

ALMA MATER STUDIORUM · UNIVERSITÀ DI BOLOGNA

Scuola di Scienze
Corso di Laurea in Fisica

Analisi delle previsioni mensili mediante il modello GLOBO (ISAC-CNR)

Relatore:
Prof. Andrea Buzzi

Presentata da:
Luca Cantarello

Correlatore:
Dott. Daniele Mastrangelo

Sessione II
Anno Accademico 2013/2014

A papà

“Weather forecasting is one of the most success stories in this book, a case of man and machine joining forces to understand and sometimes anticipate the complexity of nature”

Nate Silver
The signal and the noise

Indice

Sommario	8
1 Introduzione	11
2 Le previsioni mensili	13
2.1 La peculiarità delle previsioni mensili	13
2.2 Le previsioni di ensemble	14
2.2.1 Utilizzo e rappresentazioni delle previsioni di ensemble	15
2.3 Lo sviluppo storico delle previsioni mensili	17
2.3.1 I lavori degli anni 80 e 90	18
2.3.2 Simulazioni nel passato: reforecasting e reanalisi	19
2.3.3 Il panorama attuale	20
2.4 La verifica di una previsione	20
2.4.1 Gli score per un campo di previsione	21
2.4.2 Gli score probabilistici	23
2.4.3 Gli esiti dei lavori degli Anni 80 e 90	26
2.5 Il modello dell'ECMWF	35
2.5.1 Progressi e performance del modello dell'ECMWF	36
3 Le previsioni mensili presso il CNR-ISAC	41
3.1 I tre modelli sviluppati presso il CNR-ISAC	41
3.2 GLOBO per il medio-termine	42
3.2.1 Rappresentazione delle coordinate	43
3.2.2 Dinamica e schemi di parametrizzazione	43
3.2.3 Condizioni iniziali	45
3.3 GLOBO per le previsioni mensili	45
3.3.1 Struttura e specifiche	45
3.3.2 Il bias del modello GLOBO per previsioni mensili	47
4 Bias e RMSE del modello GLOBO	51
4.1 Simulazioni e metodo di lavoro	51

4.1.1	Reanalisi e reforecast	51
4.1.2	Bias	52
4.1.3	RMSE	53
4.2	Risultati	54
4.2.1	Bias	54
4.2.2	RMSE	55
5	Conclusioni	75
	Bibliografia	77
	Appendice	79
	Ringraziamenti	101

Sommario

In questo lavoro sono presentate le principali caratteristiche delle previsioni meteorologiche mensili, nonché il progresso scientifico e storico che le ha coinvolte e le tecniche adibite alla loro verifica. Alcune di queste tecniche sono state applicate al fine di valutare ed analizzare l'errore sistematico (o bias) e l'RMSE di temperatura a 850 hPa (T850), altezza geopotenziale a 500 hPa (Z500) e precipitazioni cumulate del modello GLOBO, utilizzato presso l'Istituto per le Scienze dell'Atmosfera e del Clima del Consiglio Nazionale delle Ricerche per formulare previsioni mensili.

I risultati mostrano la progressione temporale dell'errore, che aumenta nelle prime due settimane di integrazione numerica fino a stabilizzarsi tra la terza e la quarta. Ciò mostra che il modello, persa l'influenza delle condizioni iniziali, raggiunge un suo stato che, per quanto fisiologicamente distante da quello osservato, tende a stabilizzarsi e a configurarsi quindi come sistematico (eventualmente facilitandone la rimozione in fase di calibrazione delle previsioni). Il bias di T850 e Z500 presenta anomalie negative prevalentemente lungo le zone equatoriali, e vaste anomalie positive sulle aree extra-tropicali; quello delle precipitazioni mostra importanti sovrastime nelle zone continentali tropicali. La distribuzione geografica dell'RMSE (valutato solo per T850 e Z500) riscontra una generale maggiore incertezza nelle zone extra-tropicali, specie dell'emisfero settentrionale e nei mesi freddi.

1 Introduzione

*“Perhaps some day in the dim future it will be possible to advance the computations faster than the weather advances and at a cost less than the saving to mankind due to the information gained.
But that is a dream.”*

L.F.Richardson
Weather prediction by numerical process

Da quando nel 1922 Lewis Fry Richardson pubblicò il suo *Weather prediction by numerical process*, la meteorologia si è lasciata molta strada alle spalle. Da quel primo precoce tentativo di integrare le equazioni che descrivono i moti atmosferici al fine di prevederne lo stato futuro (un tedioso lavoro che nei primi anni Venti non poteva che essere “manuale”, senza il supporto oggi indispensabile di computer ed elaboratori), si è giunti in meno di un secolo ai modelli numerici sviluppati dai centri di calcolo, in grado di formulare previsioni che possono spingersi fino a molti giorni nel futuro, operare a diverse risoluzioni (dall’intero pianeta a ristrette aree geografiche), mappare molteplici variabili meteorologiche a svariate quote della troposfera.

Alla base delle moderne previsioni del tempo ci sono i cosiddetti modelli numerici di previsione meteorologica (che la lingua inglese ha riassunto nell’acronimo NWP: Numerical Weather Prediction). Il loro scopo è quello di riprodurre il più fedelmente possibile il comportamento dell’atmosfera negli stati successivi a quelli (conosciuti con la migliore precisione e affidabilità) in cui sono stati inizializzati.

Questi modelli lavorano principalmente su griglie di punti¹ poste su più livelli verticali, per ciascuno dei quali svolgono l’integrazione numerica delle equazioni della fluidodinamica e della termodinamica, che viene oggi affidata a computer e processori dotati di ingenti capacità di calcolo. Essendo inoltre l’atmosfera un sistema altamente caotico, le leggi che ne regolano i moti sono fortemente dipendenti dalle condizioni iniziali: una piccola variazione nei dati di inizializzazione può generare profonde differenze nelle soluzioni e dunque nell’evoluzione dello stato dell’atmosfera previsto dal modello. La naturale ed intrinseca imprecisione contenuta in tutte le misure fisiche, assieme con la necessità di interpolare certi tipi di dati altrimenti non campionabili in modo diretto (per esempio al

¹Esiste anche una classe di modelli numerici alternativi – detti *spettrali* – che sfruttano l’omonimo metodo per la rappresentazione dei campi meteorologici lungo le dimensioni orizzontali

di sopra degli oceani), rende impossibile disporre di condizioni iniziali perfette; ciò implica la presenza inevitabile di un certo grado di incertezza sui dati di inizializzazione, che si ripercuote a sua volta sulla previsione finale. Nei modelli meteorologici è presente poi un'ulteriore duplice fonte di errore: le equazioni che vengono integrate non hanno una soluzione analitica e dunque vengono approssimate numericamente; inoltre sono introdotte parametrizzazioni e semplificazioni, nell'impossibilità di riprodurre alcuni fenomeni fisici con lo stesso grado di complessità della realtà.

La propagazione nel tempo e cioè nella fase di integrazione numerica delle due fonti di errore (l'incertezza sulle condizioni iniziali e quella dovuta alla struttura del modello) assieme con l'intrinseca caoticità del sistema pongono dei severi limiti temporali alle previsioni meteorologiche, il cui grado di affidabilità decresce man mano che procede il tempo di integrazione. In un celebre lavoro di E.N.Lorenz (Lorenz 1963, [10]) tale limite temporale fu fissato ad una soglia teorica di due settimane.

Nella pratica, il campo delle previsioni deterministiche (così sono definite quelle legate all'integrazione nel tempo di equazioni date le condizioni iniziali) riesce oggi a produrre risultati di affidabilità accettabile entro i dieci giorni di previsione, nonostante i progressi raggiunti negli ultimi decenni nel campo del calcolo numerico e il miglioramento dei modelli nella descrizione dei processi microfisici e delle condizioni iniziali. Tutto ciò che supera questo range temporale sconfinava in un terreno nel quale il modello si spinge tanto in là nel tempo da perdere qualunque memoria delle condizioni iniziali con le quali è stato inizializzato.

Tuttavia, a partire dagli anni Ottanta, in molti si sono cimentati nel superamento di quella soglia, aprendo la strada a studi e ricerche su quelle che vengono definite *previsioni mensili* e *previsioni stagionali*, per le quali negli anni sono stati elaborati appositi sistemi di previsione ([5], [4], [13], [1]).

In questo lavoro ci si concentrerà esclusivamente sulle prime: nel capitolo 2 se ne fornirà una descrizione generale con l'aiuto di qualche esempio concreto (modelli per previsioni mensili attualmente operativi), la presentazione delle principali tecniche con le quali vengono elaborate (previsioni di ensemble), nonché gli strumenti statistici necessari al fine di valutarne la predicibilità (reforecasting, skill score).

Nel capitolo 3 verrà poi illustrato il modello GLOBO, elaborato dall'Istituto per le Scienze dell'Atmosfera e del Clima del Consiglio Nazionale delle Ricerche e dal 2010 utilizzato anche per l'elaborazione di previsioni mensili riferite all'Italia e all'Europa. Nel capitolo 4 verranno calcolati il bias (o errore sistematico) e l'RMSE di tale modello, tramite la tecnica del reforecasting, con il supporto di un apposito linguaggio di programmazione. Nel 5, ovvero nelle conclusioni, saranno infine discussi i risultati.

2 Le previsioni mensili

2.1 La peculiarità delle previsioni mensili

A seconda che una previsione meteorologica dipenda dalle condizioni iniziali o da quelle al contorno, sono definiti due tipi di *predicibilità* (Lorenz 1975, [9]):

- *predicibilità del I tipo*: è riferita alla dipendenza dalle condizioni iniziali e dunque riguarda le previsioni a breve e medio termine, per le quali non è possibile ottenere risultati affidabili oltre il limite teorico delle due settimane (Lorenz 1963, [10]);
- *predicibilità del II tipo*: è riferita alla dipendenza dalle condizioni al contorno e dunque riguarda le previsioni a lungo termine (mensili, stagionali, ...) per le quali si rende necessario un approccio probabilistico (cfr. §2.2) o statistico, che faccia emergere i pattern e le configurazioni che caratterizzano l'andamento medio del tempo atmosferico nel periodo considerato (Shukla 1981, [14]).

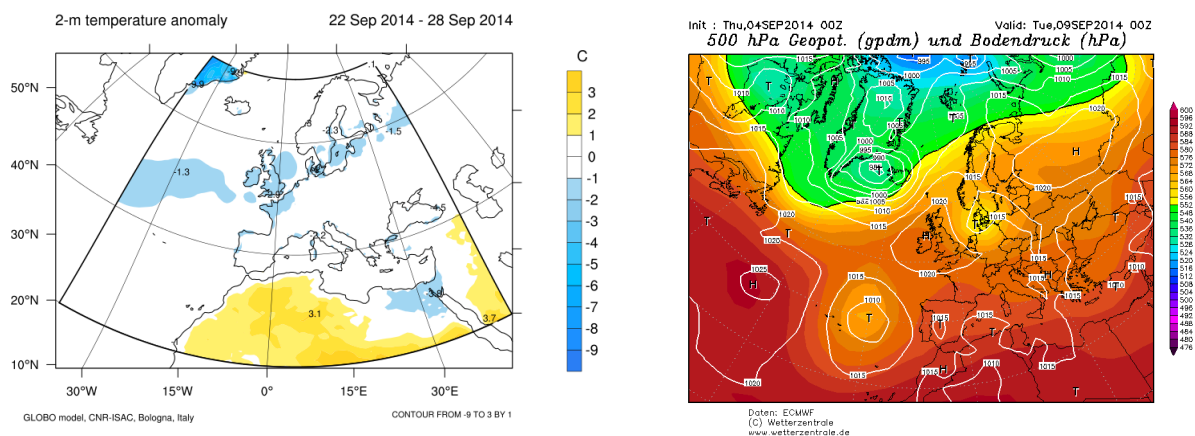
Per quanto detto, quindi, le previsioni meteorologiche mensili si distinguono da quelle a breve o medio termine per la scala spazio-temporale che le caratterizza e per la metodologia usata (probabilistica invece di deterministica). A distanza di trenta giorni, infatti, nessuno potrà mai formulare una previsione sulla temperatura massima o sulla quantità di neve o di pioggia previste in un certo luogo. Sarà invece più naturale esprimersi in termini di probabilità di anomalie (di temperatura, pressione, precipitazioni) riferite a certe aree (ad esempio l'Italia, l'Europa settentrionale, il Mediterraneo centrale), dove per anomalia si intende la differenza tra il valore della previsione p (o del dato osservato o) e quello della media climatologica, c (il diverso approccio nei due tipi di previsione è ben mostrato dalla figura (2.1)):

$$p' = p - c \tag{2.1}$$

$$o' = o - c \tag{2.2}$$

Sotto quest'ultimo aspetto le previsioni mensili hanno più a che vedere con le previsioni stagionali o multi-stagionali, benché rispetto a queste trenta giorni restino un periodo di tempo troppo breve per poter sfruttare la variabilità temporale di alcune grandezze, come il contenuto di calore degli oceani e dunque l'andamento della loro temperatura (Sea Surface Temperature, SST), che ha cicli di variazione più lunghi e più lenti. Tuttavia, alle previsioni mensili sono di supporto fonti alternative di predicibilità, fornite da fenomeni atmosferici caratterizzati da una scala di variabilità temporale interstagionale che negli anni sono stati studiati ed approfonditi. Tra le più utilizzate e conosciute vi sono l'oscillazione di Madden-Julian (Madden-Julian Oscillation, MJO), le condizioni iniziali della stratosfera, l'innevamento del suolo e l'oscillazione nord-atlantica (North Atlantic Oscillation, NAO) ([18]).

Figura 2.1: Raffronto tra una previsione a medio termine e una a lungo termine



(a) Esempio di rappresentazione di una previsione mensile: anomalia prevista per il campo di temperatura a 2 metri. *Fonte: isac.cnr.it*

(b) Esempio di rappresentazione di una previsione a medio termine: campo previsto dell'altezza geopotenziale a 500 hPa e della pressione al suolo. *Fonte: wetterzentrale.de*

L'interesse e l'attenzione rivolti allo sviluppo delle previsioni mensili sono frutto soprattutto di questa particolare condizione "intermedia" che le contraddistingue, questo appartenere ad un sorta di "terra di nessuno", caratterizzata da un range temporale troppo lungo per essere oggetto di previsioni deterministiche e troppo breve per applicare le tecniche di previsione stagionale.

2.2 Le previsioni di ensemble

Il problema della sensibilità dei modelli atmosferici alle condizioni iniziali può essere affrontato dal punto di vista statistico allo scopo di migliorarne il potere predittivo nel lungo o lunghissimo termine (cfr. §7.6 Ensemble forecasting, [20]).

Se si considera, per fare un esempio concreto, una singola previsione deterministica per l'intero emisfero boreale a due giorni del campo di temperatura al suolo, questa può essere intesa come un singolo campione estratto da un insieme di infiniti campi dello stesso tipo, tutti potenzialmente frutto di quella stessa previsione. Si può dunque attribuire un significato probabilistico alle singole emissioni di un modello, ovvero pensandole come membri appartenenti ad una distribuzione di probabilità.

Se ora si trasportano tali concetti alla meccanica classica, si può pensare di descrivere lo stato dell'atmosfera come un punto che si muove nel tempo in uno spazio n-dimensionale, nel quale ciascun grado di libertà rappresenti il campo di una certa variabile atmosferica

ad un certa quota (temperatura al suolo, pressione sul livello del mare, temperatura a 850 hPa, eccetera). In una visione di questo tipo, le condizioni iniziali sono descritte dal punto nell'istante di tempo $t_0 = 0$. Come detto, però, le condizioni iniziali non sono mai conosciute con precisione assoluta: vi sarà sempre un certo grado di incertezza a caratterizzarle. È dunque lecito pensarle non tanto come un punto isolato nello spazio delle fasi, quanto piuttosto come una densità di probabilità attorno ad esso.

Tale cambio di paradigma incide anche sulla natura stessa dei modelli meteorologici: quelle che vengono fatte evolvere nel tempo non sono più condizioni che hanno a che vedere soltanto con la dinamica e la termodinamica dell'atmosfera, bensì acquisiscono il significato di vere e proprie distribuzioni di probabilità, che vengono deformate nel tempo dalle leggi della fluido-dinamica. Per questo motivo, le previsioni frutto di questo nuovo approccio furono battezzate *previsioni dinamiche stocastiche* da colui che se ne occupò per la prima volta (E. Epstein, 1969). La difficoltà storica da subito evidente con questo nuovo tipo di previsioni fu data dall'intrattabilità analitica delle equazioni dinamiche stocastiche che dovevano essere risolte per elaborarle. La ricerca di un metodo che permettesse di integrarle è parte integrante della storia delle previsioni ensemble (si veda, a riguardo, quanto proposto da Leith sul metodo Monte Carlo, [15]).

Il principio sul quale queste si basano è dunque quello di estrarre un campione finito di condizioni iniziali dalla distribuzione di probabilità che ne descrive l'incertezza, seguendo poi l'evoluzione nel tempo delle singole traiettorie nello spazio delle fasi. Dal punto di vista pratico questo si traduce in complesse tecniche statistiche e analitiche che vengono utilizzate per perturbare le condizioni iniziali "reali", al fine di ottenere quelle usate per l'inizializzazione dei vari membri dell'ensemble.

Nelle prime fasi dell'integrazione numerica si osserva come tutti i run dell'ensemble si assomiglino, e si discostino poco l'uno dall'altro. Man mano però che la previsione procede nel tempo questi tendono a sparpagliarsi, fino a rappresentare, in alcuni casi, condizioni atmosferiche radicalmente diverse tra loro (vedi figura 2.2).

Oltre che per testare la sensibilità alle diverse condizioni iniziali, le previsioni di ensemble sono pensate per valutare anche l'incidenza sulle performance del modello di certi aspetti della sua struttura interna (per esempio il passo di griglia, l'orografia, le parametrizzazioni dei fenomeni microfisici), che talvolta distinguono un membro da un altro.

Generalmente il run inizializzato con le migliori condizioni iniziali è di controllo, mentre gli altri sono chiamati perturbazioni.

2.2.1 Utilizzo e rappresentazioni delle previsioni di ensemble

Le previsioni di ensemble possono essere sfruttate in svariati modi.

Una prima possibilità è quella di mediare tutte le perturbazioni allo scopo di ottenere una previsione più accurata delle singole corse modellistiche, nonché del run di controllo. Lo stesso Epstein si accorse infatti che la previsione ottenuta dalla media di tutte le

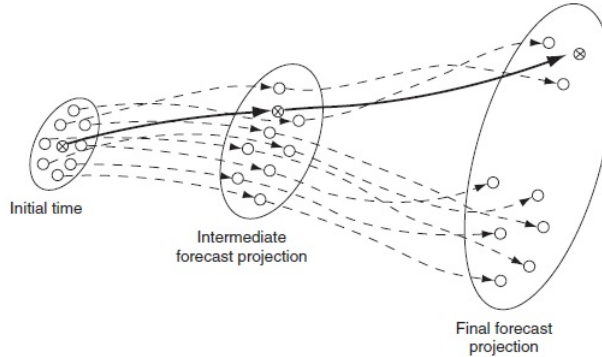


Figura 2.2: Rappresentazione schematica in due dimensioni di una previsione di ensemble. *Fonte immagine: Wilks 2011, [20]*

corse modellistiche dell'ensemble differiva da quella ottenuta dal run inizializzato con la condizione iniziale media, e in generale notò che le condizioni iniziali “ufficiali” non erano garanzia di migliori performance del run modellistico.

La spiegazione di questo comportamento risiede a ben vedere nella non linearità delle equazioni fisiche coinvolte nei modelli meteorologici, per le quali vale che:

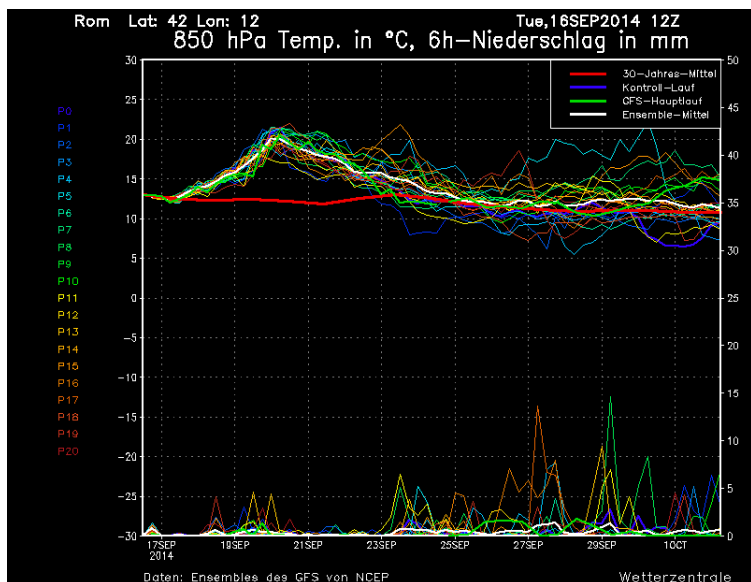
$$\frac{1}{N} \cdot \sum_{k=1}^N f(x_k) \neq f\left(\frac{1}{N} \cdot \sum_{k=1}^N x_k\right) \quad (2.3)$$

cioè la funzione calcolata nel valor medio non coincide con il valor medio della funzione. Tuttavia, la media ensemble non è sempre garanzia di buona previsione. Vi sono alcuni casi in cui le varie perturbazioni tendono a raggrupparsi in sottogruppi (come si può notare ancora nella figura 2.2). In queste situazioni la media complessiva non rappresenta in modo soddisfacente nessuna delle possibili configurazioni previste. In tali circostanze può essere utile utilizzare tecniche tipo quella del *clustering* (Ferranti et al, 2014, [6]) che, raggruppando tra loro le varie perturbazioni che si assomigliano, permette di dedurre i pattern atmosferici principali da cui poter formulare una previsione.

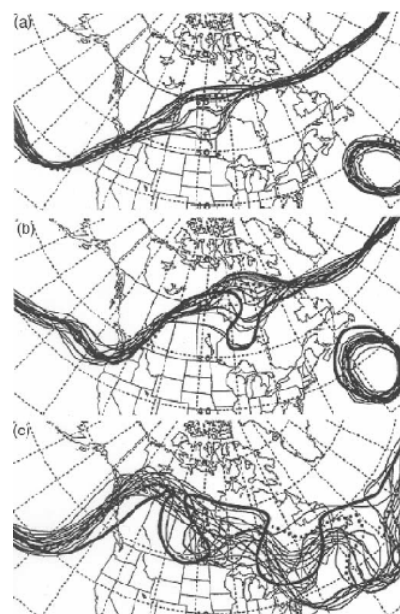
Per sua stessa natura, ovvero poiché permette di affiancare alla media delle previsioni la loro distribuzione di probabilità, la tecnica dell'ensemble fornisce informazioni sul grado di predicibilità del periodo in esame: più divergono i vari membri (o meno è possibile raggrupparli in cluster ben definiti), più sarà difficile ricavare informazioni affidabili sul tempo atmosferico nel lungo termine.

Un altro vantaggio delle previsioni di ensemble è la loro versatilità nelle rappresentazioni grafiche (si veda la figura 2.3). Sovrapponendo su mappe o grafici le isolinee dei campi (di temperatura, altezza geopotenziale, eccetera) previsti dai vari run di un ensemble si ottengono i cosiddetti *spaghetti*. Scegliendo invece di affiancarli in un'unica immagine

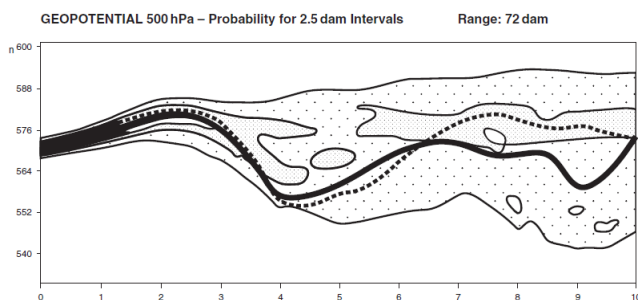
Figura 2.3: Possibili rappresentazioni grafiche di una previsione ensemble: (a) meteogramma a spaghetti, (b) mappa a spaghetti, (c) grafico plume, (d) mappa a francobolli



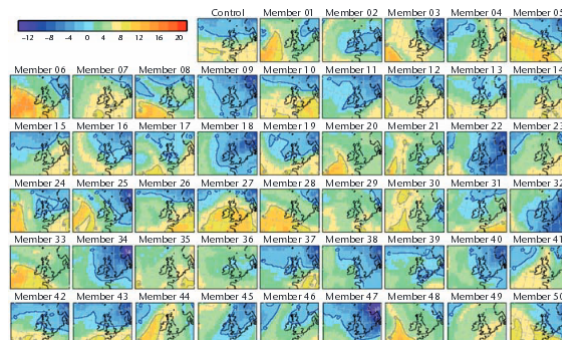
(a) Fonte: *wetterzentrale.de*



(b) Fonte: *Wilks 2011, [20]*



(c) Fonte: *Wilks 2011, [20]*



(d) Fonte: *WMO-No. 1091, [16]*

si ottengono le cosiddette *mappe a francobolli*. Altri esempi di rappresentazione grafica sono i meteogrammi ensemble e i cosiddetti grafici *plume*.

2.3 Lo sviluppo storico delle previsioni mensili

Lo sviluppo e l'interesse per le previsioni mensili sono cominciati a partire dagli anni Ottanta, agevolati dall'avvento di un deciso miglioramento nel potere computazionale

dei computer. Questo fatto permise infatti di estendere le corse dei modelli numerici oltre i cinque giorni (che allora consisteva ancora nel limite massimo di predicibilità) e di introdurre le previsioni di ensemble (cfr. §2.2).

Benché fossero già state fatte numerose simulazioni da parte di ricercatori e scienziati, il primo modello globale a produrre previsioni numeriche oltre i cinque giorni entrò in funzione nel 1982, mentre nel dicembre del 1988 divenne operativo il primo ensemble a trenta giorni di nove membri ([12]). Tuttavia, le previsioni mensili sono rimaste soltanto un campo di ricerca oggetto di simulazioni ed esperimenti riferiti al passato, almeno fino agli anni Duemila (cfr. §2.3.2). Dall'inizio del nuovo millennio, invece, i principali centri previsionali del mondo hanno introdotto modelli numerici che effettuano quotidianamente (o comunque periodicamente) previsioni a lungo o lunghissimo termine.

2.3.1 I lavori degli anni 80 e 90

Il gennaio del 1977 è ricordato dalla climatologia statunitense a causa di una intensa invernata che colpì il Nord America orientale. Un'ondata di freddo eccezionale interessò a più riprese e per l'intero mese buona parte degli Stati Uniti orientali. Persino in Florida, tra il 19 e il 20 gennaio, la neve si spinse fino alla città Miami e molte località batterono i loro record di temperatura minima, rimasti imbattuti almeno fino al 2012¹. Alcune zone dell'Ohio ricordano quello stesso periodo come uno dei più rigidi di sempre: le temperature medie mensili di gennaio '77 risultarono di anche 8-10 gradi centigradi al di sotto della norma (Miyakoda et al. 1983, [5]).

Accadimenti di tale portata spinsero molti meteorologi ad analizzare e ad approfondire l'analisi di quell'evento. In particolare, venne osservato il persistere durante tutto il mese di un forte anticiclone di blocco tra l'Oceano Pacifico nord-orientale, l'Alaska, e il nord-ovest dell'America Settentrionale. Una configurazione con quelle caratteristiche costituiva un ottimo punto di partenza per chi volesse interrogarsi sul se e quanto si fosse in grado di predire allora i caratteri di durata e persistenza nel tempo, anche dopo molti giorni di integrazione, di un pattern di quel tipo.

Uno dei primi studi in assoluto (Miyakoda et al. 1983, [5]) consistette infatti in una serie di simulazioni con lo scopo di testare le capacità di due modelli numerici dell'epoca (uno a griglia e uno spettrale) di predire quell'evento di blocking con un mese di anticipo. Lo studio prevedeva l'utilizzo di tre diverse versioni delle condizioni iniziali del 1° gennaio del 1977 e di due diverse implementazioni per la descrizione dei fenomeni fisici al di sotto del passo di griglia. Proprio il miglioramento nelle tecniche di parametrizzazione risultarono determinanti nel successo di quel lavoro (i risultati di questo e dei successivi studi sono riportati più avanti: cfr. §2.4.3).

Alcuni anni dopo, il modello considerato affidabile nel primo lavoro fu usato come base per incardinare una serie di nuove simulazioni (Miyakoda et al. 1986, [4]), con maggiore

¹<http://www.srh.noaa.gov/images/mfl/news/SnowSouthFlorida35th.pdf>

attenzione alla variazione delle condizioni iniziali. Lo studio riguardò tutti i mesi di gennaio dal 1977 al 1983 e di fatto confermò i risultati già ottenuti nella prima ricerca. In quegli anni un altro lavoro (Shukla 1981, [14]) aveva incluso il gennaio del '77 in alcune simulazioni numeriche a lunghissimo termine (fino a 60 giorni). Anche in quel caso i risultati mostrarono una discreta capacità predittiva del modello preso in esame entro i trenta giorni di integrazione numerica, soprattutto in relazione alla scelta di scale spazio-temporali opportune (cfr. §2.1).

Tra la fine degli anni Ottanta e i primi anni Novanta altri esperimenti vennero effettuati con i modelli numerici elaborati presso i principali centri previsionali: il French Weather Service compì delle simulazioni per i mesi invernali tra il 1983 e il 1990 ([13]); il Centro Europeo per le Previsioni Meteorologiche a Medio e Lungo Termine (*European Centre for Medium-Range Weather Forecast*, ECMWF) effettuò previsioni di 31 giorni col suo modello allora operativo per il medio-termine per due giorni consecutivi ogni mese dall'aprile del 1985 al gennaio del 1989 ([19]); il centro meteorologico britannico (United Kingdom Meteorological Office, UKMO) intraprese tra il 1982 e il 1986 svariate integrazioni numeriche di 40 giorni con un modello globale a 11 livelli verticali ([1]); il Canada Climate Centre (CCC) elaborò previsioni numeriche a sei mesi per tutti i gennaio dal 1979 al 1986 con un modello a bassa risoluzione (Boer 1993, [11]).

2.3.2 Simulazioni nel passato: reforecasting e reanalisi

A dispetto di tutti questi ambiziosi lavori, di fatto, fino alla fine degli anni Novanta nessuno aveva ancora realizzato una vera previsione mensile con il significato che le attribuiamo oggi. Ovvero, fino ad allora nessun centro di ricerca, né alcun ente meteorologico nazionale, aveva mai diffuso previsioni a trenta giorni nel futuro (ovvero nella cosiddetta modalità *real-time*). Tutte le ricerche elencate si riferiscono infatti a simulazioni di natura sostanzialmente sperimentale, non operativa, nelle quali l'obbiettivo era comprendere i limiti (o le abilità) dei modelli numerici nel rappresentare lo stato dell'atmosfera oltre il limite di predicibilità. Dunque, quelle che venivano elaborate erano piuttosto previsioni *a posteriori*, al tempo chiamate *hindcast*, prodotte allo scopo di confrontarle con il reale (e già noto) comportamento dell'atmosfera.

Tale genere di simulazioni è ancora alla base non solo delle normali attività sperimentali di ricerca, ma anche dei sistemi previsionali a lungo termine, come quello mensile. In quest'ultimo contesto, le simulazioni di *hindcast* vengono meglio definite con il nome di *reforecast* (da *retrospective forecast*, Hamill 2004, [8]). Ce se ne serve, per esempio, soprattutto per calcolare e correggere eventuali bias ed errori sistematici (come verrà fatto più avanti, cfr. §4) del modello operativo utilizzato, al fine di monitorare, valutare e migliorarne costantemente le performance.

Per confrontare le previsioni del modello con l'andamento reale dell'atmosfera e nel-

l'ottica di fornire ai reforecast le condizioni iniziali, è importante avere a disposizione un archivio di campi meteorologici del passato, su griglie regolari, che definisca lo stato dell'atmosfera su tutto il globo. Questo tipo di informazioni sono oggi fornite in gran quantità dai principali centri di calcolo del mondo e sono chiamate *reanalisi*.

Un semplice archivio delle osservazioni meteo (dati satellitari, di precipitazioni, temperatura, umidità, pressione) non sarebbe infatti adatto ad essere utilizzato come fonte per i dati di inizializzazione di un modello, in quanto queste hanno spesso il non banale problema di essere incomplete e disomogenee nello spazio e nel tempo. Le reanalisi sono invece vere e proprie ricostruzioni modellistiche del tempo passato e, oltre a colmare le lacune delle serie storiche, hanno il pregio di fornire i campi su griglia, anche di ulteriori parametri fisici non direttamente osservabili. Vengono infatti elaborate a loro volta grazie all'aiuto dei modelli numerici girati con peculiari schemi di assimilazione dei dati osservativi, e forniscono la migliore ricostruzione del clima passato a partire dai dati reali a disposizione.

2.3.3 Il panorama attuale

Anche se non possono essere annoverati come vere e proprie previsioni mensili, le ricerche citate poco fa hanno costituito un lungo e fondamentale lavoro preparatorio che permette oggi a molteplici centri previsionali e di ricerca nel mondo di elaborare ed emettere periodicamente previsioni mensili (o stagionali).

Oltre all'Istituto di Atmosfera e del Clima del CNR (il cui modello GLOBO è descritto nel prossimo capitolo), forniscono oggi previsioni mensili o stagionali l'ECMWF (cfr. §2.5), gli statunitensi del National Centre of Environmental Prediction (NCEP), con un modello chiamato CPSv2 operativo dal 2008, e i giapponesi della Japan Meteorological Agency.

2.4 La verifica di una previsione

Parte integrante del lavoro di produzione di una qualunque previsione meteorologica è il processo di verifica che ne segue. Valutare quanto bene ha performato un modello, o - più in generale - quanto è stata precisa una previsione è qualcosa che serve innanzitutto a determinarne la qualità e l'affidabilità.

Benché ci siano molte tecniche e molti modi per valutare se una previsione è soddisfacente o meno, il principio generale alla base del processo di verifica è ovviamente il confronto tra la previsione formulata e lo stato dell'atmosfera osservato.

Sia per i lavori citati in precedenza (2.3), sia nei risultati qui presentati per il modello GLOBO (cfr. §4), verranno utilizzate alcune di queste tecniche per la valutazione della qualità delle previsioni, nonché dei modelli che vengono utilizzati.

Uno degli strumenti a cui si ricorre più spesso nella verifica di una previsione è il calcolo dei cosiddetti *skill score*, ovvero dei “punteggi” il cui valore dovrebbe fornire un’informazione quantitativa sulla qualità di una previsione meteorologica. Gli skill score sono solitamente espressi in termini di miglioramento percentuale rispetto ad una previsione assunta come riferimento, quale può essere la media climatica, una previsione di persistenza (cioè assumere che il tempo atmosferico non cambi rispetto al periodo precedente), o anche una previsione casuale (riferita all’occorrenza statistico-climatica di un determinato fenomeno osservato).

Nel caso più generale, uno skill score SS in termini di punti percentuali è definito come:

$$SS = \frac{A - A_{ref}}{A_{perf} - A_{ref}} \cdot 100\% \quad (2.4)$$

dove A_{perf} è il valore dell’accuratezza che dovrebbe raggiungere una previsione perfetta e A_{ref} quello della previsione presa come riferimento. I due valori estremi $SS = 0\%$ e $SS = 100\%$ si ottengono rispettivamente per i casi $A = A_{ref}$ e $A = A_{perf}$. SS può assumere anche valori negativi, in tal caso significa che l’accuratezza della previsione è valutata come inferiore rispetto a quella presa come riferimento.

Un secondo importante aspetto di cui tener conto nella fase di verifica è l’incidenza e il peso del *bias* (o errore sistematico) che affligge un modello, definito come lo scostamento di una previsione (p) per un certo campo meteorologico (ψ) dall’effettivo comportamento dell’atmosfera (o):

$$B^\psi = p - o \quad (2.5)$$

Valutare l’andamento spaziale e temporale di tale fattore permette di conoscere preventivamente eventuali sovrastime e sottostime di un modello, correggendole opportunamente. Per il calcolo del bias si ricorre solitamente al calcolo di una sorta di “clima del modello” attraverso un gran numero di simulazioni di reforecast, al fine di confrontarlo con il “clima reale” rappresentato dalle reanalisi.

2.4.1 Gli score per un campo di previsione

I prodotti dei modelli numerici sono, nella gran parte dei casi, rappresentazioni spaziali delle principali variabili meteorologiche alle varie quote, quelle che nella sezione 2.2 abbiamo chiamato *campi*. Questi vengono ricavati a partire dalle griglie che la maggior parte dei modelli sfrutta per integrare le equazioni della dinamica, e che diventano indispensabili nel processo di verifica². Per ciascun elemento di queste matrici spaziali è infatti possibile raffrontare quanto previsto con quanto osservato, e da ciò ricavare il valore degli skill score che seguono.

²Nel caso dei modelli spettrali, che lavorano per così dire “nel continuo”, avviene in un certo senso il processo inverso: i campi prodotti vengono riportati sulle griglie di cui si servono già nella fase di parametrizzazione dei fenomeni fisici su scala inferiore al passo di griglia

L'errore quadratico medio, MSE L'errore quadratico medio (*Mean Squared Error*, MSE) è il criterio più comune per valutare l'accuratezza di una previsione formulata per il campo di una variabile meteorologica. Calcolare l'MSE consiste nel fare la media spaziale delle singole differenze elevate al quadrato tra il valore previsto e quello osservato per ciascun punto della griglia:

$$MSE = \frac{1}{N} \cdot \sum_{n=1}^N (p_n - o_n)^2 \quad (2.6)$$

dove N sta per il numero totale dei punti di griglia, o_n per il valore osservato e p_n quello previsto.

Spesso al posto dell'errore quadratico medio ne viene calcolata la radice, RMSE (*Root Mean Squared Error*):

$$RMSE = \sqrt{MSE} \quad (2.7)$$

L'RMSE ha il vantaggio di essere della stessa unità di misura della variabile di previsione, ed è quindi più facilmente interpretabile in termini di ordine di grandezza dell'errore.

Talvolta può essere utile voler valutare l'accuratezza, invece che del singolo, di un insieme N di campi di previsione. In tal caso si ricorre alla forma generica per gli skill score vista in precedenza (2.4):

$$SS = \frac{\sum_{i=1}^N MSE(i) - \sum_{i=1}^N MSE_{ref}(i)}{0 - \sum_{i=1}^N MSE_{ref}(i)} \quad (2.8)$$

L'Anomaly Correlation, AC L'anomalia di correlazione (*Anomlay Correlation*, AC) è un'altra possibile misura di come sono legati tra loro il campo previsto e quello osservato. Questo score è definito a partire dalle definizioni di anomalia già citate (2.1 e 2.2), e di fatto lo si può considerare come un coefficiente di correlazione lineare (che dunque copre tutti i valori da +1 a -1). In letteratura se ne trovano due diverse versioni.

La prima è la *Centered Anomaly Correlation*, AC_C , definita nel modo seguente e calcolata per tutti i punti n della griglia del modello:

$$AC_C = \frac{\sum_{n=1}^N (p'_n - \bar{p}') (o'_n - \bar{o}')}{\sqrt{\sum_{n=1}^N (p'_n - \bar{p}')^2 \cdot \sum_{n=1}^N (o'_n - \bar{o}')^2}} \quad (2.9)$$

dove \bar{p}' e \bar{o}' indicano i valori mediati su tutta la griglia di N punti.

La seconda forma di AC è la *Uncentered Anomaly Correlation*, AC_U . Questa differisce dalla 2.9 in quanto i valori medi calcolati sull'intera mappa non vengono sottratti:

$$AC_U = \frac{\sum_{n=1}^N p'_n \cdot o'_n}{\sqrt{\sum_{n=1}^N (p'_n)^2 \sum_{n=1}^N (o'_n)^2}} \quad (2.10)$$

L'anomalia di correlazione è detta talvolta anche *Pattern Correlation*, in quanto serve a rintracciare le somiglianze tra campo previsto e campo osservato, in termini di scarto dalla media climatologica.

Né l' AC_C , né l' AC_U , , tuttavia, sono sufficienti a valutare complessivamente la qualità di una previsione. Il motivo è che entrambi, essendo dei coefficienti di correlazione, danno sì informazioni su quanto bene siano legate tra loro due variabili (in questo caso quanto l'anomalia prevista e quella osservata siano linearmente correlate), ma non tengono in alcun conto il bias, ovvero la presenza di eventuali errori sistematici. In altre parole, pensando di voler graficare le condizioni atmosferiche previste ed osservate, l'AC sarà in grado di dirci quanto bene i punti del grafico possono essere interpolati da una retta, ma non fornirà nessuna informazione sulla pendenza della retta. Per questo motivo, in molti lavori, l'AC viene fornita insieme con l'MSE (o l'RMSE) che fornisce invece una stima quantitativa diretta dell'accuratezza di una previsione.

Anche nel caso dell'AC, è possibile generalizzare al caso di un insieme di campi di previsione. Tuttavia, invece di sfruttare la (2.4), si preferisce confrontare il valore dell'*Anomaly Correlation* con i valori di riferimento $AC_{ref} = 0.5$ oppure $AC_{ref} = 0.6$. Questi valori sono scelti in virtù del fatto che, come mostrato da Murphy ed Epstein (1989), a $AC_C = 0.5$ corrisponde un valore di $MSE = 0$, mentre per $AC_C = 0.6$, MSE vale 0.2.

2.4.2 Gli score probabilistici

Oltre all'MSE e all'AC, esistono una serie di altri skill score di natura più strettamente probabilistica, ugualmente prodotti per valutare l'accuratezza delle previsioni e largamente sfruttati nei lavori che se ne occupano. La loro definizione viene solitamente fatta a partire dalla *tabella di contingenza* (figura 2.4), che schematizza i quattro possibili casi semplificati di verifica di una previsione raggruppandoli in: eventi previsti che si verificano (a), eventi previsti che non si verificano (b), eventi non previsti che si verificano (c), eventi non previsti che non si verificano (d).

Dalla tabella di contingenza è possibile definire alcuni scalari che torneranno utili nelle successive definizioni degli skill score:

- L'*Hit Rate*, H , definito come il rapporto tra il numero delle previsioni corrette per gli eventi previsti e il numero totale di volte che quell'evento si è verificato:

$$H = \frac{a}{a + c} \quad (2.11)$$

- Il *False Alarm Rate*, F , definito come il rapporto tra il numero delle previsioni errate per gli eventi previsti e il numero totale di volte che quell'evento non si è verificato:

		Osservazione		
		Si	No	
Previsione	Si	a	b	a+b
	No	c	d	c+d
		a+c	b+d	N = a+b+c+d

Figura 2.4: La tabella di contingenza serve a classificare una o più previsioni in base al loro successo o fallimento

$$F = \frac{b}{b+d} \quad (2.12)$$

Il diagramma ROC Il diagramma ROC (*Relative Operating Characteristic* o *Receiver Operating Characteristic*) è uno strumento grafico di verifica di una previsione meteorologica. Si tratta di un diagramma bidimensionale in cui in ascissa è rappresentato il *False Alarm Rate*, F , mentre in ordinata l'*Hit Rate*, H (si veda la figura 2.5).

Nel caso del ROC occorre che la previsione sia presentata in forma probabilistica, ovvero che siano fornite le occorrenza di una data osservazione o_i fissata una certa probabilità di soglia p_i . Facendo variare gradualmente la soglia, varierà anche il numero di eventi che rientrano nelle varie categorie della tabella di contingenza (a, b, c, d) e dunque anche i valori di F ed H . Graficando le coppie di punti (F_i, H_i) si ottiene una curva in grado di descrivere la qualità di una previsione. I due casi limite sono quelli per cui la curva è una retta inclinata di 45° (previsione casuale) o una spezzata che coincide con il limite alto e sinistro del diagramma ($H = 1$ e $F = 0$, ovvero una previsione perfetta).

È possibile (e talvolta pratico) riassumere le informazioni fornite dal diagramma in un unico scalare, e la scelta è solitamente quella di calcolare l'area A sottesa alla curva, che dunque oscillerà tra i valori $A_{random} = \frac{1}{2}$ e $A_{perf} = 1$.

In alternativa, è sempre possibile calcolare lo score secondo la consueta definizione della 2.4:

$$SS_{ROC} = \frac{A - A_{random}}{A_{perf} - A_{random}} \quad (2.13)$$

Il Ranked Probability Score, RPS

Questo skill score è pensato per la verifica di previsioni probabilistiche che coinvolgono più di due categorie di eventi, uscendo dalla dicotomia sì-no rappresentata nella tabella di contingenza (2.4). Solitamente ciò comporta la creazione di alcune “fasce” di probabilità nella formulazione della previsione, che corrispondono ad altrettante componenti per gli eventi osservati. Al verificarsi dell’evento oggetto della previsione, la componente dell’osservazione corrispondente alla fascia in cui quella rientra assume valore 1, mentre le altre il valore 0.

Si possono in questo modo calcolare le cosiddette previsioni (P_j) e osservazioni (O_j) cumulative:

$$P_j = \sum_{n=1}^j p_n \quad j = 1, \dots, N \quad (2.14a)$$

$$O_j = \sum_{n=1}^j o_n \quad j = 1, \dots, N \quad (2.14b)$$

dove N indica il numero totale delle fasce di probabilità (e degli eventi osservati), e dunque per definizione vale che $P_N = 1$ e $O_N = 1$. Il *Ranked Probability Score*, RPS , è la somma in quadratura delle differenze tra le componenti delle fasce di previsione e quelle degli eventi osservati:

$$RPS = \sum_{j=1}^N (P_j - O_j)^2 \quad (2.15)$$

Volendo estendere il concetto ad un insieme di m previsioni (cioè di m coppie previsione-evento), sarà sufficiente mediare:

$$\overline{RPS} = \frac{1}{m} \sum_{k=1}^m RPS_k \quad (2.16)$$

Anche per l’ RPS è possibile calcolare lo skill score con la consueta definizione (2.4), usando come riferimento le probabilità ricavate dalla climatologia:

$$SS_{RPS} = \frac{\overline{RPS} - \overline{RPS}_{clim}}{0 - \overline{RPS}_{clim}} \quad (2.17)$$

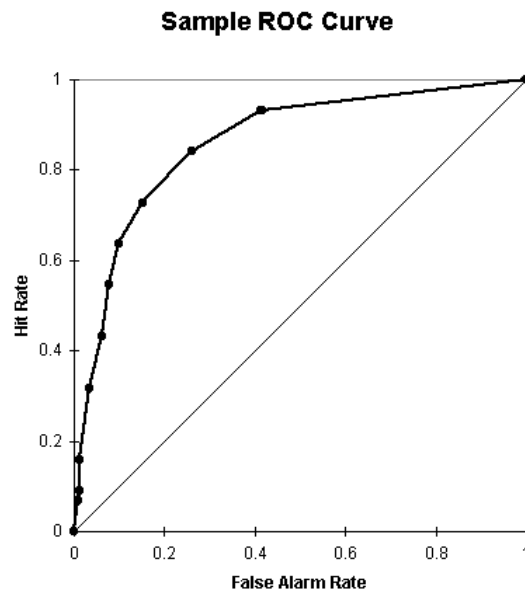


Figura 2.5: Un esempio di diagramma ROC

Nel caso (semplificato) in cui $N = 2$ l' RPS è chiamato *Brier Score*, BS :

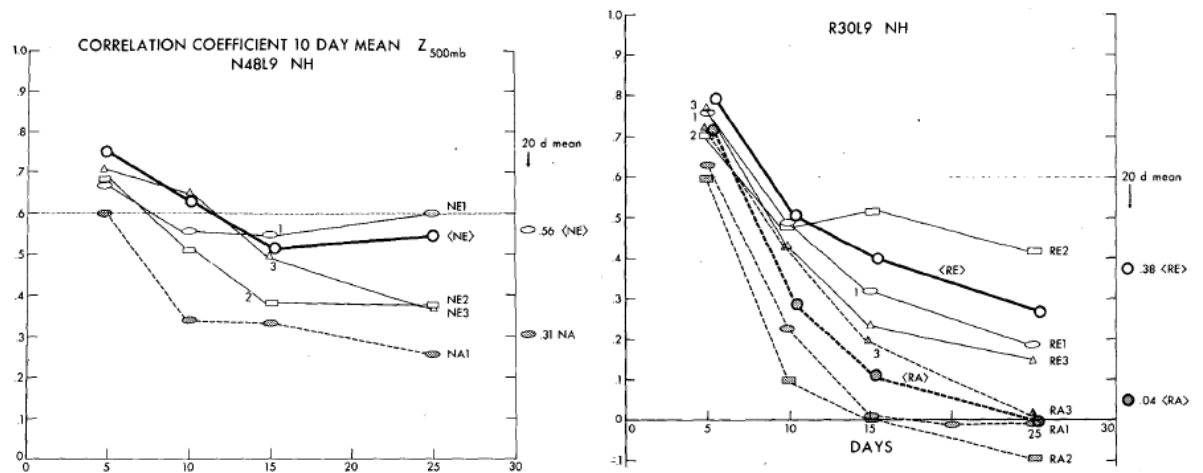
$$BS = \frac{1}{m} \sum_{k=1}^m (p_k - o_k)^2 \quad (2.18)$$

2.4.3 Gli esiti dei lavori degli Anni 80 e 90

È possibile a questo punto presentare brevemente i risultati dei lavori precedentemente citati. Come si potrà notare, ce ne sono di più e meno incoraggianti.

Il grado di successo delle ricerche elencate varia molto a seconda del modello utilizzato e del numero di simulazioni effettuate. Grande evidenza assume il fatto che la predicibilità non sia un parametro omogeneo ed uguale su tutta la terra, e di come la rimozione dell'errore sistematico dei modelli sia in grado di migliorare significativamente le performance delle simulazioni.

Figura 2.6: Anomaly Correlation delle varie simulazioni per il campo di altezza geopotenziale a 500 hPa. Fonte: *Miyakoda et al. 1983*, [5]

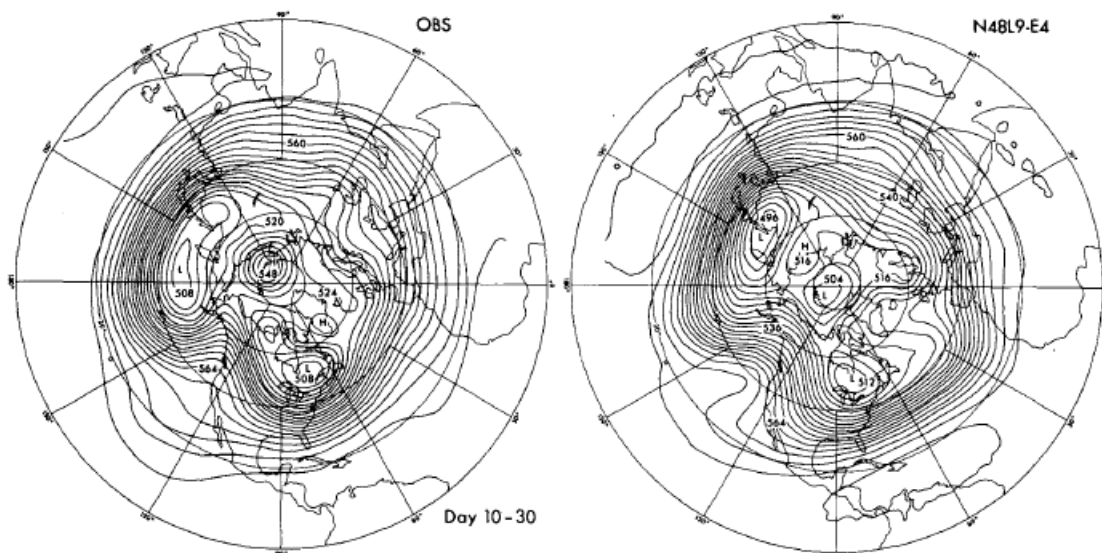


(a) Simulazioni effettuate tramite il modello a griglia (b) Simulazioni effettuate tramite il modello spettrale

Miyakoda et al. (1983) Si tratta di uno dei primi e principali lavori sulle previsioni mensili.

La figura (2.6) riporta l'andamento in funzione del giorno di previsione dell'Anomaly Correlation (Equazione 2.9) tra il campo previsto e quello osservato, riferita ad una altezza geopotenziale di 500 hPa e mediata su dieci giorni per ciascuno dei modelli utilizzati.

Figura 2.7: Raffronto tra il campo di gpt a 500 hPa previsto e osservato per il modello più performante (NE1 della figura 2.6) Fonte: Miyakoda et. al (1983), [5]



Come si può notare, ci sono grosse differenze tra le performance del modello spettrale (R30L9) e di quello a griglia (N48L9), e quest'ultimo risulta certamente migliore (si confrontino <NE> con <RE> e <RA>, sul lato destro dei due grafici). Si osserva inoltre l'importanza della scelta e della qualità dello schema di parametrizzazione per i fenomeni fisici al di sotto del passo di griglia (le due versioni usate in questo lavoro sono indicate dai suffissi -A ed -E). Quest'ultimo sembra proprio uno dei punti dirimenti della simulazione, a detta degli stessi autori:

"After considerable frustration, we finally decided to use an expensive, i.e., high resolution finite difference GCM with improved subgrid scale (SGS) processes [quello indicato della sigla NE]. The resulting prediction turned out to be good, without using anomalous external forcing. The most important factor for our success in this case was the inclusion of improved SGS process into the GCM"

Vi sono infine le classiche differenze di comportamento del modello a seconda della scelta dei dati di inizializzazione (indicati nei grafici dalle cifre 1, 2 o 3).

In figura (2.7) sono riportate le mappe del campo dell'altezza geopotenziale a 500 hPa così come osservata e prevista dal modello.

Quanto alla capacità di predire il lungo periodo sottomedio, nella figura (2.8) vi è un raffronto tra le anomalie riferite alla quota di 850 hPa degli ultimi venti giorni del gennaio 1977, osservate e previste dal modello.

Miyakoda et al. (1986) In questo lavoro (che costituisce una sorta di continuazione del precedente) sono eseguite 24 simulazioni riguardanti tutte mesi di gennaio. Il modello utilizzato è analogo a quello dello studio pilota (Miyakoda et al. 1983). Le caratteristiche

Figura 2.8: Raffronto tra le anomalie previste e quelle osservate riferite alla quota di 850 hPa degli ultimi 20 giorni del gennaio 1977. *Fonte: Miyakoda et al. (1983), [5]*

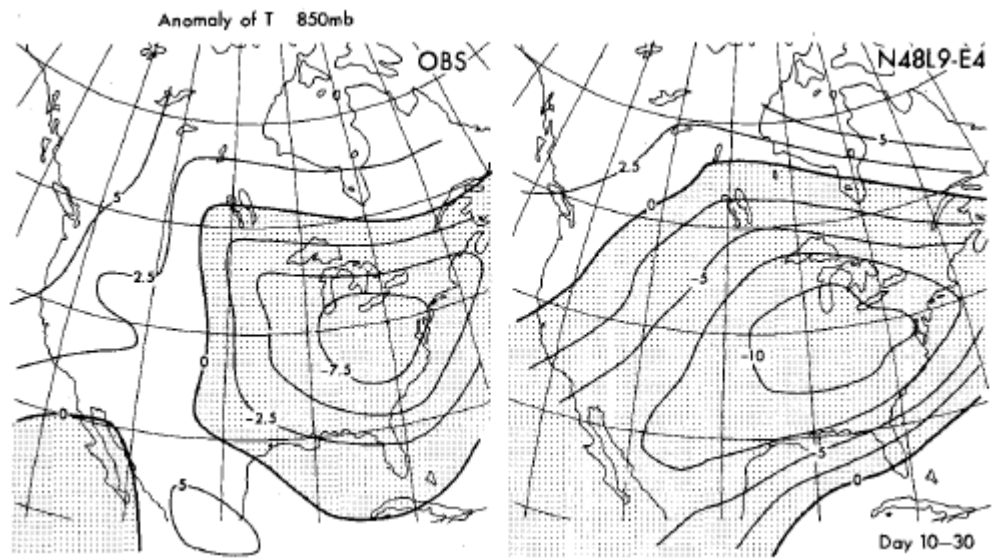
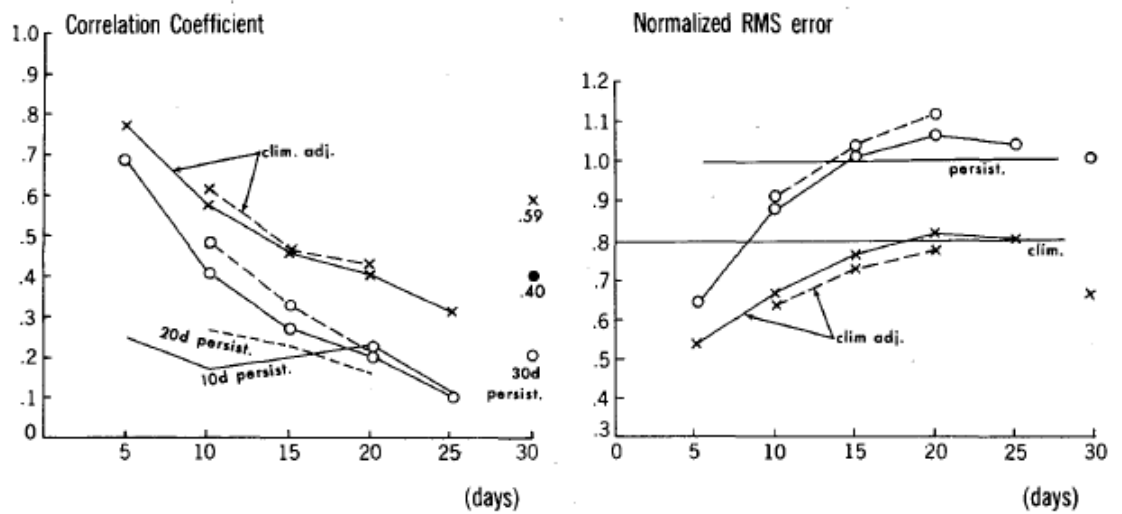


Figura 2.9: AC e RMSE (normalizzato ad una previsione di persistenza) mediati su tutti gli ensemble di tutte le simulazioni prima (o) e dopo (x) la correzione per l'errore sistematico. *Fonte: Miyakoda et al. (1986), [4]*



dei dati di inizializzazione sono riassunti nella figura (2.10).

Per quel che riguarda i risultati, la ricerca mette bene in evidenza l'importanza di correggere il bias (chiamato qui e in altri lavori anche *climate drift*) prodotto dalle simulazioni, che incide mediamente in modo significativo sulla predicibilità dei periodi presi in esame (si veda la figura 2.9).

Nel lavoro viene anche sottolineato come gli anni presi in considerazione si discostino a loro volta dalla media climatica, mostrando una particolare propensione agli eventi di blocking.

Shukla (1981) In questo studio sono presentati i risultati di 9 simulazioni da 60 giorni di integrazione ciascuno, secondo lo schema riportato in figura (2.11).

Per valutare la capacità predittiva del modello scelto viene valutata, attraverso la varianza, la significatività statistica del rapporto tra la variabilità tra i vari run di controllo (inizializzati con condizioni iniziali molto diverse tra loro) - chiamata F_1 - e quella interna alle varie perturbazioni degli ensemble (inizializzati con condizioni iniziali molto simili tra loro, prodotte da perturbazioni casuali applicate a quelle del run ufficiale)

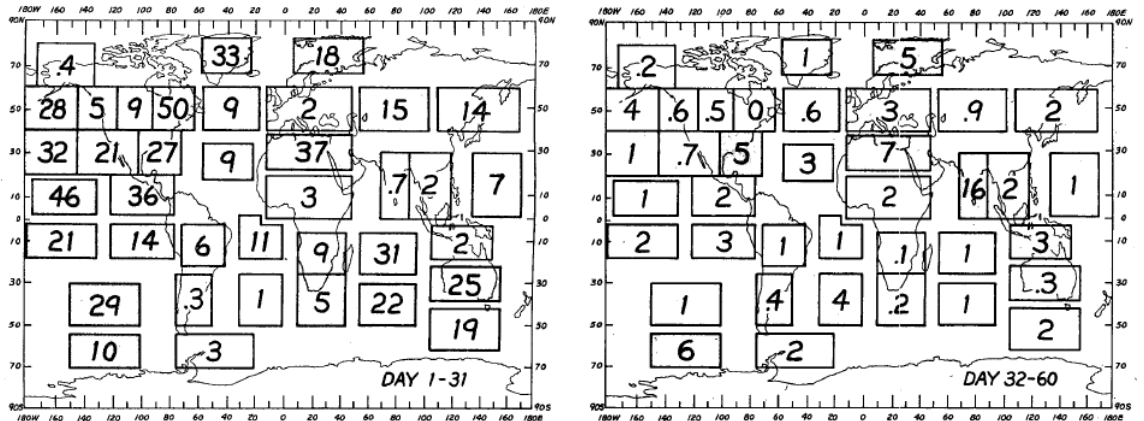
number	initial time	Level III data		
		A	B	C
1	OOGMT 1 JAN 1977	GFDL-1		
	"		NMC	
	OOGMT 2 JAN 1977			GFDL-1
2	12GMT 31 DEC 1977	GFDL-1		
	OOGMT 1 JAN 1978		NMC	
	12GMT 1 JAN 1978			GFDL-1
3	OOGMT 1 JAN 1979	GFDL-1		
	"		NMC	
	"			ECMWF
4	OOGMT 16 JAN 1979	GFDL-1		
	"		NMC	
	"			ECMWF
5	OOGMT 1 JAN 1980	GFDL-2		
	"		NMC	
	"			ECMWF
6	12GMT 1 JAN 1981	GFDL-2		
	OOGMT 1 JAN 1981		NMC	
	"			ECMWF
7	OOGMT 1 JAN 1982	GFDL-2		
	"		NMC	
	"			ECMWF
8	OOGMT 1 JAN 1983	GFDL-2		
	"		NMC	
	"			ECMWF

Figura 2.10: Dati di inizializzazione utilizzati per le simulazioni in Miyakoda et al. (1986). GFDL indica le condizioni iniziali elaborate presso i Laboratori di Geofisica Fluido-Dinamica di Princeton, ECMWF quelle presso il Centro Europeo, NMC quelle presso la *National Meteorological Centre*. Fonte: Miyakoda et al. (1986), [4]

	Initial condition	1 Jan 1975 $j = 1$	1 Jan 1976 $j = 2$	1 Jan 1977 $j = 3$
Control	$i = 1$	C_{11}	C_{21}	C_{31}
Random perturbation	$i = 2$	C_{12}	C_{22}	C_{32}
Random perturbation	$i = 3$	C_{13}	C_{23}	
Random perturbation	$i = 4$		C_{24}	

Figura 2.11: Dati di inizializzazione utilizzati per le simulazioni in Shukla (1981). Fonte: Shukla (1981), [14]

Figura 2.12: Rappresentazione spaziale di F diviso per macroaree. Fonte: Shukla (1981), [14]



(a) Valori di F per i primi 31 giorni di integrazione (b) Valori di F per il secondo mese di integrazione numerica (32-60)

- chiamata F_2 :

$$F = \frac{F_1}{F_2} \cdot \frac{N - J}{J - 1}$$

dove N è il numero totale di simulazioni e J indica il numero di date di inizializzazione prese in considerazione. Il valore di F è considerato significativo al 95% per valori superiori a 5,1.

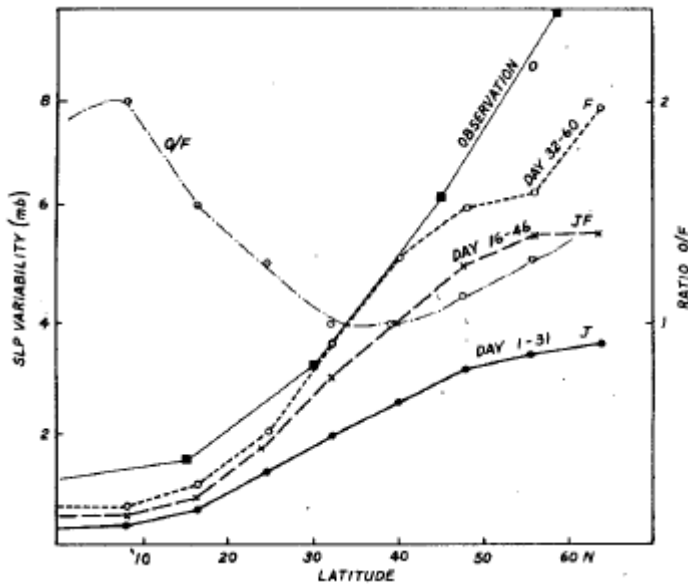
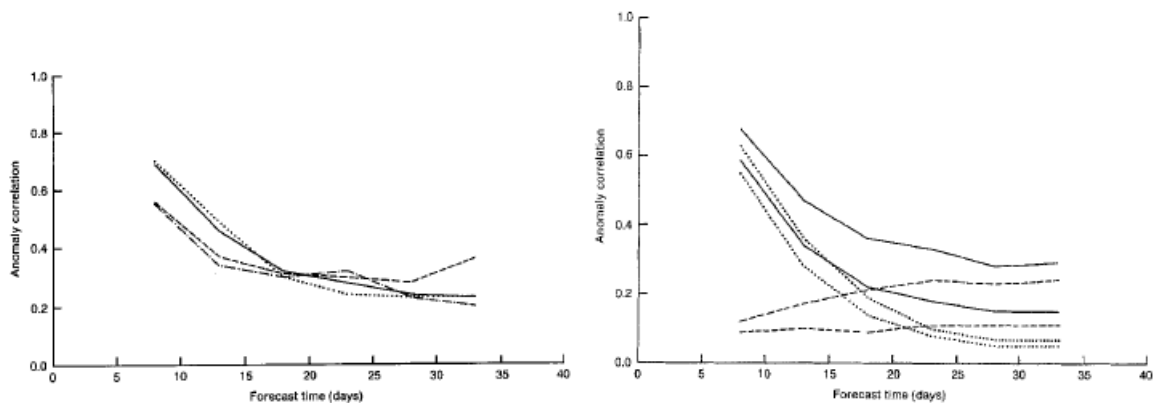


Figura 2.13: Valori dell'RMSE mediati lungo la latitudine per vari intervalli temporali

Figura 2.14: AC mediata su 15 giorni per tutte le simulazioni. *Fonte: Murphy e Dickinson 1989, ([1])*



(a) AC della pressione sul livello del mare divisa per stagioni: (—) inverno, (...) primavera, (- - -) e dopo (sopra) la rimozione del bias (—). (...) inverno, (- - -) autunno. Bias già rimosso. (b) AC dell'altezza gpt dei 500 hPa, prima (sotto) e dopo (sopra) la rimozione del bias (—). (...) rappresenta l'AC calcolata rispetto ad una previsione casuale, (- - -) il relativo incremento della prestazione.

Il valore di F diviso per macroaree riferito ai periodi 1-32 giorni e 32-60 giorni è riportato nella figura (2.12). In figura (2.13) sono infine riportati anche i vari valori dell'RMSE.

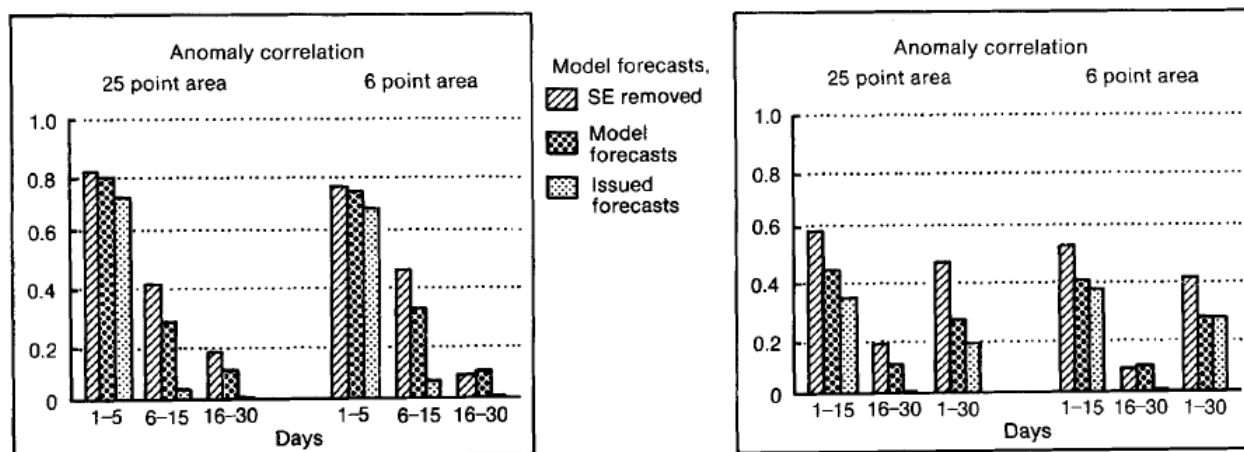
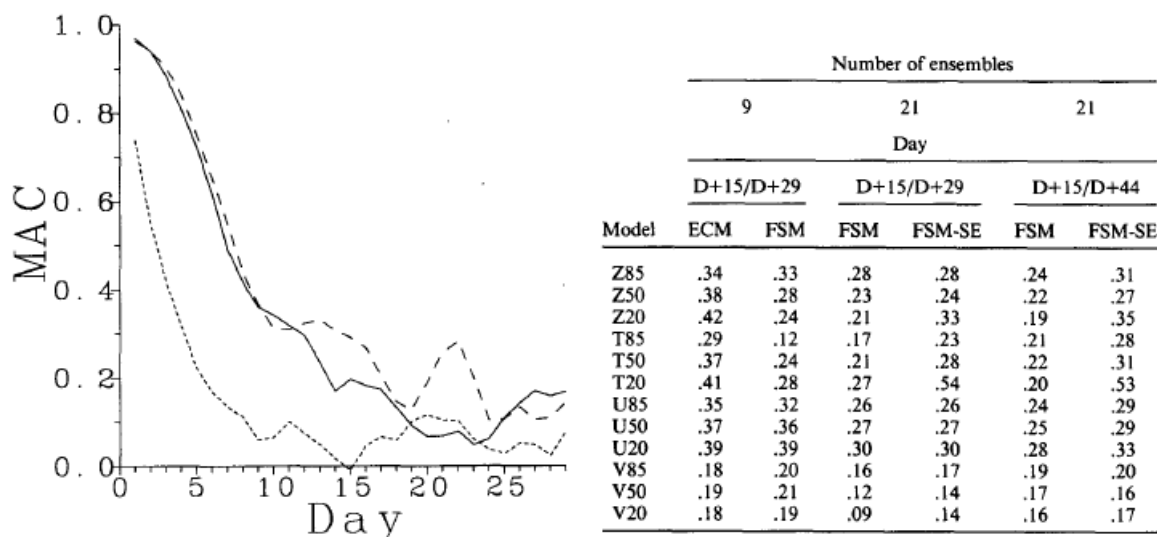


Figura 2.15: Confronto per l'area della Gran Bretagna tra l'AC mediata sulle 48 simulazioni della ricerca (prima e dopo la correzione per il bias) e quella relativa al sistema di previsioni mensili attivo presso UKMO. *Fonte: Murphy e Dickinson (1989), [1]*

Figura 2.16: Valori dell'Anomaly Correlation mediati su tutte le simulazioni, calcolati per l'emisfero settentrionale. Fonte: Déqué e Royer (1992), [13]



(a) AC in funzione del tempo dell'altezza dei gpt a 500 hPa. (—) rappresenta le performance del modello francese, (- - -) quelle dell'ECMWF, (...) quelle di una previsione di persistenza. (b) Valori dell'AC calcolati per il modello francese (FSM) e quello del Centro Europeo (ECM), divisi per numero di simulazioni (9 o 21), e calcolati per svariati campi meteorologici: di temperatura (T), altezza geopotenziale (Z), vento zonale (U) e meridionale (V) alle quote di 200, 500 e 850 hPa. D indica l'intervallo temporale di integrazione numerica.

Murphy e Dickinson (1989) La ricerca presenta gli esiti di 64 simulazioni (16 per ogni stagione dell'anno) da 40 giorni ciascuna, effettuate presso l'UKMO con un modello numerico di circolazione globale a 11 livelli verticali.

In figura (2.14) sono rappresentate le performance del modello relative all'emisfero boreale, mentre in figura (2.15) quelle relative all'area della Gran Bretagna, su cui lo studio si concentra, confrontando le performance delle simulazioni con quelle di un altro modello per previsioni mensili (basato però su previsioni a medio termine e tecniche di previsione di stampo statistico) in uso presso l'UKMO.

Anche in questo caso viene sottolineato (ed è evidente dalle immagini) come la rimozione dell'errore sistematico infici positivamente sulla previsione.

Déqué e Royer (1992) In questo lavoro sono presentati i risultati di un ciclo di 21 simulazioni numeriche di 44 giorni ciascuna, svolte tutte per periodi invernali, con un modello spettrale ad alta risoluzione prodotto dal France Weather Service agli inizi degli

anni Novanta. Una parte dei risultati ottenuti (nove in tutto) sono stati confrontati anche con le simulazioni effettuate presso l'ECMWF.

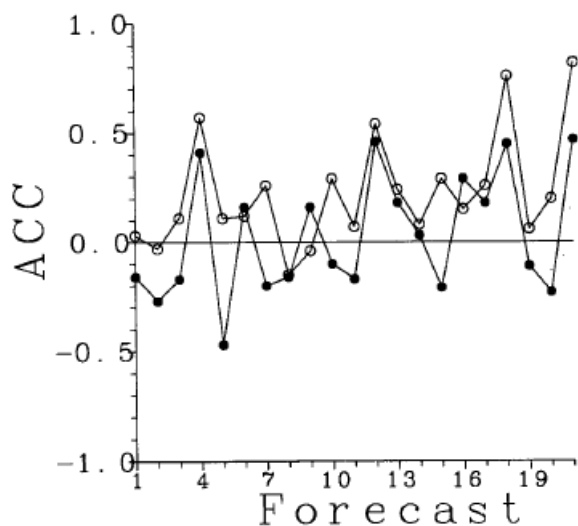
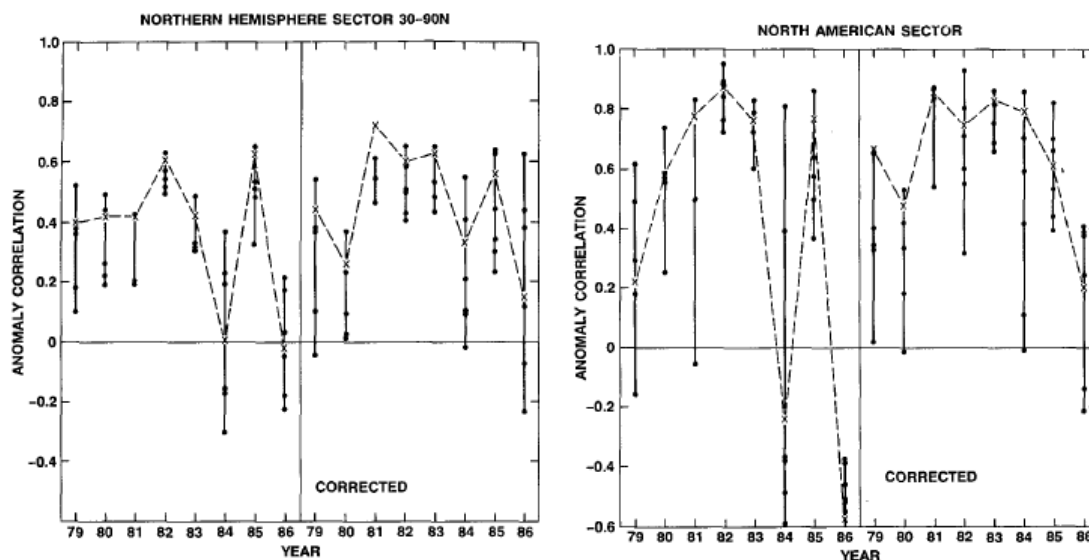


Figura 2.17: ACC per l'emisfero settentrionale dell'altezza geopotenziale a 500 hPa per ciascuna delle 21 simulazioni, riferita alla persistenza dei trenta giorni antecedenti (●) e dopo la correzione per l'errore sistematico (○). *Fonte: Déqué e Royer (1992), [13]*

Figura 2.18: Anomaly Correlation per l'altezza geopotenziale a 500 hPa prima (semi metà sinistra) e dopo (semi metà destra) la correzione per l'errore sistematico. *Fonte: Boer (1993), [11]*



(a) AC per l'emisfero settentrionale. I valori dell'AC per le singole simulazioni sono indicati con un punto mentre i valori medi con una ×.
 (b) AC per il nord America. I valori dell'AC per le singole simulazioni sono indicati con un punto mentre i valori medi con una ×.

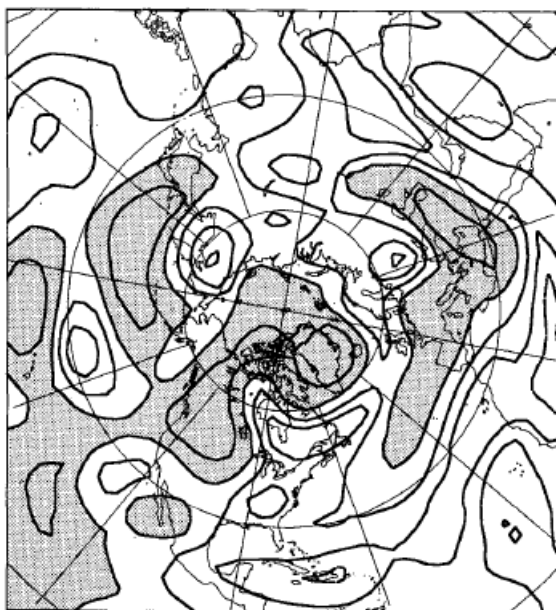


Figura 2.19: Distribuzione spaziale (emisfero settentrionale) dell'AC dell'altezza geopotenziale dei 500 hPa mediata su tutte le simulazioni. Ogni linea rappresenta un incremento di 0.2, le aree ombreggiate sono quelle per le quali $AC > 0.4$.

In figura (2.16a) sono mostrati i valori dell'AC mediata sulle nove simulazioni messe a confronto con quelle dell'ECMWF per una serie di variabili meteorologiche. Inoltre, sono tabellati i valori mediati su tutte e 21 le simulazioni prima e dopo la correzione per l'errore sistematico e divise per intervallo di integrazione numerica. In figura (2.16b) sono invece stati graficati nel tempo. Si noti anche in questo lavoro l'incidenza dell'errore sistematico sui risultati (figura 2.17).

Boer (1993) Questo studio affronta in maniera estesa la questione degli errori nelle previsioni a lungo termine e presenta i risultati relativi a 42 simulazioni, tutte inizializzate nel mese di gennaio (dal 1979 al 1986), realizzate presso il Canadian Climate Centre con un modello spettrale a bassa risoluzione dotato di 10 livelli verticali. La parte interessante ai nostri fini è quella iniziale concernente l'errore sistematico.

In figura (2.18) sono presentati i risultati ottenuti in termini di Anomaly Correlation per l'emisfero settentrionale e per il Nord America riferiti all'altezza geopotenziale a 500 hPa, prima e dopo la correzione per l'errore sistematico. Quest'ultima fornisce a dire il vero un miglioramento soltanto modesto delle performance del modello, e comunque più incisiva sulle previsioni riguardanti il Nord America.

In figura (2.19) l'AC mediata su tutte le simulazioni è rappresentata nella sua disposizione spaziale.

2.5 Il modello dell'ECMWF

Il Centro Europeo per le Previsioni Meteorologiche a Medio e Lungo Termine (ECMWF), con sede a Reading, in Gran Bretagna, è uno dei più rinomati enti ad occuparsi di modelli e previsioni meteorologiche nel mondo. Negli anni Duemila è stato anche uno dei primi centri di calcolo ad attivare un modello per previsioni mensili (*Monthly Forecasting System*).

Alla fine degli anni Ottanta presso l'ECMWF erano state effettuate alcune simulazioni, ma con risultati poco soddisfacenti ([19]). Il tema è stato ripreso soltanto nel nuovo Millennio, e a partire dal marzo del 2002 è entrato in funzione in via sperimentale un modello numerico spettrale per previsioni mensili, fatto girare una volta ogni due settimane per oltre due anni. Nell'ottobre del 2004, accertatene le soddisfacenti performance ([19]), divenne operativo e la frequenza delle corse resa settimanale (dal novembre del 2011 è stata ulteriormente incrementata a bisettimanale).

Da allora il modello dell'ECMWF è stato modificato numerose volte, e nei più vari aspetti: nella fisica, nella risoluzione (sia orizzontale che verticale), nel periodo scelto per il reforecast, nel database per le reanalisi. Nel 2008, inoltre, è stato unito al più vasto *Ensemble Prediction System* (EPS) e da allora sfrutta un sistema che gli permette di variare la risoluzione col procedere dell'integrazione.

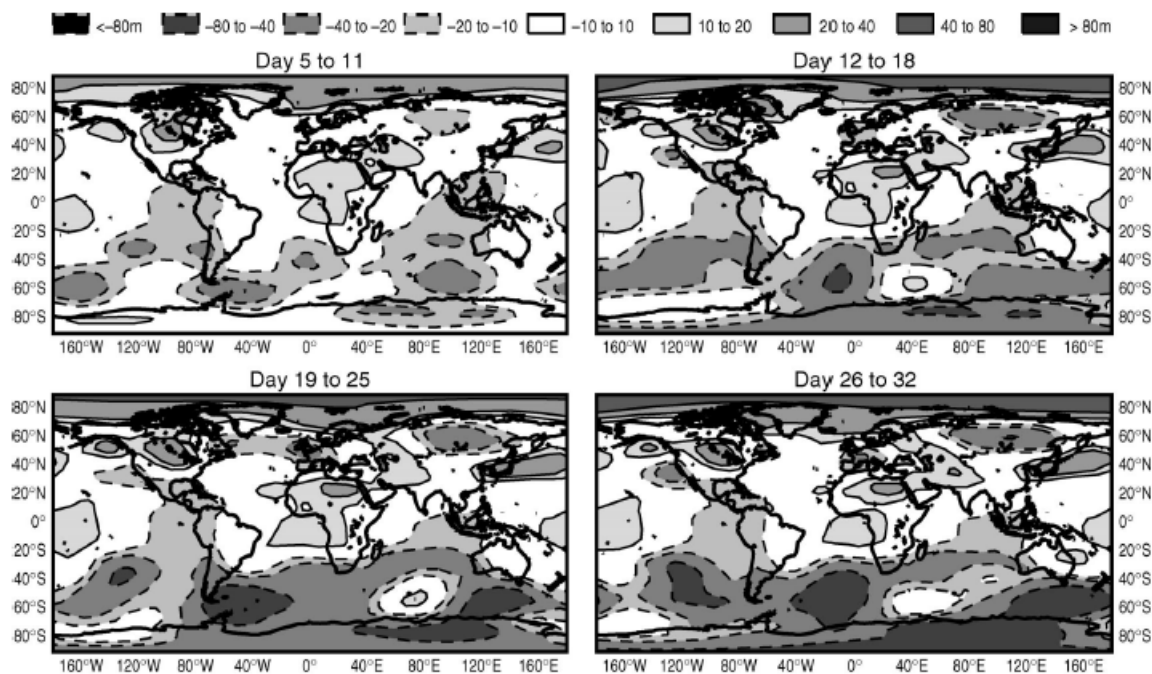


Figura 2.20: Bias dell'altezza geopotenziale a 500 hPa diviso per intervallo temporale. Reanalisi del dataset ERA-40. Fonte: Vitart 2004, [19]

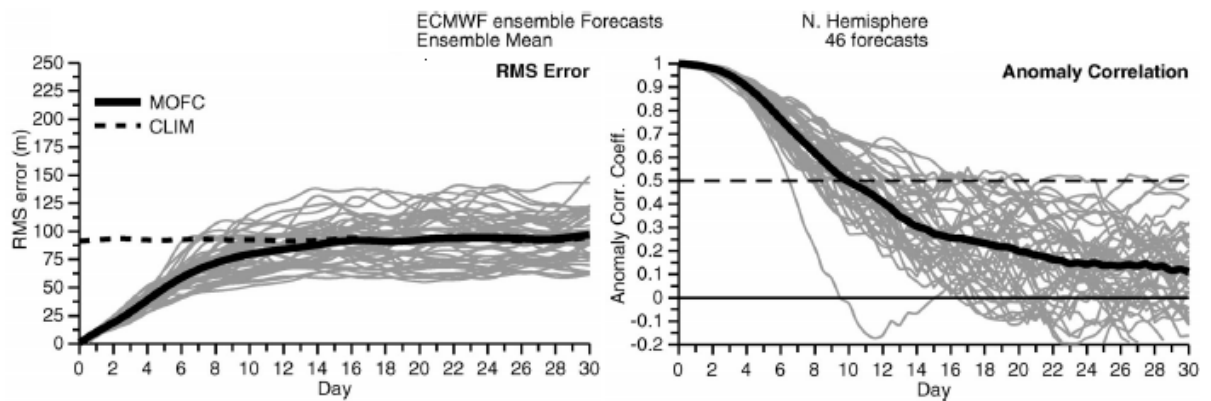


Figura 2.21: AC e RMSE dell'altezza geopotenziale a 500 hPa in funzione del tempo. Ciascuna delle 45 simulazioni effettuate è rappresentata da una linea grigia, quella nera raffigura la media su tutti i casi. *Fonte: Vitart (2004), [19]*

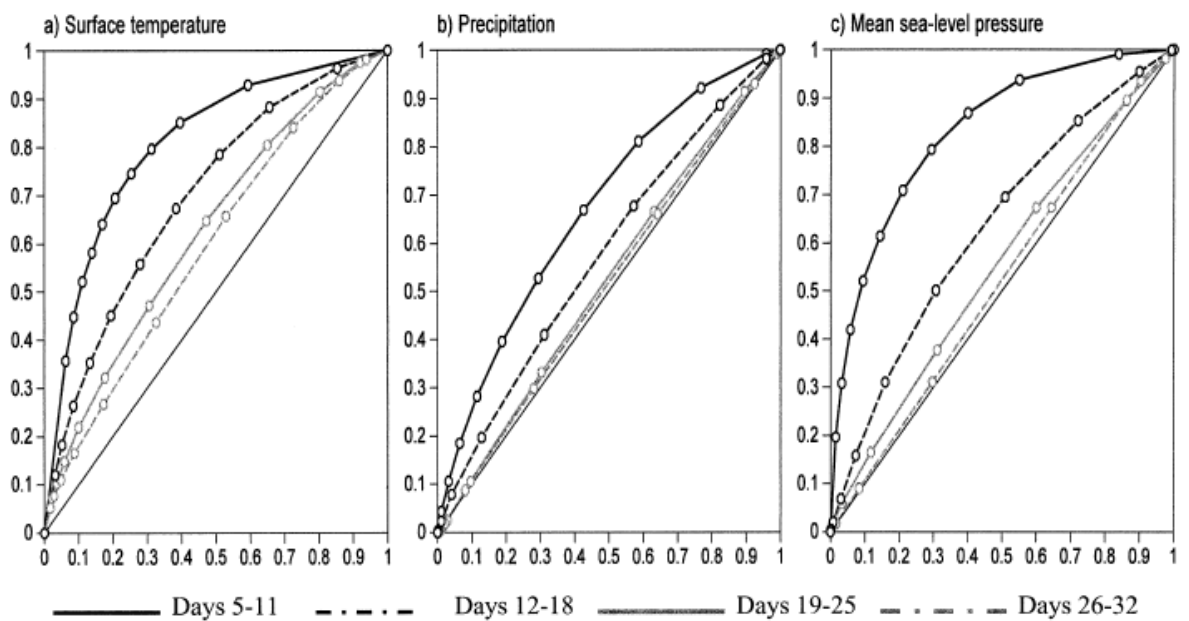


Figura 2.22: Diagrammi ROC della probabilità che i valori settimanali medi di temperatura al suolo (a), precipitazioni (b) e pressione sul livello del mare (c) ricadano nel tercile superiore (ovvero siano sopramedia). Valori calcolati per l'emisfero settentrionale (al di sopra del 30°N).

2.5.1 Progressi e performance del modello dell'ECMWF

Il primo modello utilizzato in modalità *real-time* dall'ECMWF per le previsioni mensili è descritto in (Vitart 2004 [19]). Lì è riportata anche una analisi della predicibilità riferita

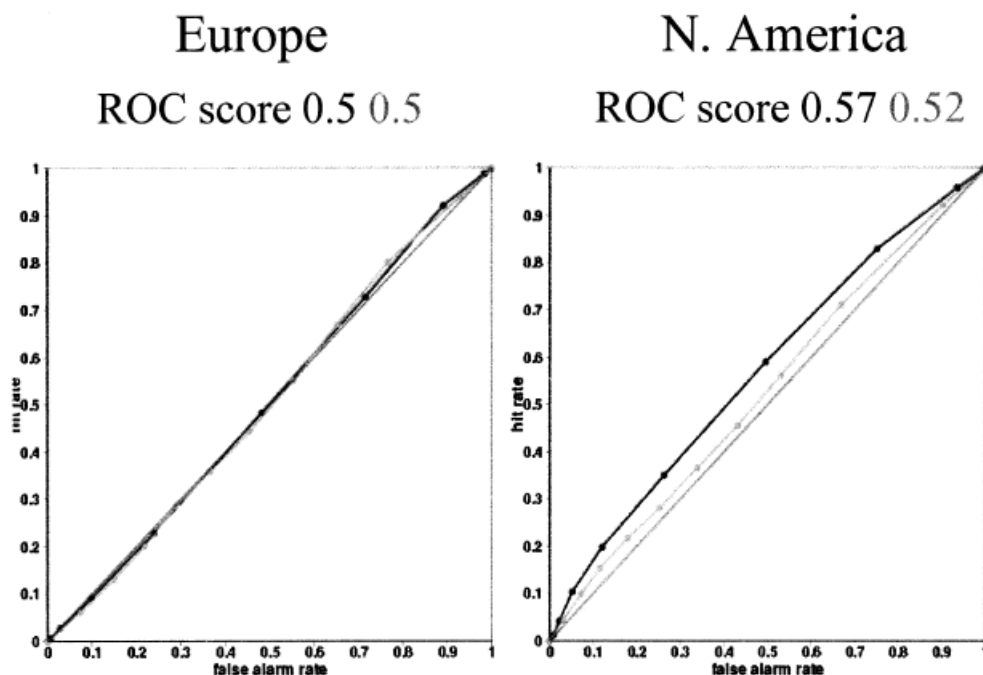


Figura 2.23: Diagrammi ROC della probabilità che la temperatura al suolo mediata nell'intervallo temporale 19-32 giorni ricada nel terzile superiore (ovvero sia sopramedia) per l'Europa e il Nord America.

alle 45 previsioni effettuate in real-time tra il marzo del 2002 e il dicembre del 2003, ovvero nel periodo in cui il modello era ancora nella sua fase sperimentale.

Il modello, di tipo spettrale (1.125°gradi di risoluzione), a 40 livelli verticali, era formato da un ensemble di 51 membri (uno di controllo e 50 perturbazioni) e si spingeva fino a 32 giorni di integrazione numerica. Utilizzava un accoppiamento oceano-atmosfera e veniva fatto girare una volta ogni due settimane.

La climatologia del modello era calcolata a partire da un ensemble di 5 membri ciascuno di 32 giorni di integrazione, inizializzati nei 12 anni antecedenti e nella stessa data (stesso giorno, stesso mese) della previsione in real-time.

In figura (2.20) è riportato il bias riferito ad una altezza geopotenziale di 500 hPa per ciascuna delle 4 settimane di integrazione a partire dal quinto giorno. In figura (2.21) sono invece rappresentati l'AC e l'RMSE (sempre per il geopotenziale a quella quota).

In figura (2.22) sono invece rappresentati i ROC diagram riferiti a temperatura al suolo, precipitazioni e pressione sul livello del mare per ciascuna settimana di integrazione mediati su tutto l'emisfero settentrionale (ad esclusione della fascia tropicale).

Nel lavoro viene sottolineata anche la forte dipendenza della predicibilità dalla zona geografica. In questo senso, il Nord America è più predicibile dell'Europa (figura 2.23).

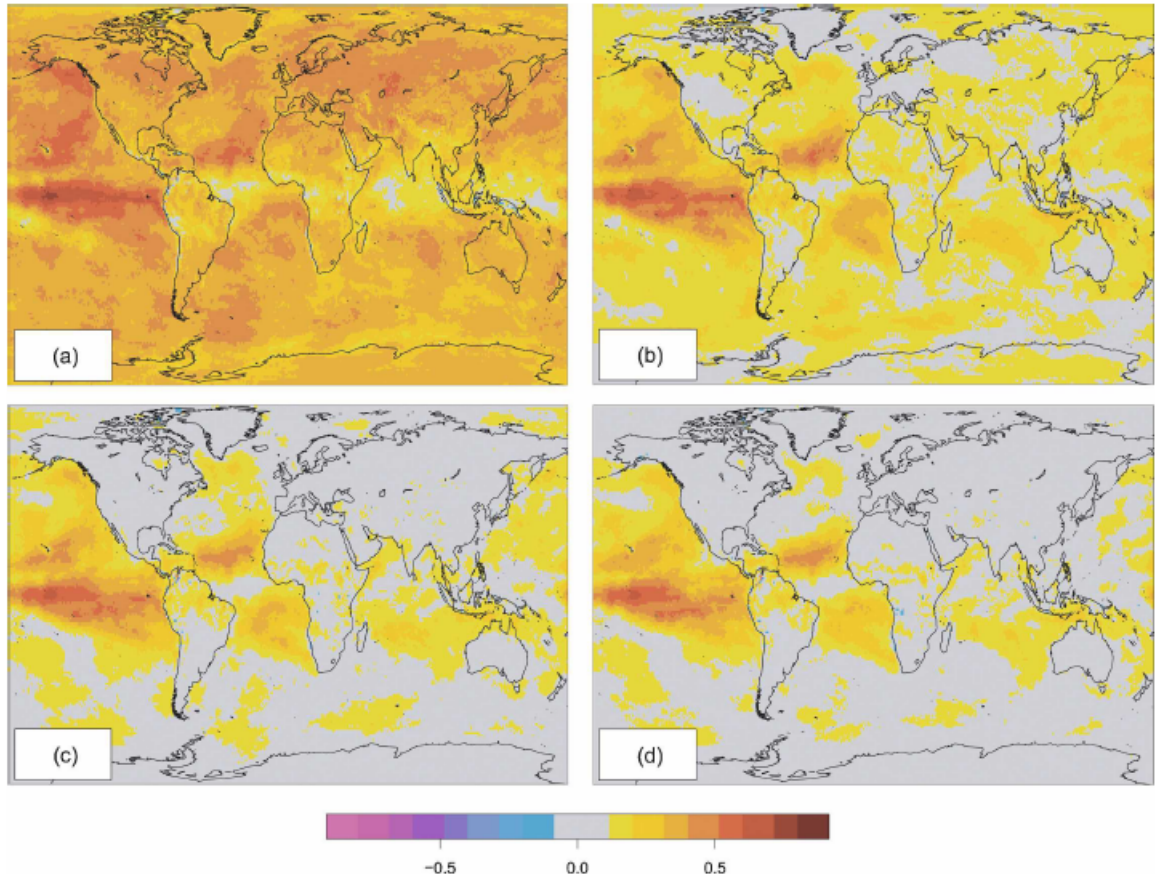


Figura 2.24: SS_{RPSS_D} della temperatura sulla superficie medio annuale suddiviso per settimana di previsione: (a) prima settimana, (b) seconda, (c) terza, (d) quarta. *Fonte: Weigel et al (2008), [2]*

In generale, i buoni risultati ottenuti nella predicibilità fino a 18 giorni (periodo in cui il modello performa meglio della climatologia e della persistenza), e per alcune zone anche fino a fine corsa, spinsero l'ECMWF a rendere operativo il modello a partire dall'ottobre del 2004.

In un lavoro successivo (Weigel et al. 2008, [2]) sono state studiate le performance del modello nella previsione della temperatura sfruttando gli score probabilistici (cfr. §2.4), in particolare l' SS_{RPS} , in una versione modificata come segue, chiamata *Discrete Ranked Probability Score*, SS_{RPSS_D} :

$$SS_{RPSS_D} = 1 - \frac{\overline{RPS}}{\overline{RPS}_{clim} + \frac{D_0}{M}} \quad D_0 = \frac{K^2 - 1}{6K} \quad (2.19)$$

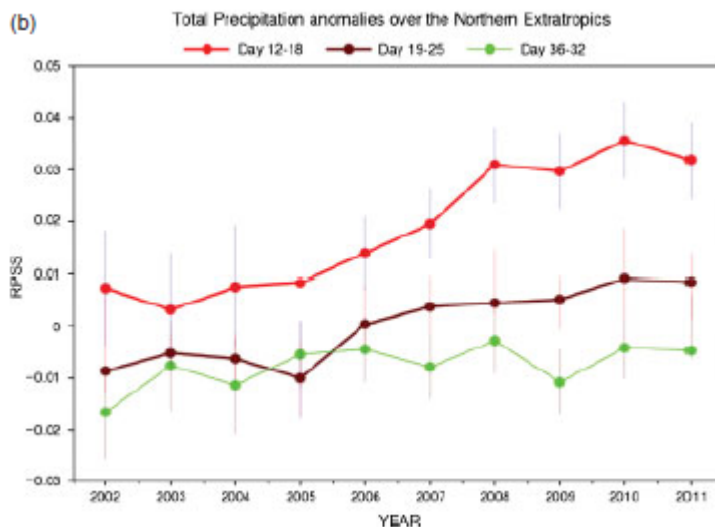
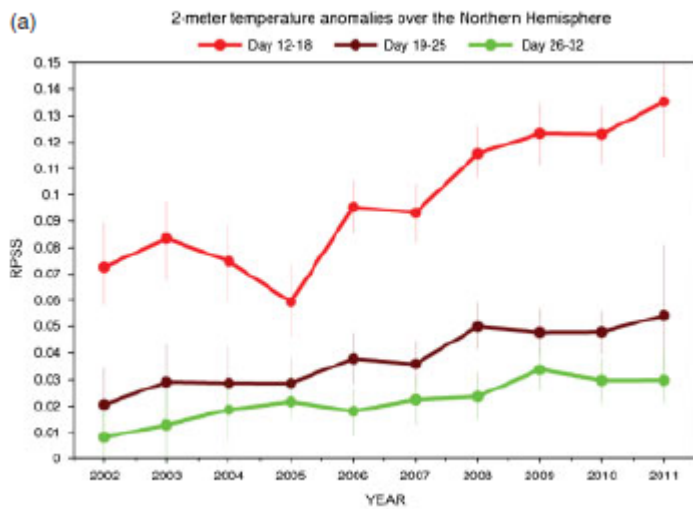


Figura 2.25: Evoluzione dell'RPSS relativo all'anomalia della temperatura sulla superficie (a) e a quella relativa alle precipitazioni (b) su tutto l'emisfero settentrionale, durante il periodo di attività (fino al 2011) del modello per previsioni mensili sviluppato presso l'ECMWF. *Fonte: Vitart (2014), [18]*

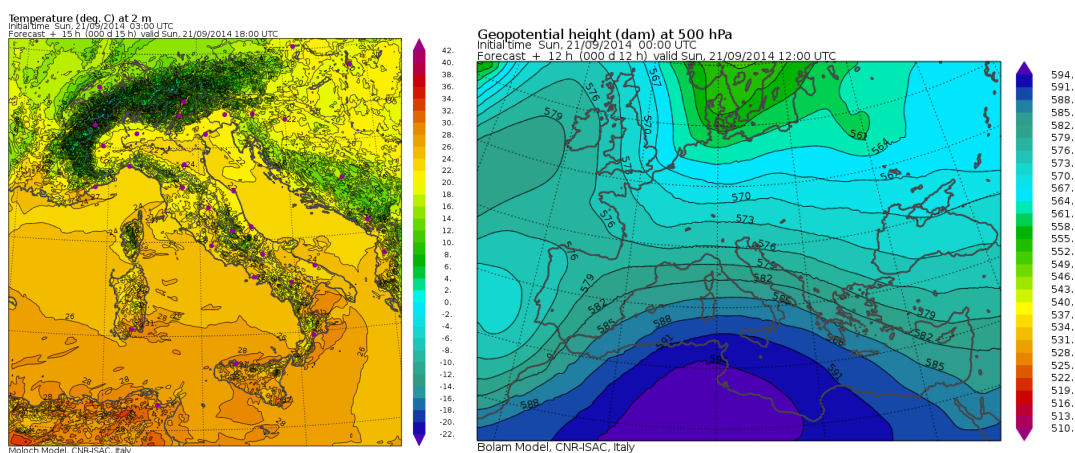
dove K sta per il numero di sottogruppi in cui è divisa la previsione, ed M la dimensione dell'ensemble. La convenienza nell'uso di questa versione dell'RPS è di renderla meno dipendente dalla dimensione dell'ensemble (che è piccolo - soltanto 5 membri - nel caso del calcolo della media climatica).

Nella figura (2.24) sono rappresentati i valori del sovradescritto score.

In un ultimo e recente lavoro (Vitart 2014, [18]), vengono riassunte le performance mediate su tutto l'emisfero settentrionale per la temperatura al suolo e le precipitazioni, sempre tramite l' SS_{RPS} (figura 2.25).

3 Le previsioni mensili presso il CNR-ISAC

Figura 3.1: Elaborazioni prodotte dai modelli operativi presso il CNR-ISAC



(a) Un campo di temperatura al suolo previsto dal modello MOLOCH (b) Un campo di altezza geopotenziale a 500 hPa previsto da BOLAM

3.1 I tre modelli sviluppati presso il CNR-ISAC

Agli inizi degli anni Novanta, l'Istituto di Scienze dell'Atmosfera e del Clima (ISAC) del Consiglio Nazionale delle Ricerche (CNR), ha avviato lo sviluppo di un modello numerico a griglia (*Bologna Limited Area Model*, BOLAM, figura 3.1b), attualmente operante su una griglia di risoluzione di 8.3 km e 50 livelli verticali, in grado di fornire previsioni fino a tre giorni per l'area europea-mediterranea e pensato per descrivere i fenomeni di mesoscala. Da allora altri due modelli sono stati sviluppati a partire da BOLAM presso lo stesso istituto e sono oggi operativi: MOLOCH e GLOBO.

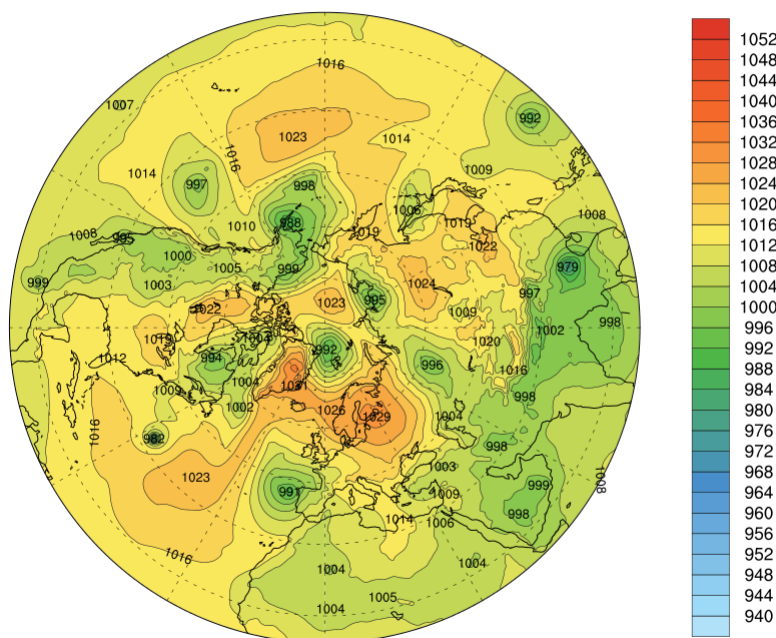
MOLOCH (figura 3.1a) è un modello ad alta risoluzione (ad oggi operante su una griglia di passo di 1.55 km), non idrostatico¹, costruito allo scopo di migliorare la previsione

¹Un modello si dice idrostatico quando sfrutta il fatto che, in prima approssimazione, la colonna verticale sia in equilibrio idrostatico. In tal caso vale l'equazione:

$$\delta p = -\rho \cdot g \cdot \delta z \quad (3.1)$$

che lega la variazione infinitesima della quota (δz) a quella della pressione atmosferica (δp), tramite la densità dell'aria (ρ) e l'accelerazione di gravità (g). La condizione idrostatica è un'approssimazione

Pressure reduced to MSL (hPa)
 Initial time Mon, 15/09/2014 00:00 UTC
 Forecast time + 48 h Valid Wed, 17/09/2014 00:00 UTC



GLOBO model, CNR-ISAC, Bologna, Italy

Figura 3.2: Una elaborazione del modello GLOBO, in uso presso il CNR-ISAC, per il campo di pressione al suolo sull'emisfero settentrionale.

dei fenomeni convettivi e su microscala. Opera come BOLAM su 50 livelli verticali e da quello ricava le proprie condizioni iniziali e al contorno. Si spinge fino a 48 ore di previsione ed è centrato sull'Italia.

GLOBO (figura 3.2) è stato sviluppato per ultimo, negli anni Duemila. Fornisce attualmente previsioni fino a 6 giorni per tutto l'emisfero settentrionale con un passo di griglia di 24 km e 60 livelli verticali. A partire dal 2010 l'ISAC si è servito di tale modello anche per elaborare previsioni meteorologiche mensili.

3.2 GLOBO per il medio-termine

Le caratteristiche del modello GLOBO ora presentate sono descritte più estesamente in (Malguzzi et al., 2011, [7]). Ivi si trovano anche i risultati del lavoro di verifica sulle performance del modello, effettuato per il periodo agosto 2009 - gennaio 2011.

valida soltanto nel caso in cui il passo di griglia orizzontale di un modello sia sufficientemente grande.

3.2.1 Rappresentazione delle coordinate

GLOBO opera su un dominio sferico le cui coordinate orizzontali sono rappresentate dalla latitudine e dalla longitudine (λ, θ) .

Il sistema di coordinate verticali ha una struttura ibrida in cui una coordinata isentropica modulata dall'orografia, σ , muta gradualmente con l'altitudine in una coordinata di pressione, P :

$$P = P_{ref} \cdot \sigma - (P_{ref} - P_{sur}) \cdot \sigma^\alpha \quad (3.2)$$

dove P_{sur} è la pressione sul livello del mare, P_{ref} un valore di pressione preso come riferimento e α un parametro che incide sulla rapidità con cui σ tende verso P . Si noti che per $\sigma = 1$ vale che $P = P_{sur}$, mentre per $\sigma = 0$ si ottiene $P = 0$, ovvero la coordinata verticale del modello è limitata da una superficie σ sia in cima che alla base.

3.2.2 Dinamica e schemi di parametrizzazione

GLOBO è un modello idrostatico (vale dunque la 3.1) che integra le classiche equazioni primitive che regolano i moti atmosferici. Le principali variabili prognostiche da esso prodotte sono le due componenti della velocità orizzontale del vento (u e v), la pressione sul livello del mare (P_{sur}) e la temperatura virtuale (T_V)².

Le equazioni del momento lungo l'orizzontale definite dal modello sono:

$$\frac{du}{dt} - uv \cdot \frac{\tan \theta}{a} - 2\Omega \sin \theta \times v = \frac{-R_d T_V}{a \cos \theta} \frac{P_{sur} \sigma^\alpha}{P_{ref} \sigma - (P_{ref} - P_{sur}) \sigma^\alpha} \frac{\partial}{\partial \lambda} \ln P_{sur} - \frac{1}{a \cos \theta} \frac{\partial \Phi}{\partial \lambda} \quad (3.4)$$

$$\frac{dv}{dt} - u^2 \cdot \frac{\tan \theta}{a} - 2\Omega \sin \theta \times u = \frac{-R_d T_V}{a} \frac{P_{sur} \sigma^\alpha}{P_{ref} \sigma - (P_{ref} - P_{sur}) \sigma^\alpha} \frac{\partial}{\partial \theta} \ln P_{sur} - \frac{1}{a} \frac{\partial \Phi}{\partial \theta} \quad (3.5)$$

dove R_d è la costante dei gas perfetti per l'aria secca, a e Ω rispettivamente il raggio e la velocità angolare terrestri, Φ l'altezza geopotenziale. Quest'ultima è calcolata come segue:

$$\frac{\partial \Phi}{\partial \ln \sigma} = -R_d T_V \cdot \frac{P_{ref} - \alpha(P_{ref} - P_{sur}) \sigma^{\alpha-1}}{P_{ref} - (P_{ref} - P_{sur}) \sigma^{\alpha-1}} \quad (3.6)$$

Attraverso l'integrazione dell'equazione di continuità:

$$\frac{\partial}{\partial t} \left(\frac{\partial P}{\partial \sigma} \right) + \frac{\partial}{\partial x} \left(u \frac{\partial P}{\partial \sigma} \right) + \frac{\partial}{\partial y} \left(v \frac{\partial P}{\partial \sigma} \right) + \frac{\partial}{\partial \sigma} \left(\dot{\sigma} \frac{\partial P}{\partial \sigma} \right) = 0 \quad (3.7)$$

²La temperatura virtuale è una grandezza fisica definita dalla termodinamica, ed è la temperatura a cui una particella d'aria secca si dovrebbe trovare per avere la stessa densità a parità di pressione per una corrispondente particella d'aria umida. È definita come:

$$T_V = \frac{T}{1 - \frac{e}{p} \cdot (1 - \epsilon)} \quad (3.3)$$

dove T è la temperatura dell'aria, e la pressione parziale esercitata dal vapor acqueo, p la pressione atmosferica ed ϵ il rapporto tra la costante dei gas perfetti per l'aria secca e quella per l'aria umida.

sono invece ricavate le equazioni per la velocità verticale ($\dot{\sigma}$) e la pressione sul livello del mare:

$$\frac{\partial P_{sur}}{\partial t} = - \int_0^1 C d\sigma \quad (3.8)$$

$$\dot{\sigma} \frac{\partial P}{\partial \sigma} = -\sigma^\alpha \frac{\partial P_{sur}}{\partial t} - \int_0^\sigma C d\sigma \quad (3.9)$$

dove C è definita come:

$$C = P_{ref}(1 - \alpha\sigma^{\alpha-1})C_1 + \alpha\sigma^{\alpha-1}C_2$$

con

$$C_1 = \frac{1}{a \cos \theta} \left(\frac{\partial u}{\partial \lambda} + \frac{\partial}{\partial \theta}(v \cos \theta) \right) \quad C_2 = \frac{1}{a \cos \theta} \left(\frac{\partial}{\partial \lambda}(u P_{sur}) + \frac{\partial}{\partial \theta}(v P_{sur} \cos \theta) \right)$$

Fanno inoltre parte della struttura dinamica del modello l'equazione termodinamica (dove C_p è il calore specifico a pressione costante e ω la coordinata verticale definita attraverso il gradiente di pressione):

$$\frac{dT_V}{dt} = \frac{R_d T_V}{C_p} \frac{\omega}{P} \quad (3.10)$$

nonché le leggi di conservazione dell'umidità specifica e delle varie specie dell'acqua (neve, grandine, pioggia, eccetera).

GLOBO è dotato inoltre di molteplici schemi di parametrizzazione per la modellizzazione dei fenomeni fisici di cui le equazioni della fluido-dinamica e della termodinamica non possono tenere conto da sole. Gli aspetti principali a cui provvedono tali schemi sono:

- la rappresentazione dei fenomeni che coinvolgono il cosiddetto *Surface Layer*, SL, e il suolo (modellizzato con un modello di suolo a 7 livelli), e di come questi sono influenzati dalla presenza di boschi, copertura nevosa, ostacoli non risolti dal passo di griglia;
- la parametrizzazione del *Planetary Boundary Layer*, PBL (ovvero lo strato della troposfera a diretto contatto con il suolo) e dei fenomeni legati alla turbolenza che lo riguardano;
- i processi convettivi e quelli che riguardano la microfisica dei processi di precipitazione, sia su larga scala che al di sotto del passo di griglia;
- i processi legati al flusso di radiazione solare e all'interazione tra questa, le nubi e gli aerosol atmosferici;
- i processi legati al *gravity wave drag*, ovvero l'interazione tra onde di gravità e moto esplicito.

3.2.3 Condizioni iniziali

Le condizioni iniziali di GLOBO sono fornite dalle reanalisi giornaliere (0000 UTC) del *Global Ensemble Forecasting System* (GEFS), elaborate presso il NCEP e successivamente interpolate sia orizzontalmente sia verticalmente sulla griglia e sui livelli verticali del modello.

I dati sull'orografia sono forniti dal dataset della NOAA, mentre quelli riguardanti la tipologia del suolo e le caratteristiche della vegetazione sono ottenuti dal *Global Land Cover Facility* dell'Università del Maryland.

Il modello viene attualmente lanciato una volta al giorno e fornisce previsioni nel futuro fino a 6 giorni (144 ore).

3.3 GLOBO per le previsioni mensili

A partire dal 2010 il CNR-ISAC, avviando una collaborazione con il Dipartimento della Protezione Civile Nazionale (DPCN), ha deciso di utilizzare il modello GLOBO anche per produrre previsioni mensili probabilistiche sulle anomalie di temperatura e precipitazioni per l'Italia e l'Europa. Il progetto è stato avviato ufficialmente nell'agosto del 2010, e da allora viene emesso mensilmente un bollettino³. Nel marzo del 2014 il modello ha subito alcune importanti modifiche: il passo di griglia è stato ridotto a 0.8° di longitudine per 0.56° di latitudine (da 1.0° x 0.75°) e i livelli verticali portati a 54. Inoltre, il numero di membri dell'ensemble è passato da 32 a 40.

3.3.1 Struttura e specifiche

La struttura dinamica e gli schemi di parametrizzazione fisica di GLOBO esposti nella precedente sezione (cfr. §3.2) continuano ad essere validi per il modello utilizzato per formulare previsioni mensili.

Nell'ottica di produrre una previsione a lungo termine viene tuttavia introdotto un debole accoppiamento atmosfera-oceano tramite la modellizzazione dell'evoluzione delle temperature superficiali dei mari (SST):

$$\frac{\partial SST}{\partial t} = \frac{F_{net}}{H\rho C} - \gamma \cdot (SST - T_{clim} - T_{an}) \quad (3.11)$$

dove T_{clim} è la media climatologica (calcolata su 20 anni di osservazioni fornite dalle reanalisi dell'ECMWF ERA-Interim), T_{an} l'anomalia osservata al momento dell'inizializzazione (e successivamente decrementata del 15% ogni mese), F_{net} il contributo netto dovuto ai flussi di turbolenza e radiazione, ρC la capacità termica dell'acqua marina,

³reperibile all'indirizzo http://www.isac.cnr.it/dinamica/projects/forecast_dpc/month_en.htm

2-m temperature

6 Oct 2014 - 20 Oct 2014

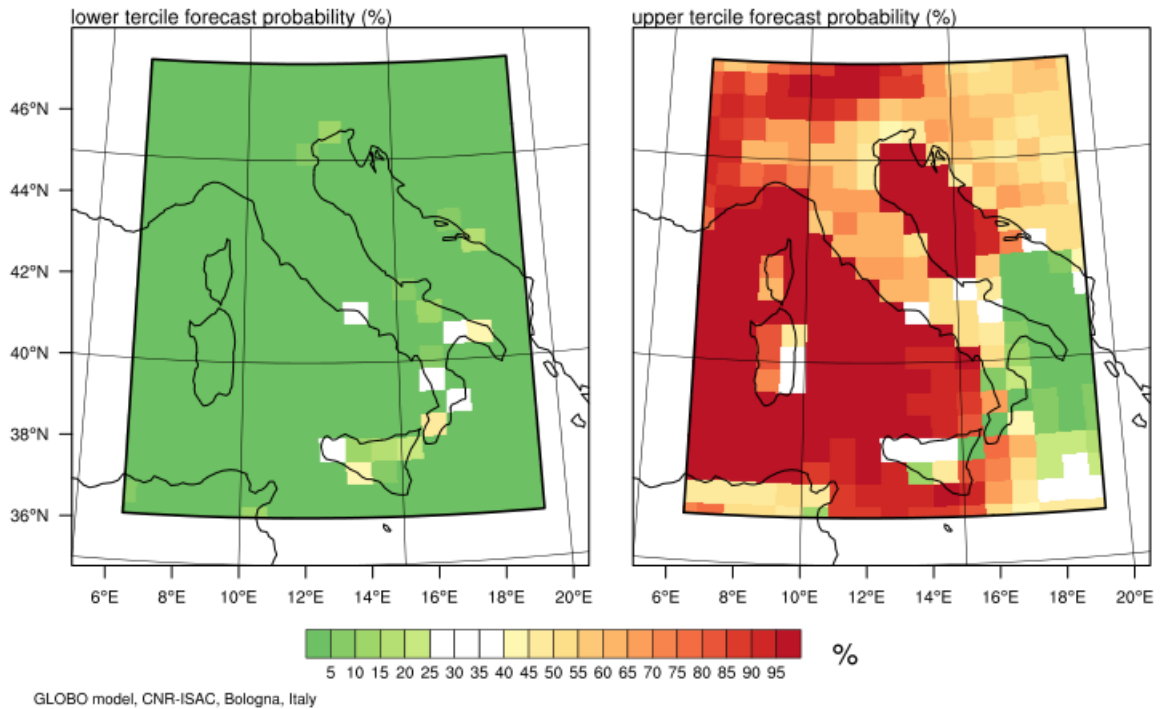


Figura 3.3: Esempio di elaborazione del modello per previsioni mensili GLOBO: probabilità che la temperatura al suolo ricada nel tercile inferiore (a sinistra) o superiore (a destra), nel periodo 6-20 ottobre 2014, per ciascun punto di griglia sul territorio italiano.

H una costante indicante lo spessore dello strato rimescolato dell'oceano, e infine γ una costante dimensionata di valore $2,3 \text{ giorni}^{-1}$.

Il modello tiene conto anche della copertura di ghiaccio marino, inizializzata sempre mediante le ERA-Interim e le cui variazioni nel tempo durante la fase di integrazione numerica seguono lo stesso rate della media climatologica.

Le previsioni mensili elaborate da GLOBO sfruttano la tecnica dell'ensemble (cfr. §2.2). Il modello viene fatto girare una volta alla settimana, producendo per ogni data di inizializzazione 40 corse deterministiche, 10 per ciascuno degli orari sinottici di riferimento (00 UTC, 06 UTC, 12 UTC, 18 UTC). Le condizioni iniziali sono fornite tutte dal *Global Ensemble Forecasting System* (GEFS) del NCEP.

Il modello produce previsioni circa l'anomalia (calcolata come nelle equazioni 2.1 e 2.2)

rispetto alla media climatica di riferimento 1981-2010 per le seguenti variabili:

- altezza geopotenziale a 500 hPa;
- pressione sul livello del mare;
- temperatura alla quota di 850 hPa;
- temperatura al suolo;
- precipitazioni cumulate.

3.3.2 Il bias del modello GLOBO per previsioni mensili

L'importanza della calibrazione di un modello, ovvero della valutazione dell'errore sistematico da cui è afflitto (cfr. §2.4) permette di apportare utili correzioni nella fase di post-processing e dunque di migliorarne conseguentemente le performance.

Uno studio dell'errore sistematico del modello per previsioni mensili GLOBO è tra gli obiettivi espliciti di questo lavoro (cfr. §1). Tuttavia, un precedente lavoro svolto presso il CNR-ISAC (Mastrangelo et al. 2012, [3]) si è già occupato della questione nel periodo in cui era ancora operativa la vecchia versione del modello (aggiornato nel marzo del 2014). Verranno dunque ora descritte il metodo perseguito e i risultati ottenuti in quella ricerca, che serviranno da linee guida per impostare il lavoro che verrà affrontato nel prossimo capitolo.

Nello studio sopra citato furono scelti come periodo climatico di riferimento i ventuno anni compresi tra il 1° gennaio 1989 e il 15 dicembre 2009, mentre il database ERA-Interim dell'ECMWF forniva i dati per le rispettive reanalisi. All'interno di questo intervallo di tempo venivano lanciate 21 simulazioni inizializzate alle 00 e altre 21 alle 12 UTC, ciascuna della durata di 32 giorni, il giorno 1 e il giorno 15 di ogni mese (chiamati *calendar day*), ottenendo così un ensemble di 42 membri. Si ebbero in questo modo un totale di 1008 simulazioni utili a determinare il clima del modello per il periodo in previsione, da cui poter calcolare e rimuovere l'errore sistematico.

Definito a questo punto $\Psi_{i,y,h}(\lambda, \theta, \nu)$ come il valore previsto per una certa variabile meteorologica Ψ nel punto di griglia di coordinate (λ, θ) , dopo $\nu = 0, \dots, 31$ giorni di integrazione numerica dal reforecast inizializzato con i dati dell'ora $h = 1$ (0000 UTC), 2 (1200 UTC) del calendar day C_i (con $i = 1, \dots, 24$) dell'anno $y = 1, \dots, 21$, l'errore commesso dal modello era dato da:

$$E_{i,y,h}(\lambda, \theta, \nu) = \Psi_{i,y,h}(\lambda, \theta, \nu) - \Psi^T_{C_i+\nu,y,h}(\lambda, \theta) \quad (3.12)$$

dove Ψ^T definisce il corrispettivo valore di confronto ottenuto dalle reanalisi dell'ECMWF.

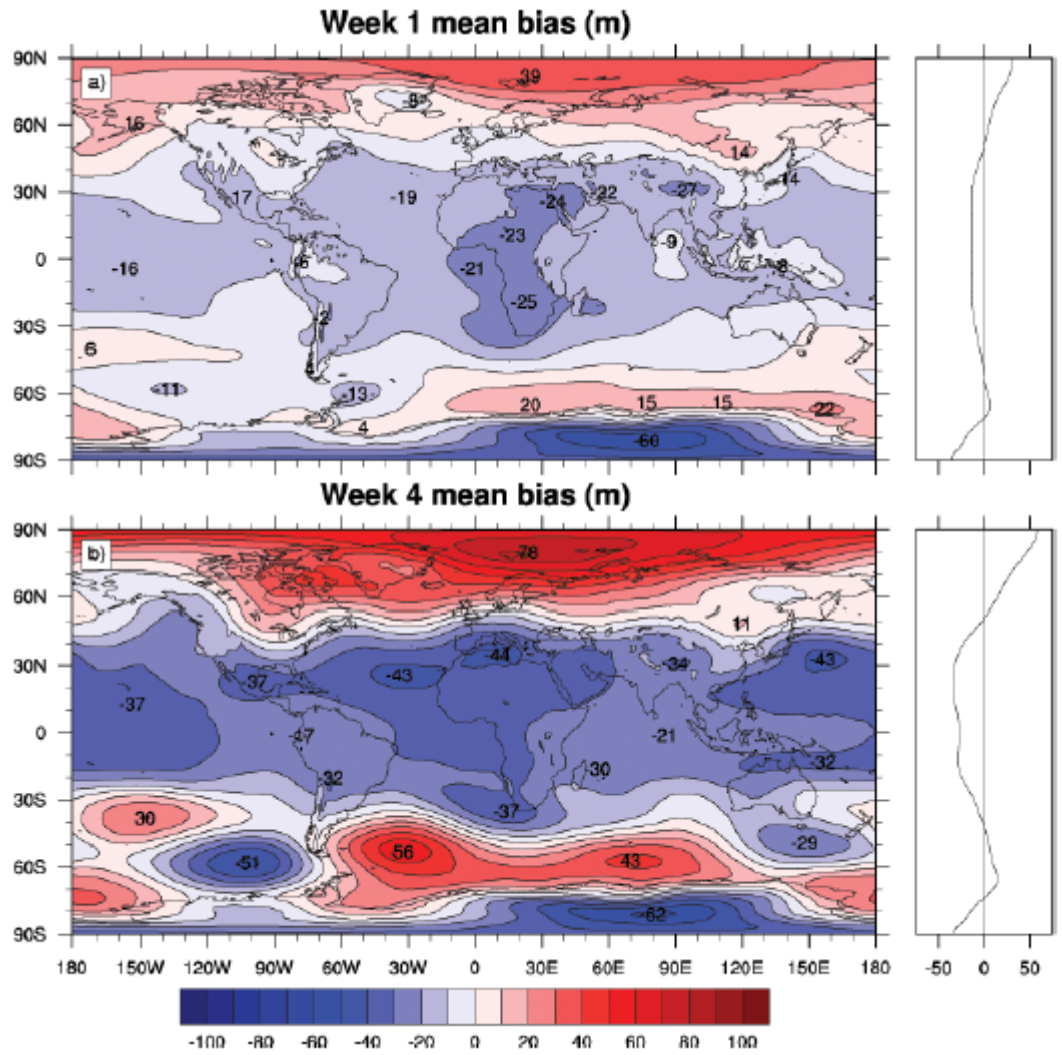


Figura 3.4: bias globale per la prima e la quarta settimana di integrazione numerica, calcolato secondo l'equazione 3.14 per l'altezza geopotenziale ai 500 hPa

Una stima preliminare del bias veniva dunque ottenuta attraverso:

$$B^{\Psi}_i(\lambda, \theta, \nu) = \frac{1}{42} \sum_{y=1}^{21} \sum_{h=1}^2 E_{i,y,h}(\lambda, \theta, \nu) \quad i = 1, \dots, 24 \quad (3.13)$$

Tuttavia, il bias calcolato come nella 3.13 (in quanto funzione del calendar day) è soggetto sia alla variabilità stagionale (tra due calendar day lontani nel tempo) sia a quella sinottica e su piccola scala (tra due calendar day vicini tra loro). Nel caso di un modello per previsioni mensili, e dunque nell'ottica di far risaltare le configurazioni su larga scala,

era importante che nella fase di rimozione del bias non venissero introdotti segnali su piccola scala, poiché questo avrebbe apportato ulteriore rumore al modello. Per questo motivo venne introdotto un peso gaussiano, allo scopo di introdurre una media sui calendar day adiacenti e filtrare in questo modo i pattern su scale inferiori a quelle di interesse:

$$B^\Psi(\lambda, \theta, \nu, c) = \frac{1}{W} \sum_{i=1}^{24} B^\Psi_i(\lambda, \theta, \nu) e^{-\left(\frac{|c-C_i|}{\Delta}\right)^2} \quad (3.14)$$

dove W è la somma dei singoli pesi e Δ un parametro che viene fissato al fine di ottenere la migliore calibrazione.

In figura (3.4) è riportato il bias così calcolato (mediato su tutto il periodo di reforecasting) della prima e della quarta settimana riferito all'altezza geopotenziale dei 500 hPa.

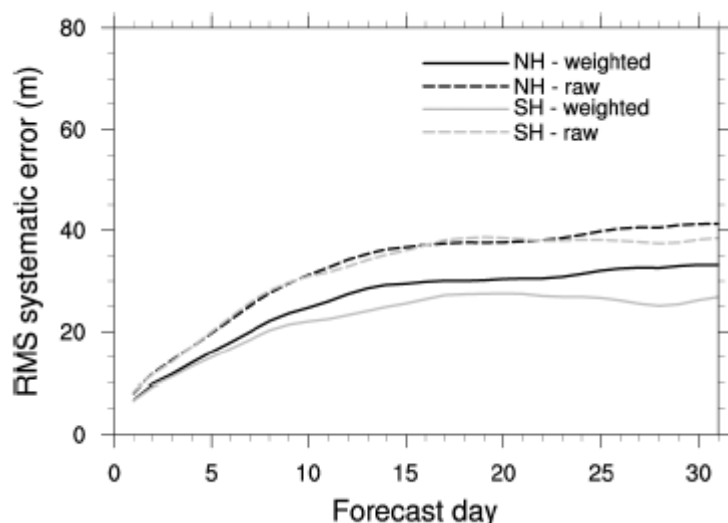
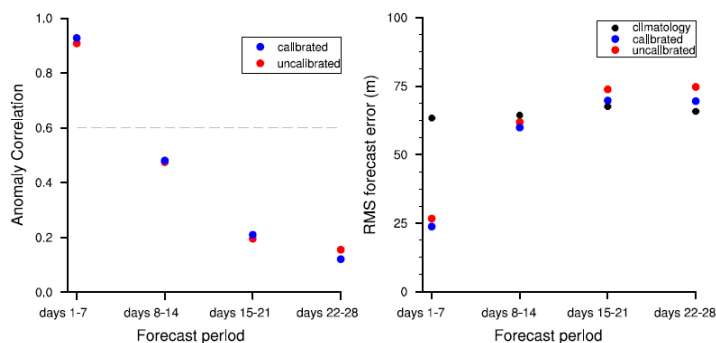


Figura 3.5: RMSE emisferico, calcolato tramite bias rozzo (eq. 3.13, linea tratteggiata) e pesato (eq. 3.14, linea continua)



Oltre al bias, in (Mastrangelo et al., 2012,[3]) è stato analogamente calcolato l'RMSE per l'emisfero settentrionale (da 20°N a 90°N) e meridionale (da -20°S a -90°S). Si veda in tal proposito la figura (3.5).

Infine, fu fatta anche una verifica preliminare delle performance del modello in modalità real-time, esaminando un campione di previsioni di 13 mesi (da agosto 2010 ad agosto 2011 compresi). Il calcolo di AC_U e $RMSE$ relativi sono riportati in figura (3.6).

Figura 3.6: RMSE e AC della verifica real-time per l'emisfero settentrionale dell'altezza geopotenziale ai 500 hPa.

4 Bias e RMSE del modello GLOBO

Nel seguente capitolo viene presentato il fulcro del presente lavoro: il calcolo del bias e dell'RMSE del modello GLOBO per previsioni mensili. Nella prima sezione sono espone la metodologia usata e le caratteristiche delle simulazioni trattate. Nella seconda sono invece prima commentati e poi riportati i grafici e le mappe relative ai risultati ottenuti.

4.1 Simulazioni e metodo di lavoro

La valutazione delle performance del modello trattata qui riguarda tre dei cinque campi forniti da GLOBO in modalità di previsione mensile: le precipitazioni cumulate, l'altezza geopotenziale ai 500 hPa (in seguito spesso riassunta dalla sigla Z500) e la temperatura alla quota di 850 hPa (in sigla: T850). Si è scelto inoltre di riferirsi alla media climatologica trentennale di riferimento più recente (1981-2010).

Per quanto concerne il trattamento dei dati e la loro elaborazione grafica, si è sfruttato un linguaggio di programmazione particolarmente versatile e adatto a trattare grandi moli di dati meteo-climatologici: NCL (Ncar Command Language)¹.

Invece, per quanto riguarda le simulazioni di reforecast e le reanalisi necessarie per la valutazione delle performance di GLOBO (cfr. §2.3.2), verranno chiarite di seguito le caratteristiche valide nel presente caso.

4.1.1 Reanalisi e reforecast

Per quel che riguarda le reanalisi si sono sfruttate le rielaborazioni ERA-Interim fornite dal Centro Europeo per le Previsioni Meteorologiche a Breve e Medio Termine (ECM-WF).

Per i reforecast si sono invece utilizzate 2190 simulazioni effettuate mediante il modello GLOBO. Queste ultime coprono il periodo trentennale compreso tra il 1° gennaio 1981 e il 31 dicembre del 2010 e sono costituite da simulazioni tutte inizializzate alle 00 UTC, di durata di 35 giorni ciascuna, con una cadenza di integrazione di dodici ore (00 UTC e 12 UTC) e distanziate di 5 giorni l'una dall'altra (per un totale di 73 *calendar day* uniformemente distribuiti nell'arco dell'anno). Tali simulazioni sono state elaborate all'interno di un progetto di collaborazione con l'IGI (Italian Grid Infrastructure), come esposto in Cesini et al. 2013, [17].

Nella fase procedurale, a causa delle differenti risoluzioni delle griglie dei campi di reanalisi (1°x 1°) e di reforecast (0.80°x 0.56°) si è dovuto ricorrere al regridding tramite interpolazione bilineare, riportando il passo di griglia del modello GLOBO a quello delle ERA-Interim. Dunque, tutte le elaborazioni qui riportate sono calcolate su campi di tale

¹<http://www.ncl.ucar.edu/>

risoluzione (1° di latitudine per 1° di longitudine).

4.1.2 Bias

La definizione del bias applicata nel presente lavoro è definita a partire dall'errore E , differenza tra i valori del campo previsto e di quello osservato:

$$E_{i,y,h}(\lambda, \theta, \nu) = \Psi_{i,y,h}(\lambda, \theta, \nu) - \Psi^T_{C_i+\nu,y,h}(\lambda, \theta) \quad (4.1)$$

dove C_i rappresenta il *calendar day* con $i = 1, \dots, 73$, λ e θ sono le coordinate spaziali (rispettivamente latitudine e longitudine), $y = 1, \dots, 30$ l'anno in cui si trova il *calendar day*, ν e h rispettivamente il giorno e l'orario di previsione ($\nu = 0, \dots, 35$, $h = 1$ (0000 UTC), $h = 2$ (1200 UTC)). Ψ e Ψ^T sono ancora rispettivamente i valori forniti dalle simulazioni dei reforecast e quelli delle reanalisi.

Dunque, il bias preso come riferimento per le successive trattazioni è definito come la media trentennale del valore di E :

$$B^\Psi_{i,h}(\lambda, \theta, \nu) = \frac{1}{30} \sum_{y=1}^{30} E_{i,y,h}(\lambda, \theta, \nu) \quad i = 1, \dots, 24 \quad (4.2)$$

A partire da tale valore, per ciascuna delle variabili di previsione considerate (precipitazioni cumulate, Z500, T850), sono stati mappati i valori del bias per le prime quattro settimane di integrazione, W (ovvero riferiti ai vari intervalli di integrazione, ν_W : $\nu_1 = 0, \dots, 6$, $\nu_2 = 7, \dots, 13$ e così via), sia mediati sull'intero anno (cioè sulla totalità dei *calendar day*):

$$B^\Psi(\lambda, \theta, W) = \frac{1}{73} \frac{1}{7} \frac{1}{2} \sum_{i=1}^{73} \sum_{\nu_W} \sum_h B^\Psi_{i,h}(\lambda, \theta, \nu) \quad (4.3)$$

sia sulle singole stagioni², S :

$$B^\Psi(\lambda, \theta, W, S) = \frac{1}{s} \frac{1}{7} \frac{1}{2} \sum_{i_S} \sum_{\nu_W} \sum_h B^\Psi_{i,h}(\lambda, \theta, \nu) \quad (4.4)$$

dove i_S sono i *calendar day* della relativa stagione ed s è la loro quantità.

Si tenga conto che nel calcolo del bias relativo alle precipitazioni cumulate, essendo queste prodotte dal modello come valori di accumulo delle antecedenti 12 ore e presentando quindi possibili discontinuità, il valore relativo alle diverse settimane di integrazione numerica è da intendersi come bias riferito a somme di valori cumulati e non come medie temporali.

²Si sfrutteranno, da qui in avanti, le tipiche abbreviazioni: JJA per indicare i mesi dell'estate boreale (June, July, August), SON quelli dell'autunno, DJF dell'inverno, MAM della primavera.

4.1.3 RMSE

Una seconda fase del lavoro ha riguardato il calcolo dell'RMSE (*Root Mean Squared Error*; cfr. §2.4.1). In tale frangente ci si è limitati a valutazioni per la temperatura in quota (T850) e per l'altezza geopotenziale (Z500).

Si sono prodotti sia dei campi spaziali, sia delle serie temporali riferite ad alcune aree geografiche specifiche: Emisfero Boreale (da 20°N a 90°N), Emisfero Australe (da 20°S a 90°S), zona Euro-Atlantica (da 30°N a 70°N e da 20°W a 40°E), e zona del Nord Pacifico (da 25°N a 70°N e da 160°E a 70°W).

Per quanto concerne l'RMSE areale, si è applicata la definizione dell'equazione 2.6, al variare dell'area considerata A , e si è partiti dal valore medio settimanale del bias:

$$RMSE(W, A) = \frac{1}{s} \sum_{i_s} RMSE_{i_s}(W, A) \quad (4.5)$$

con:

$$RMSE_{i_s}(W, A) = \sqrt{\frac{1}{\Lambda \cdot \Theta} \frac{1}{\sum_{\Lambda} w_{\lambda}} \sum_{\lambda \in A} \sum_{\theta \in A} w_{\lambda} B^{\Psi}(\lambda, \theta, W)^2} \quad (4.6)$$

dove $\Lambda \cdot \Theta$ indica il numero totale di di griglia all'interno dell'area A , e dove sono stati introdotti dei pesi lungo la latitudine (w_{λ}) per tener conto, nel computo di una media spaziale, delle variazioni di dimensioni della griglia indotte dalle coordinate sferiche.

Oltre al valore medio settimanale, si è prodotta una serie temporale anche per i singoli valori giornalieri (grafici bla bla).

In secondo luogo, per quanto riguarda i campi spaziali, l'RMSE di per sé (equazione 2.6) non è uno score probabilistico, e dunque non sarebbe adatto a valutare simulazioni su periodi così lunghi. Tuttavia, non avendo a che fare con vere e proprie simulazioni di ensemble (e quindi non potendo ragionare in termini probabilistici), ricorriamo lo stesso all'RMSE cercando di aumentarne la significatività statistica. Avendo a disposizione molte simulazioni distribuite nel tempo, facciamo delle medie temporali delle varie integrazioni sui punti di griglia, sfruttando dunque una versione dell'RMSE mediata sul tempo e non sullo spazio.

L'RMSE mediato sull'intero anno, per ciascun punto di griglia, al variare della settimana di integrazione ($W = 1, 2, 3, 4$) è stato dunque definito come:

$$RMSE(\lambda, \theta, W) = \sqrt{\frac{1}{73} \frac{1}{7} \frac{1}{2} \sum_{i=1}^{73} \sum_{\nu_W} \sum_h B^{\Psi}_{i,h}(\lambda, \theta, \nu)^2} \quad (4.7)$$

ed è raffigurato nelle figure bla bla.

L'RMSE mediato sulle singole stagioni, sempre per ciascun punto di griglia, e sempre al

variare della settimana di integrazione, è stato invece calcolato come valore medio sui vari calendar day stagionali, i_S , in numero s :

$$RMSE(\lambda, \theta, W, S) = \frac{1}{s} \sum_{i_S} RMSE_{i_S}(\lambda, \theta, W) \quad (4.8)$$

con

$$RMSE_{i_S}(\lambda, \theta, W) = \sqrt{\frac{1}{14} \sum_{\nu_W} \sum_h B^{\Psi}_{i_S, h}(\lambda, \theta, \nu)^2} \quad (4.9)$$

e dove S è il numero di *calendar day* della stagione considerata.

4.2 Risultati

Vengono ora riportati e commentati i principali risultati ottenuti, esposti poi nei grafici e nelle mappe a seguire.

4.2.1 Bias

Una considerazione generale valida per tutti i successivi commenti è che l'errore sistematico tende ad aumentare man mano che l'integrazione numerica procede. Tuttavia, superate le prime due settimane di integrazione, i valori di bias ed il pattern tendono a stabilizzarsi facendo emergere quindi la componente sistematica dell'errore, come necessario nell'ambito della previsione mensile. Questo rispecchia quanto ci si aspetta da un modello numerico dell'atmosfera integrato su queste scale temporale: superati i primi 10-15 giorni di simulazione, le componenti a scala sinottica del moto sono ormai svincolate dalle condizioni iniziali e, mediando su un grande numero di simulazioni, non dovrebbero lasciare impronta sui campi risultanti, lasciando emergere le componenti a più grande scala dell'errore sistematico.

Precipitazioni cumulate Le figure (4.1 - 4.4) ritraggono l'evoluzione del bias medio annuale delle precipitazioni cumulate espresso come medie settimanali sul periodo di previsione. Si può notare come il modello sovrastima le precipitazioni su tutta l'area continentale equatoriale, ovvero sulle zone tropicali del Sud America, dell'Africa e del sud-est asiatico, mentre le sottostima sulle zone oceaniche. Le precipitazioni, quindi il bias, risultano più abbondanti anche in prossimità delle zone con rilievi orografici importanti, ma in generale in tutta la fascia interessata dall'attività convettiva tropicale. Precipitazioni più abbondanti dell'osservato colpiscono anche l'Africa subsahariana e le zone più a nord. Questo dato, assieme alla presenza di precipitazioni abbondanti anche nelle aree forestali del Sud America e dell'Africa sin dalla prima settimana di integrazione, suggerisce una probabile influenza dei processi umidi del suolo. In generale il modello performa bene

sulle aree continentali dell'emisfero settentrionale: sul Nord America, come sull'Europa e sull'Asia orientale. Una vasta e duratura (sebbene leggera) sottostima coinvolge gran parte del territorio russo.

Z500 Il bias dell'altezza di geopotenziale alla quota isobarica di 500 hPa (4.9 - 4.12) mostra un pattern caratterizzato da: aree negative nella fascia intertropicale e aree immediatamente adiacenti, aree positive sulle medie latitudini settentrionali e polo Nord; un pattern più irregolare nell'emisfero Sud. In particolare, estese aree di sovrastima sono presenti nell'emisfero settentrionale, con i massimi su Siberia e Canada, e a ridosso dell'Antartide. È presente inoltre un'ulteriore zona di forte bias positivo nel Pacifico Settentrionale, che però non assume la stessa posizione durante tutto l'anno (si vedano le immagini 5.13-5.16 in appendice): durante il trimestre invernale (DJF) ha un massimo a sud del mare di Bering, mentre in quello caldo (JJA) si sposta verso il continente asiatico, a largo del Giappone e della penisola della Kamcatka. Va osservato che la presenza di un gradiente positivo di bias tra medie latitudini e Polo Nord, indica la presenza di correnti zonali occidentali meno intense di quanto riprodotto dalle rianalisi. Per quanto concerne le anomalie negative, queste si concentrano su Africa, Sud America, e nella zona euro-mediterranea. Il bias sul Polo Sud tende a cambiare segno nel corso delle 4 settimane.

T850 ³ Il bias della temperatura alla quota isobarica di 850 hPa (4.5 - 4.8) assume un andamento ben diverso tra alte e basse latitudini in entrambi gli emisferi ed è grossolanamente distribuito come quello dell'altezza di geopotenziale alla quota di 500 hPa. Le anomalie negative prevalgono nelle zone tropicali, mentre quelle positive sono concentrate tutte nei dintorni dei poli, e su gran parte dell'emisfero settentrionale. Le anomalie positive più intense interessano l'area della Siberia e del Canada, mentre un importante bias negativo coinvolge la zona dell'Africa tropicale settentrionale e la penisola araba: quest'ultimo risulta fortemente in relazione con la sovrastima delle precipitazioni su quelle aree. Sulle parti meridionali dell'Oceano Pacifico e Atlantico, sono presenti sin dalla prima settimana due aree di bias positivo. Nel corso del periodo di simulazione tendono a regredire leggermente mentre delle aree di bias negativo si sviluppano più a sud. Queste aree mostrano una minore corrispondenza con il bias di Z500.

4.2.2 RMSE

Anche per la progressione settimanale dell'RMSE vale quanto esposto per il bias: man mano che il tempo di integrazione procede, aumentano gli errori sulla previsione.

³Si tenga conto che, per ragioni orografiche, sono state mascherate tutte le aree della superficie terrestre per le quali GLOBO prevedeva una altezza rispetto al livello del mare uguale o superiore ai 1800 metri.

Z500 La distribuzione spaziale dei valori di RMSE per l'altezza geopotenziale (4.17 - 4.20) ricalca quella del bias. Le zone di maggiore incertezza risultano essere ancora una volta le alte latitudini dell'emisfero settentrionale e le zone antartiche in quello meridionale (con i massimi a ridosso del continente, nella parte meridionale dell'Oceano Indiano). Si ripete a sua volta anche qui la forte incertezza sul Pacifico Settentrionale e con la stessa variazione di posizione tra trimestre caldo (JJA) e freddo (DJF). Le zone equatoriali sono quelle con valori di RMSE più piccoli.

L'area di bias negativo della temperatura (particolarmente intensa nel semestre MAM-JJA) nell'Africa tropicale settentrionale e nella penisola arabica sembra essere correlata ad un aumento dell'RMSE dell'altezza geopotenziale sull'area sahariana nello stesso periodo (figure 5.5-5.8 in Appendice).

T850 Anche in questo caso ritroviamo la stessa distribuzione spaziale delle anomalie già riscontrate per il bias (4.13 - 4.16). Ritroviamo così elevati valori dell'RMSE nella zona dell'Africa equatoriale Settentrionale e nella penisola arabica, lì dove era presente un forte bias negativo per la T850. Vi è poi un'estesa anomalia nell'emisfero settentrionale con i massimi nella zona polare, in Canada, Siberia, e nella parte asiatica dell'Oceano Pacifico Settentrionale. Le aree equatoriali risultano quelle meno interessate da incertezza, con valori al di sotto del grado fino alla quarta settimana su gran parte del sud-est asiatico, l'Oceano Indiano e la parte settentrionale del Sud America.

Serie temporali Nei grafici (4.21 - 4.36) sono mostrate le serie temporali giornaliere e settimanali dell'RMSE per la temperatura in quota e l'altezza geopotenziale, riferite alle varie aree geografiche di riferimento scelte (Emisfero boreale, Emisfero Australe, Europa, area del Nord Pacifico). Si nota (soprattutto in quelli riferiti alla Z500) come i picchi di maggiore incertezza si abbiano nelle stagioni fredde (DJF per l'area europea e del Nord Pacifico, JJA per l'emisfero meridionale), come conseguenza di una maggiore dinamicità atmosferica.

D'altro canto, si nota invece come in Europa la previsione riguardo la temperatura in quota sia più incerta durante l'estate (JJA), con valori medi di RMSE più elevati. Nell'area del Nord Pacifico i valori più alti si riscontrano invece durante l'inverno (DJF). Dalla forma delle curve descrittive l'andamento dell'RMSE (molto pendenti nelle prime due settimane, quasi orizzontali nella terza e nella quarta) si può dedurre una volta in più il progredire dell'errore del modello, che si stabilizza una volta superato il limite di predicibilità, ovvero dopo aver perso memoria delle condizioni iniziali con cui è stato inizializzato.

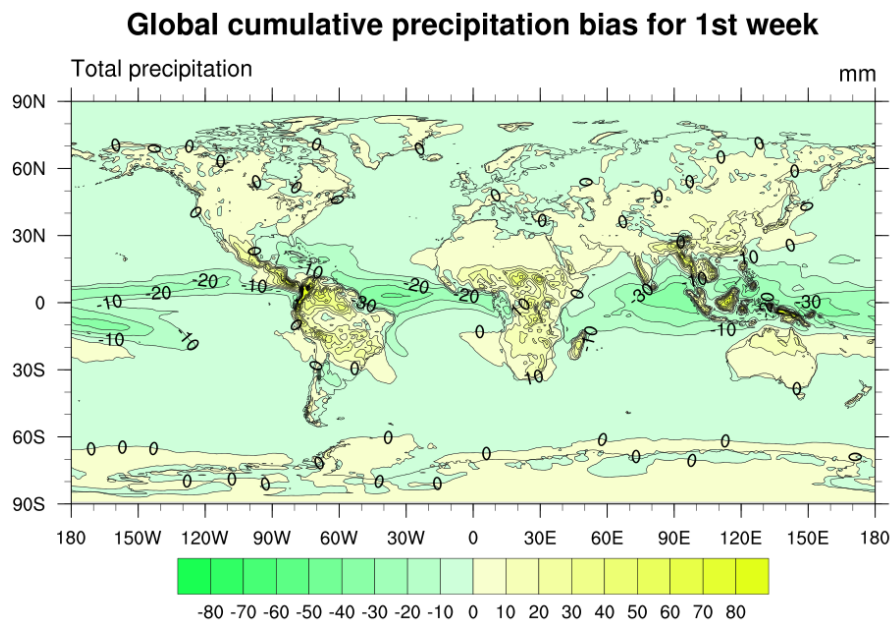


Figura 4.1: Bias globale medio annuale per le precipitazioni cumulate medie annuali, mediato sulla 1^a settimana

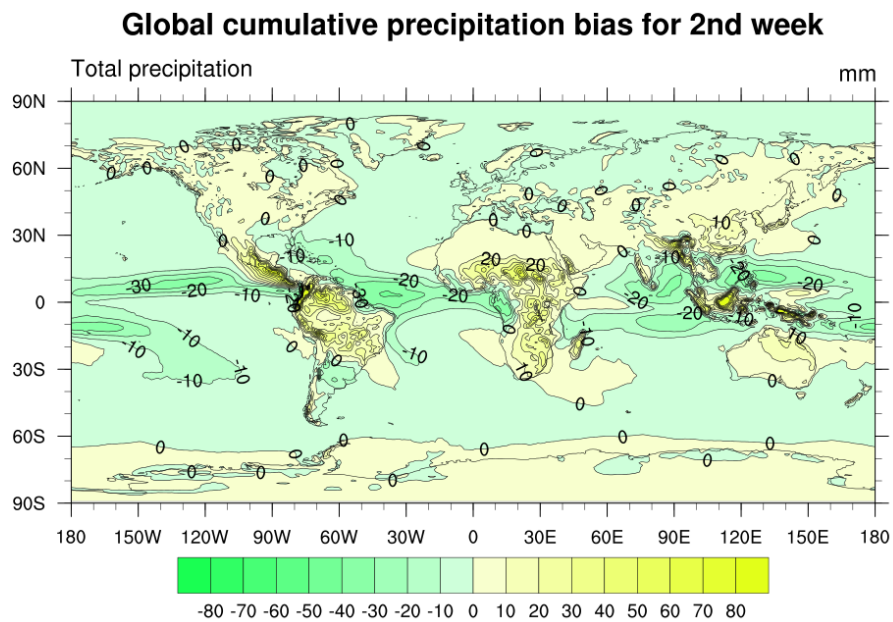


Figura 4.2: Bias globale medio annuale per le precipitazioni cumulate medie annuali, mediato sulla 2^a settimana

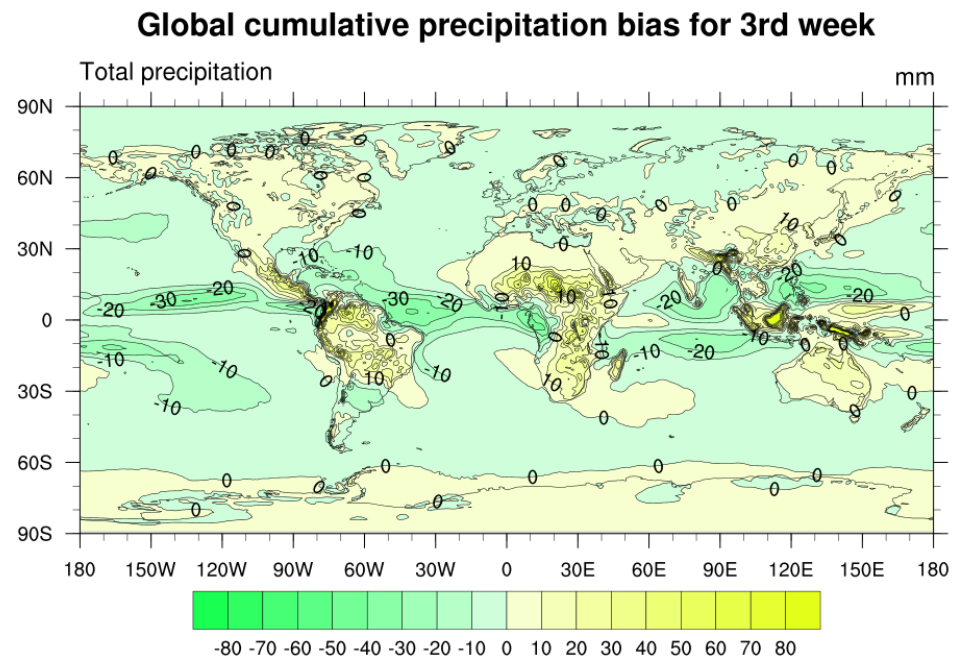


Figura 4.3: Bias globale medio annuale per le precipitazioni cumulate medie annuali, mediato sulla 3^a settimana

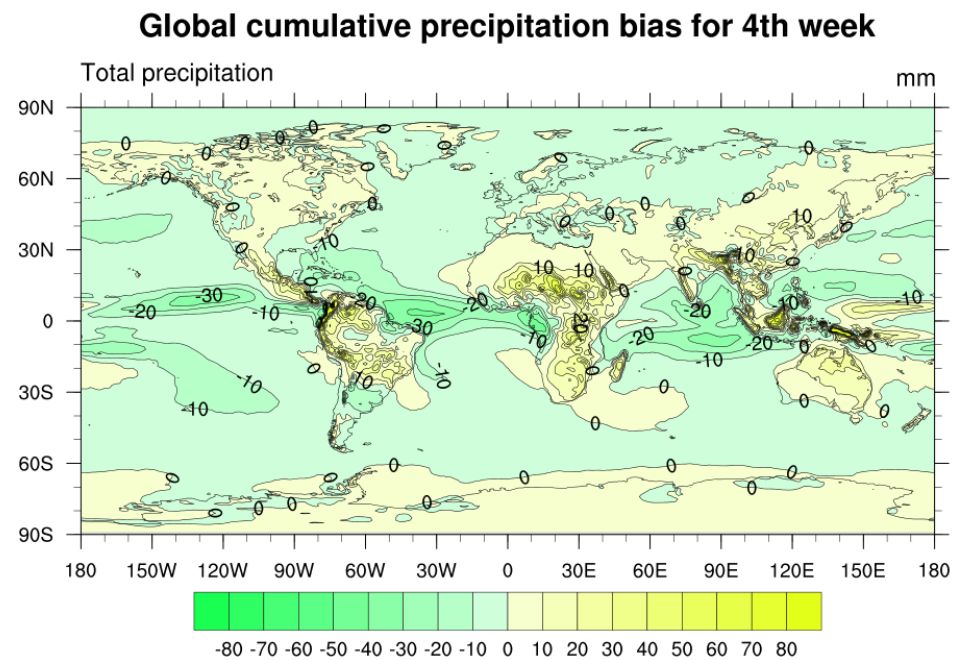


Figura 4.4: Bias globale medio annuale per le precipitazioni cumulate, mediato sulla 4^a settimana

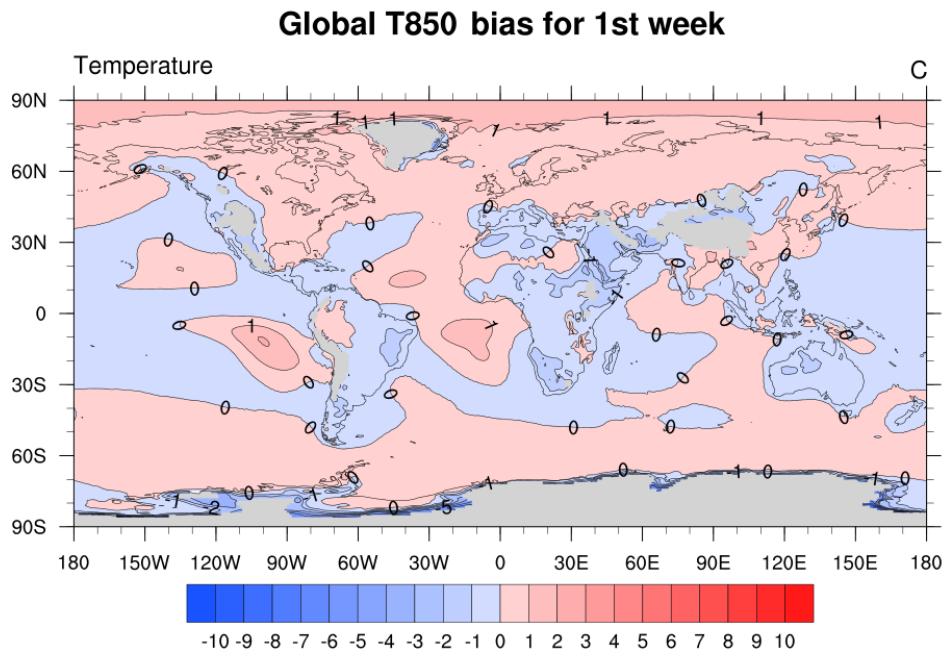


Figura 4.5: Bias globale medio annuale per la temperatura a 850 hPa, mediato sulla 1^a settimana

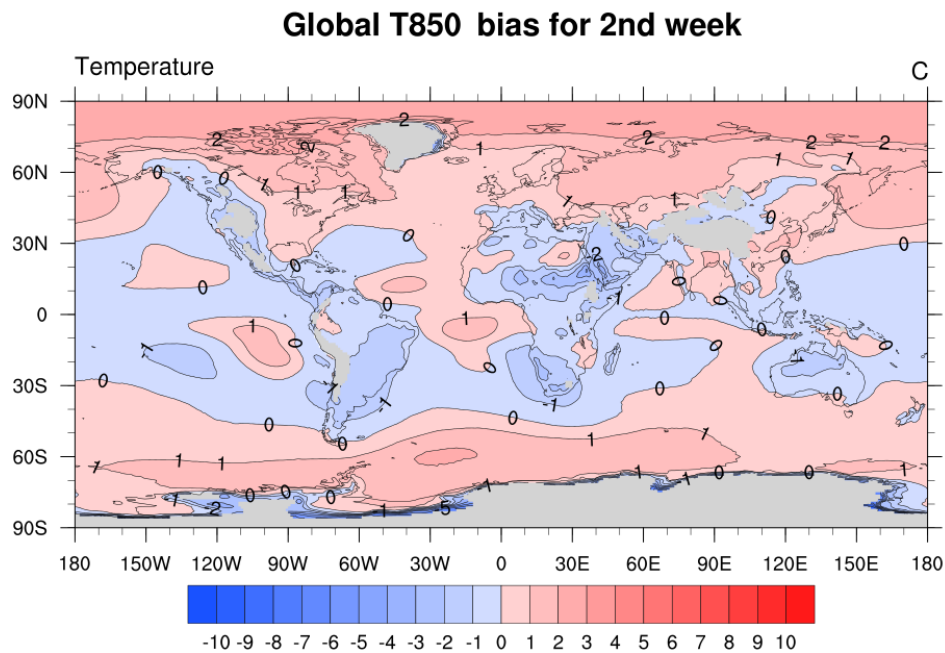


Figura 4.6: Bias globale medio annuale per la temperatura a 850 hPa, mediato sulla 2^a settimana

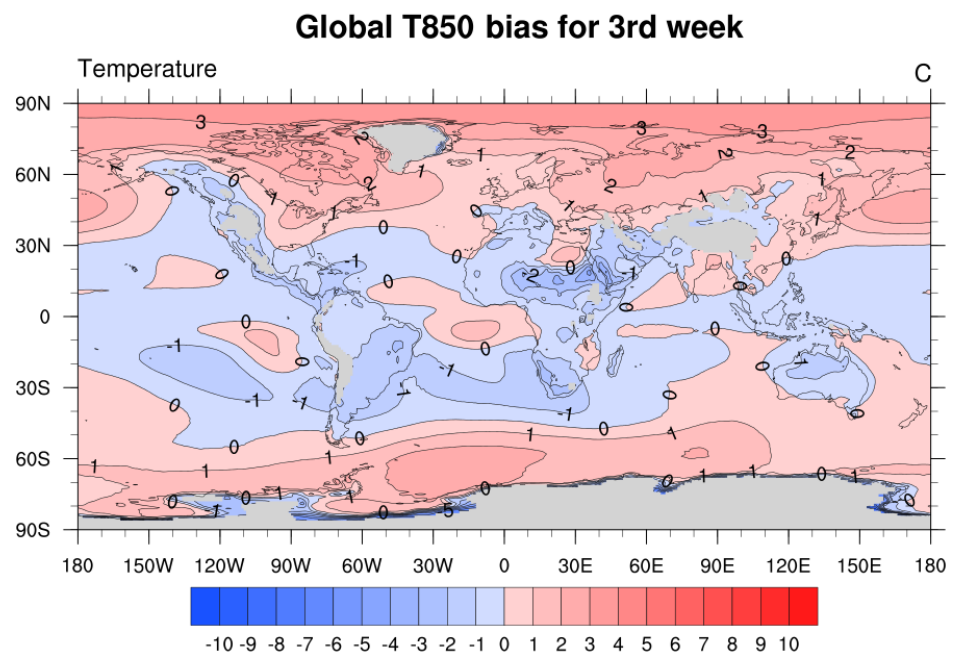


Figura 4.7: Bias globale medio annuale per la temperatura a 850 hPa, mediato sulla 3^a settimana

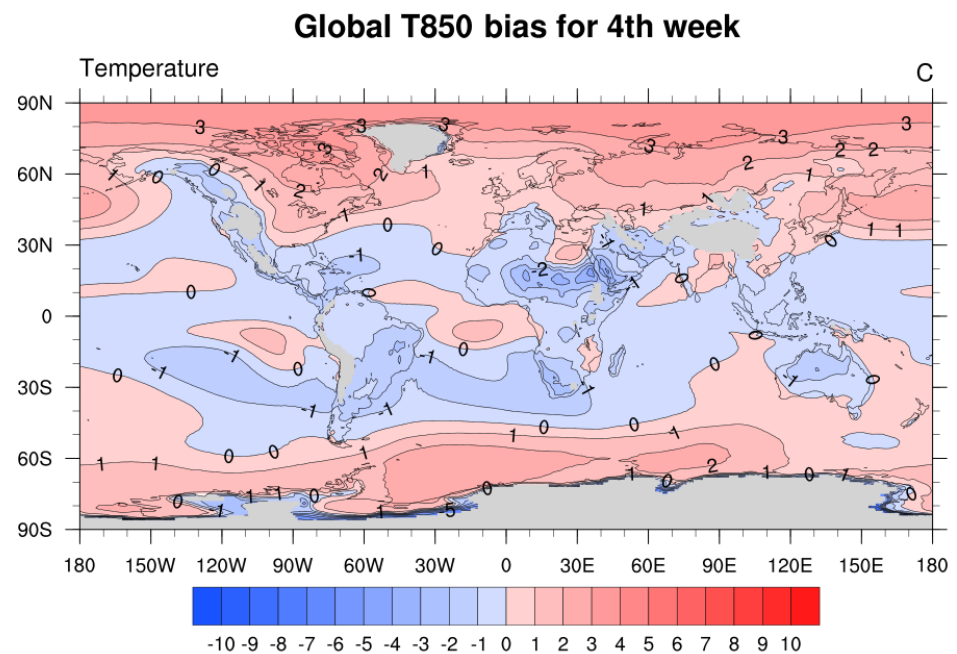


Figura 4.8: Bias globale medio annuale per la temperatura a 850 hPa, mediato sulla 4^a settimana

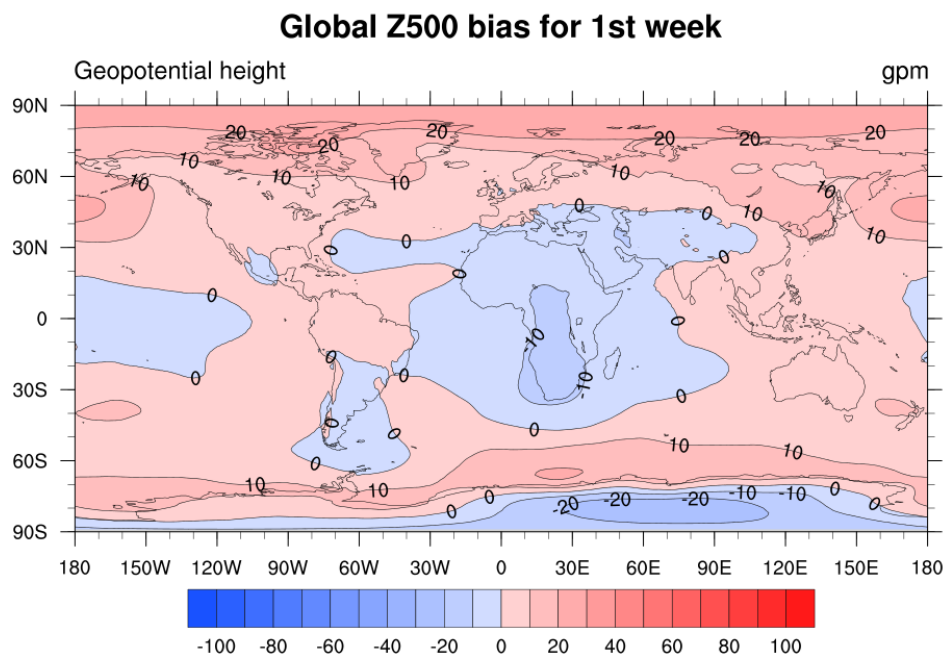


Figura 4.9: Bias globale medio annuale per l'altezza geopotenziale a 500 hPa, mediato sulla 1^a settimana

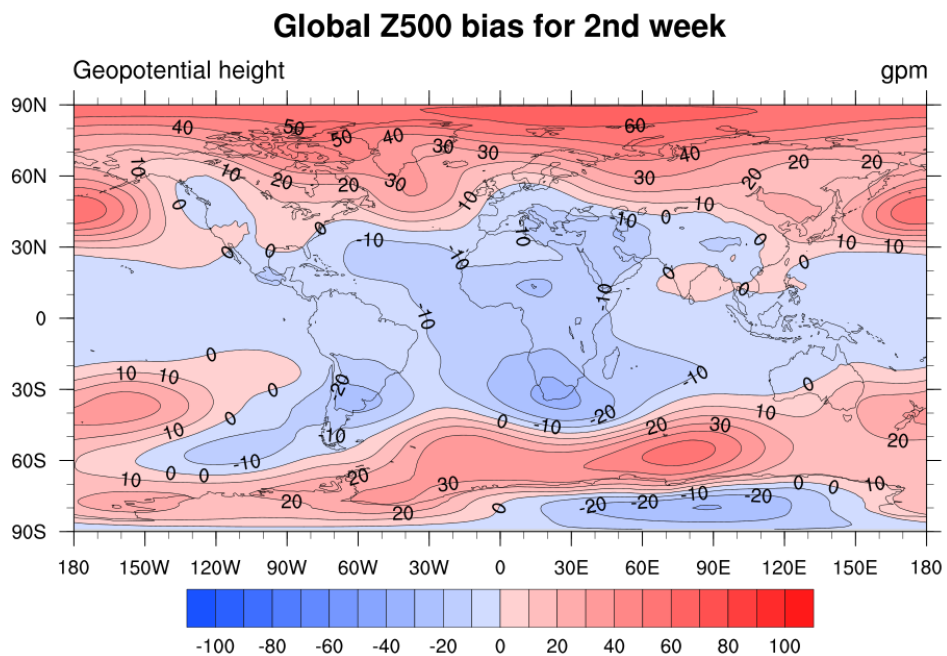


Figura 4.10: Bias globale medio annuale per l'altezza geopotenziale a 500 hPa, mediato sulla 2^a settimana

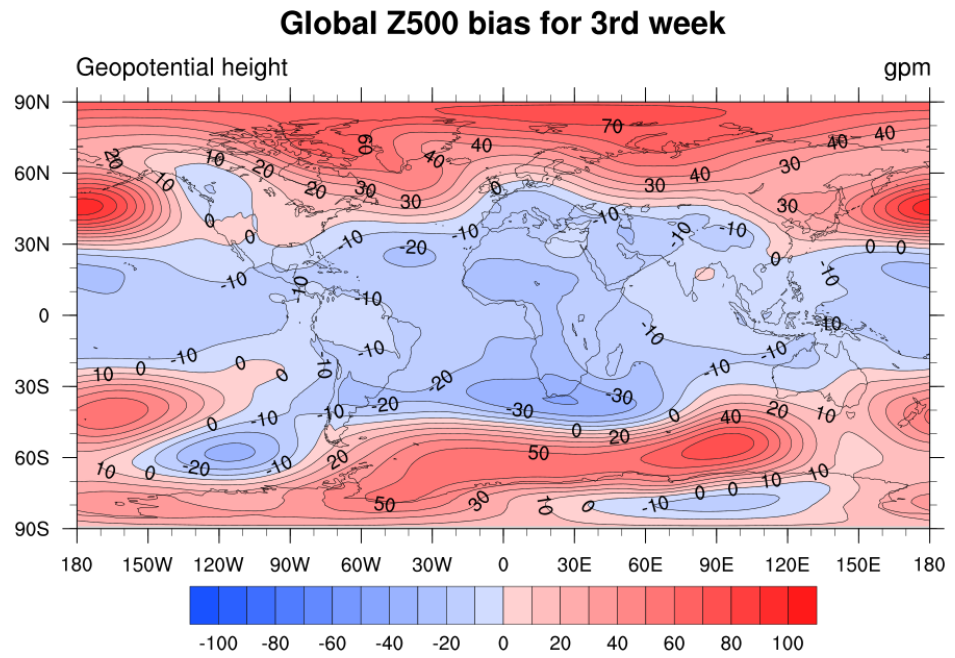


Figura 4.11: Bias globale medio annuale per l'altezza geopotenziale a 500 hPa, mediato sulla 3^a settimana

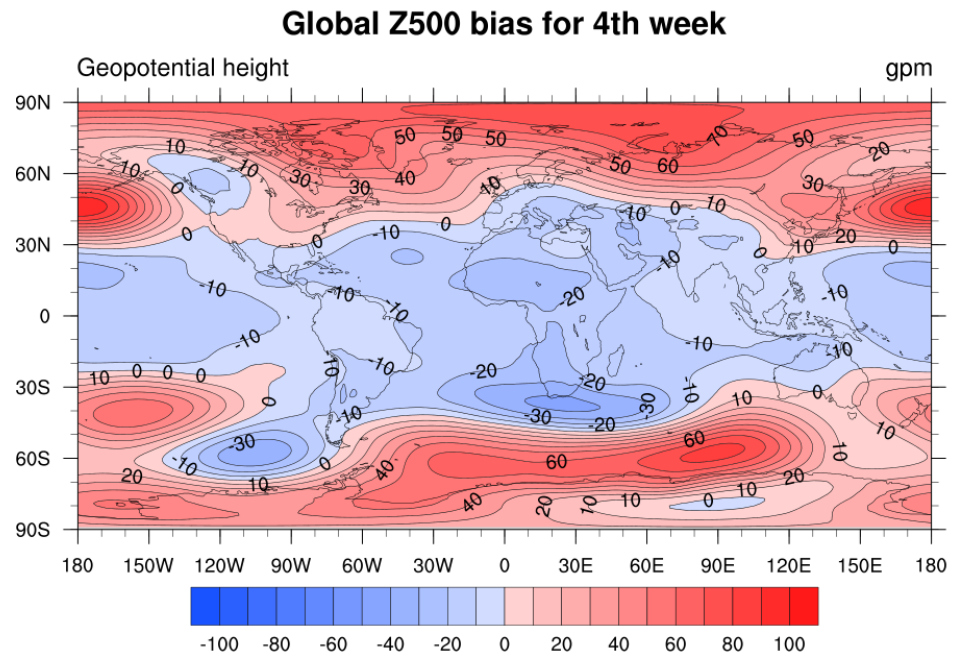


Figura 4.12: Bias globale medio annuale per l'altezza geopotenziale a 500 hPa, mediato sulla 4^a settimana

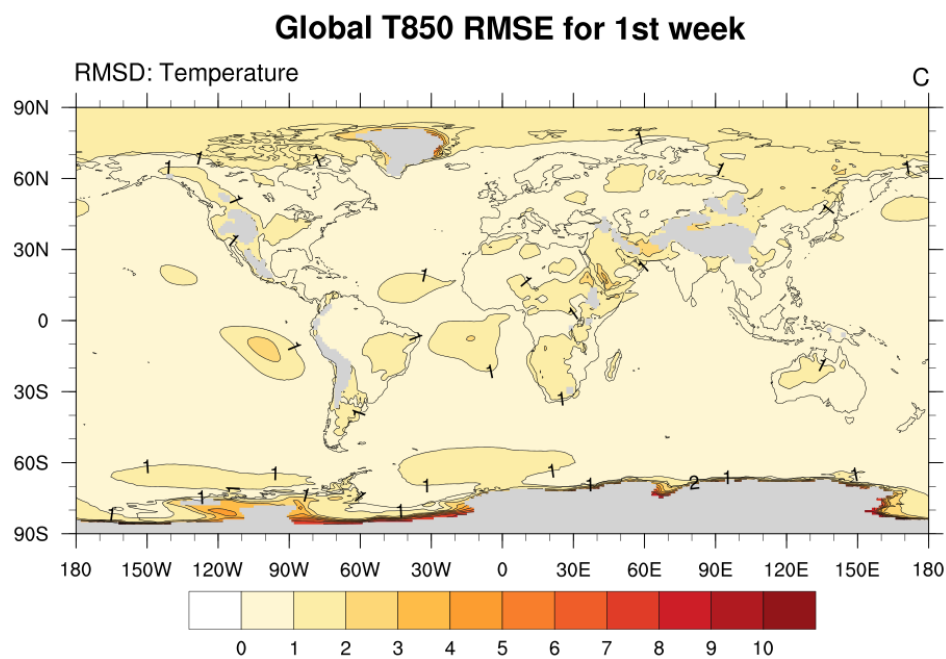


Figura 4.13: RMSE globale medio annuale per la temperatura a 850 hPa, mediato sulla 1^a settimana

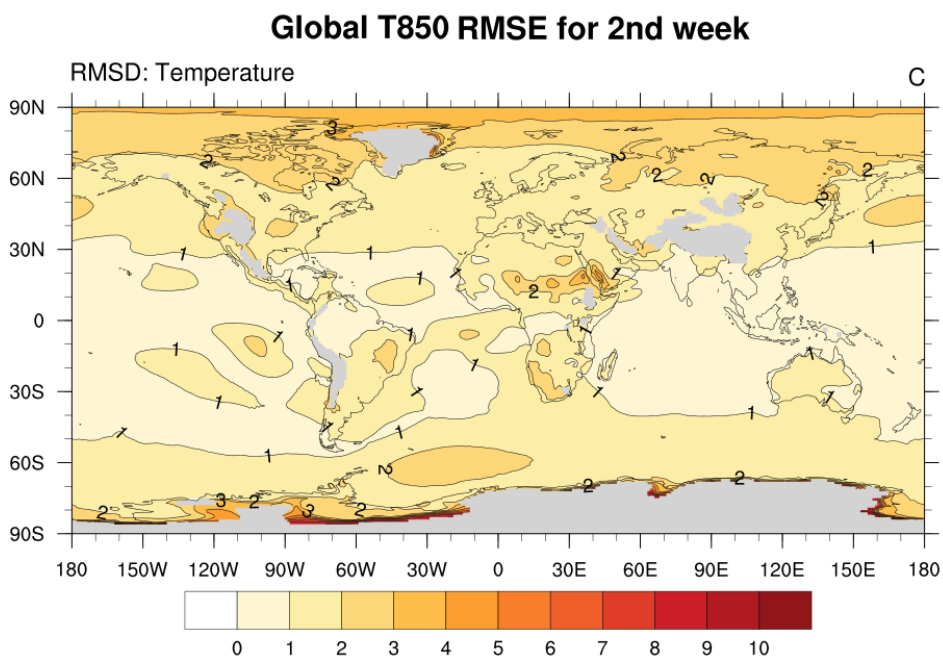


Figura 4.14: RMSE globale medio annuale per la temperatura a 850 hPa, mediato sulla 2^a settimana

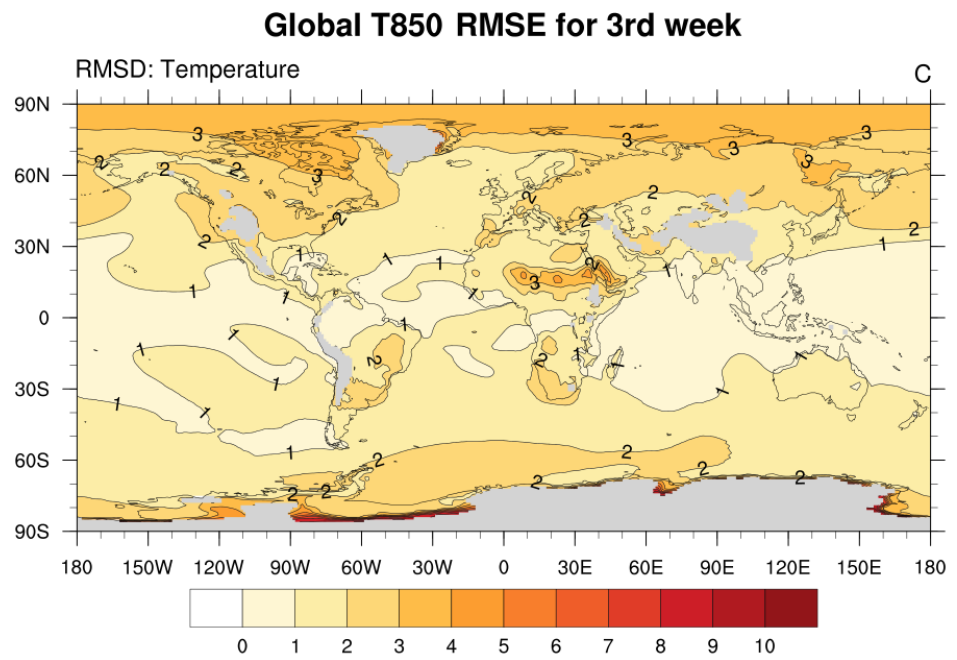


Figura 4.15: RMSE globale medio annuale per la temperatura a 850 hPa, mediato sulla 3^a settimana

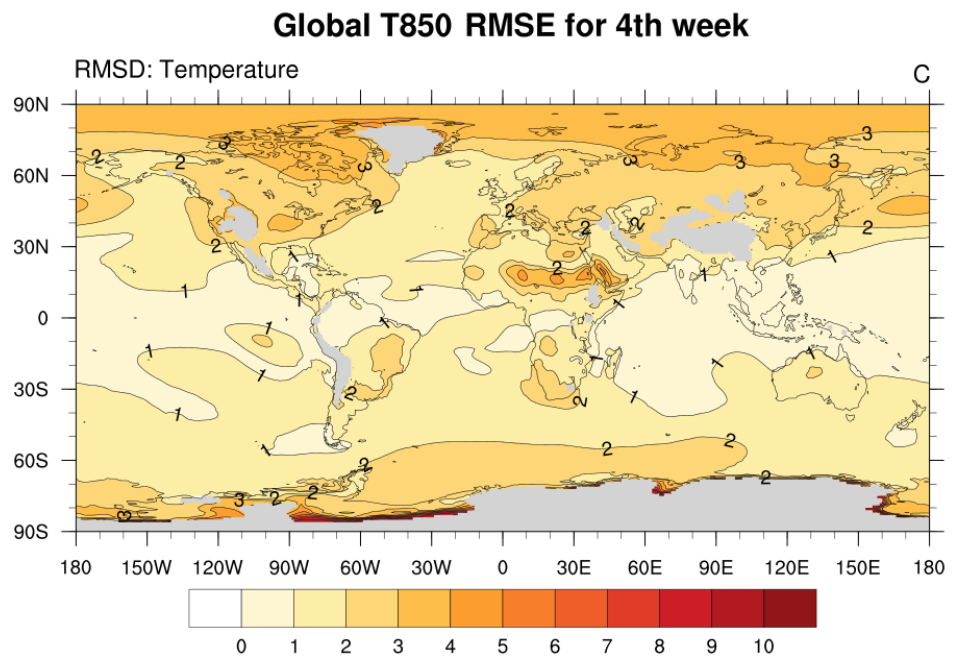


Figura 4.16: RMSE globale medio annuale per la temperatura a 850 hPa, mediato sulla 4^a settimana

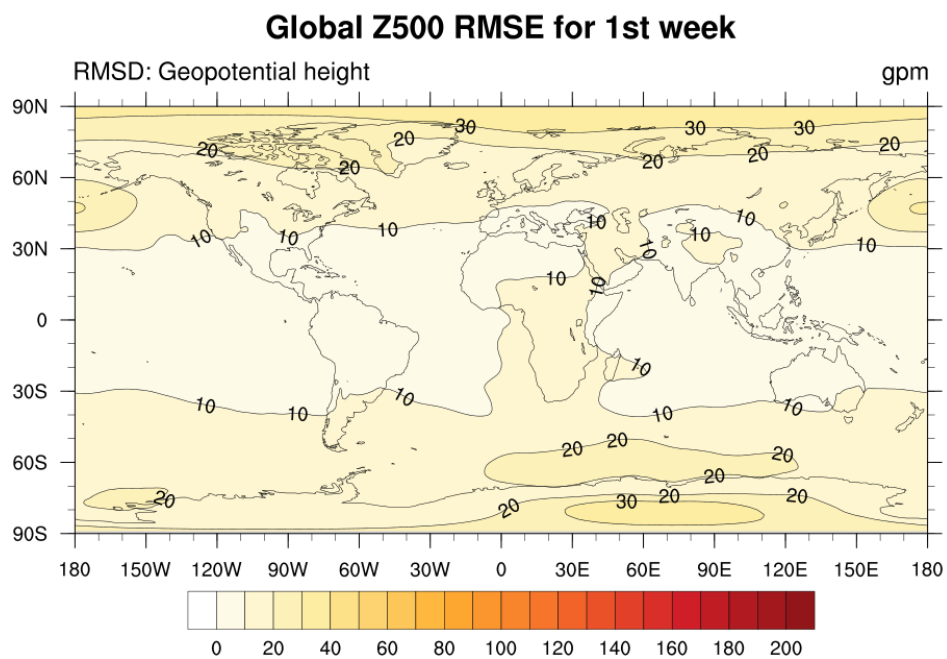


Figura 4.17: RMSE globale medio annuale per l'altezza geopotenziale a 500 hPa mediato sulla 1^a settimana

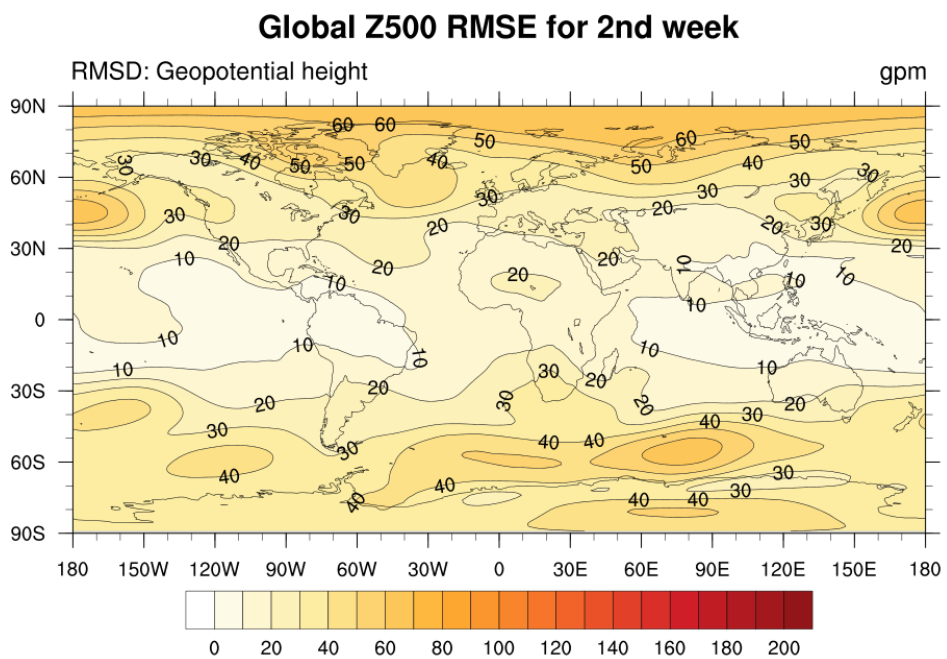


Figura 4.18: RMSE globale medio annuale per l'altezza geopotenziale a 500 hPa mediato sulla 2^a settimana

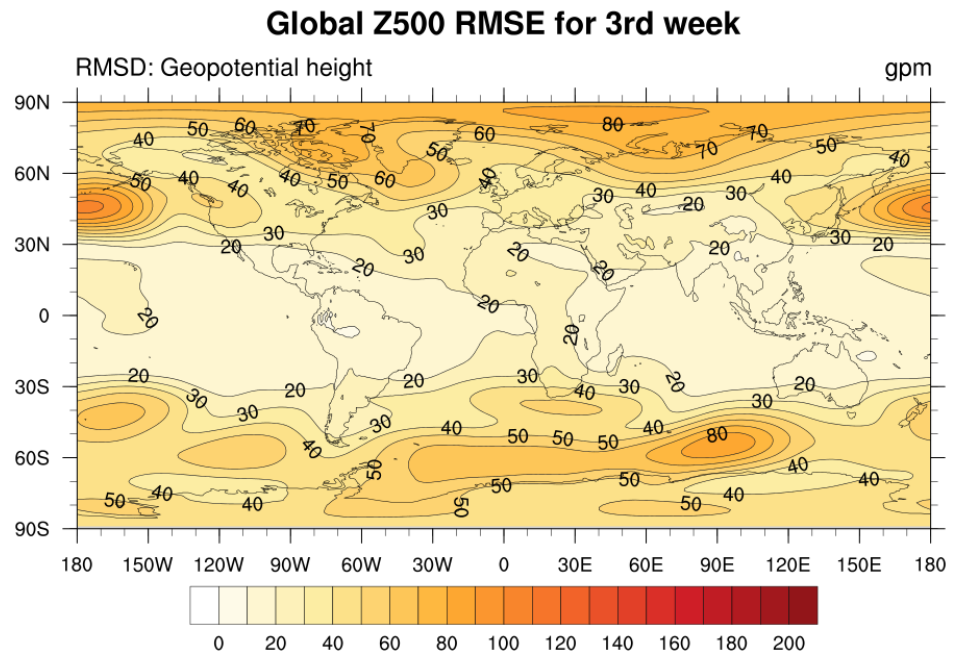


Figura 4.19: RMSE globale medio annuale per l'altezza geopotenziale a 500 hPa mediato sulla 3^a settimana

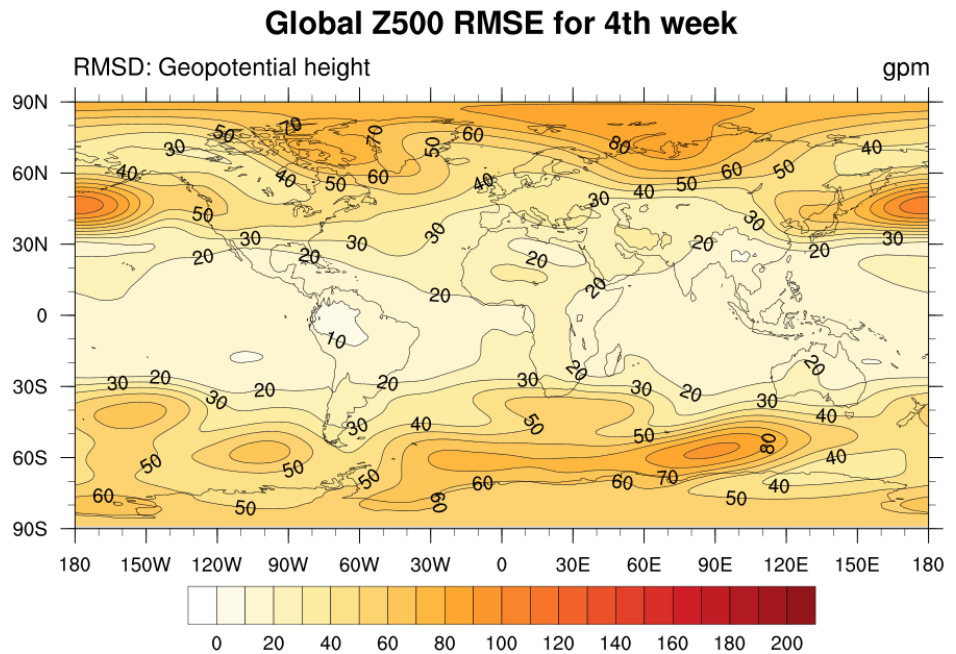


Figura 4.20: RMSE globale medio annuale per l'altezza geopotenziale a 500 hPa mediato sulla 4^a settimana

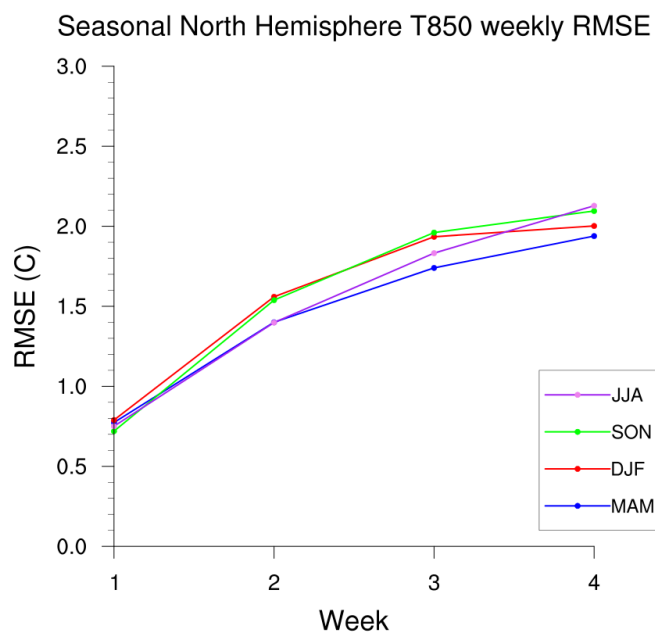


Figura 4.21: RMSE medio settimanale della temperatura a 850 hPa (C) mediato sull'emisfero settentrionale extra-tropicale per stagione di inizializzazione

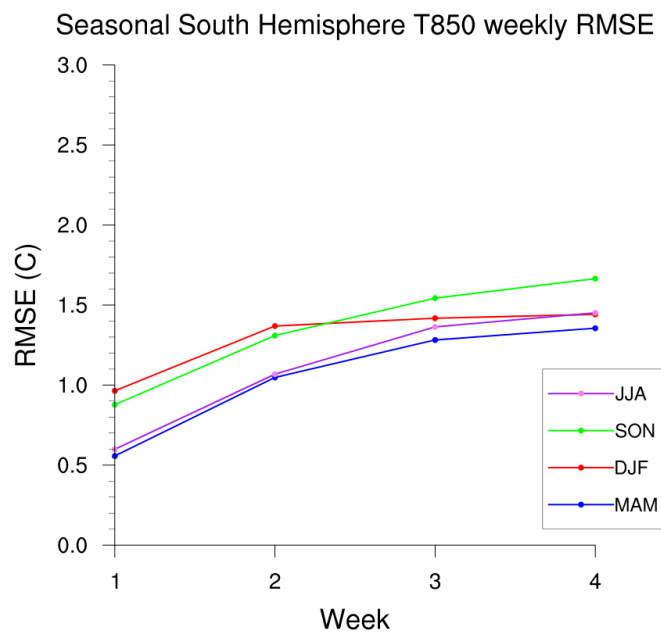


Figura 4.22: RMSE medio settimanale della temperatura a 850 hPa (C) mediato sull'emisfero meridionale extra-tropicale per stagione di inizializzazione

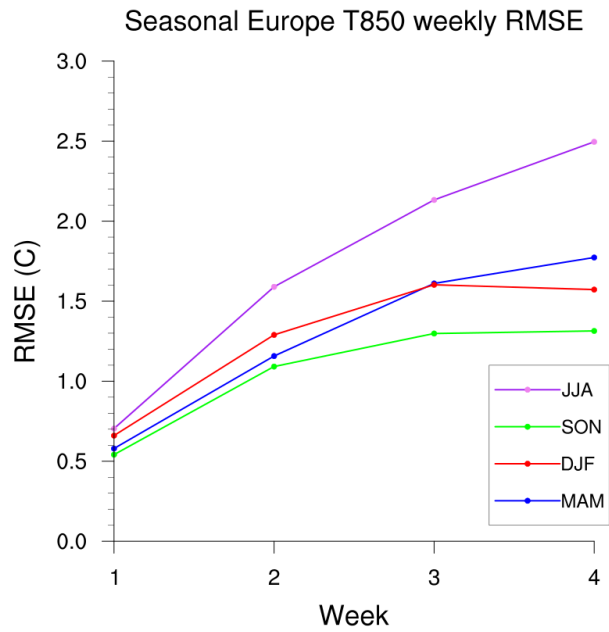


Figura 4.23: RMSE medio settimanale della temperatura a 850 hPa (C) mediato sull'area europea per stagione di inizializzazione

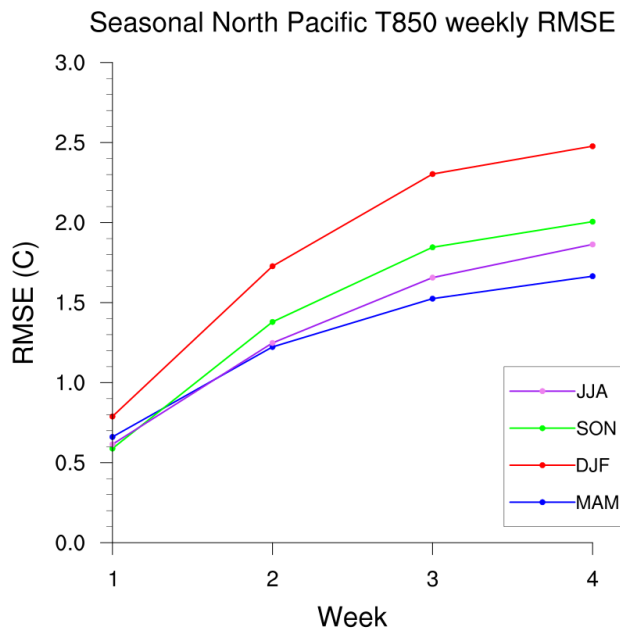


Figura 4.24: RMSE medio settimanale della temperatura a 850 hPa (C) mediato sull'area del Nord Pacifico per stagione di inizializzazione

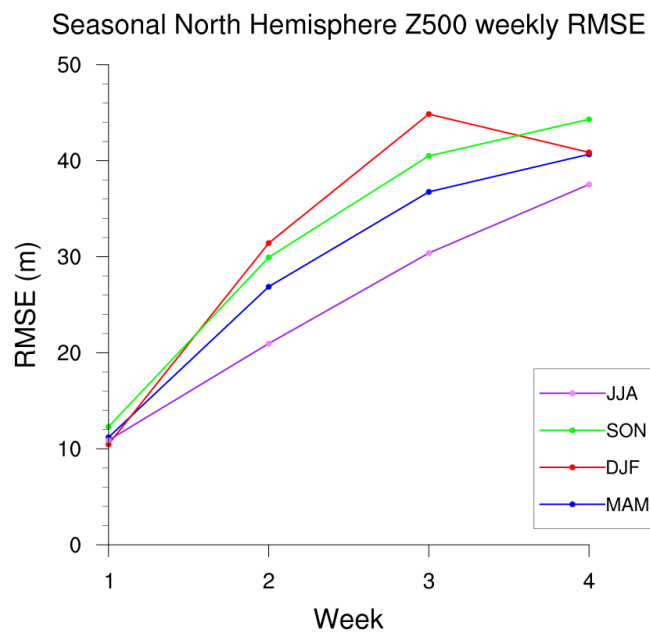


Figura 4.25: RMSE medio settimanale dell'altezza geopotenziale a 500 hPa (m) mediato sull'emisfero settentrionale extra-tropicale per stagione di inizializzazione.

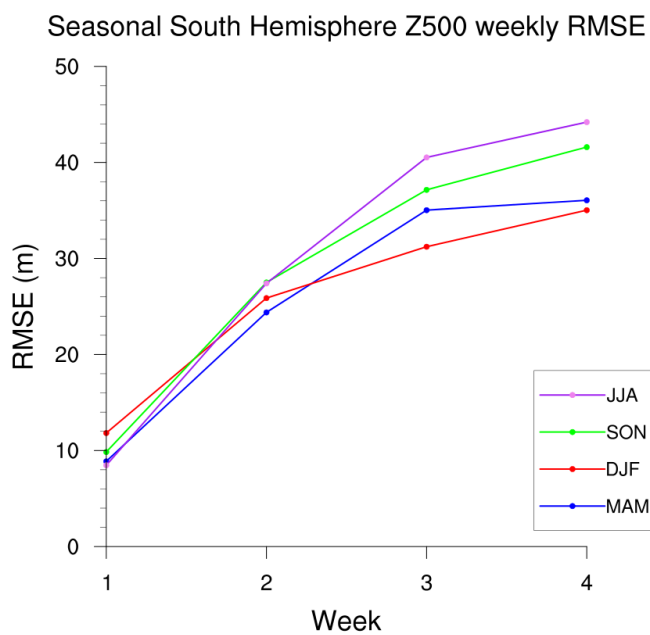


Figura 4.26: RMSE medio settimanale dell'altezza geopotenziale a 500 hPa (m) mediato sull'emisfero meridionale extra-tropicale per stagione di inizializzazione

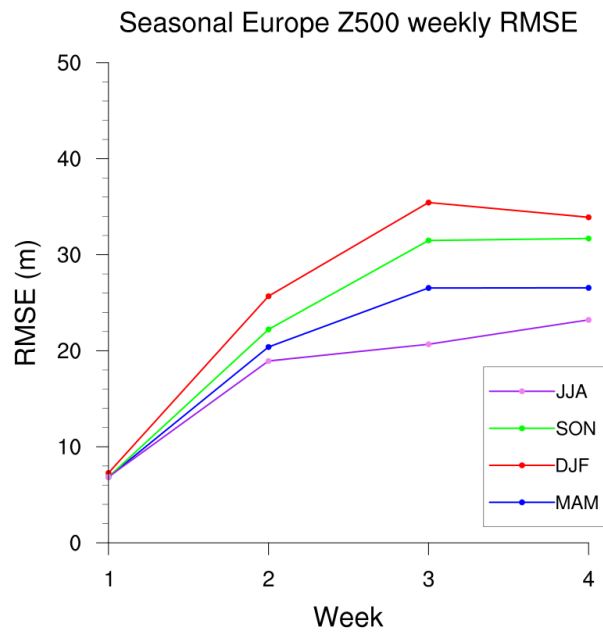


Figura 4.27: RMSE medio settimanale dell'altezza geopotenziale a 500 hPa (m) mediato sull'area europea per stagione di inizializzazione

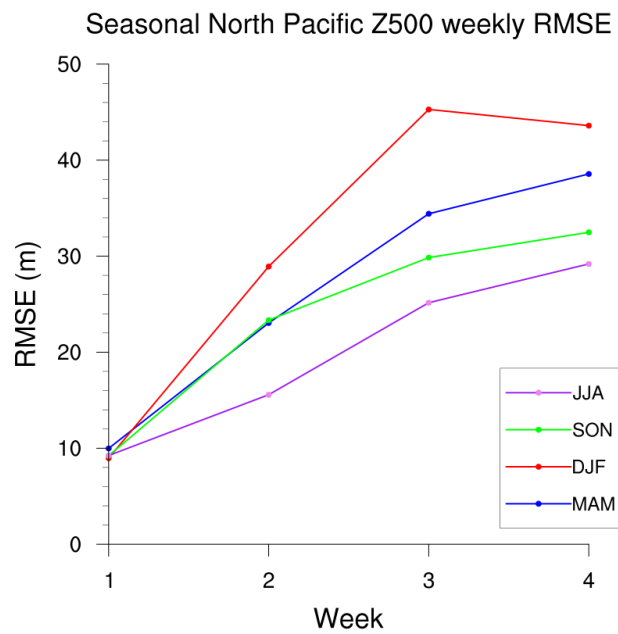


Figura 4.28: RMSE medio settimanale dell'altezza geopotenziale a 500 hPa (m) mediato sull'area del Nord Pacifico per stagione di inizializzazione

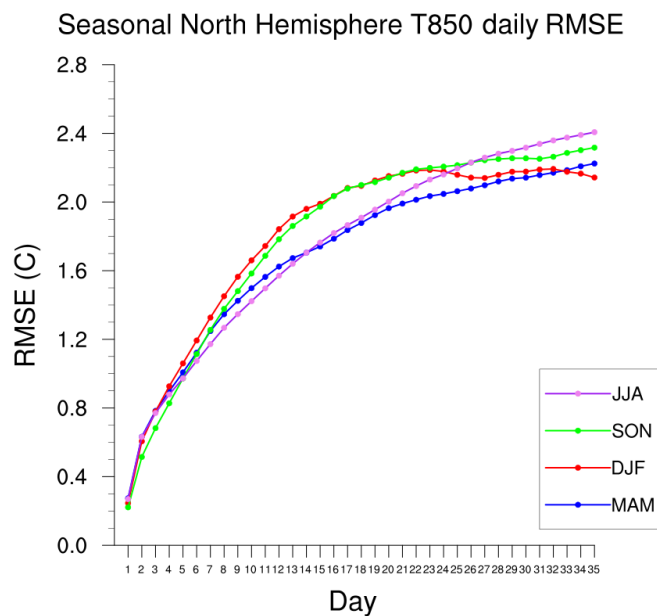


Figura 4.29: RMSE medio giornaliero della temperatura a 850 hPa (C) mediato sull'emisfero settentrionale per stagione di inizializzazione

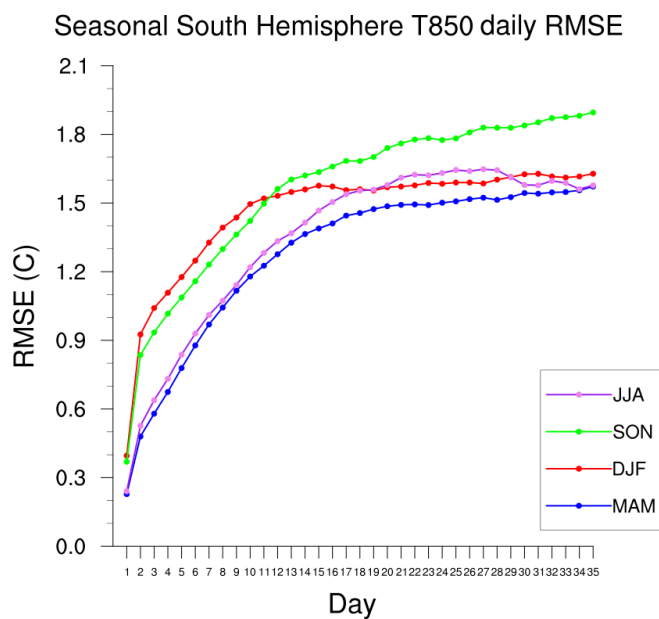


Figura 4.30: RMSE medio giornaliero della temperatura a 850 hPa (C) mediato sull'emisfero meridionale per stagione di inizializzazione.

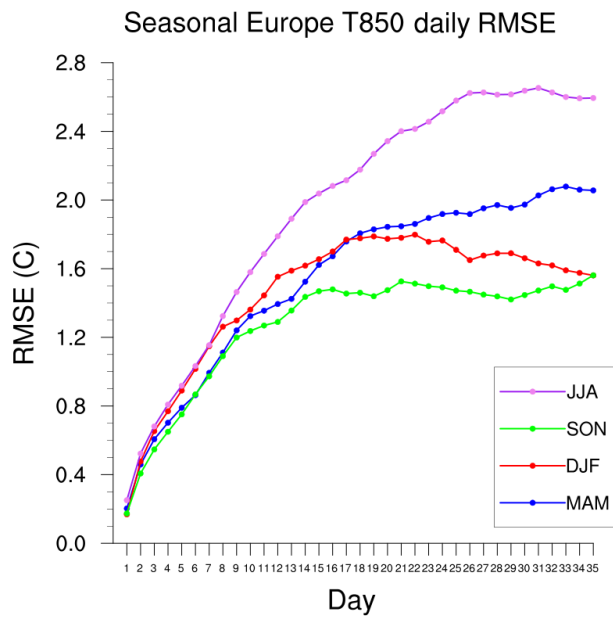


Figura 4.31: RMSE medio giornaliero della temperatura a 850 hPa (C) mediato sull'area europea per stagione di inizializzazione

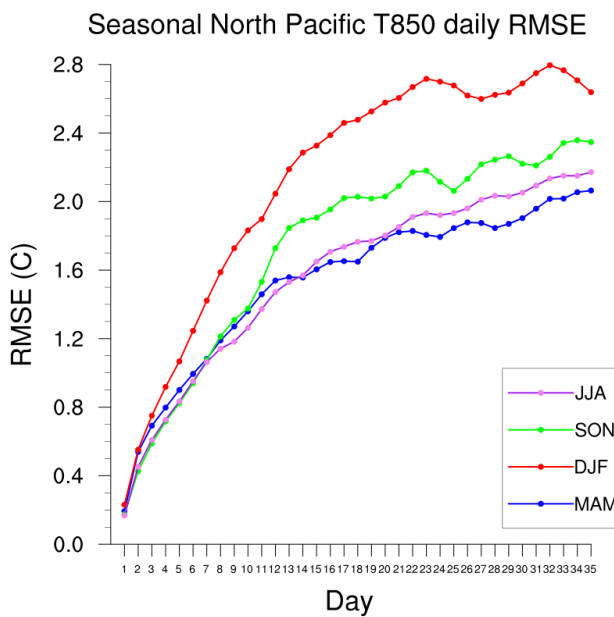


Figura 4.32: RMSE medio giornaliero della temperatura a 850 hPa (C) mediato sull'area del Nord Pacifico per stagione di inizializzazione

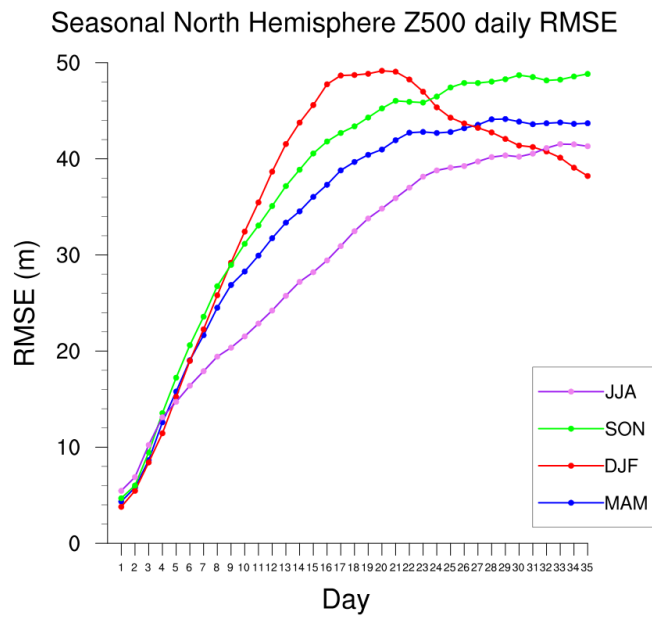


Figura 4.33: RMSE medio giornaliero dell'altezza geopotenziale a 500 hPa (m) mediato sull'emisfero settentrionale per stagione di inizializzazione

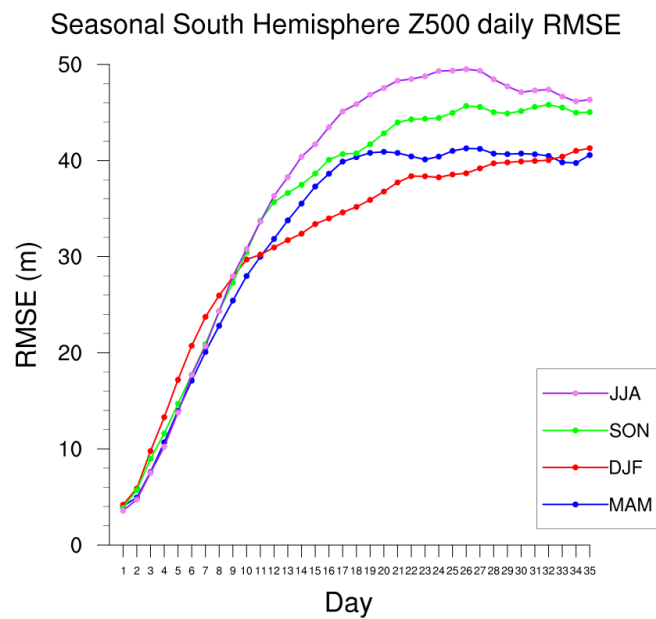


Figura 4.34: RMSE medio giornaliero dell'altezza geopotenziale a 500 hPa (m) mediato sull'emisfero meridionale per stagione di inizializzazione

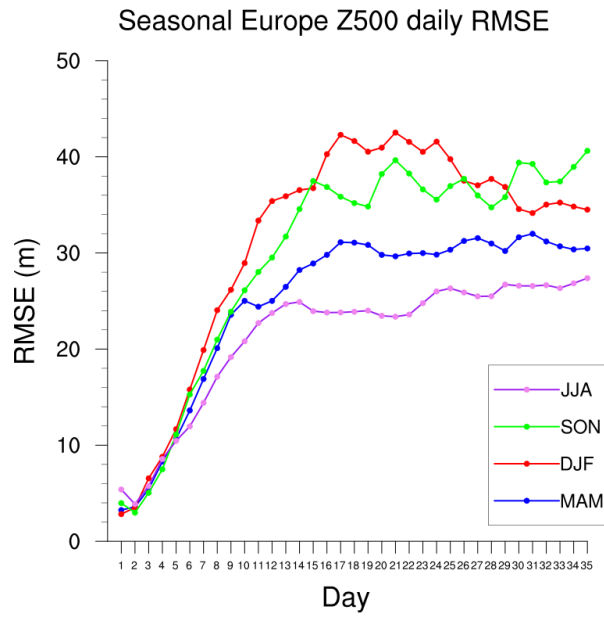


Figura 4.35: RMSE medio giornaliero dell'altezza geopotenziale a 500 hPa (m) mediato sull'area europea per stagione di inizializzazione

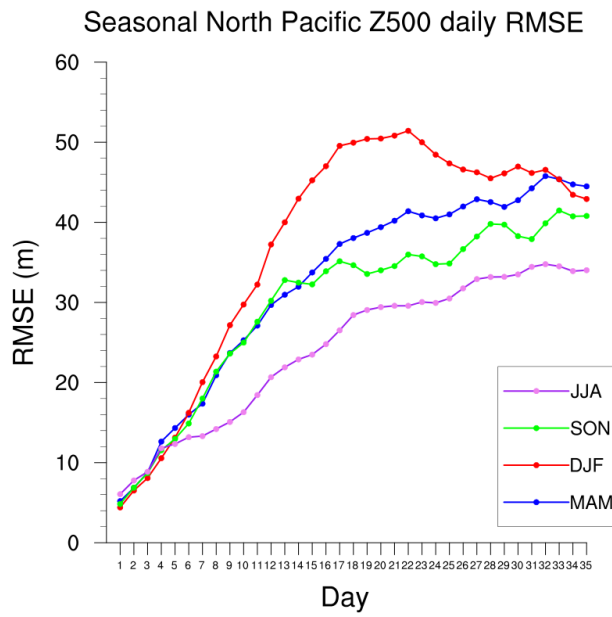


Figura 4.36: RMSE medio giornaliero dell'altezza geopotenziale a 500 hPa (m) mediato sull'area del Nord Pacifico per stagione di inizializzazione

5 Conclusioni

Le previsioni mensili richiedono sistemi previsionali costituiti, oltre che dalle comuni previsioni sullo stato dell'atmosfera futuro (i forecast), anche da una diversa classe di previsioni che sono inizializzate e si esauriscono nel passato: i reforecast. Questi ultimi sono prodotti allo scopo di essere confrontati con il reale comportamento dell'atmosfera, solitamente ricreato attraverso le reanalisi (cfr. §2.3.2). Dal confronto tra tempo previsto ed osservato è possibile studiare l'incidenza dell'errore sistematico di un modello, nonché le sue performance o il suo comportamento in alcuni casi particolari (come ad esempio specifiche condizioni iniziali).

A tal proposito, un'analisi del bias e dell'RMSE del modello GLOBO – utilizzato dal 2010 presso il CNR-ISAC per la formulazione di previsioni mensili – è stata effettuata grazie al confronto tra le reanalisi (ERA-Interim) del Centro Europeo per le Previsioni Meteorologiche a Breve e Medio Termine (ECMWF) e una serie di 2190 simulazioni di reforecast. Da queste (ciascuna della durata di 35 giorni, lanciate singolarmente a distanza di cinque giorni l'una dall'altra: dunque non di tipo probabilistico o ensemble) è stato poi ricavato il clima trentennale (1981-2010) del modello. L'applicazione delle medie temporali ha permesso di mettere in risalto i pattern su larga scala degli score, che andavano a stabilizzarsi nella seconda parte del periodo di integrazione numerica (terza e quarta settimana). Ciò rispecchia la componente sistematica dell'errore compiuto dal modello: componente che, una volta nota, è possibile rimuovere per ottenere performance migliori nella fase di forecast.

Il bias delle precipitazioni cumulate mostra importanti sovrastime nelle aree equatoriali continentali (ovvero in corrispondenza delle aree coperte da foreste pluviali) e sottostime al di sopra degli oceani e dei mari delle stesse latitudini. Il bias della temperature in quota (T850) e dell'altezza geopotenziale (Z500) presenta estese anomalie positive alle alte e medie latitudini dell'emisfero settentrionale, con massimi su Siberia, Canada e Pacifico Settentrionale, nonché a ridosso dell'Antartide. Le anomalie negative sono concentrate sulle zone tropicali e una anomalia negativa particolarmente intensa che riguarda soprattutto la T850 interessa l'area dell'Africa tropicale settentrionale.

L'RMSE ricalca a grandi linee le anomalie del bias, a causa della stretta correlazione tra i due score: le zone polari e sub-polari di entrambi gli emisferi, il Nord Pacifico e (specie per quel che riguarda la T850) l'area a cavallo tra Africa sub-sahariana e penisola arabica, si confermano regioni di particolare difficoltà per la previsione mensile da parte di GLOBO. Lo studio dell'andamento dell'RMSE in funzione del tempo di integrazione (giornaliero e settimanale) mediato su alcune macro-aree (Emisfero nord e sud, Europa, Nord Pacifico) e riferito alle varie stagioni (DJF, MAM, JJA, SON) ha permesso di evidenziare la variazione delle anomalie nell'arco dell'anno solare: ad esempio il semestre freddo (nei rispettivi emisferi) risulta essere quello con i valori di RMSE di Z500 più elevati a causa della maggiore dinamicità atmosferica.

Lo studio dei processi all'origine di bias e RMSE qui mostrati richiederebbe una trattazione più approfondita che, tuttavia, non rientra negli obiettivi del presente lavoro.

Bibliografia

- [1] J.M.Murphy e A.Dickinson. «Extended-range prediction experiments using an 11-level GCM». In: *Meteorology and Atmospheric Physics* 40 (1989), pp. 61–83.
- [2] A.P.Weigel et al. «Probabilistic verification of monthly temperature forecasts». In: *Monthly Weather Review* 136 (dic. 2008), pp. 5162–5182.
- [3] D.Mastrangelo et al. «First outcomes from the CNR-ISAC monthly forecasting system». In: *Advances in Science & Research*. 8. European Meteorological Society. Copernicus Publications, 2012, pp. 77–82.
- [4] K.Miyakoda et al. «One-month forecast experiment without anomaly boundary forcings». In: *Monthly Weather Review* 114 (dic. 1986), pp. 2363–2401.
- [5] K.Miyakoda et al. «Simulation of a blocking event of January 1977». In: *Monthly Weather Review* 111 (apr. 1983), pp. 846–869.
- [6] L.Ferranti et al. «Flow-dependent verification of the ECMWF ensemble over the Euro-Atlantic sector». In: *Q.J.R. Meteorol. Soc.* (2014).
- [7] P.Malguzzi et al. «The Meteorological Global Model GLOBO at the ISAC-CNR of Italy Assessment of 1.5 Yr of Experimental Use for Medium-Range Weather Forecasts». In: *Weather and Forecasting* 26 (dic. 2011), pp. 1045–1055.
- [8] T.Hamill et al. «Ensemble Reforecasting: Improving Medium-Range Forecast Skill Using Retrospective Forecasts». In: *Monthly Weather Review* 132 (giu. 2004), pp. 1434–1447.
- [9] E.N.Lorenz. «Climatic predictability». In: *The Physical Basis of Climate and Climate Modelling*. A cura di GARP Publication Series. Vol. 16. World Meteorological Organisation, 1975, pp. 132–136.
- [10] E.N.Lorenz. «Deterministic nonperiodic flow». In: *Journal of Atmospheric Sciences* 20 (mar. 1963), pp. 130–141.
- [11] G.J.Boer. «Systematic and random errors in an extended-range forecasting experiment». In: *Monthly Weather Review* 121 (gen. 1993), pp. 173–188.

- [12] M.S.J. Harrison. «Long-range forecasting since 1980: empirical and numerical prediction out to one month for the United Kingdom». In: *Weather* 50.12 (dic. 1995), pp. 440–449.
- [13] M.Déque e J.F.Royer. «The skill of extended-range extratropical winter dynamical forecasts». In: *Journal of climate* 5 (nov. 1992), pp. 1346–1356.
- [14] J.Shukla. «Dynamical predicatability of monthly means». In: *Journal of the atmospheric sciences* 38.12 (dic. 1981), pp. 2547–2572.
- [15] C.E. Leith. «Theoretical skill of Monte Carlo forecasts». In: *Monthly Weather Review* 102.60 (giu. 1974), pp. 409–418.
- [16] World Meteorological Organisation. *Guidelines on Ensemble Prediction Systems and Forecasting*. 2012.
- [17] Proceedings of science, cur. *Porting workflows based on small and medium parallelism applications to the Italian Grid Infrastructure*. IGI. Mar. 2013.
- [18] Frédéric Vitart. «Evolution of ECMWF sub-seasonal forecast skill score». In: *Q.J.R. Meteorological Society* (2014).
- [19] Frédéric Vitart. «Monthly forecasting at ECMWF». In: *Monthly Weather Review* 132 (dic. 2004), pp. 2761–2779.
- [20] Daniel S. Wilks. *Statistical methods in the atmospheric sciences*. A cura di ELSEVIER. 3^a ed. Academic Press, 2011.

Appendice

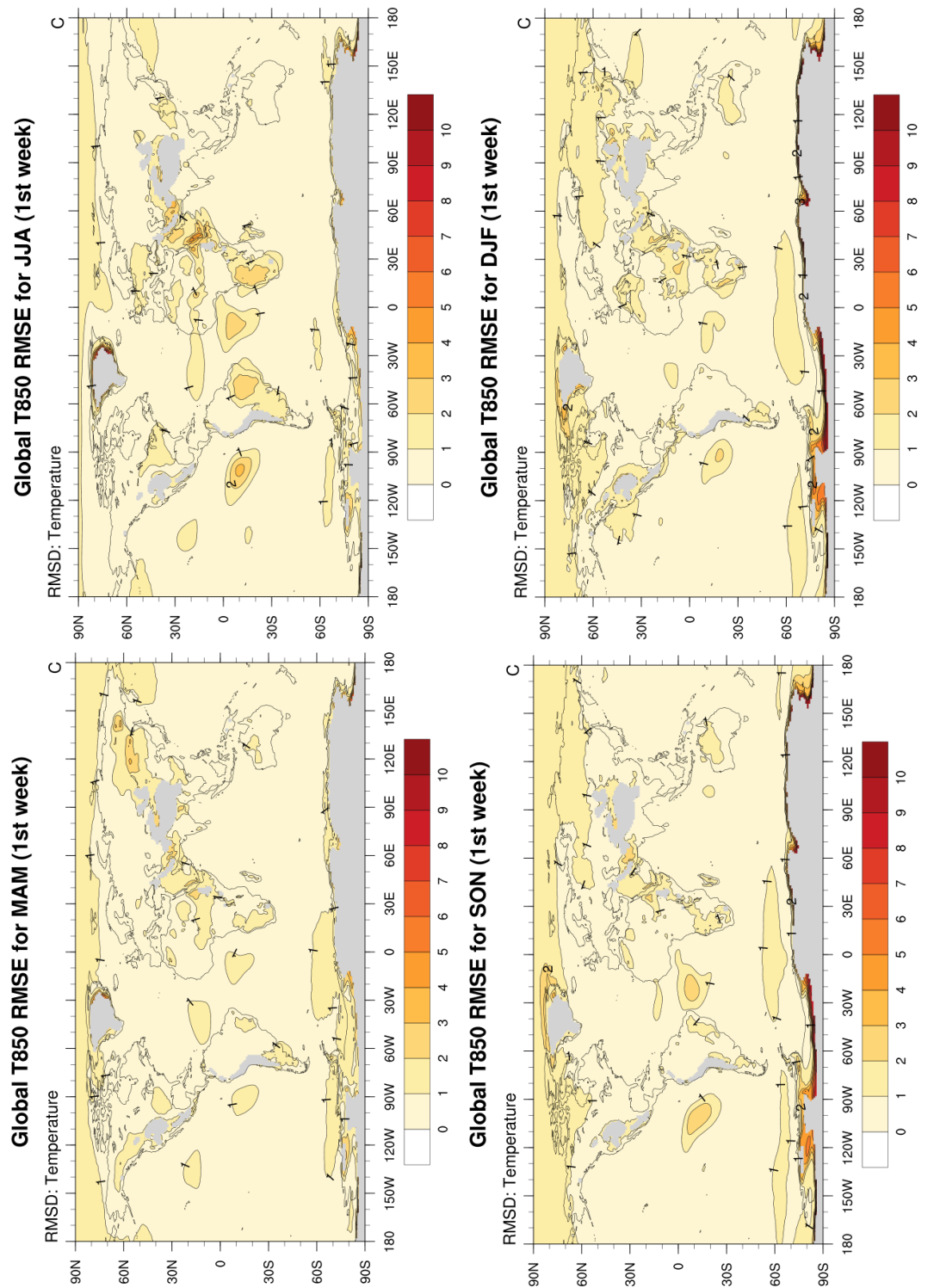


Figura 5.1: RMSE globale medio stagionale per la temperatura a 850 hPa mediato sulla 1^a settimana

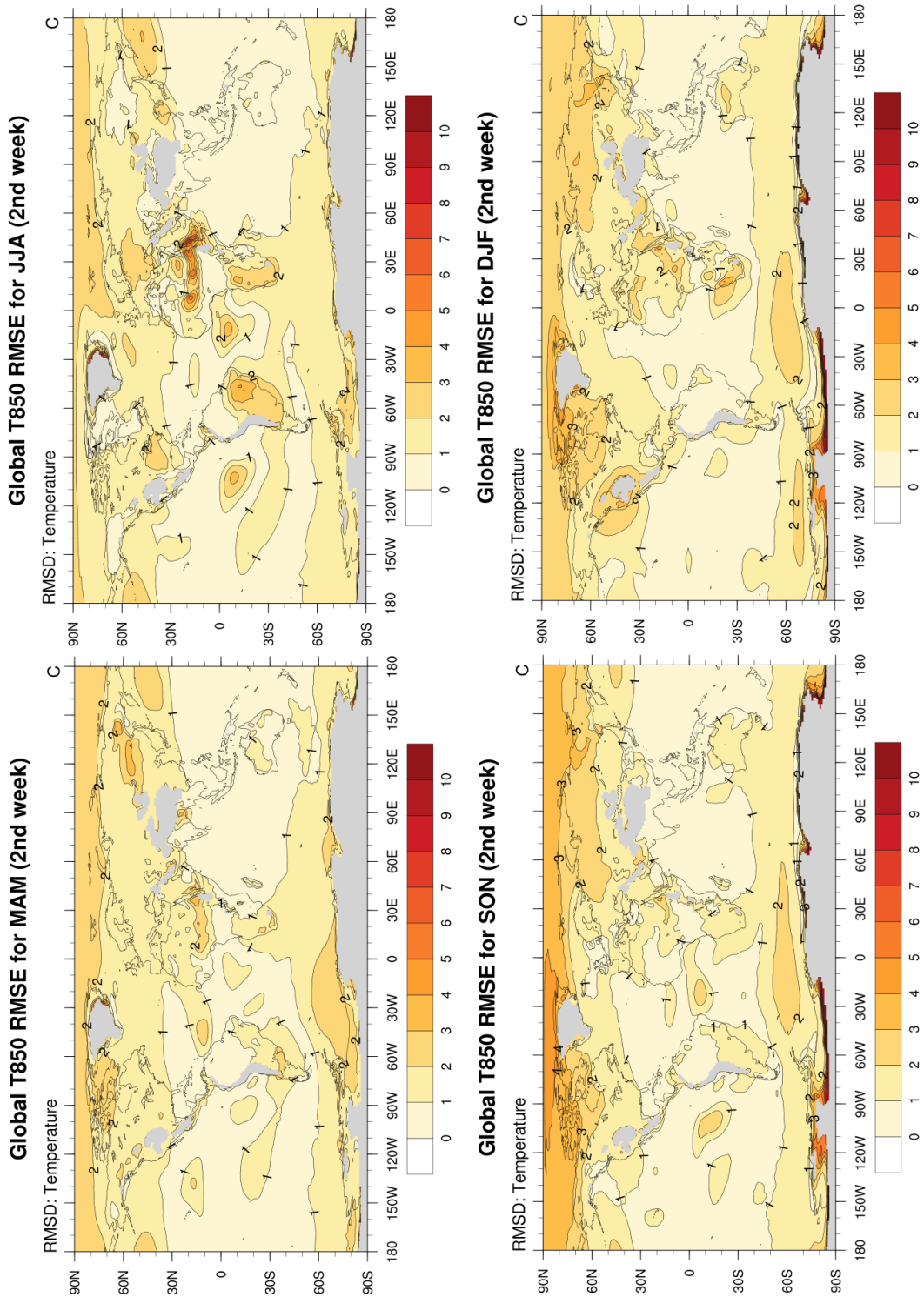


Figura 5.2: RMSE globale medio stagionale per la temperatura a 850 hPa mediato sulla 2^a settimana

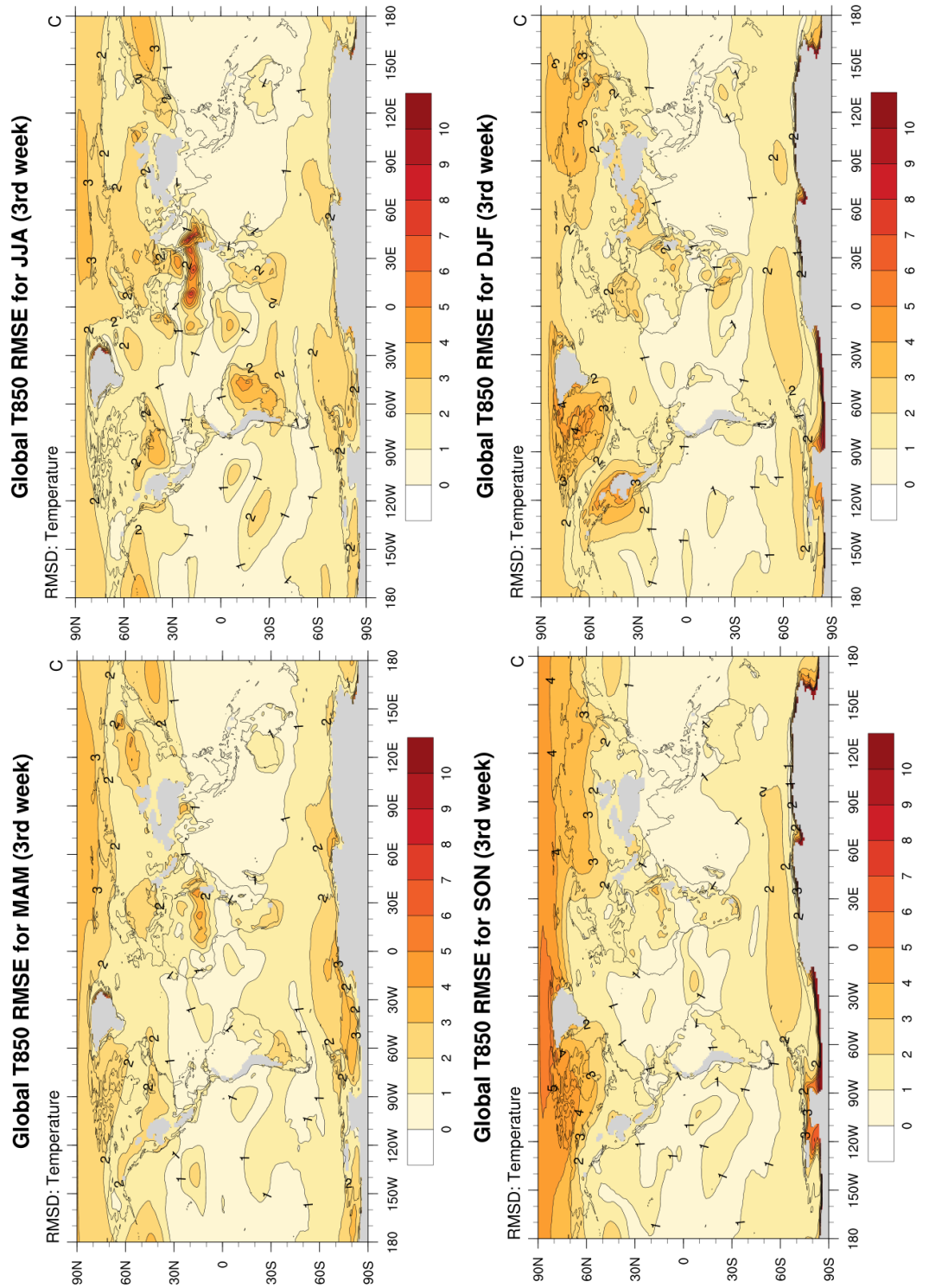


Figura 5.3: RMSE globale medio stagionale per la temperatura a 850 hPa mediato sulla 3^a settimana

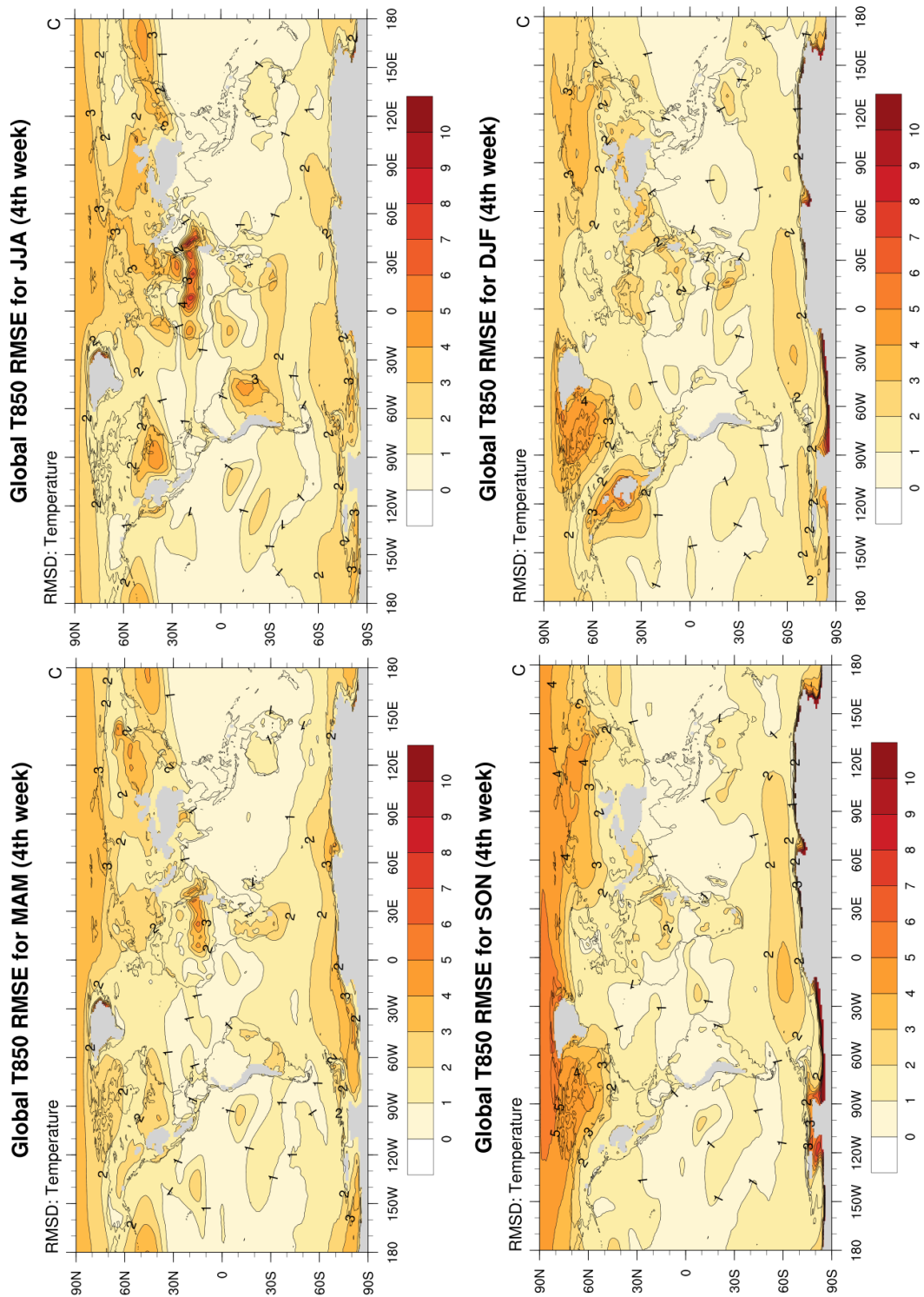


Figura 5.4: RMSE globale medio stagionale per la temperatura a 850 hPa mediato sulla 4^a settimana

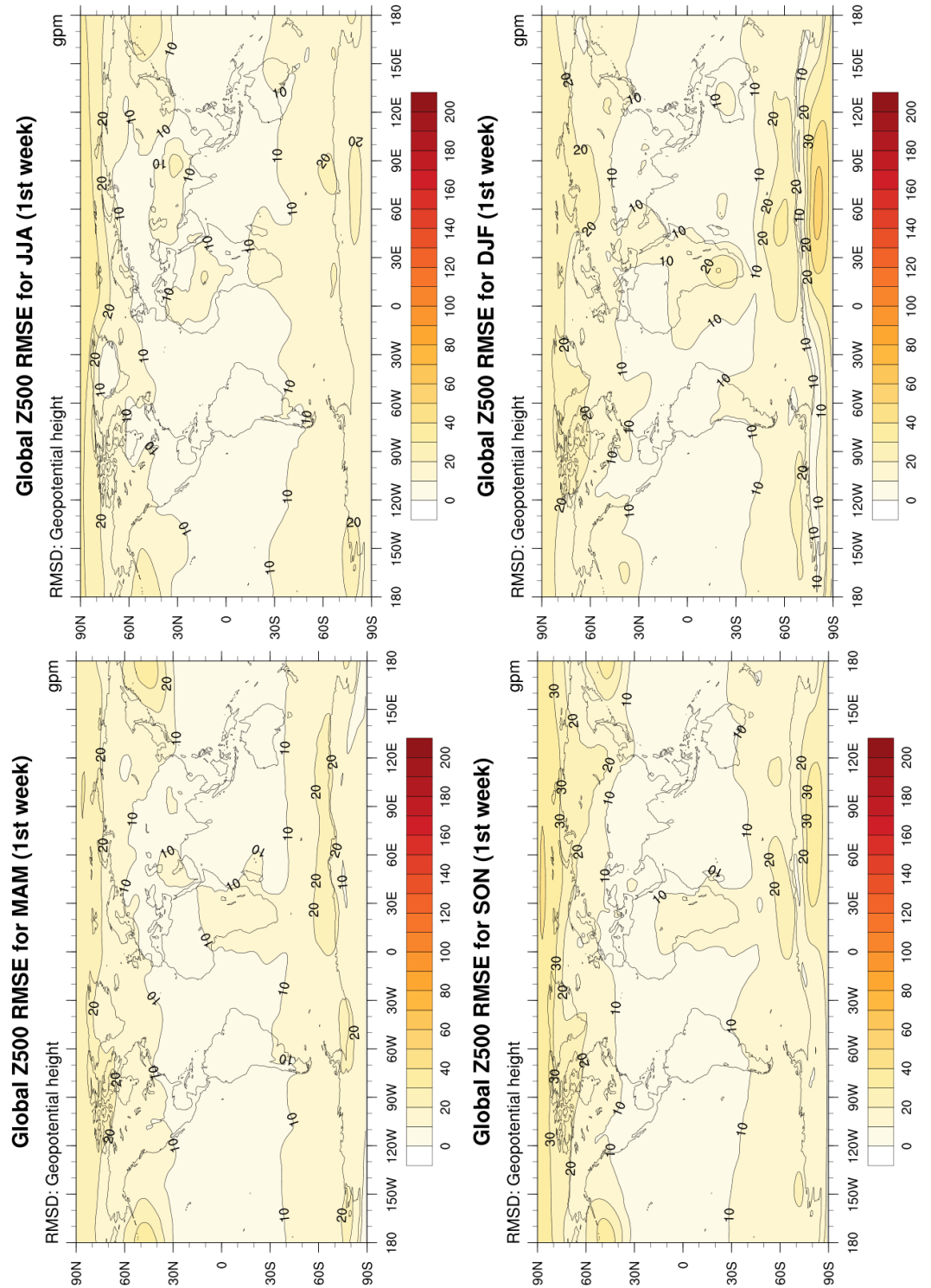


Figura 5.5: RMSE globale medio stagionale per l'altezza geopotenziale a 500 hPa mediato sulla 1^a settimana

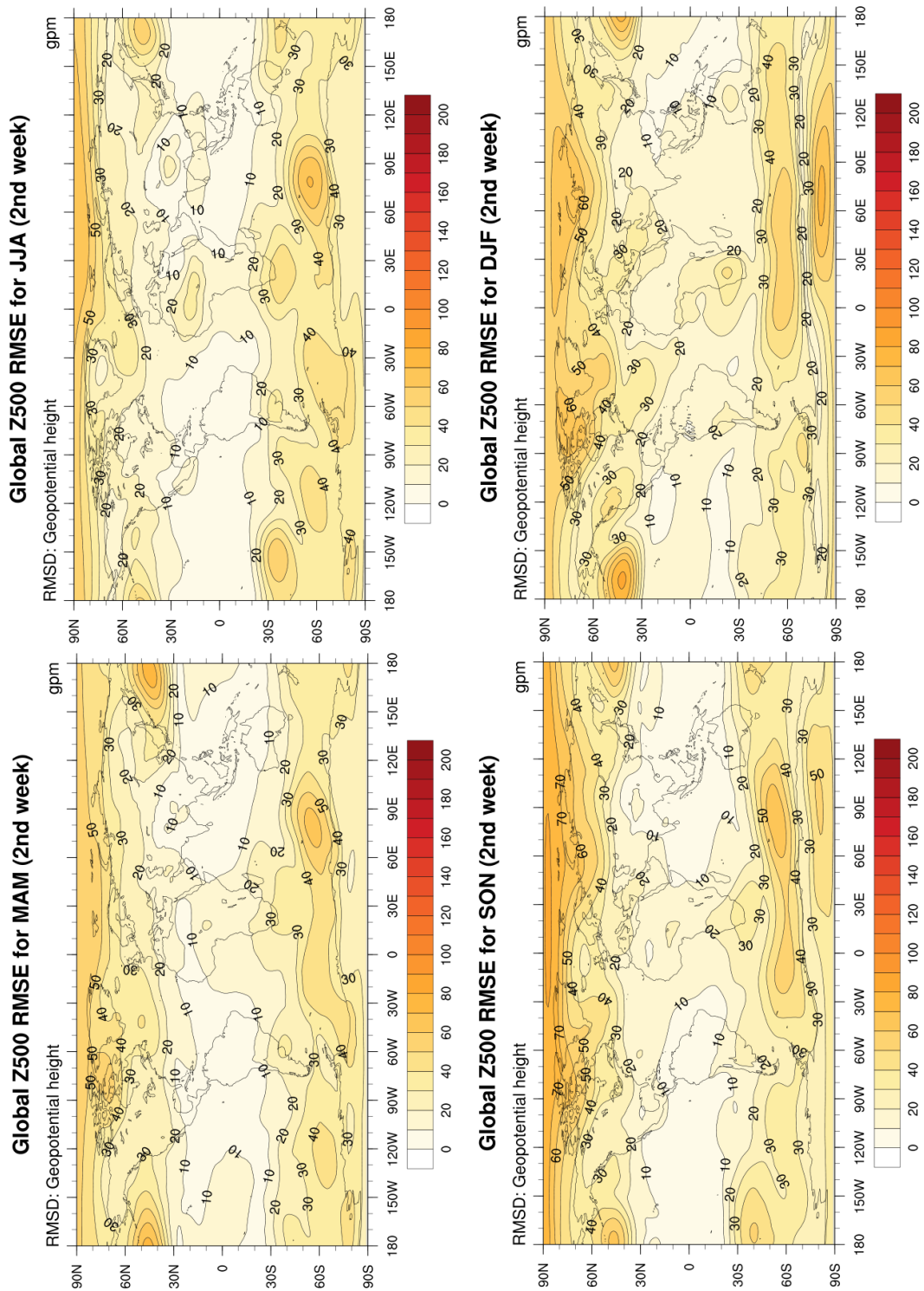


Figura 5.6: RMSE globale medio stagionale per l'altezza geopotenziale a 500 hPa mediato sulla 2^a settimana

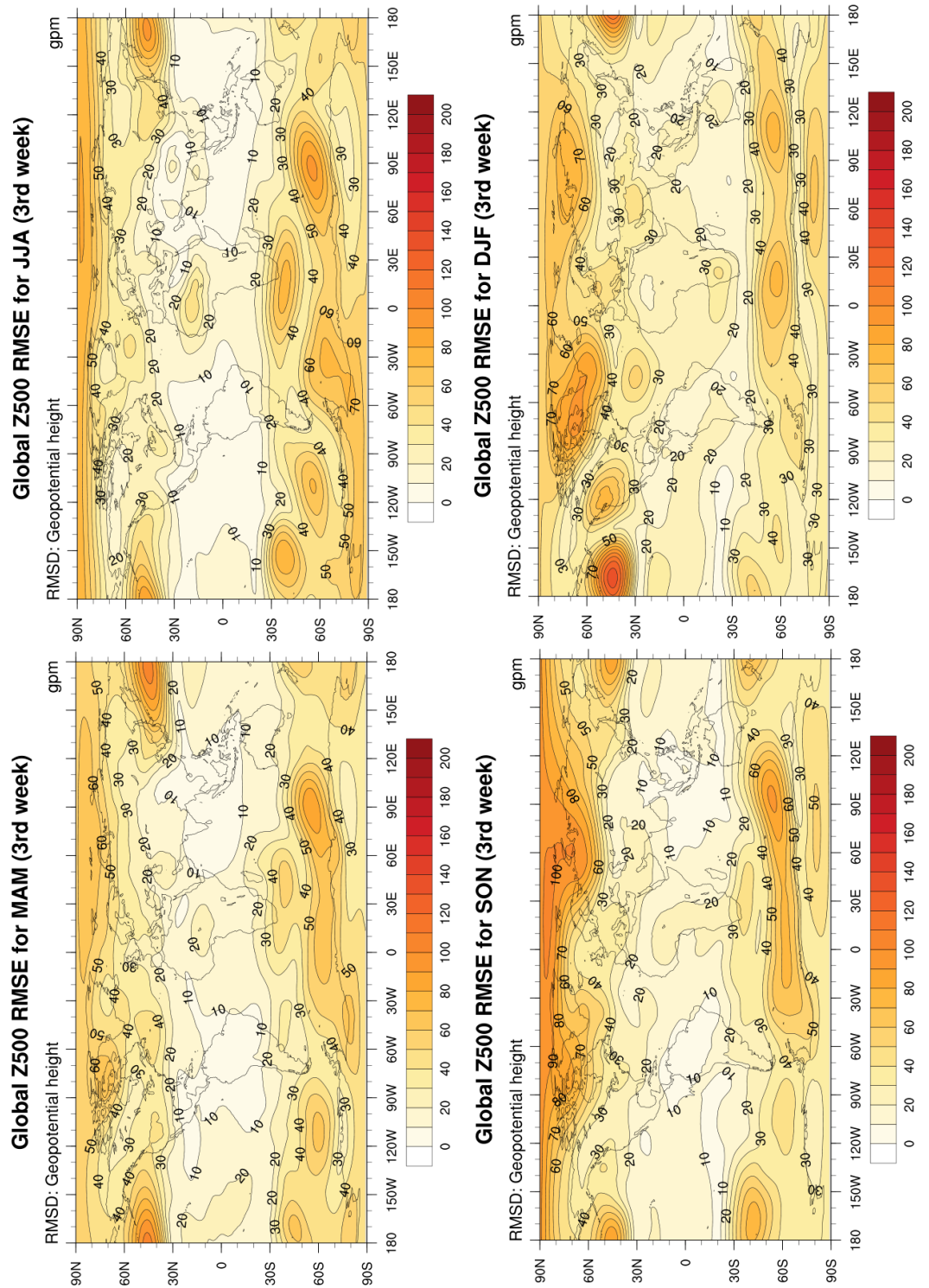


Figura 5.7: RMSE globale medio stagionale per l'altezza geopotenziale a 500 hPa mediato sulla 3^a settimana

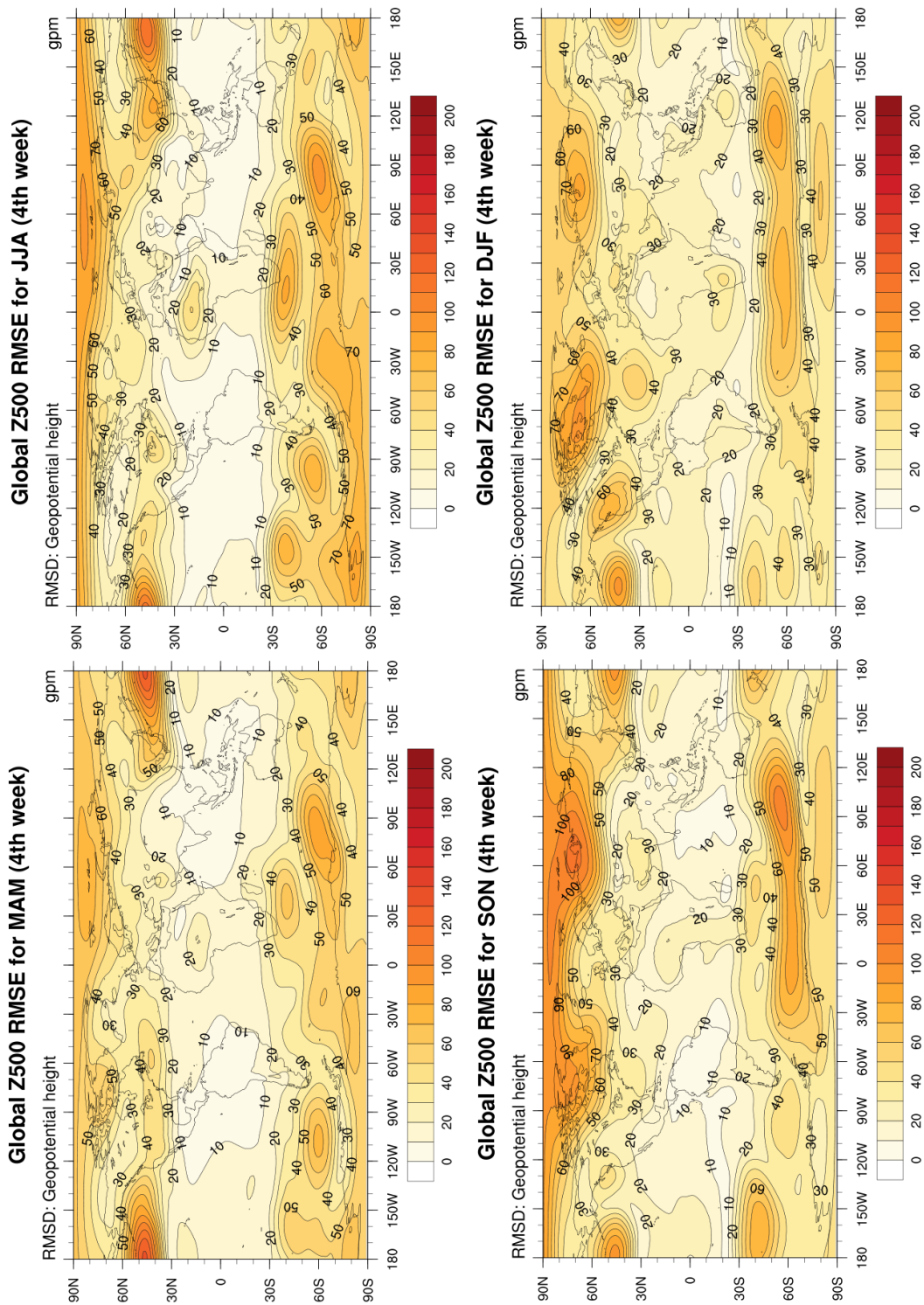


Figura 5.8: RMSE globale medio stagionale per l'altezza geopotenziale a 500 hPa mediato sulla 4^a settimana

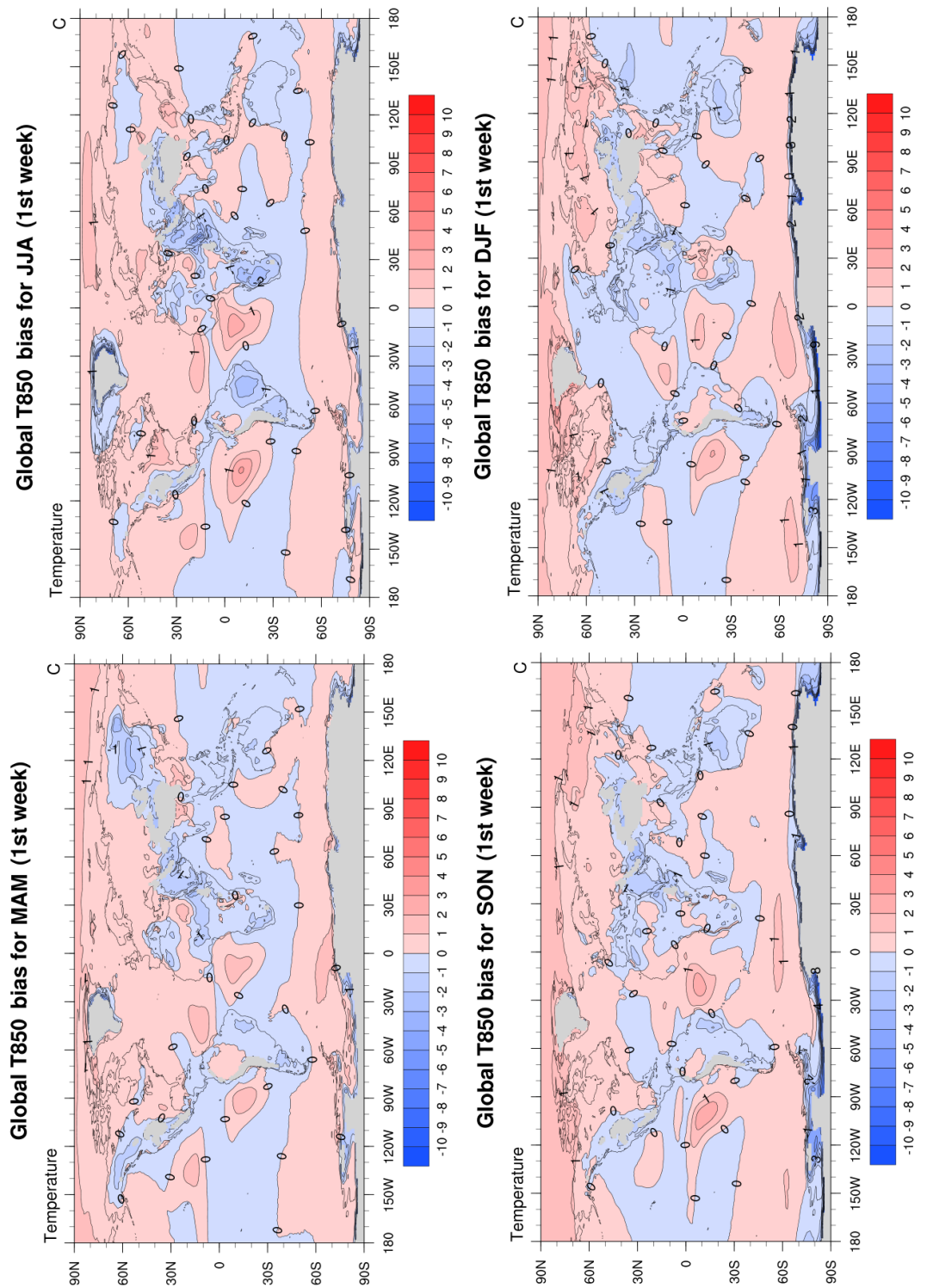


Figura 5.9: Bias globale medio stagionale per la temperatura a 850 hPa mediato sulla 1^a settimana

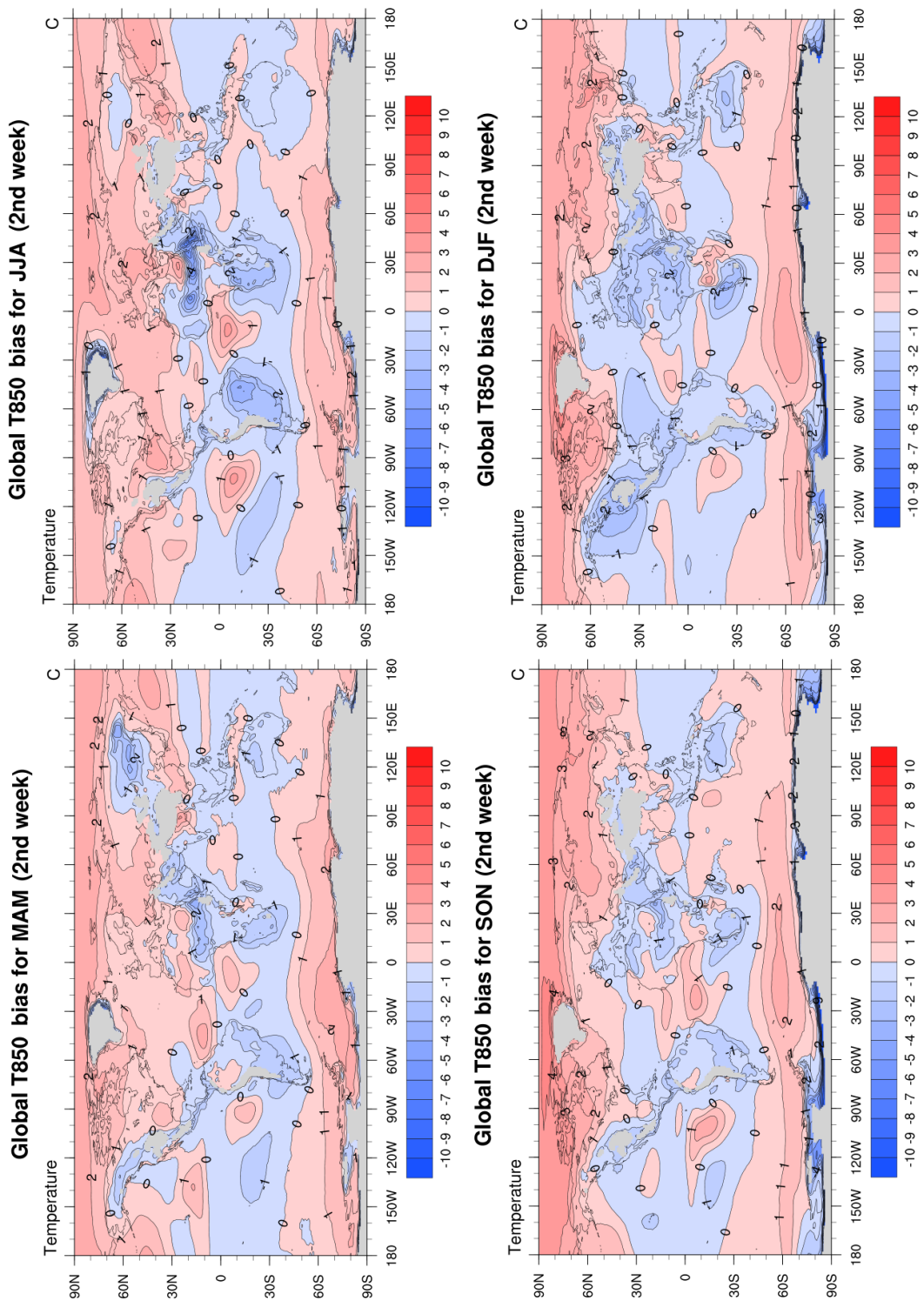


Figura 5.10: Bias globale medio stagionale per la temperatura a 850 hPa mediato sulla 2^a settimana

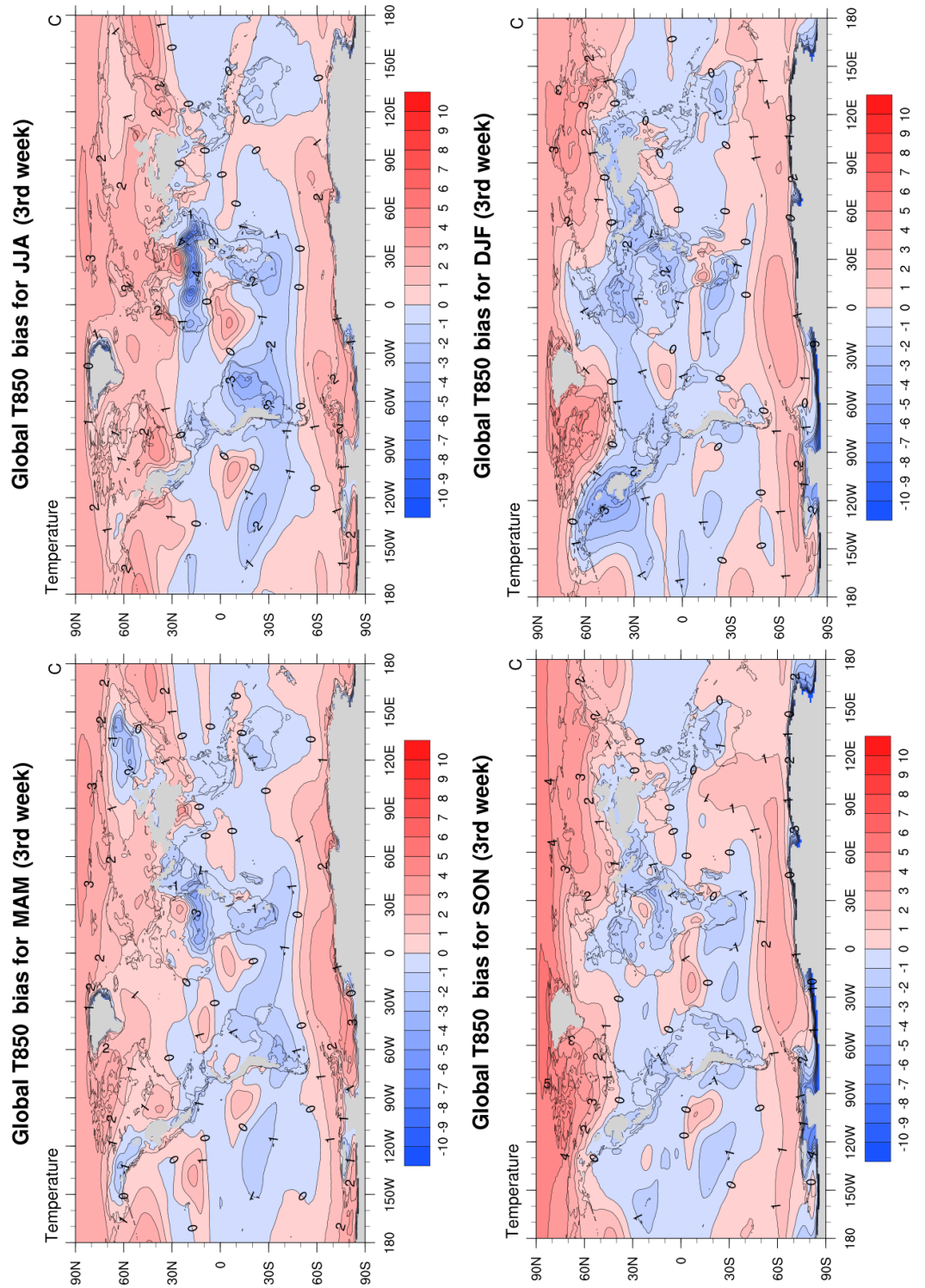


Figura 5.11: Bias globale medio stagionale per la temperatura a 850 hPa mediato sulla 3^a settimana

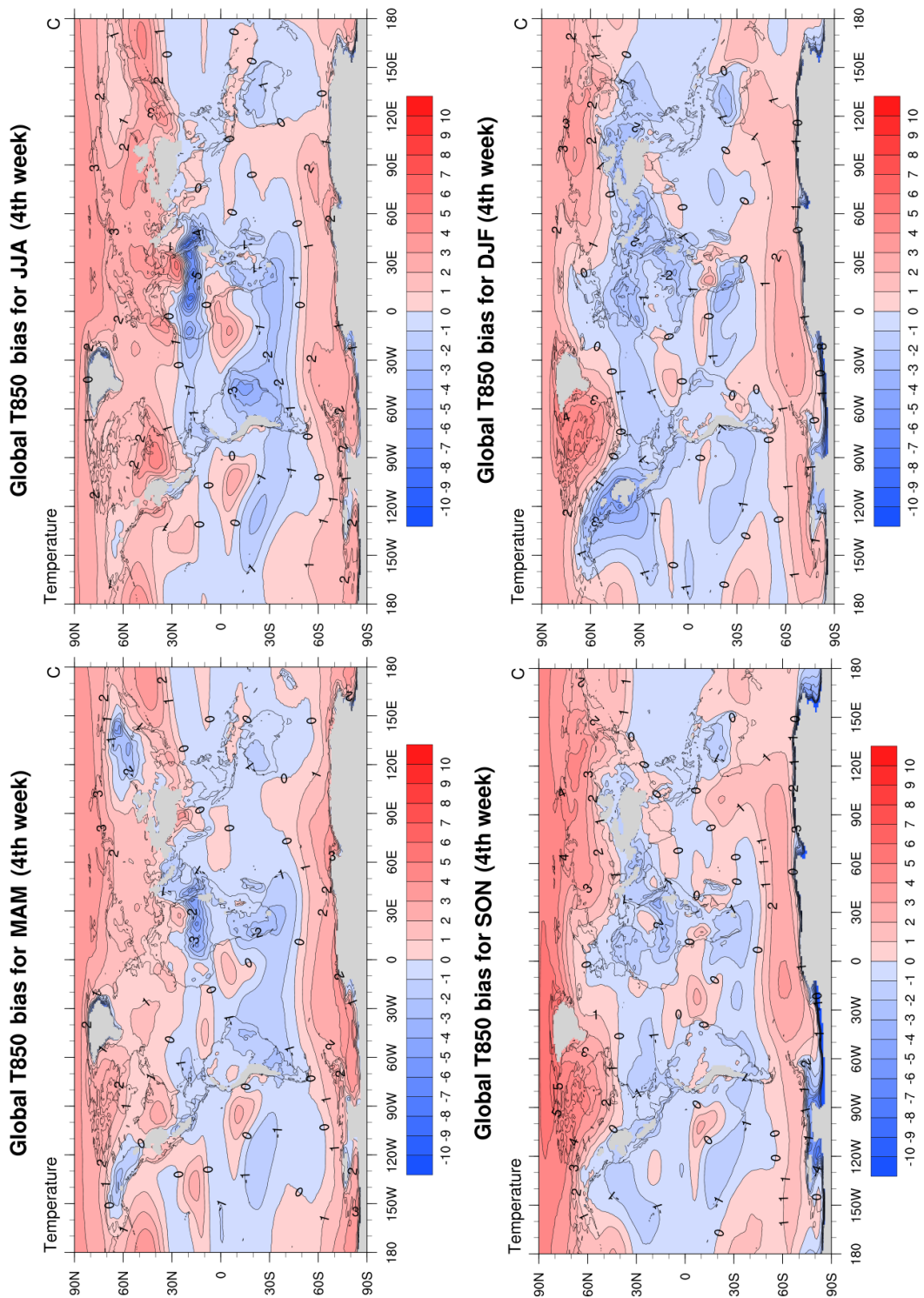


Figura 5.12: Bias globale medio stagionale per la temperatura a 850 hPa mediato sulla 4^a settimana

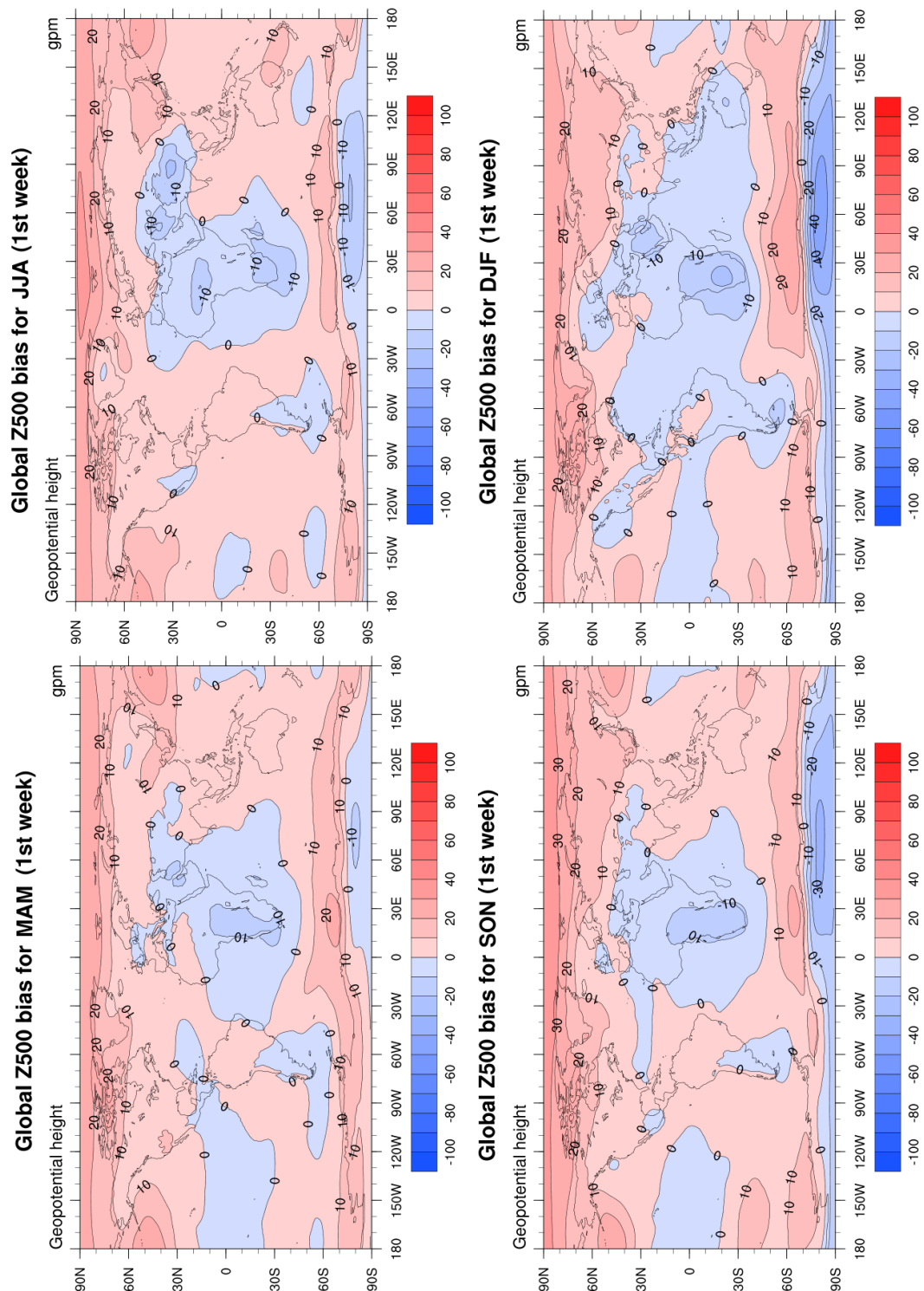


Figura 5.13: Bias globale medio stagionale per l'altezza geopotenziale a 500 hPa mediato sulla 1^a settimana

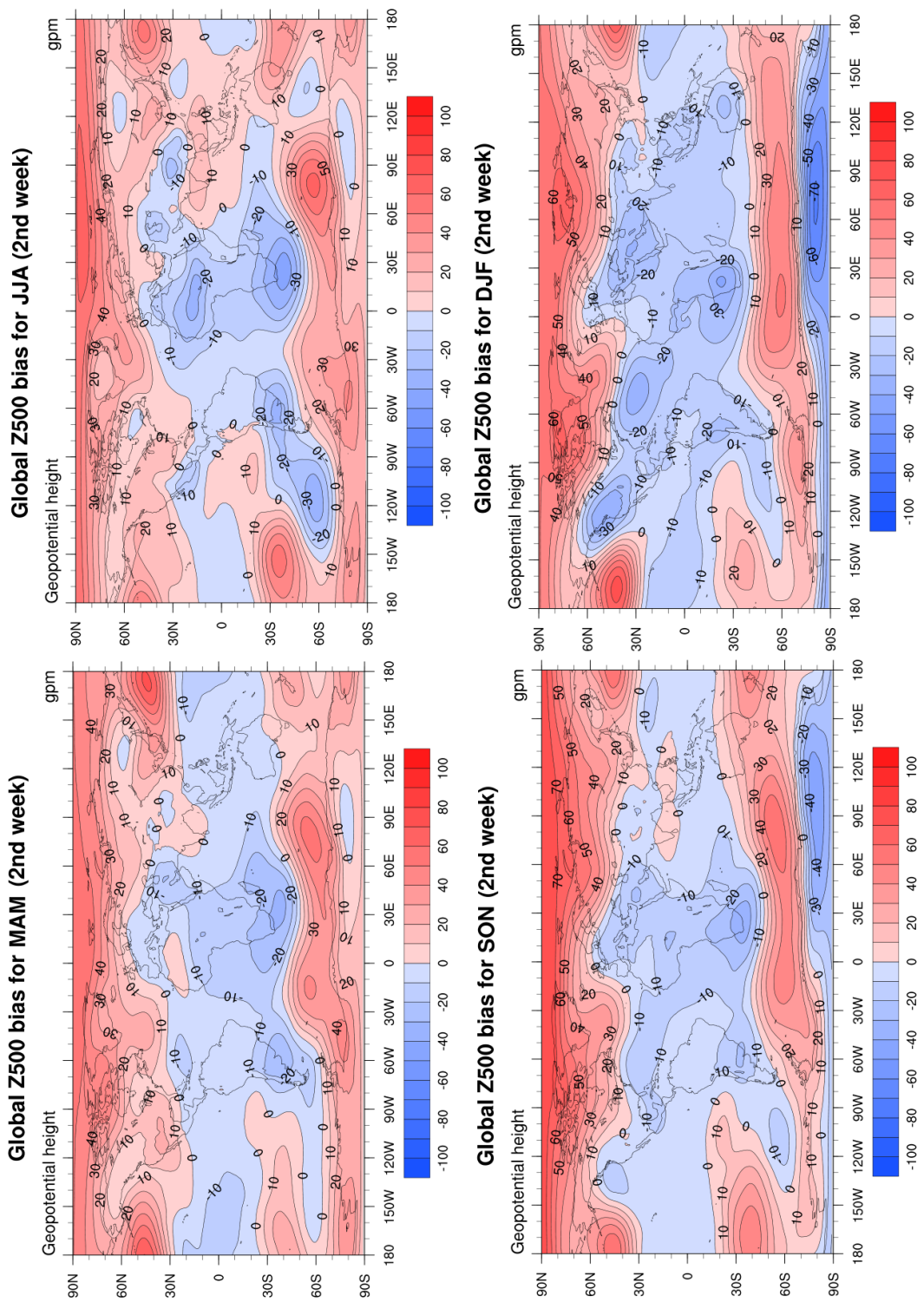


Figura 5.14: Bias globale medio stagionale per l'altezza geopotenziale a 500 hPa mediato sulla 2^a settimana

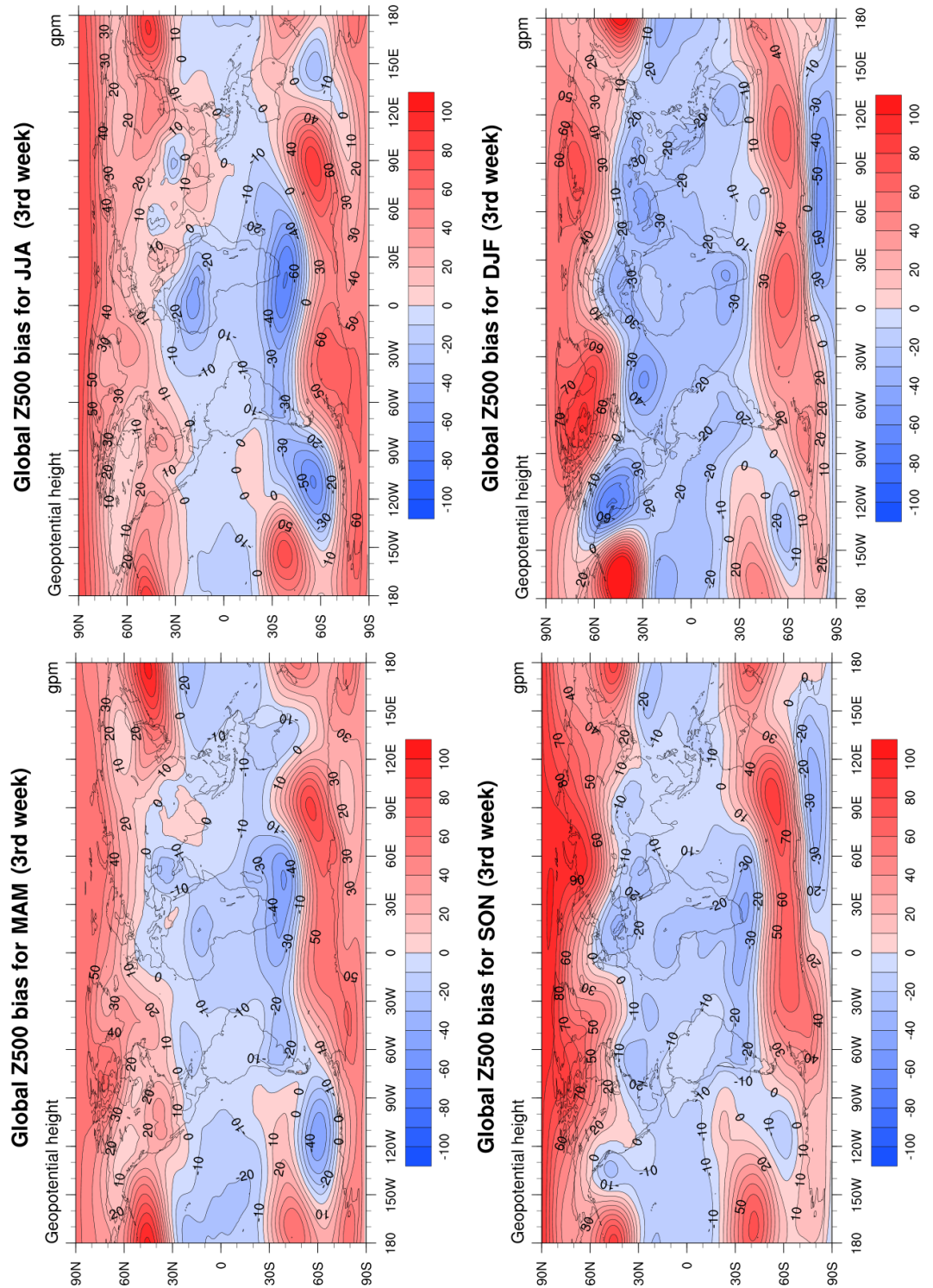


Figura 5.15: Bias globale medio stagionale per l'altezza geopotenziale a 500 hPa mediato sulla 3^a settimana

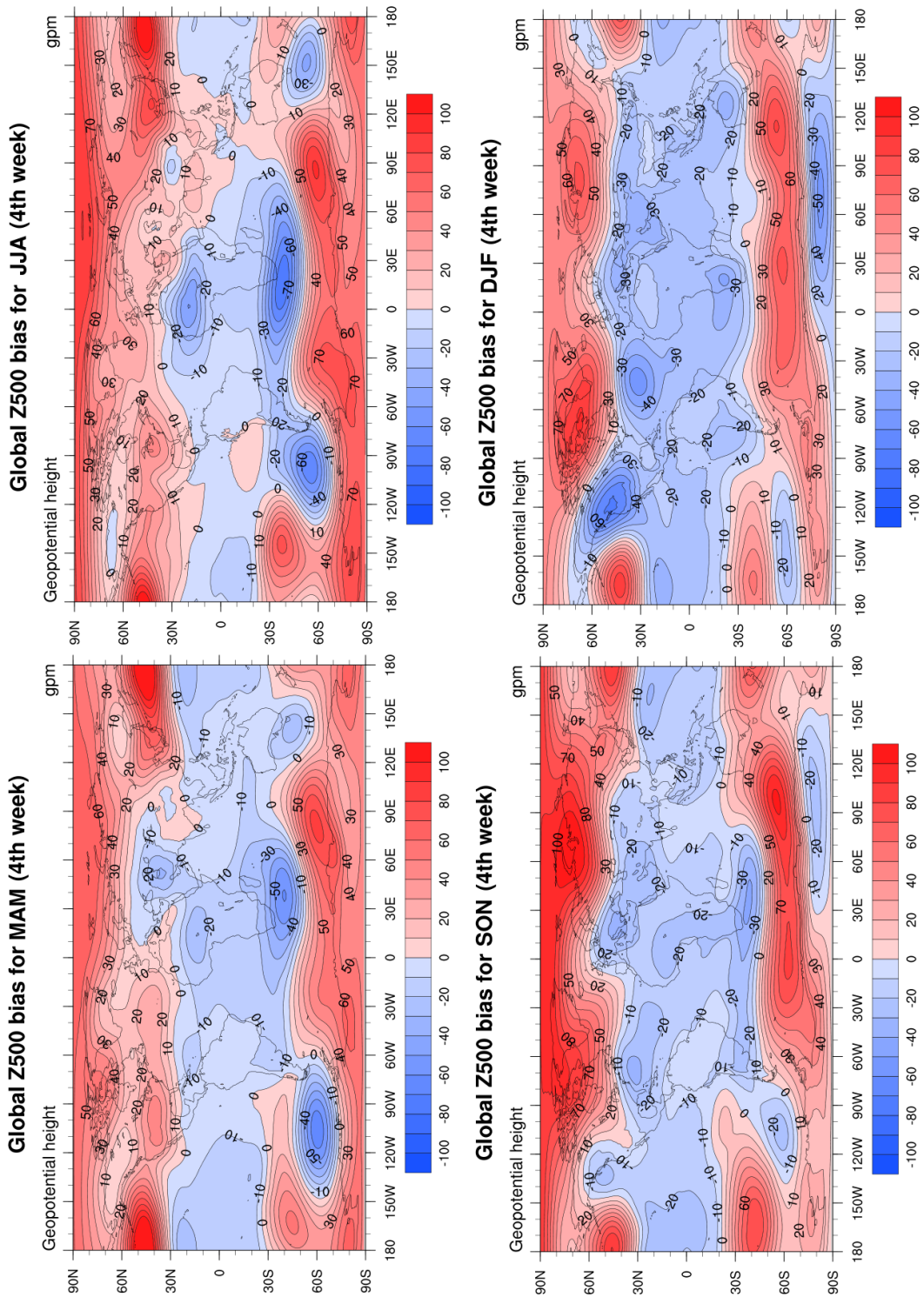


Figura 5.16: Bias globale medio stagionale per l'altezza geopotenziale a 500 hPa mediato sulla 4^a settimana

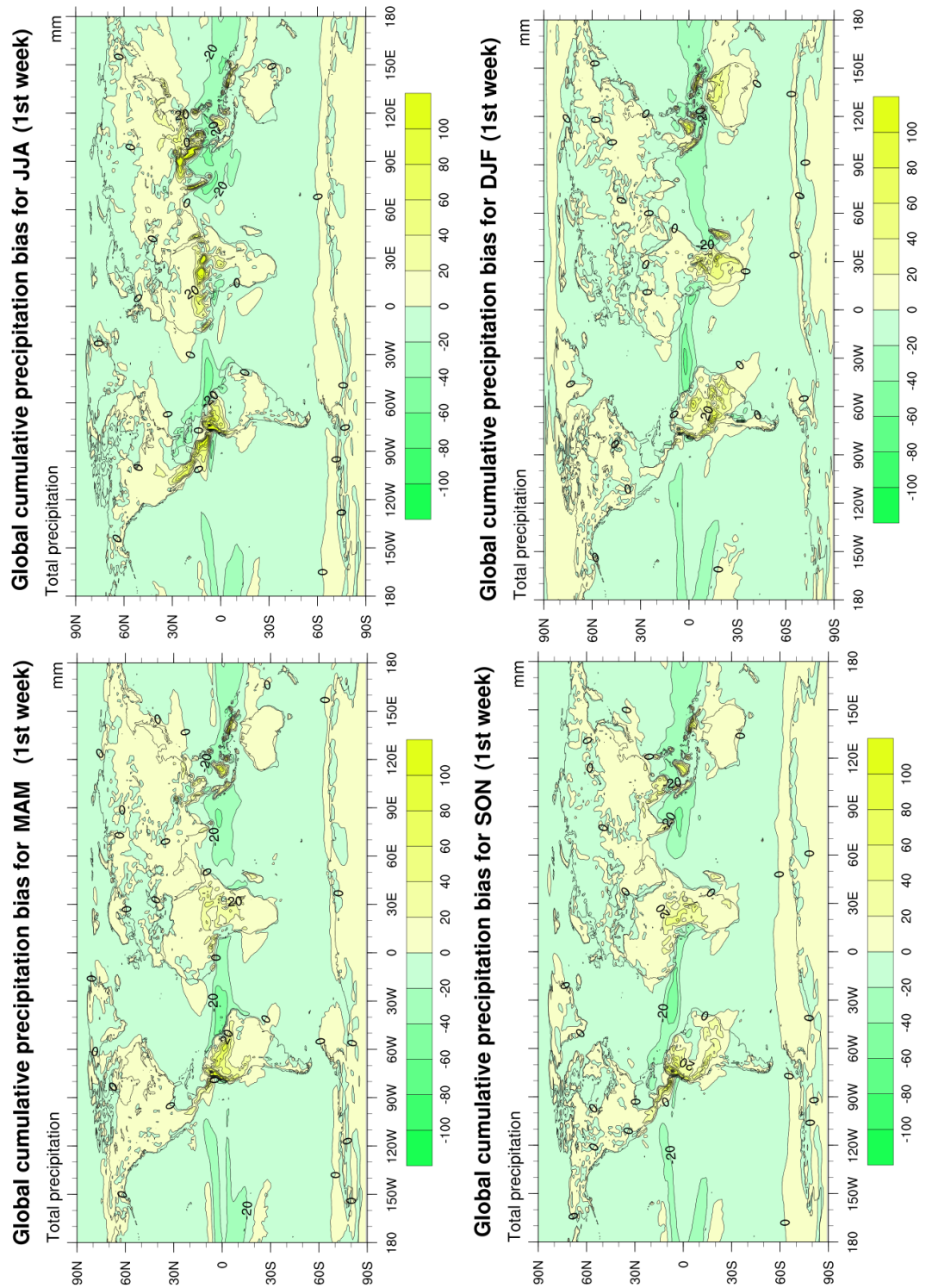


Figura 5.17: Bias globale medio stagionale per le precipitazioni cumulate, mediato sulla 1^a settimana

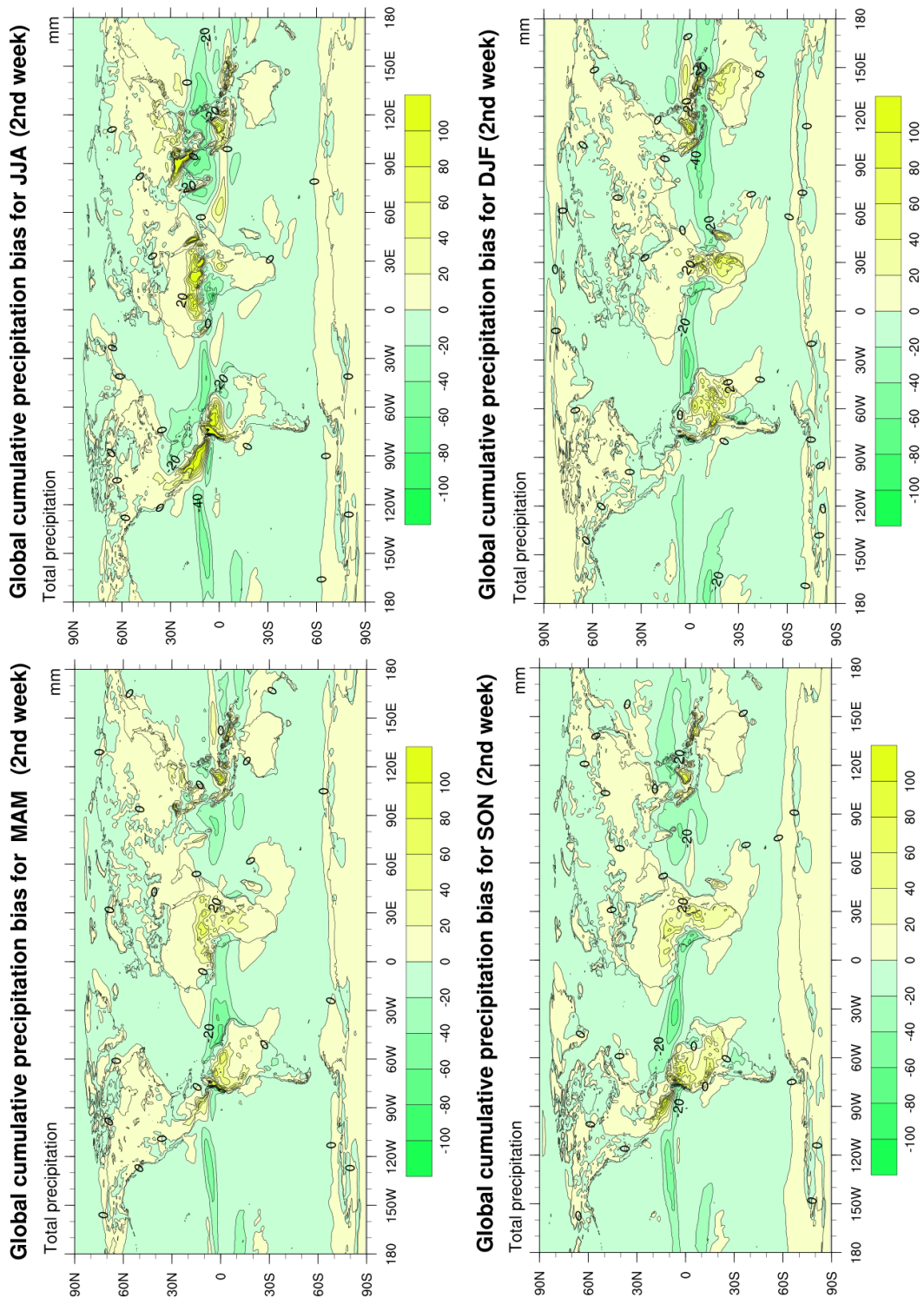


Figura 5.18: Bias globale medio stagionale per le precipitazioni cumulate, mediato sulla 2^a settimana

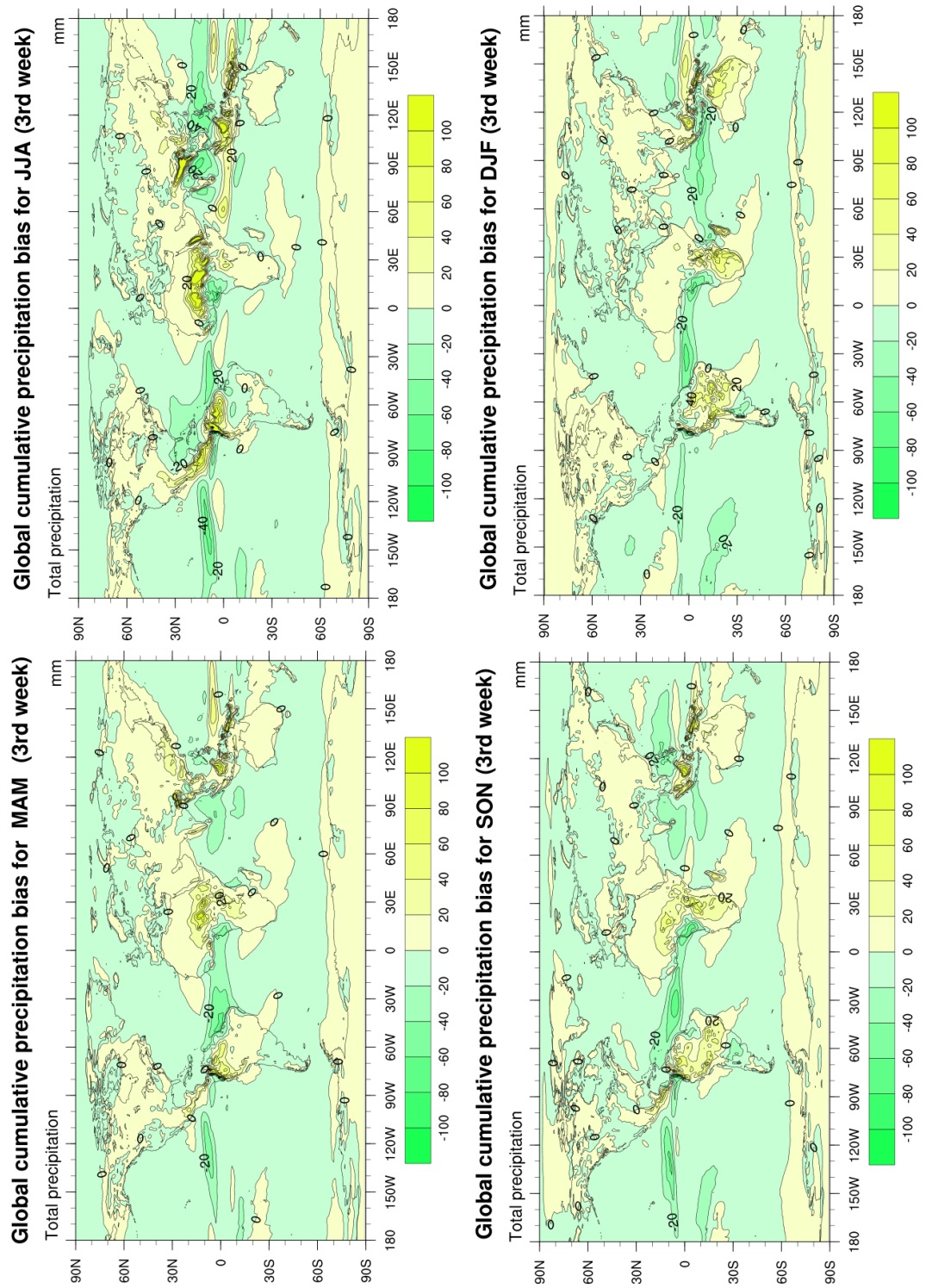


Figura 5.19: Bias globale medio stagionale per le precipitazioni cumulate, mediato sulla 3^a settimana

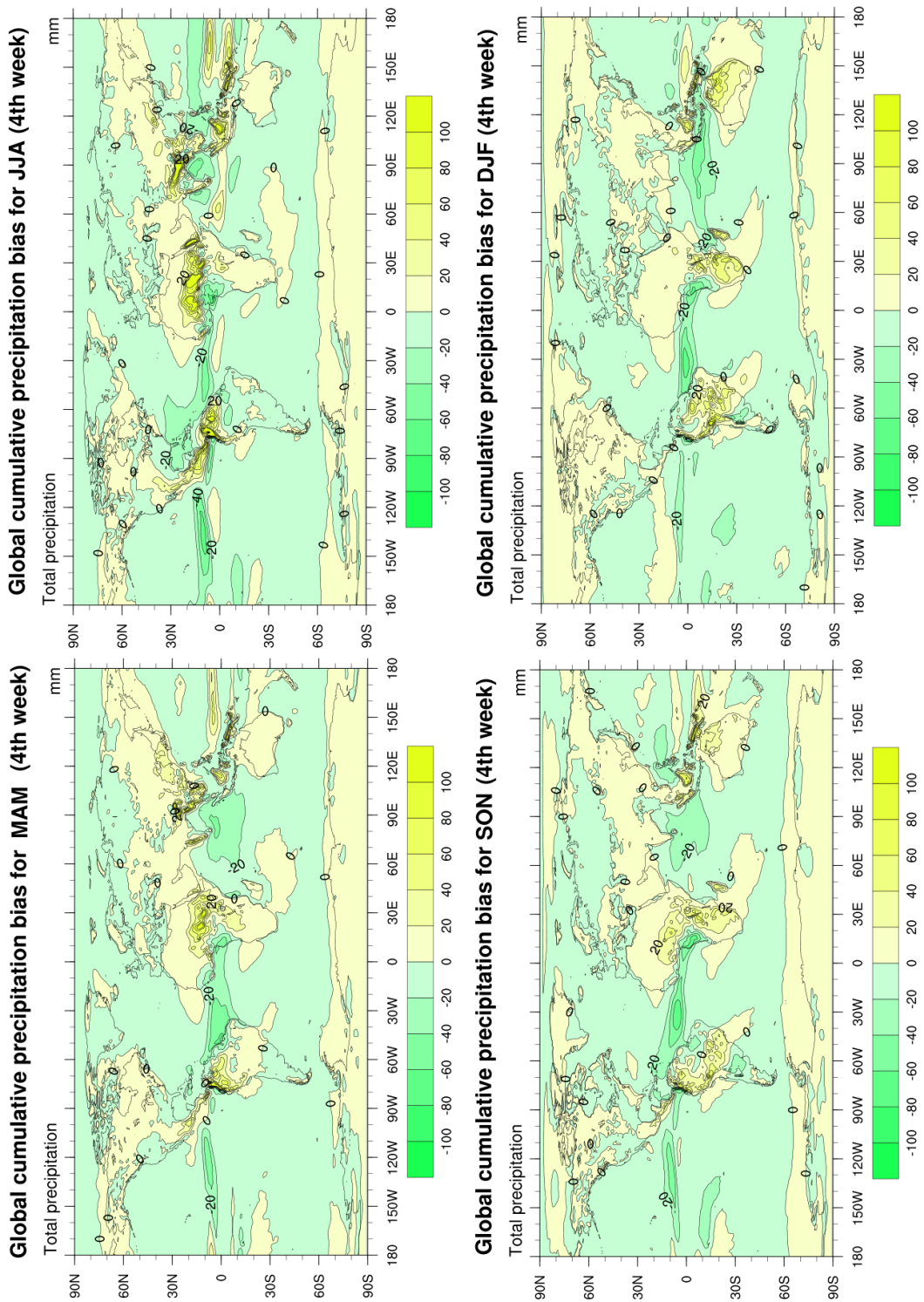


Figura 5.20: Bias globale medio stagionale per le precipitazioni cumulate, mediato sulla 4^a settimana

Ringraziamenti

“[...] and thanks for all the fish”

D. Adams

A dire *grazie* la mamma ce lo insegna quando ancora siamo piccini. Assieme ai *prego*, ai *per piacere*, rientra in quel codice di buone maniere e convivenza civile che se oggi è considerato solo una costruzione o un'ipocrisia è solo perché ci siamo incattiviti noi, col tempo.

Ma comunque.

Questo lavoro, le cose che ci trovate dentro, sono il prodotto di tre mesi di lavoro e allo stesso tempo il risultato (la logica conseguenza, la conclusione) di tre anni di studio. Entrambi (i tre anni, i tre mesi) hanno i loro grazie specifici.

Quindi.

Quanto ai tre mesi.

Se c'è una gerarchia in queste cose, il primo grazie è proprietà esclusiva ed indiscussa di Daniele Mastrangelo e del professor Andrea Buzzi, rispettivamente il correlatore e il relatore di questa tesi, la quale senza il loro supporto, il loro tempo, la loro pazienza, non sarebbe mai esistita.

A Daniele in particolare il mio grazie per aver sopportato la periodica occupazione delle sue giornate (e del suo ufficio), e per aver avuto la buona volontà di correggere le cose che scrivevo, il più delle volte da schiaffi o censura.

Quanto ai tre anni.

Un grazie alla mia famiglia, a mia madre a mio fratello, per essere una cosa sola anche quando c'è di mezzo l'ubiquità (Bologna, Padova, Baltimora).

Un grazie gigantesco a Matteo, Riccardo, Luca, Maurizio per essere stati il sale, il bello, la gioia, il piacere di questo triennio. E per aver costituito, insieme, il più invincibile e glorioso gruppo di studio della storia.

Agli altri amici dell'università, e in particolare: a Lorenzo, per condivedere con me ogni santo giorno il 92T10 del codice fiscale, e a Sara, per essere uno dei suoi fieri *amichetti*

stronzetti.

Un grazie ai transalpini, Silvia e Santiago, anche se sarebbe stato bello avervi qui.

Un grazie a Giacomo, Carlo, Elena, Angela: i più bei ritorni a Padova nel weekend sono sempre quelli in cui c'è qualcosa da fare con loro.

Un grazie *clamoroso* ai miei amici LG: Stefania e Stefano, Lorenzo e Lorenza. E a Pietro, per la più bella previsione meteo in nota vocale a tre cadenze (veneta, romagnola, pesarese) della storia.

Un grazie a quelli che metteranno da parte il lavoro, la scuola, lo studio, tutto quello che avrebbero di meglio da fare per venire qui oggi, e festeggiare con me: io temo non abbiate la benché minima idea di quello che state facendo.

Un grazie a quelli che “ehi! Ti sei scordato di me nei ringraziamenti della tesi!”. Lo so, e nel caso: non l'ho fatto apposta :-)