

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

Statistica per le Tecnologie e le Scienze



RELAZIONE FINALE

**Cambiamento Climatico e Venti Estremi: un'Applicazione
del Downscaling Statistico a Diversi Scenari Climatici nel
Nord Italia**

Relatore Prof.ssa Giuliana Cortese
Dipartimento di Scienze Statistiche

Laureando: Arturo Bevilacqua
Matricola N. 1197816

Anno Accademico 2022/2023

Indice

Introduzione.....	1
1. Il Downscaling Climatico.....	3
1.1 Scale dei Moti Geofisici.....	4
1.2 Downscaling Statistico.....	5
1.2.1 Concetti Fondamentali.....	5
1.2.2 Perfect Prognosis o Model Output Statistics?.....	6
1.2.3 Vantaggi.....	7
2. Perfect Prognosis.....	9
2.1 Assunzioni.....	9
2.1.1 Il Cambiamento Climatico.....	10
2.2 Metodi.....	10
2.2.1 Modelli Regressivi.....	10
2.2.2 Metodo degli Analoghi.....	12
2.3 Il predittando Y.....	13
2.4 I predittori X.....	14
2.4.1 EOF/PCA.....	15
2.5 Limitazioni dei metodi PP.....	16
3. I Dati.....	17
3.1 Osservazioni locali.....	17
3.2 ERA-5.....	19
3.3 Proiezioni climatiche GCM.....	20
3.3.1 Scenari SSP.....	20
4. Applicazione.....	23
4.1 Software.....	23
4.2 Verifica dell'adeguatezza del GCM.....	23
4.2.1 Test di Kolmogorov-Sminov a due campioni.....	25
4.3 Selezione del metodo di downscaling.....	25
4.3.1 Modello Gamma.....	26
4.3.2 Validazione.....	27
4.4 Downscaling degli scenari SSP.....	28
Conclusioni.....	31
Bibliografia.....	33

Introduzione

Quando parliamo di cambiamento climatico facciamo comunemente riferimento al riscaldamento globale, ma ci sono molti altri aspetti, o “sintomi”, meno evidenti. Considerando il sistema Terra nel suo insieme viene naturale ipotizzare che un cambiamento della temperatura possa infatti causare o a sua volta essere causato da altri effetti. Alcuni di questi possono essere di altrettanta rilevanza in termini di rischio per la popolazione: alluvioni, tempeste violente, siccità, maree eccezionali, ecc.

È possibile, ad esempio, che il cambiamento climatico abbia un effetto sui venti intensi localizzati? Se sì, di che tipo?

Attraverso questa tesi si cercherà di rispondere a queste domande attraverso le tecniche che ricadono sotto il nome di “downscaling statistico”.

Nel Capitolo 1 definiremo il downscaling climatico e lo contestualizzeremo all’interno del panorama climatologico attuale, distinguendone le due tipologie: downscaling dinamico e statistico. In seguito ci focalizzeremo sul downscaling statistico, gettandone le basi teoriche e differenziandolo in *Perfect Prognosis* (PP) e *Model Output Statistics* (MOS).

Nel Capitolo 2 svilupperemo più a fondo il concetto di *Perfect Prognosis* analizzandone le assunzioni, delineando i metodi statistici che lo caratterizzano e i requisiti delle variabili coinvolte.

Passeremo poi, con il Capitolo 3, ad una descrizione accurata dei diversi dataset utilizzati per effettuare il downscaling, portando come esempio le osservazioni delle raffiche massime giornaliere di vento in Veneto. Prima dell’effettiva applicazione del downscaling alle proiezioni future, nel Capitolo 4 valuteremo la validità delle ipotesi del downscaling per i nostri dati e confronteremo le performance di diversi modelli su dati osservati. Una volta selezionato il modello migliore per identificare cambiamenti nella distribuzione dei venti più intensi lo applicheremo alle proiezioni future per ottenere una previsione a distanza di diversi decenni.

1. Il Downscaling Climatico

La maggior parte delle proiezioni climatiche moderne sono basate sui Modelli di Circolazione Globale o GCM (*Global Circulation Model*), ovvero modelli fisici che, date delle condizioni meteorologiche iniziali e al contorno, risolvono numericamente le equazioni fluidodinamiche che governano i cambiamenti dell'atmosfera terrestre e/o degli oceani, simulando così gli andamenti futuri di questi ultimi (Flato, et al., 2013).

Tali modelli, iniziati a svilupparsi già a partire dagli anni '50 (Phillips, 1956), grazie alla crescente disponibilità di potenza computazionale hanno raggiunto nei decenni complessità sempre maggiori (Flato, et al., 2013). Ad oggi i GCM più dettagliati producono simulazioni con risoluzioni spaziali di circa $1^\circ \times 1^\circ$, con celle tra i 100 e i 200 km di ampiezza (Figura 1). Sebbene tali risoluzioni siano sufficienti allo studio dei fenomeni meteorologici su scala planetaria e/o sinottica, quando rivolgiamo il nostro interesse su scale inferiori, che si tratti di una regione ristretta o di un insieme di singoli punti (es. stazioni meteorologiche) necessitiamo di informazioni più dettagliate.

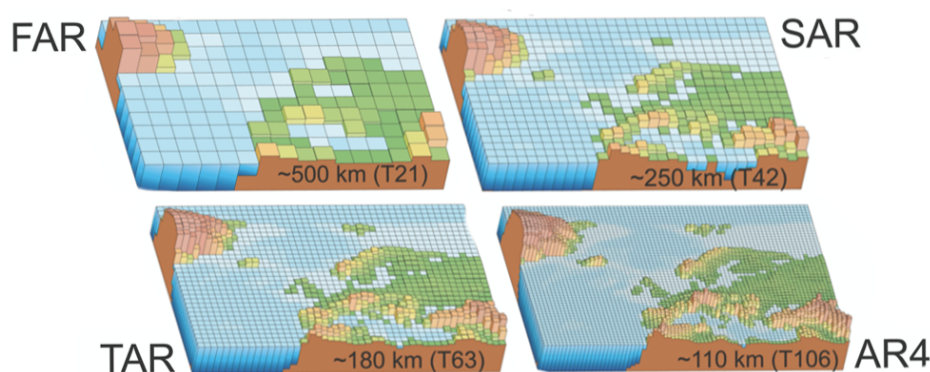


Figura 1 – Risoluzione spaziale delle diverse generazioni di GCM usati nei rapporti IPCC: FAR (1990), SAR(1996), TAR (2001) e (2007).

È in questo contesto che si definisce il concetto di downscaling climatico, come “riscaldamento” delle informazioni disponibili da una griglia più ampia ad una griglia più fitta o, più spesso, a singoli punti (Maraun & Widmann, 2018).

1.1 Scale dei Moti Geofisici

In meteorologia dinamica si è soliti dividere i fenomeni atmosferici e oceanici su diverse scale spaziali in quanto questi possono assumere caratteristiche molto diverse a seconda della dimensione di riferimento.

Si distinguono le seguenti scale dei moti geofisici:

- **Scala planetaria:** macro-movimenti di ampiezza continentale o superiore, mediati su tempi lunghi superiori al mese;
- **Scala sinottica:** moti di dimensione variabile tra centinaia e migliaia di chilometri, mediati su tempi giornalieri. Spiegano gran parte della variabilità meteorologica;
- **Mesoscala:** moti dalle decine alle centinaia di chilometri, come fenomeni locali dovuti alla circolazione atmosferica, quali la brezza termica, o eventi estremi come uragani, alluvioni e tempeste;
- **Microscala:** moti inferiori alla decina di chilometri, come turbolenze e nubi convettive.

Il downscaling, in teoria, potrebbe dunque essere applicato per passare da una qualsiasi scala superiore ad una inferiore; tuttavia, nella pratica viene utilizzato principalmente dai fenomeni su scala sinottica, simulati adeguatamente dai GCM, ai fenomeni di mesoscala, poiché questi ultimi rappresentano spesso gli avvenimenti di maggior interesse nella vita della popolazione (Maraun & Widmann, 2018).

1.2 Downscaling Statistico

È possibile dividere i moderni metodi di downscaling in due categorie principali: downscaling dinamico e downscaling statistico.

Il downscaling dinamico si basa sullo sviluppo e utilizzo di Modelli Climatici Regionali o RCM (*Regional Climate Model*). Come i GCM, gli RCM simulano numericamente i processi climatici tramite la fluidodinamica, bensì su scala regionale e considerando le caratteristiche orografiche locali (come vallate, catene montuose minori, laghi ecc.), che altrimenti verrebbero ignorate dai GCM. Le proiezioni su scala locale vengono dunque ottenute simulando dagli RCM e ponendo come condizioni al contorno i valori restituiti dai GCM. Date queste caratteristiche, il downscaling dinamico fornisce dati su una griglia di punti che ricopre tutta la superficie di studio.

Nel downscaling statistico le relazioni tra i moti di ampia scala e le variabili d'interesse di piccola scala vengono invece ricercate empiricamente sulle osservazioni passate. Tali relazioni vengono poi applicate agli scenari futuri simulati dai GCM per ottenere le previsioni su scala locale. Poiché le osservazioni provengono normalmente da singole stazioni non equi spaziate sul territorio d'interesse, il downscaling statistico viene comunemente usato per proiettare le informazioni GCM sulle singole stazioni locali.

Ai fini di questo studio si approfondirà di seguito solamente il secondo approccio, ossia il downscaling statistico.

1.2.1 Concetti Fondamentali

Il downscaling statistico è basato sull'assunzione che esista una relazione stocastica tra i moti di larga scala e quelli locali. Tale relazione, nota in letteratura come funzione di downscaling (Maraun & Widmann, 2018), viene rappresentata nella forma di un modello statistico $F(\cdot)$. Si può dunque considerare la distribuzione di una variabile locale d'interesse Y in funzione di determinati predittori X di ampia scala,

$$Y \sim F(X) \tag{1}$$

Ciò permette di caratterizzare i metodi di downscaling statistico in base a come viene formulata la funzione di downscaling $F(\cdot)$ e a come vengono trattate le realizzazioni dei predittori X in fase di calibrazione del modello¹.

1.2.2 Perfect Prognosis o Model Output Statistics?

È infatti possibile distinguere tra due diversi approcci a seconda dell'origine dei valori utilizzati per la stima dei coefficienti del modello: se i valori delle X provengono da osservazioni si parla di metodi *Perfect Prognosis (PP)* (Klein, et al., 1959), mentre se provengono da simulazioni passate di modelli numerici (GCM) si parla di metodi *Model Output Statistics (MOS)* (Glahn & Lowry, 1972). A prescindere dal metodo utilizzato, il modello così ottenuto viene in fine applicato alle simulazioni future dei predittori per ottenere le proiezioni della variabile d'interesse a livello locale (Figura 2).

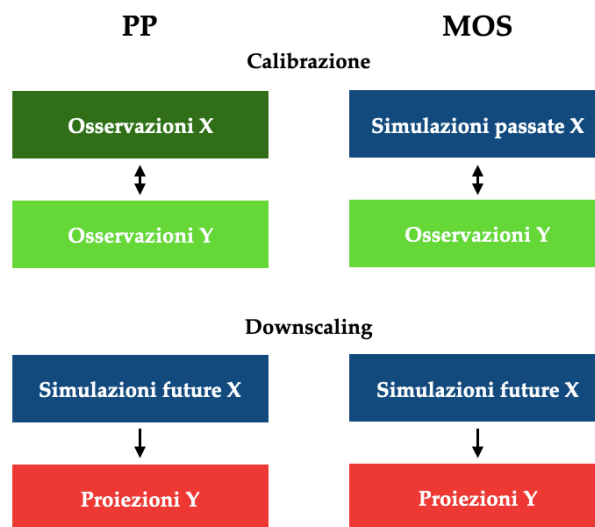


Figura 2 – Differenze tra Perfect Prognosis e Model Output Statistics. Le caselle verdi indicano la provenienza dei dati da osservazioni, le caselle blu da simulazioni GCM e le caselle rosse rappresentano il prodotto finale del downscaling. (Maraun & Widmann, 2018)

¹Nella letteratura del downscaling si indicano con “calibrazione” tipicamente i metodi di stima dei parametri del modello.

Nei metodi MOS la calibrazione viene effettuata tra le osservazioni della variabile risposta di piccola scala e delle simulazioni passate GCM. Poiché i GCM però simulano degli scenari “plausibili” (non per forza quelli che effettivamente poi si sono verificati) delle variabili meteorologiche in modo continuativo, non c’è una stretta corrispondenza temporale tra i due. Si predilige dunque studiare in questo caso le differenze nelle caratteristiche distributive a lungo termine (e.g. annuali) tra GCM e osservazioni. Tale tipo di calibrazione, nella pratica, risulta efficace solamente nel caso in cui la risposta e i predittori condividano la stessa grandezza fisica. Per questo motivo le tecniche MOS vengono comunemente considerate come mere correzioni del bias (Maraun & Widmann, 2018).

Nei metodi PP, il fatto che la calibrazione venga invece effettuata sulle osservazioni sia dei predittori che della risposta permette di utilizzare variabili anche non omogenee tra loro, ovvero misurate su diverse grandezze fisiche. Infatti, poichè le osservazioni esistono in un dato istante t sia per il predittore che per il predittando, si presuppone che queste siano, anche solo in minima parte, correlate tra di loro e che quindi contengano informazione utile a spiegare il rapporto tra grande e piccola scala anche su intervalli piccoli (e.g. giornalieri).

In questa tesi ci si è limitati ad analizzare e ad applicare tecniche appartenenti a quest’ultimo tipo di downscaling statistico, la *Perfect Prognosis*.

1.2.3 Vantaggi

Negli ultimi decenni è stato provato con diversi studi che il downscaling statistico permette di ottenere previsioni con abilità pressoché equivalenti a quelle del downscaling dinamico ma con un rilevante risparmio in termini di costo computazionale (fino a 100 volte meno) (Roux, et al., 2018; Scott Spak, 2007), rendendolo in molti casi il metodo preferibile.

2. Perfect Prognosis

Le prime applicazioni di metodi PP risalgono alla seconda metà degli anni '50 quando Klein et al. (1959), a pochi anni dallo sviluppo dei primi modelli numerici per la circolazione del meteo, sfruttarono le serie storiche a loro disposizione per fare inferenza sulle relazioni statistiche tra le variabili di circolazione globale e i cambiamenti locali. Ad oggi, come detto in precedenza, i metodi PP sono i più utilizzati per effettuare il downscaling.

2.1 Assunzioni

Nell'approccio PP il fatto che l'origine dei predittori X non sia la stessa tra calibrazione (osservazioni) e downscaling (simulazioni) rende questo metodo vulnerabile ad errori di bias. Qualsiasi distorsione nelle simulazioni GCM verrebbe infatti trasmessa ai risultati finali del downscaling, i quali saranno dunque coerenti con il GCM ma saranno tanto meno consistenti con la realtà quanto più il GCM è distorto. Il termine "*perfect*" si riferisce dunque proprio al fatto che nell'applicare il downscaling alle simulazioni si sta assumendo che queste siano capaci di rappresentare perfettamente i processi reali. Ne segue che non solo è necessario selezionare predittori che spieghino la maggior parte della variabilità regionale della risposta, ma anche individuare un GCM che li simuli adeguatamente nella regione d'interesse.

In fine la forma funzionale del modello deve tenere in considerazione le caratteristiche fisiche delle variabili di grande e piccola scala (e.g. il modello deve fornire valori plausibili della risposta anche in presenza di nuovi valori dei predittori, ossia fuori dal range osservato durante la calibrazione).

Se tali condizioni, dette anche assunzioni di stazionarietà, sono soddisfatte, cioè tutti i predittori rilevanti sono inclusi nel modello e la loro influenza sulla risposta è accuratamente modellata anche al di fuori degli stati osservati; allora il modello è valido per il clima futuro.

2.1.1 Il Cambiamento Climatico

Tutte le assunzioni viste sopra fanno affidamento al fatto che si mantenga una certa stazionarietà nelle condizioni climatiche future e che quanto successo in passato si possa ripetere. Tuttavia, è globalmente riconosciuta l'attuale situazione di cambiamento climatico verso il riscaldamento globale.

Tali assunzioni vanno quindi riviste di conseguenza se si vuole effettuare il downscaling sulle proiezioni climatiche future regionali².

È necessario dunque incorporare nel modello le variabili che determinano un cambiamento della risposta regionale dovuto al cambiamento climatico.

2.2 Metodi

Di seguito vengono riportati i principali metodi di *perfect prognosis* downscaling utilizzati per modellare la funzione di downscaling.

2.2.1 Modelli Regressivi

La scelta più immediata per modellare la relazione tra fenomeni di larga scala e di piccola scala sono i modelli regressivi.

Nel caso più semplice, in cui la variabile risposta si assume distribuita secondo una distribuzione normale, può essere utilizzato un modello di regressione lineare,

$$Y_t = \beta_0 + \beta_1 X_{t1} + \dots + \beta_M X_{tM} + \varepsilon_t \quad (2)$$

in cui una serie storica della risposta Y_t viene espressa come una combinazione lineare di M serie storiche dei predittori X_{t1}, \dots, X_{tM} e un errore casuale $\varepsilon_t \sim \mathcal{N}(0, \sigma^2)$. Si ricorda inoltre che, poiché la linearità nei modelli di regressione lineare è intesa nei parametri β_i , è possibile effettuare determinate trasformazioni non lineari delle variabili (e.g. esponenziale, logaritmo, radice ecc.) purché queste conservino la linearità nei parametri.

² Nel caso di previsioni di breve termine il cambiamento climatico può essere ignorato in quanto i suoi effetti su scala giornaliera sono pressoché irrilevanti. In futuro sarà sufficiente ricalibrare il modello.

Spesso, tuttavia, l'assunzione di normalità non è plausibile. Si considerino, per esempio, le precipitazioni giornaliere. Per la maggior parte delle località la distribuzione è fortemente asimmetrica, con molti giorni in cui i livelli di precipitazioni sono nulli o molto bassi e pochi giorni con precipitazioni intense. Per questo motivo la probabilità di pioggia e l'intensità nei giorni piovosi vengono spesso modellate separatamente. In contesti simili può essere utile l'utilizzo di modelli lineari generalizzati, in cui la distribuzione della variabile risposta Y condizionatamente ai predittori X si assume proveniente dalla famiglia esponenziale e il suo valore atteso viene espresso in funzione di una combinazione lineare delle X . Per delle serie storiche Y_t e X_{t1}, \dots, X_{tM} si può definire il modello come segue:

$$E(Y_t) = \mu_t = g^{-1}(\beta_0 + \beta_1 X_{t1} + \dots + \beta_M X_{tM}) \quad (3)$$

La funzione $g(\cdot)$, detta *funzione di collegamento*, viene definita in base alla distribuzione della risposta Y_t . Per ogni distribuzione della famiglia esponenziale esiste una funzione di collegamento canonica ottenuta dall'esponenziale della funzione di densità. Tuttavia, bisogna porre particolare attenzione ai casi in cui il dominio della funzione di collegamento non coincida con i possibili valori ammessi dal valore atteso $E(Y_t)$ ed eventualmente utilizzare funzioni non-canoniche che correggano tale incongruenza.

I modelli lineari generalizzati hanno visto un'ampia applicazione nella modellazione delle precipitazioni (Chandler & Wheeler, 2002; Yang, et al., 2005; San-Martín, et al., 2017), la probabilità di pioggia, ad esempio, può essere rappresentata tramite un modello di regressione logistica, mentre per il livello giornaliero viene spesso considerato un modello Gamma.

Quando non è fattibile effettuare assunzioni distributive un altro strumento a disposizione è la regressione quantile. Invece di stimare il valore atteso della distribuzione, nella regressione quantile vengono stimati la mediana o altri quantili d'interesse in funzione dei predittori. Tale tipo di regressione fornisce stimatori più robusti ai valori anomali rispetto alla regressione lineare.

Un'alternativa meno diffusa e piuttosto recente è la regressione non lineare mediata tramite reti neurali artificiali. L'apprendimento automatico è stato utilizzato, ad esempio, per effettuare il downscaling sulle precipitazioni in Sud Africa (Hewitson & Crane, 2002), Giappone (Uvo, et al., 2001) e Regno Unito (Haylock, et al., 2006).

In tutti i metodi regressivi di cui sopra, per ognuna delle stazioni di rilevamento della variabile risposta, viene calibrato un modello indipendente dalle altre stazioni. Per introdurre nella modellazione la dipendenza spaziale tra le osservazioni è necessario sviluppare modelli di regressione multi-sito.

Applicando tecniche di analisi multidimensionale quali il calcolo delle Componenti Principali (PC), l'Analisi della Massima Covarianza (MCA) o l'Analisi delle Correlazioni Canoniche (CCA) è infatti possibile decomporre le serie storiche delle osservazioni su più stazioni in pattern spazio-dipendenti (Bretherton, et al., 1992; Bretherton, et al., 1992).

Oltre a tenere conto della relazione spaziale tra le osservazioni, questi modelli possono portare ad un minor numero di regressori.

2.2.2 Metodo degli Analoghi

Un'altra tecnica diffusa sotto l'approccio *Perfect Prognosis* è il metodo degli analoghi, il quale si basa sul presupposto che due stati atmosferici simili – gli analoghi – possano risultare in simili condizioni meteo locali.

Dato uno stato x_t dei predittori di larga scala al tempo t , il *downscaling* viene effettuato ricercando un suo analogo x_k per il quale sia disponibile un'osservazione locale y_k . Tale osservazione funge dunque da previsione per il clima locale:

$$y_{t_{pred}} = y_k \text{ con } k : x_k = \text{analog}(x_t) \quad (4)$$

Gli analoghi dello stato x_t vengono determinati minimizzando la sua distanza, solitamente quella Euclidea (Matulla, et al., 2008), dalle osservazioni disponibili dei predittori.

$$\text{analog}(x_t) = x_k \mid d(x_t, x_k) = \min \quad (5)$$

Una comune critica al metodo degli analoghi riguarda, banalmente, l'esistenza degli stessi. Van Den Dool (1994) ha evidenziato come, a causa dell'alto numero di gradi di libertà dell'atmosfera, su scala globale non possano esistere veri e propri analoghi. Tuttavia, Zorita e von Storch (1999) sostengono che, poiché il *downscaling* viene solitamente applicato solo ad una regione limitata della Terra, gran parte della complessità di uno stato possa essere filtrata come rumore, dimostrandone inoltre l'efficacia rispetto a metodi più complessi.

In ogni caso il problema persiste per gli eventi rari, in quanto questi potrebbero essere mal rappresentati dalle osservazioni passate, o non rappresentati affatto, e quindi difficilmente identificabili con questo metodo.

2.3 Il predittando Y

La variabile risposta, o "predittando", può essere una qualsiasi variabile meteorologica d'interesse per la quale siano disponibili delle osservazioni. Tendenzialmente questo avviene per variabili quali temperatura, precipitazione o intensità del vento osservate in stazioni meteorologiche vicino alla superficie. In diversi studi d'impatto ambientale è tuttavia necessario il calcolo di indici più o meno complessi come combinazione di più fattori meteorologici. In alcuni casi è stato possibile effettuare il *dowscaling* direttamente su tali indici (Casanueva, et al., 2014).

2.4 I predittori X

Come illustrato nella sezione 2.1, i predittori devono essere scelti in modo da spiegare adeguatamente la variabilità della risposta sia sul breve che sul lungo periodo, contenere informazioni riguardo il cambiamento climatico e contemporaneamente essere simulati quanto meglio possibile dal modello GCM. A tal fine è utile selezionare una regione per i predittori ben più estesa della singola cella GCM più vicina alla stazione obiettivo (von Storch, et al., 1993), mantenendola comunque limitata per non aggiungere troppa informazione inutile (Wilby, et al., 2004). Solitamente la regione selezionata per i predittori è centrata sulle località obiettivo e sufficientemente più grande in modo da comprendere tutti i processi rilevanti per il clima locale. La centratura, tuttavia non è strettamente necessaria, in quanto particolari deviazioni potrebbero portare ad una migliore rappresentazione di processi rilevanti per il clima locale (Wilby, et al., 2004). Se ad esempio è noto che il meteo nella zona d'interesse è influenzato prevalentemente da flussi d'aria tipicamente provenienti da una specifica direzione, ampliare il dominio dei predittori verso tale direzione potrebbe migliorare la capacità di questi ultimi nel determinare la risposta locale. Poiché i valori "grezzi" dei predittori sono spesso fortemente correlati tra loro nello spazio, è possibile calcolare e sostituirli con delle trasformazioni. A tal fine si è rivelata utile l'analisi delle componenti principali (Huth, 1999), nota in climatologia anche con il nome di analisi delle Funzioni Ortogonali Empiriche o EOF (Bjornsson & Venegas, 1997).

2.5 Limitazioni dei metodi PP

Un difetto comune a tutti i metodi PP è l'inabilità a rappresentare il feedback locale. I metodi PP proiettano la circolazione di ampia scala sul clima locale, se quindi il feedback locale non altera la circolazione di larga scala il suo effetto viene ignorato. Ad esempio, un innalzamento delle temperature in una determinata zona faciliterà l'evaporazione dell'acqua dal terreno e quando questo sarà secco faciliterà un ulteriore riscaldamento. Quest'ultimo tipo di mutamento locale spesso non viene colto con i metodi PP.

Un'altra lacuna dei metodi PP è il fatto che nella maggior parte di essi, come visto in precedenza, non viene modellata né la dipendenza temporale (Y_t dipende dalle X_k solo per $t = k$) né la dipendenza spaziale (viene stimato separatamente un modello per ogni stazione obiettivo). Gli aspetti di dipendenza spazio-temporale dei fenomeni su piccola scala vengono quindi spesso considerati "ereditati" dai predittori (Maraun & Widmann, 2018).

3. I Dati

Per effettuare il *downscaling* con l'approccio PP sono necessarie tre diverse tipologie di dati (Tabella 2). Nella fase iniziale di calibrazione del modello vengono utilizzate osservazioni provenienti da stazioni meteorologiche locali e ricostruzioni storiche dei predittori di larga scala nella regione circostante. Nella fase finale di previsione viene applicato il modello ad un nuovo set di dati, ovvero una proiezione climatica ottenuta da un GCM, per valutarne l'andamento futuro.

3.1 Osservazioni locali

Le osservazioni locali sono state ricavate dal sito web dell'Agenzia Regionale per la Prevenzione e Protezione Ambientale del Veneto (ARPAV), da cui sono accessibili le misurazioni giornaliere delle principali variabili meteorologiche in oltre 200 stazioni del Veneto a partire dal 2010.

In particolare, per quanto riguarda il vento, sono presenti le medie giornaliere di velocità e direzione nonché la massima raffica di vento. Ai fini dell'analisi degli eventi estremi di vento la scelta è dunque naturalmente ricaduta sulla massima raffica giornaliera come variabile per la quale effettuare il *downscaling*.

Per ridurre in seguito il costo computazionale è stato inoltre selezionato un sottoinsieme ridotto di stazioni in modo da garantire comunque una discreta copertura del territorio regionale (Figura 3) e una percentuale di valori mancanti per stazione inferiore al 5% (Tabella 1).

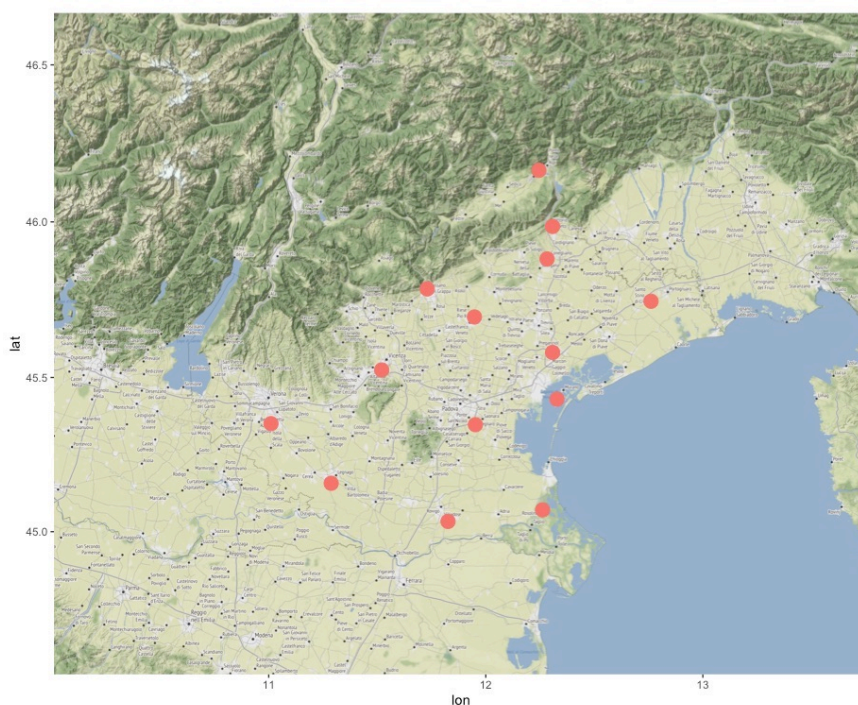


Figura 3 – Posizioni delle stazioni ARPAV da cui sono state ricavate le osservazioni delle raffiche giornaliere di vento

Tabella 1 – Informazioni geografiche e percentuale di valori mancanti per le stazioni ARPAV utilizzate

Codice	Comune	Longitudine	Latitudine	Altitudine	NA
0232	Bassano del Grappa	11.73017735	45.78504906	127m	2,11%
0264	Belluno	12.24470834	46.16467858	379m	0,84%
0102	Castelfranco Veneto	11.94778216	45.69492112	49m	2,13%
0100	Conegliano	12.28232702	45.88132477	90m	1,52%
0131	Legnago	11.28774164	45.15740933	12m	3,98%
0111	Legnaro	11.95217201	45.34734845	7m	1,90%
0227	Mogliano Veneto	12.30779083	45.5807425	5m	3,07%
0159	Portogruaro	12.76128681	45.74548888	2m	0,97%
0112	Rosolina	12.26177816	45.07114136	-2m	1,50%
0231	Rovigo	11.8259685	45.03340149	2m	3,14%
0252	Venezia	12.32809643	45.42996715	18m	0,55%
0127	Buttapietra	11.01059728	45.35084342	39m	2,57%
0451	Vicenza	11.52039317	45.52434984	29m	1,60%
0240	Vittorio Veneto	12.30760378	45.98539025	123m	2,09%

3.2 ERA-5

Prodotto dal Centro Europeo per le Previsioni Meteorologiche a Medio Termine (ECMWF) e accessibile dal portale europeo Copernicus Climate Data Store (CDS), il dataset ERA-5 è il risultato di un'ampia rianalisi di dati storici sia di superficie che di alta quota provenienti da tutto il mondo e in continuo aggiornamento. Tale dataset rappresenta una ricostruzione completa (senza valori mancanti) delle condizioni meteorologiche mondiali dal 1959 ad oggi su una griglia con risoluzione di circa $0.25^\circ \times 0.25^\circ$ e con frequenza oraria.

Per la stima del modello le pseudo-osservazioni dei predittori sono state pertanto ricavate dai dati ERA-5 nello stesso periodo dei dati locali (2010-2022). È stato selezionato il dominio spaziale compreso tra le seguenti coordinate [$8^\circ\text{E} - 15^\circ\text{E}$; $43^\circ\text{N} - 48^\circ\text{N}$], sufficientemente ampio da comprendere i principali elementi geografici che circondano il Veneto: le Alpi, l'Alto Adriatico e la Pianura Padana. Nonostante l'ampia disponibilità di variabili, per garantire l'applicabilità del *downscaling* alle proiezioni climatiche del GCM, la scelta dei possibili predittori si è ridotta alle sole variabili che avessero un corrispettivo nel GCM (Tabella 3). I dati sono poi stati aggregati giornalmente per coincidere con la frequenza delle osservazioni.

Sebbene l'utilizzo di dati provenienti da rianalisi (e.g. ERA-5, ERA40, NCEP-NCAR) sia stato oggetto di discussione in quanto possibile fonte di incertezza aggiuntiva (Brands, et al., 2012) essi vengono comunque ampiamente utilizzati nel *downscaling* in quanto sufficientemente rappresentativi della realtà nella maggior parte del pianeta (Sterl, 2004) con riscontri di errori rilevanti solamente ai tropici (Manzanas, et al., 2015).

Sono state inoltre utilizzate le raffiche massime giornaliere ERA-5 nei punti della griglia prossimi alle stazioni per l'imputazione dei valori mancanti in queste ultime tramite regressione lineare semplice. Considerata la ridotta frazione di valori mancanti e la buona correlazione tra i due dataset (correlazione media di 0,67 tra le raffiche nelle stazioni e nei punti ERA-5 più vicini), si è ritenuto che la perdita in variabilità causata dall'imputazione tramite regressione fosse accettabile.

3.3 Proiezioni climatiche GCM

Le simulazioni future di larga scala, selezionate per il periodo 2015-2100, sono state ottenute anch'esse dal Copernicus CDS. Tra i 134 GMC presenti all'interno del progetto CMIP6 (*Coupled Model Intercomparison Project*) si è voluto utilizzare per questo studio le proiezioni del modello italiano CMCC-ES2, del Centro euro-Mediterraneo sui Cambiamenti Climatici.

3.3.1 Scenari SSP

Nel contesto del CMIP6 il Gruppo intergovernativo sul cambiamento climatico (IPCC) ha definito dei possibili "scenari socio-economici condivisi" (*Shared Socio-Economic Pathways*), ovvero delle prospettive di emissioni e di sviluppo socioeconomico globali per il ventunesimo secolo. Per questi scenari sono state infine prodotte simulazioni della risposta climatica con ognuno dei 134 modelli di circolazione globale.

In questa tesi sono stati considerati gli scenari SSP1-2.6, SSP2-4.5 ed SSP5-8.5, che possono essere descritti come segue:

- **SSP1-2.6:** vengono presi provvedimenti in favore dello sviluppo sostenibile e della protezione del clima riducendo le emissioni fino ad ottenere nel 2100 un forzante radiativo pari a $2,6 W/m^2$.
Rappresenta il caso ideale in termini di sostenibilità;
- **SSP2-4.5:** scenario a sviluppo intermedio in cui l'emissione di gas a effetto serra è arginata, ma le loro concentrazioni nell'atmosfera continuano ad aumentare nei prossimi 50 anni, giungendo ad un forzante radiativo pari a $4,5 W/m^2$ nel 2100.
Rappresenta il caso intermedio tra quelli plausibili;
- **SSP5-8.5:** non viene preso alcun provvedimento in favore della protezione del clima. Viene perpetrato un consumo intensivo di combustibili fossili causando un continuo aumento delle emissioni di gas, giungendo ad un forzante radiativo di $8,5 W/m^2$ nel 2100.
Rappresenta il caso peggiore.

Tabella 2 – Dettagli dei dataset utilizzati

Dataset	Tipo di dataset	Periodo	Frequenza	Risoluzione
ARPAV	<i>Osservazioni in-situ</i>	2010-2022	<i>Giornaliera</i>	<i>Dati puntuali</i>
ERA-5	<i>Rianalisi</i>	2010-2022	<i>Oraria</i>	0.25° x 0.25°
CMCC-ES2	<i>Proiezioni climatiche</i>	2015-2100	<i>Giornaliera</i>	1° x 1°
CMCC-ES2	<i>Simulazioni storiche</i>	2010-2015	<i>Giornaliera</i>	1° x 1°

Tabella 3 – Dettagli delle variabili

Dataset	Nome variabile	Unità di misura	Descrizione
ARPAV	vraff	<i>m/s</i>	<i>Raffica di vento massima giornaliera</i>
ERA-5	mx2t	<i>K</i>	<i>Temperatura massima a 2m dal suolo</i>
	mn2t	<i>K</i>	<i>Temperatura minima a 2m dal suolo</i>
	t2m	<i>K</i>	<i>Temperatura a 2m dal suolo</i>
	msl	<i>Pa</i>	<i>Pressione atmosferica al livello del mare</i>
	mtpr	<i>mm/s</i>	<i>Precipitazione media</i>
	u10	<i>m/s</i>	<i>Componente W-E del vento a 10m dal suolo</i>
	v10	<i>m/s</i>	<i>Componente S-N del vento a 10m dal suolo</i>
	w10*	<i>m/s</i>	<i>Intensità del vento a 10m dal suolo</i>
CMCC-ES2	tasmax	<i>K</i>	<i>Temperatura massima vicino alla superficie</i>
	tasmin	<i>K</i>	<i>Temperatura minima vicino alla superficie</i>
	tas	<i>K</i>	<i>Temperatura vicino alla superficie</i>
	psl	<i>Pa</i>	<i>Pressione atmosferica al livello del mare</i>
	pr	<i>mm/s</i>	<i>Precipitazione media</i>
	sfcWind	<i>m/s</i>	<i>Intensità del vento vicino alla superficie</i>

* non presente nel dataset originale. Ottenuta come $w10 = \sqrt{u10^2 + v10^2}$

4. Applicazione

4.1 Software

Per l'applicazione dei metodi di *downscaling* presentati nella sezione 2.2 si sono sfruttate le librerie R del progetto Climate4R (Iturbide, et al., 2019) sviluppato dal Santander Meteorology Group.

Climate4R comprende quattro librerie, utili per ognuna delle diverse fasi del *downscaling*: *loadR* per l'importazione dei dataset nell'ambiente R, *transformR* per la manipolazione dei dati su griglia, *downscaleR* per l'implementazione delle tecniche di *downscaling* e infine *visualizeR* per la produzione dei grafici.

4.2 Verifica dell'adeguatezza del GCM

Come spiegato nella sezione 2.1 le previsioni ottenute con i metodi PP si basano sull'assunzione che il GCM selezionato sia capace di simulare "perfettamente" i predittori su larga scala. Per valutare questa ipotesi si è dunque effettuato un test di Kolmogorov-Smirnov (KS) a due campioni con livello di significatività del 5% tra le simulazioni storiche CMCC-ES2 e i dati ERA-5 per ognuno dei predittori nel periodo 2010-2015.

Dopo aver effettuato un'interpolazione dei dati CMCC-ES2 sulla griglia ERA-5 con il criterio del punto più vicino (*nearest-neighbour*) è stata calcolata la statistica test KS per ognuno dei punti della griglia.

Dai risultati raccolti nella Figura 3 si può concludere che il modello GCM rappresenta adeguatamente solo la temperatura in superficie (*tasmin*, *tasmax* e *tas*) mentre le simulazioni di pressione (*psl*), precipitazione (*pr*) e velocità del vento (*sfcWind*) sono molto lontane dalla realtà su tutta l'area analizzata.

Tuttavia, centrando le variabili la situazione migliora notevolmente per quanto riguarda la pressione (*psl*), mentre con la standardizzazione l'unica variabile ancora mal rappresentata rimane solamente la precipitazione.

Si è deciso dunque di escludere la precipitazione (pr) dalla lista dei predittori e standardizzarli per la successiva fase di definizione dei modelli.

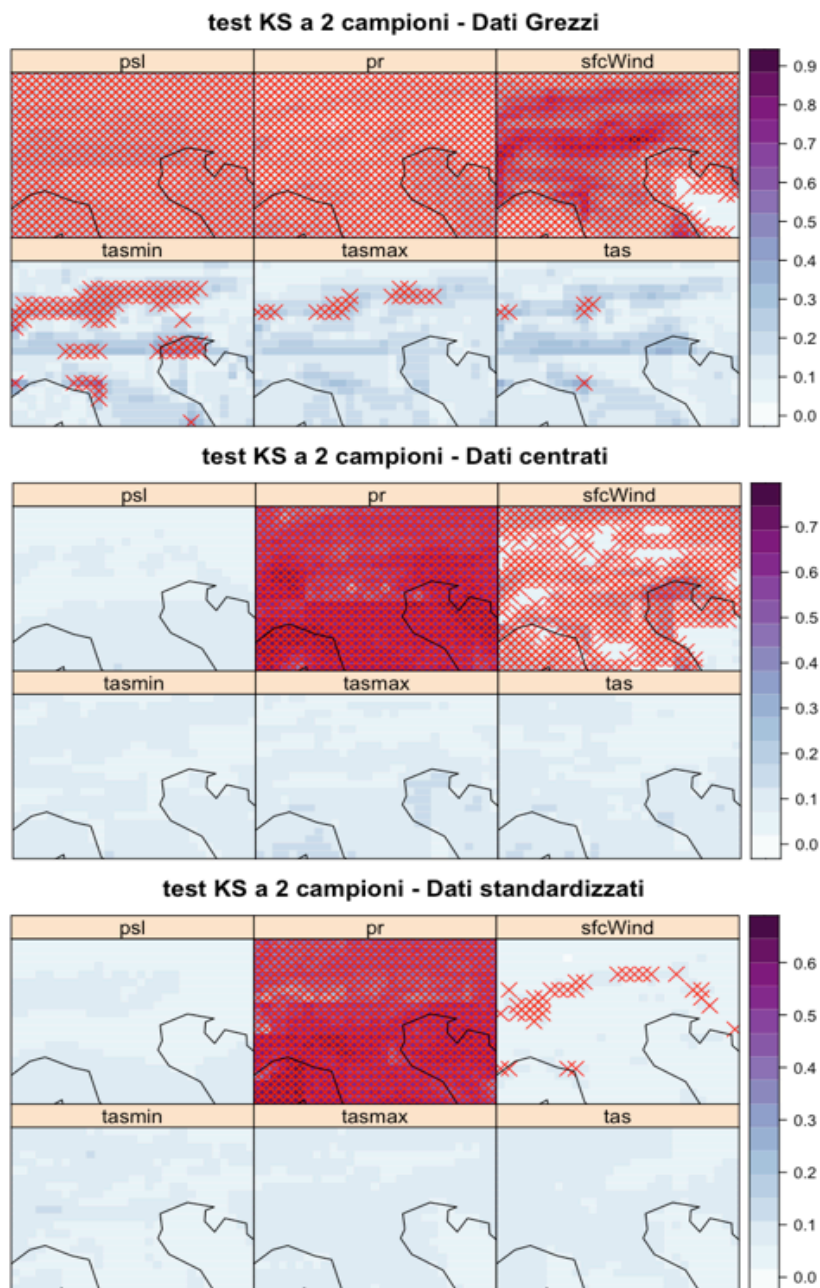


Figura 4 – Valori della statistica KS per i predittori. A colori più scuri corrisponde una maggiore differenza nelle distribuzioni campionarie dei dati CMCC-ES2 ed ERA5 tra il 2010 e il 2015. Le X rosse indicano i punti dove è stata rifiutata l'ipotesi nulla di uguaglianza delle distribuzioni.

4.2.1 Test di Kolmogorov-Sminov a due campioni

Il test KS è un test non parametrico per confrontare l'ipotesi nulla che i due campioni provengano dalla stessa distribuzione contro l'alternativa che abbiano due distribuzioni differenti. La statistica test KS è definita come la massima differenza assoluta tra le funzioni di ripartizione empiriche dei due campioni:

$$D_{n,m} = \sup_x |F_{1,n}(x) - F_{2,n}(x)|$$

Si può notare che poiché le funzioni di ripartizione empiriche assumono valori in $[0, 1]$ lo stesso vale per la statistica test KS, la quale porta a rifiutare l'ipotesi nulla per valori "grandi" (prossimi a 1).

4.3 Selezione del metodo di downscaling

Si sono volute confrontare per diverse combinazioni di predittori (Tabella 4) la regressione lineare generalizzata (indicata in seguito come GLM) e il metodo degli analoghi (ANALOG) al fine di trovare il modello più appropriato ad effettuare in fine il *downscaling* delle raffiche massime giornaliere di vento per gli scenari climatici futuri. Il procedimento e i criteri riprendono quelli tracciati dall'esperimento internazionale VALUE, volto a fornire le linee guida per la valutazione delle tecniche di *downscaling* (Maraun, et al., 2015) (Bedia, et al., 2020).

Nella Tabella 4 si può vedere come, nei casi in cui siano stati usati i predittori solamente standardizzati, si sia deciso di escludere da questi o la temperatura media (tas) o la massima e la minima (tasmax, tasmin). Questo perché le tre variabili sono, per ovvie ragioni, fortemente correlate tra di loro.

Per i GLM sono stati inoltre presi i predittori da zone progressivamente più ampie attorno ai punti delle stazioni: punto più vicino, 9 punti più vicini, 25 punti più vicini; oppure le componenti principali di questi (calcolate su tutta l'area disponibile). Le componenti principali sono state calcolate o singolarmente per predittore (PCA per variabile) oppure anche sulla matrice combinata dei dati

(PCA combinata). La soglia per la selezione delle PCA da includere è stata impostata al 90% della variabilità dei predittori.

4.3.1 Modello Gamma

In letteratura la distribuzione di Weibull si è spesso dimostrata la più idonea nella modellazione del vento (Conradsen, et al., 1984) (Tuller & Brett, 1984:), a maggior ragione se l'interesse è verso i valori massimi giornalieri (*block maxima*), in quanto, secondo la teoria dei valori estremi (Beirlant, et al., 2004) questi tenderebbero ad una distribuzione appartenente alla famiglia delle GEV (distribuzione generalizzata dei valori estremi), di cui la Weibull fa parte.

Tuttavia, in altri studi sulla velocità del vento sono stati ottenuti buoni risultati anche tramite la regressione lineare generalizzata, grazie alla flessibilità della distribuzione Gamma e alla sua più facile implementazione (Yan, et al., 2002).

Si è dunque optato per un modello Gamma-GLM, con funzione di collegamento logaritmica per garantire valori positivi della risposta.

Tabella 4 – Impostazioni dei predittori sperimentate

Modello	Prova	Predittori	Punti griglia	Trasformazione
GLM	G1	tasmax, tasmin, psl, sfcWind	<i>più vicino</i>	<i>Standardizzazione</i>
	G2	tas, psl, sfcWind	<i>più vicino</i>	<i>Standardizzazione</i>
	G3	tasmax, tasmin, psl, sfcWind	<i>9 più vicini</i>	<i>Standardizzazione</i>
	G4	tas, psl, sfcWind	<i>9 più vicini</i>	<i>Standardizzazione</i>
	G5	tasmax, tasmin, psl, sfcWind	<i>25 più vicini</i>	<i>Standardizzazione</i>
	G6	tas, psl, sfcWind	<i>25 più vicini</i>	<i>Standardizzazione</i>
	G7	tasmax, tasmin, tas, psl, sfcWind	<i>tutti</i>	<i>PCA combinate</i>
	G8	tasmax, tasmin, tas, psl, sfcWind	<i>tutti</i>	<i>PCA per variabile</i>
ANALOG	A1	tasmax, tasmin, psl, sfcWind	<i>tutti</i>	<i>Standardizzazione</i>
	A2	tas, psl, sfcWind	<i>tutti</i>	<i>Standardizzazione</i>
	A3	tasmax, tasmin, tas, psl, sfcWind	<i>tutti</i>	<i>PCA combinate</i>
	A4	tasmax, tasmin, tas, psl, sfcWind	<i>tutti</i>	<i>PCA per variabile</i>

4.3.2 Validazione

Per la selezione del modello migliore è stata utilizzata la convalida incrociata, la quale consiste nel dividere iterativamente il dataset in parti complementari, una dedicata alla stima del modello (*training set*) e una dedicata alla validazione (*validation set*). Nello specifico sono stati partizionati i dati (ARPAV ed ERA5) 13 volte: ad ogni iterazione si sono stimati i coefficienti del modello escludendo dal dataset uno dei 13 anni. Ognuno dei 13 modelli così stimati è dunque stato usato per simulare i dati nel proprio anno mancante. Queste simulazioni sono state quindi unite in un'unica serie storica per effettuare un confronto con le osservazioni e valutare la capacità predittiva del modello.

Visto l'interesse per la valutazione degli effetti del cambiamento climatico sui venti estremi nel lungo periodo, l'attenzione non è stata posta sulla capacità del modello di rappresentare l'evoluzione giornaliera del vento, ma piuttosto su alcune sue caratteristiche distributive ritenute rilevanti.

Per ogni stazione sono stati quindi calcolati i seguenti indici di confronto tra serie prevista e serie osservata: differenza relativa delle medie, rapporto delle varianze, radice dell'errore quadratico medio (RMSE), indice di correlazione di Pearson e differenza tra i 98° centili. I valori ottenuti sono poi stati raccolti in dei violin-plot (Figura 5): ogni "violino" corrisponde ad un modello e rappresenta, in modo simile ad un boxplot, la distribuzione tra le 14 stazioni dell'indice in questione; il colore ne indica la media.

Si notano subito delle differenze piuttosto nette tra i modelli GLM e i modelli ANALOG: nonostante la media delle previsioni ANALOG si discosti solo del 4% circa dal valore reale, i modelli di regressione rappresentano molto meglio il fenomeno in media; gli ANALOG al contrario si dimostrano migliori nel catturare la variabilità del fenomeno. Dal RMSE si può vedere che l'errore di previsione è minore nel caso dei GLM e che questi forniscono previsioni più correlate alle osservazioni rispetto agli ANALOG. Dal 98° centile si vede in fine come i metodi migliori nel rappresentare l'estremità destra della distribuzione siano gli analoghi.

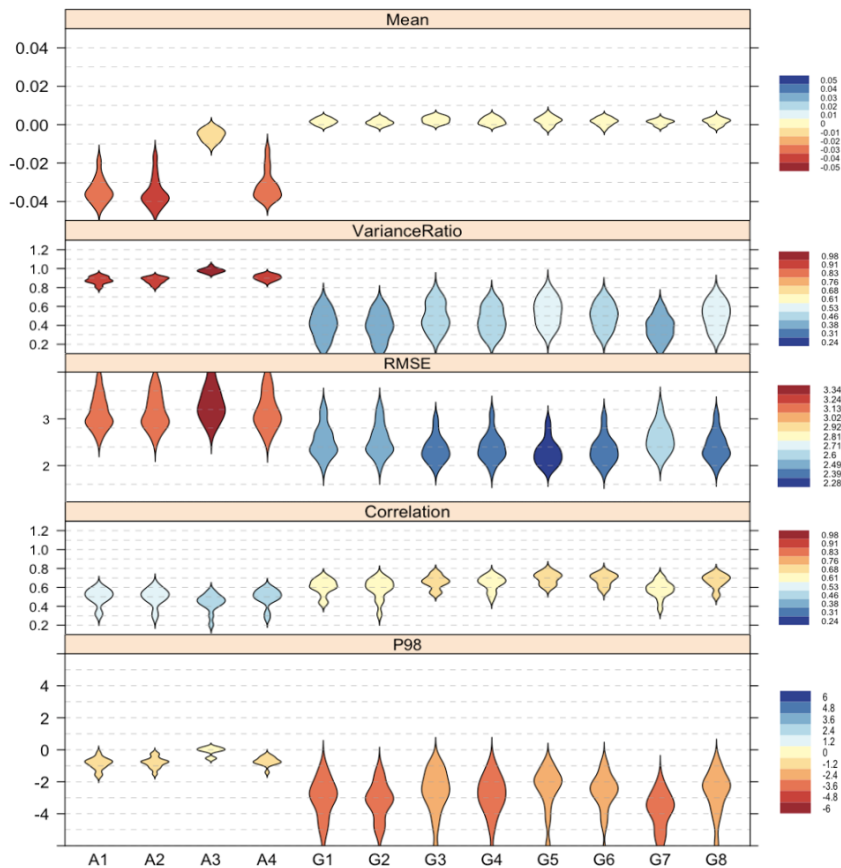


Figura 5 – Violin-plot delle misure di performance dei modelli

Fatte queste considerazioni ed essendo interessati all'andamento futuro dei soli valori alti della risposta, si è optato per il modello A3 per procedere con il *downscaling*.

4.4 Downscaling degli scenari SSP

Prima di applicare il modello alle simulazioni future, ricordando le considerazioni fatte nel paragrafo 4.2 e la scelta di standardizzare i predittori durante la calibrazione del modello, è stato necessario standardizzare le simulazioni future utilizzando media e varianza delle simulazioni passate (che nel test KS si erano rivelate sufficientemente simili in distribuzione alle osservazioni). Dopodiché, sommando la media delle osservazioni e moltiplicando per la deviazione standard delle stesse, si è fatto in modo che

l'inizio della serie simulata dei predittori fosse coerente con le ultime osservazioni. Applicando quindi il metodo di *downscaling* si sono ottenute le proiezioni dei tre scenari SSP per ognuna delle 14 stazioni nel periodo 2015-2100. Per valutare l'andamento della coda superiore della distribuzione, e quindi l'evoluzione dei venti più forti, sono stati calcolati i valori massimi annui e le frequenze assolute annue di giorni con più di 25m/s di vento (ritenuta essere una soglia oltre la quale si possano cominciare a riscontrare danni minori in alcune località). Le serie così ottenute sono riportate nella Figura 6: si può vedere in nero la serie calcolata sulle osservazioni fino al 2022, mentre in blu, verde e rosso, rispettivamente, gli scenari SSP1-2.6 (migliore), SSP2-4.5 (intermedio) e SSP5-8.5 (peggiore).

Nella Figura 7 si evidenzia invece il cambiamento nel quantile campionario di livello 0.98 tra i decenni 2040-2050 (prima riga) e 2090-2100 (seconda riga) e il decennio di riferimento 2010-2020, per ognuno degli scenari SSP (colonne). A valori grandi corrispondono aumenti del quantile e dell'intensità degli eventi rari, mentre a valori piccoli corrisponde un calo del quantile e un'attenuazione degli eventi più intensi.

Dal primo grafico si può vedere che non sembrerebbero esserci sostanziali cambiamenti imputabili al cambiamento climatico. Infatti, le serie storiche simulate non presentano alcuna tendenza evidente. Tuttavia, andando ad osservare le stazioni separatamente attraverso il secondo grafico, sembrerebbe consolidarsi un lieve pattern divergente tra le stazioni più vicine al mare, dove gli eventi intensi sembrerebbero stabili o in leggera diminuzione, rispetto a quelle nell'entroterra e in prossimità delle montagne, dove si evidenzia un leggero aumento dell'intensità, con in particolare due stazioni (in corrispondenza di Vicenza e Vittorio Veneto) dove l'aumento potrebbe essere più sostenuto.

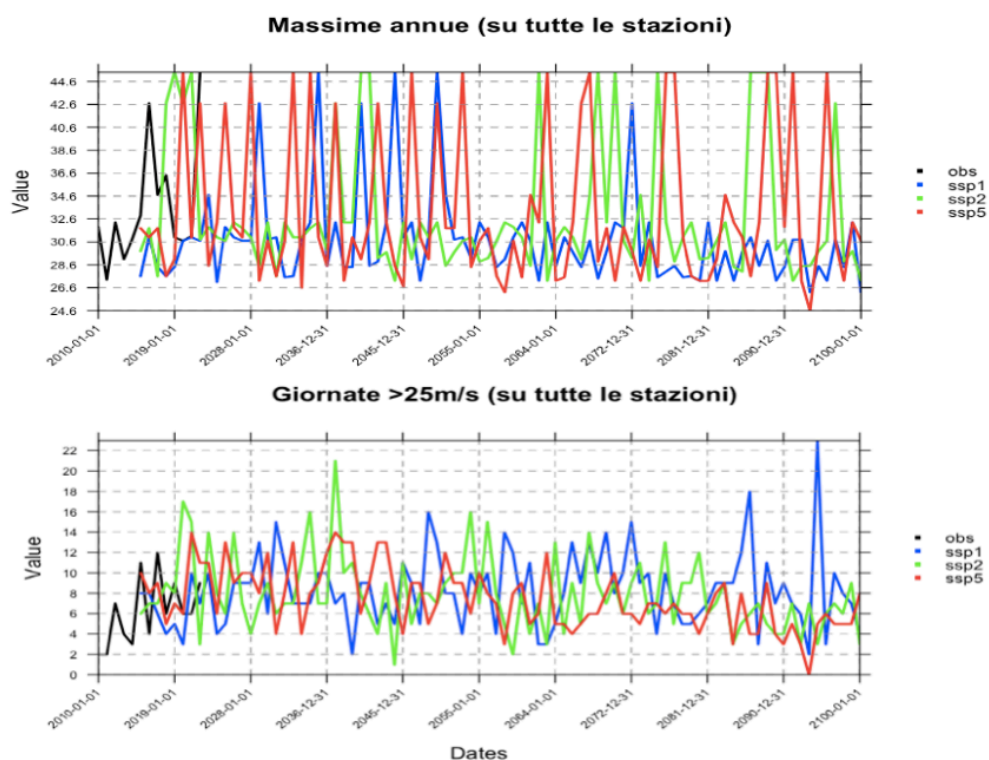


Figura 6 – Statistiche annue calcolate sulle proiezioni 2015-2100 delle raffiche giornaliere di vento.

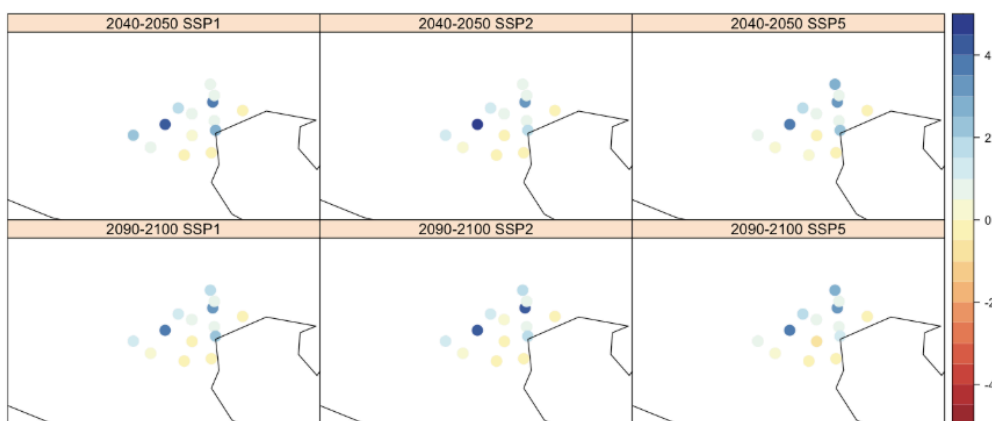


Figura 7 - Differenza dei quantili di livello 0.98 tra il decennio 2010-2022 e le proiezioni nei decenni 2040-2050 (prima riga) e 2090-2100 (seconda riga) per stazione.

Conclusioni

Il *downscaling* statistico, applicato alle proiezioni climatiche per il Veneto fino al 2100, non ha evidenziato un effetto complessivo del cambiamento climatico sugli eventi di vento forte. Tuttavia, guardando i risultati per singole località si è identificata una possibile tendenza decrescente nella zona costiera e, viceversa, di intensificazione nella pianura interna.

Sebbene il metodo degli analoghi sia sembrato il più appropriato tra i modelli testati, per sua costruzione non è capace di restituire valori della risposta diversi dal campione utilizzato durante la calibrazione del modello. Questa caratteristica potrebbe aver portato ad una sottostima delle code della distribuzione. Per ovviare a questo problema, in futuro potrebbe essere d'interesse sviluppare un modello più avanzato, sfruttando possibilmente la distribuzione generalizzata dei valori estremi (GEV), più adatta ai dati di tipo *block maxima*.

Un ulteriore miglioramento potrebbe essere effettuato sotto il punto di vista dei predittori. In questo studio ci si è limitati a temperatura (massima, minima e media), pressione e vento superficiali giornalieri; tuttavia, cambiando la frequenza temporale di riferimento (da valori giornaliere a valori mensili) o il GCM, sono disponibili molte altre variabili più specifiche e misurate a quote diverse, le quali possono apportare ulteriore informazione.

Bibliografia

- Rummukainen, M., 2010. State-of-the-art with regional climate models. *WIREs Climate Change*, pp. 82-96.
- Klein, W. H., Lewis, B. M. & Enger, I., 1959. Objective prediction of five-day mean temperatures during winter. *Journal of Atmospheric Sciences*, pp. 672–682.
- Ramon, J. et al., 2021. A perfect prognosis downscaling methodology for seasonal prediction of local-scale wind speeds. *Environmental Research Letters*, 16(5).
- Maraun, D. & Widmann, M., 2018. *Statistical Downscaling and Bias Correction for Climate Research*. s.l.:Cambridge University Press.
- Glahn, H. R. & Lowry, D. A., 1972. The Use of Model Output Statistics (MOS) in Objective Weather Forecasting. *Journal of Applied Meteorology and Climatology*, 11(8), pp. 1203-1211.
- Roux, R. L. et al., 2018. Comparison of statistical and dynamical downscaling results from the WRF model. *Environmental Modelling & Software*, Volume 100, pp. 67-73.
- Scott Spak, T. H. B. L. R. G., 2007. A comparison of statistical and dynamical downscaling for surface temperature in North America. *Journal of Geophysical Research: Atmospheres*, 112(D8).
- Chandler, R. E. & Wheeler, H. S., 2002. Analysis of rainfall variability using generalized linear models: A case study from the west of Ireland. *Water Resources Research*, 38(10), pp. 1-11.
- Yang, C. et al., 2005. Simulation and downscaling models for potential evaporation. *Journal of Hydrology*, 302(1-4), pp. 239-254.
- San-Martín, D. et al., 2017. Reassessing Model Uncertainty for Regional Projections of Precipitation with an Ensemble of Statistical Downscaling Methods. *Journal of Climate*, 30(1), pp. 203-223.
- Hewitson, B. & Crane, R. G., 2002. Self-Organizing Maps: Applications to synoptic climatology. *Climate Research*, 22(1), pp. 13-26.
- Uvo, C. B. et al., 2001. Statistical atmospheric downscaling for rainfall estimation in Kyushu Island, Japan. *Hydrology and Earth System Science*, 5(2), pp. 259–271.
- Haylock, M. R. et al., 2006. Downscaling heavy precipitation over the United Kingdom: a comparison of dynamical and statistical methods and their future scenarios. *International Journal of Climatology*, 26(10), pp. 1397-1415.
- Zorita, E., Kharin, V. & Von Storch, H., 1992. The Atmospheric Circulation and Sea Surface Temperature in the North Atlantic Area in Winter: Their Interaction and Relevance for Iberian Precipitation. *Journal of Climate*, 5(10), pp. 1097-1108.
- Bretherton, C. S., Smith, C. & Wallace, J. M., 1992. An Intercomparison of Methods for Finding Coupled Patterns in Climate Data. *Journal of Climate*, 5(6), pp. 541–560.
- Phillips, N. A., 1956. The general circulation of the atmosphere: A numerical experiment. *Quarterly Journal of the Royal Meteorological Society*, 82(352), pp. 123-164.

- Goosse, H., 2015. *Climate System Dynamics and Modelling*. s.l.: Cambridge University Press.
- Brands, S., Gutiérrez, J. M., Herrera, S. & Cofiño, A. S., 2012. On the Use of Reanalysis Data for Downscaling. *Journal of Climate*, 25(7), pp. 2517–2526.
- Manzanas, R. et al., 2015. Statistical Downscaling in the Tropics Can Be Sensitive to Reanalysis Choice: A Case Study for Precipitation in the Philippines. *Journal of Climate*, 28(10), pp. 4171–4184.
- Sterl, A., 2004. On the (In)Homogeneity of Reanalysis Products. *Journal of Climate*, 17(19), pp. 3866–3873.
- Maraun, D., Widmann, M. & Gutiérrez, J. M., 2019. Statistical downscaling skill under present climate conditions: A synthesis of the VALUE perfect predictor experiment. *International Journal of Climatology*, Volume 39, pp. 3692–3703.
- Benestad, R., 2016. *Downscaling Climate Information*, Oxford: Oxford Research Encyclopedia of Climate Science.
- Matulla, C. et al., 2008. Influence of similarity measures on the performance of the analog method for downscaling daily precipitation. *Climate Dynamics*, Volume 30, pp. 133–144.
- Zorita, E. & Storch, H. v., 1999. The Analog Method as a Simple Statistical Downscaling Technique: Comparison with More Complicated Methods. *Journal of Climate*, 12(8), pp. 2474–2489.
- Gutiérrez, J. M. et al., 2013. Reassessing Statistical Downscaling Techniques for Their Robust Application under Climate Change Conditions. *Journal of Climate*, 26(1), pp. 171–188.
- Gutiérrez, J. M. et al., 2019. An intercomparison of a large ensemble of statistical downscaling methods over Europe: Results from the VALUE perfect predictor cross-validation experiment. *International Journal of Climatology*, 39(9), pp. 3750–3785.
- Iturbide, M. et al., 2019. The R-based climate4R open framework for reproducible climate data access and post-processing. *Environmental Modelling & Software*, Volume 111, pp. 42–54.
- Conradsen, K., Nielsen, L. B. & Prahm, L. P., 1984. Review of Weibull statistics for estimation of wind speed distributions. *Journal of Applied Meteorology and Climatology*, 23(8), pp. 1173–1183.
- Tuller, S. E. & Brett, A. C., 1984. The characteristics of wind velocity that favor the fitting of a Weibull distribution in wind speed analysis. *Journal of Applied Meteorology and Climatology*, 23(1), pp. 124–134.
- Beirlant, J. et al., 2004. *Statistics of Extremes: Theory and Applications*. s.l.:John Wiley & Sons, Ltd.
- Yan, Z. et al., 2002. An Analysis of Daily Maximum Wind Speed in Northwestern Europe Using Generalized Linear Models. *Journal of Climate*, 15(15), pp. 2073–2088.
- Flato, G. et al., 2013. Evaluation of Climate Models. In: *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge: Cambridge University Press.

Maraun, D. et al., 2015. VALUE: A framework to validate downscaling approaches for climate change studies. *Earth's Future*, Volume 3, pp. 1-14.

Bedia, J. et al., 2020. Statistical downscaling with the downscaleR package (v3.1.0): contribution to the VALUE intercomparison experiment. *Geoscientific Model Development*, 13(3), pp. 1711–1735.

Casanueva, A. et al., 2014. Statistical downscaling of climate impact indices: testing the direct approach. *Climatic Change*, 127(3-4), pp. 547–560.

von Storch, H., Zorita, E. & Cubasch, U., 1993. Downscaling of global climate change estimates to regional scales: an application to Iberian rainfall in wintertime. *Journal of Climate*, 6(6), pp. 1161– 1171.

Wilby, R. L. et al., 2004. Guidelines for use of climate scenarios developed from statistical downscaling methods. *IPCC Task Group on Data and Scenario Support for Impact and Climate Analysis (TGICA)*.

Huth, R., 1999. Statistical downscaling in central Europe: evaluation of methods and potential predictors. *Climate Research*, Volume 13, pp. 91-101.

Bjornsson, H. & Venegas, S. A., 1997. *A Manual for EOF and SVD analyses of Climatic Data*, s.l.: McGill University.

Dool, H. V. D., 1994. Searching for analogues, how long must we wait?. *Tellus A: Dynamic Meteorology and Oceanography*, 46(3), pp. 314-324.