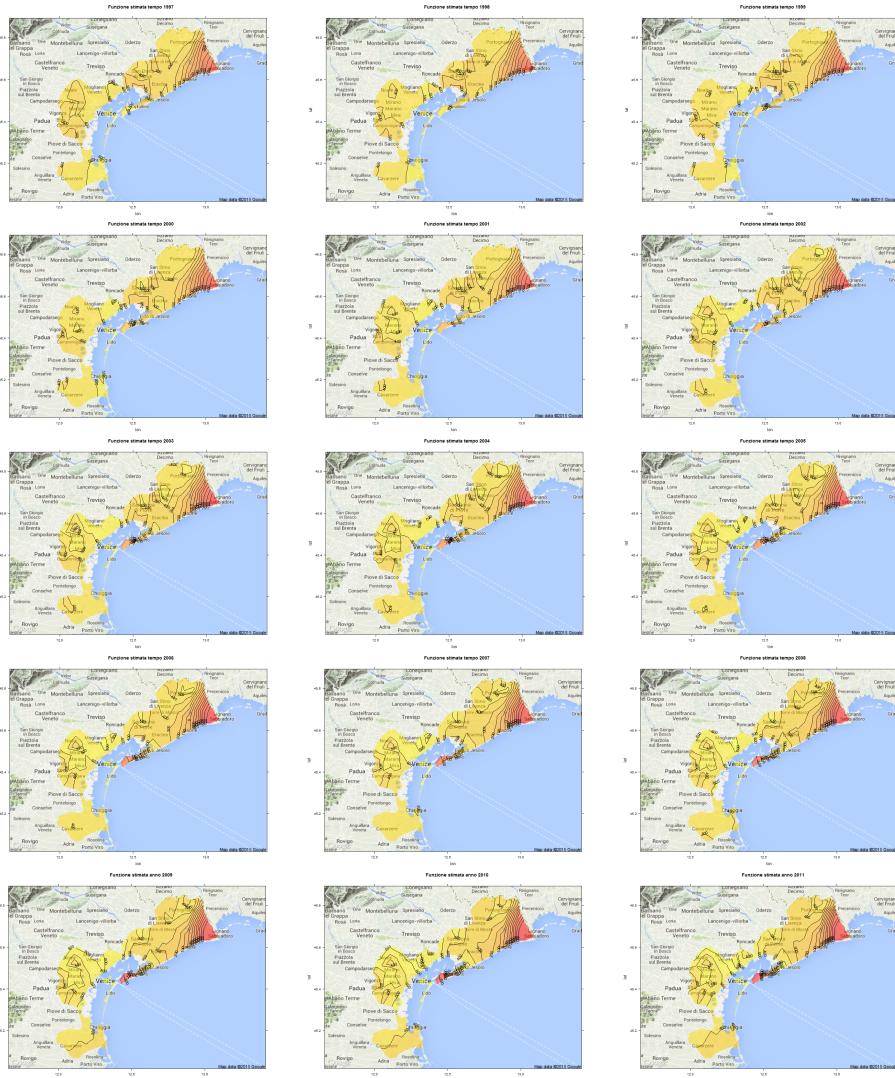


Figura 6.6: Estimated spatio-temporal field of the Venice waste data with the STSR method without covariates at fixed time instants: from year 1997 to year 2011.

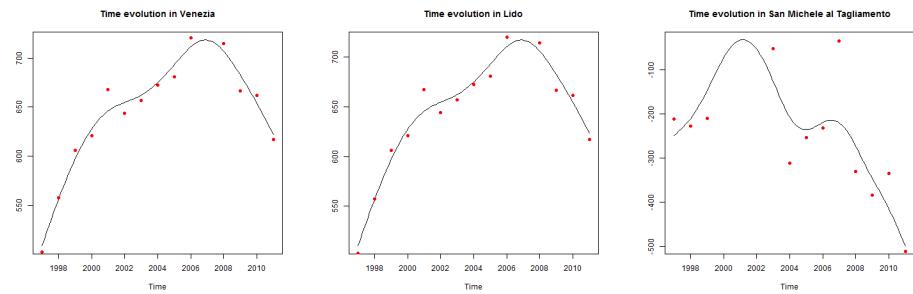


Figura 6.7: Estimated spatio-temporal field of the Venice waste data with the STSR method without covariates at three fixed spatial points: Venice, Lido and San Michele al Tagliamento.

In fig. ?? Sono riportati gli sviluppi temporali stimati in alcuni comuni dati. In particolare sono stati scelti Venezia, Jesolo e Caorle). Si può notare come la funzione stimata spieghi bene l'andamento tracciato dai dati iniziali senza cadere nell'eccessiva interpolazione.

Capitolo 7

Conclusioni e sviluppi futuri

In questo lavoro di tesi è stato analizzato nel dettaglio il modello STR-PDE nell'ambito della stima funzionale per dati varianti all'interno di un dominio spaziale e di un intervallo temporale. Il modello, che si propone di essere un'estensione del caso puramente spaziale già analizzato in letteratura, è stato sviluppato in codice R. Dal confronto con gli altri metodi e da quanto ricavato con le stime, soprattutto sul dominio a forma di C in cui è possibile conoscere il valore reale della funzione, si può concludere che i risultati prodotti sono molto buoni.

Diversa è la conclusione per le prestazioni computazionali del codice. Per semplicità computazionale le basi degli elementi finiti sono state scelte lineari e la produzione dei rifiuti è stata analizzata solamente nella provincia di Venezia, pur avendo a disposizione i dati di tutto il Veneto. Inoltre, durante l'esecuzione del codice, si è potuto notare che alcune funzioni come la minimizzazione di $GCV(\lambda)$ o il calcolo dei valori stimati ad un istante di tempo fissato (usati ad esempio per conoscere il profilo della funzione ad un certo anno) sono molto lente. Ovviamente per analisi di dataset di grosse dimensioni deve essere messa in conto una spesa di tempo elevata, ma R certamente non ha aiutato. Infatti, è noto che R non sia un linguaggio di programmazione fortemente efficiente, e questo ha caratterizzato la lentezza di esecuzione. Il più chiaro sviluppo futuro può essere l'uso di questo codice come base per lo sviluppo di un algoritmo più veloce, attraverso l'integrazione con un linguaggio di programmazione più efficiente (come il C++) nei colli di bottiglia più evidenti.

Dopo che sarà stata sviluppata l'integrazione del codice, sarà possibile garantire una analisi più agile anche per dataset di dimensioni più elevate o per elementi finiti di ordine maggiore. In questo modo si avrà a disposizione uno strumento di analisi statistica buono non solo dal punto di vista dei risultati, ma anche in termini di efficienza computazionale.

Bibliografia

- [1] Nicole H. Augustin, Verena M. Trenkel, Simon N. Wood, Pascal Lorraine, *Space-time modelling of blue ling for fisheries stock management*, Environmetrics, 24, 109–119, (2013)
- [2] Laura Azzimonti, Laura M. Sangalli, Piercesare Secchi, Maurizio Domanin, Fabio Nobile, *Blood flow velocity field estimation via spatial regression with PDE penalization*, Journal of the American Statistical Association, (2015)
- [3] Peter Craven, Grace Wahba, *Smoothing noisy data with spline functions: estimating the correct degree of smoothing by the method of generalized cross-validation*, Numerische Mathematik, 31, 377–403, (1979)
- [4] Giampiero Marra, David L. Miller, Luca Zanin, *Modelling the spatio-temporal distribution of the incidence of resident foreign population*, Statistica Neerlandica, 66, 133–160, (2012)
- [5] Timothy O. Ramsay, *Spline smoothing over difficult regions*, Journal of the Royal Statistical Society: Series B, 64, 307–319, (2002)
- [6] Laura M. Sangalli, James O. Ramsay, Timothy O. Ramsay, *Spatial spline regression models*, Journal of the Royal Statistical Society: Series B, 75, 681–703, (2013)
- [7] Simon N. Wood, Mark W. Bravington, Sharon L. Hedley, *Soap film smoothing*, Journal of the Royal Statistical Society: Series B, 70, 931–955, (2008)