
Progetto NAUSICA

Downscaling di analisi meteorologiche ad alta risoluzione sul dominio Alpe Adria

Stato di avanzamento del progetto al 01/01/2018

RAPPORTO N. 001/2018

ARPA FVG - CRMA



Centro Regionale di Modellistica Ambientale

Copyright © ARPA FVG - CRMA Copyright © ARPA FVG, 2018



This work is released under the terms of the license Creative Commons Attribution / NonCommercial / ShareAlike.

ARPA FVG-Aria-Elaborati tecnico-scientifici

Documento: ARPA FVG - CRMA

Responsabile: Anna Chiara Goglio

Sito web di riferimento: www.arpa.fvg.it

Data emissione 01/01/2018

Termini uso.

Si ricorda ai lettori di questo rapporto che tutto si può usare ma bisogna stare attenti a quello che si dice. In particolare l'ARPA FVG non è responsabile per eventuali errori di stampa o di interpretazione dei dati. In ogni caso il lettore si assume ogni responsabilità per tutto quello che dice a seguito dell'utilizzo delle informazioni riportate in questo rapporto.

Indice

1	<i>Introduzione</i>	9
1.1	<i>Obiettivi e motivazioni</i>	10
1.2	<i>Outline del report</i>	11
2	<i>Strumenti e Metodi</i>	13
2.1	<i>Modello WRF e pre-processore WPS</i>	13
2.2	<i>Condizioni al contorno: analisi ECMWF</i>	15
2.3	<i>Altri strumenti utilizzati nella pre- e post-elaborazione</i>	15
2.4	<i>Workflow manager ecFlow</i>	15
2.5	<i>Metodo di lavoro: TRAC, Jenkins e GIT</i>	16
2.6	<i>Risorse di calcolo disponibili</i>	18
3	<i>Configurazione dei run</i>	21
3.1	<i>Scalabilità del modello</i>	22
3.2	<i>Estensione del dominio regionale</i>	23
3.2.1	<i>Caratteristiche dei campi generati sul dominio nazionale</i>	25
3.3	<i>Parametri di WRF e UPP</i>	26
3.3.1	<i>Grandezze cumulate e differenziali</i>	26
3.3.2	<i>Coordinate dei campi nei file GRIB</i>	31
3.3.3	<i>Modello di suolo e aggiornamento delle SST</i>	31
3.4	<i>Valutazione della necessità di assimilare misure</i>	43
4	<i>Flussi di lavoro</i>	53
4.1	<i>Produzione del database e post-processamento degli output</i>	53
4.1.1	<i>Prodotti delle suite NAUSICA_run e NAUSICA_pp</i>	54

4.2	<i>Post-elaborazione del database</i>	57
4.2.1	<i>Prodotti della suite WRF_ana_post</i>	57
4.3	<i>Controllo qualità dei campi al suolo</i>	59
4.3.1	<i>Test statistici eseguibili</i>	60
4.3.2	<i>Variabili confrontabili</i>	63
4.3.3	<i>Definizione delle aree di analisi</i>	64
4.3.4	<i>Prodotti delle suite NAUSICA_ver e NAUSICA_nc</i>	67
4.3.5	<i>Visualizzazione dei risultati di controllo della qualità</i>	68
4.4	<i>Controllo qualità dei campi in quota</i>	69
4.4.1	<i>Test eseguibili e variabili confrontabili in NAUSICA_rds</i>	69
4.4.2	<i>Problematiche riscontrate nella metodologia utilizzata</i>	71
4.4.3	<i>Suite NAUSICA_srds</i>	75
5	<i>Risultati del progetto disponibili</i>	77
5.1	<i>File archiviati</i>	78
5.2	<i>Valutazione della qualità delle simulazioni</i>	80
5.2.1	<i>Temperatura</i>	81
5.2.2	<i>Intensità del vento</i>	81
5.2.3	<i>Umidità relativa</i>	89
5.2.4	<i>Precipitazioni</i>	89
5.2.5	<i>Precipitazioni > 10 mm/h</i>	92
5.2.6	<i>Precipitazioni > 25 mm/h</i>	92
5.2.7	<i>Pressione</i>	92
5.2.8	<i>Radiazione</i>	97
5.2.9	<i>Evapotraspirazione giornaliera</i>	98
5.3	<i>Analisi dei campi in quota</i>	100
6	<i>Sviluppi futuri del progetto</i>	107
A	<i>Namelist di WRF e WPS</i>	109
A.1	<i>Namelist WPS</i>	109
A.2	<i>Namelist WRF</i>	110
B	<i>Codici dei prodotti delle suite</i>	113
B.1	<i>Codici NAUSICA_ver</i>	113
B.2	<i>Codici NAUSICA_rds</i>	115

C	<i>Settaggio del WRF per l'aggiornamento delle SST</i>	117
D	<i>Misure a disposizione</i>	119
	D.1 <i>Misure stazioni OSMER</i>	119
	D.2 <i>Misure Evapotraspirazione (stazioni OSMER)</i>	122
	D.3 <i>Radiosondaggi</i>	123
E	<i>Struttura ed utilizzo delle suite NAUSICA_run e NAUSICA_pp</i>	125
	E.1 <i>Utilizzo della suite</i>	125
	E.2 <i>Struttura della Suite</i>	125
	E.2.1 <i>Famiglie di controllo</i>	126
	E.2.2 <i>Shell Script di creazione della suite.def</i>	127
	E.2.3 <i>File di inizializzazione della suite .ini</i>	129
	E.3 <i>Post-processing nella suite NAUSICA_run</i>	130
	E.3.1 <i>Controllo dei GRIB prodotti e delle variabili in essi contenute</i>	130
F	<i>Struttura ed utilizzo della suite WRF_ana_post</i>	131
	F.1 <i>Utilizzo della suite</i>	131
	F.2 <i>Struttura della Suite</i>	131
	F.2.1 <i>Famiglie di produzione user_%user-famiglia%</i>	131
	F.2.2 <i>Task di produzione</i>	132
	F.2.3 <i>Task di controllo</i>	132
	F.2.4 <i>Famiglia FARM_ana2fc</i>	133
	F.2.5 <i>Aggiunta del taglio del dominio</i>	133
G	<i>Struttura ed utilizzo delle suite NAUSICA_ver e NAUSICA_nc</i>	135
	G.1 <i>Utilizzo della suite NAUSICA_ver e NAUSICA_nc</i>	135
	G.2 <i>Struttura della suite NAUSICA_ver</i>	136
	G.2.1 <i>File elenco confronti</i>	136
	G.2.2 <i>Variabili confrontabili</i>	137
	G.2.3 <i>Famiglie di estrazione NAUSICA_ver_"codice"_"periodo"</i>	138
	G.2.4 <i>Eseguibili e file accessori</i>	145

<i>H</i>	<i>Struttura ed utilizzo delle suite NAUSICA_rds e NAUSICA_srds</i>	147
<i>H.0.1</i>	<i>Descrizione dei compiti dei singoli task</i>	148
<i>I</i>	<i>Grandezze cumulate e differenziali</i>	151
<i>I.0.1</i>	<i>Problema relativo alle grandezze cumulate</i>	151
<i>I.0.2</i>	<i>Calcolo delle grandezze differenziali a partire dalle cumulate</i>	152
<i>I.0.3</i>	<i>Introduzione della produzione diretta di grandezze differenziali</i>	154
<i>J</i>	<i>Conversione di coordinate</i>	155
<i>J.0.1</i>	<i>Problema relativo alla conversione in coordinate lat/lon</i>	155
<i>J.0.2</i>	<i>Problema generale relativo alla conversione in coordinate lat/lon (copygb.exe)</i>	155
	<i>Bibliografia</i>	175

1

Introduzione

Nausicaa è una figura della mitologia greca che compare nell'Odissea di Omero, il suo nome in greco antico significa "colei che brucia le navi". Omero conferisce al personaggio il significato letterario dell'amore inesperto, probabilmente uno dei primi esempi in letteratura di amore non corrisposto. Nausicaa, ritratta in Figura 1.1, è anche una figura materna per il guerriero: lo cura, lo invita a casa e gli dice al momento dell'addio, "non dimenticarmi, perché ti ho ridato la vita" indicando così il suo status di "nuova madre" nei suoi confronti [1].



Figura 1.1: Nausicaa incontra Odisseo.

L'obiettivo del progetto NAUSICA [2] era la creazione, per mezzo del *Weather Research and Forecasting (WRF) Model* [3], di un database di analisi meteorologiche ad alta risoluzione e si prefiggeva di automatizzarne l'aggiornamento tramite una catena modellistica operativa. Il database prodotto, centrato sul Friuli Venezia Giulia, è caratterizzato da una risoluzione di 2 km e, attualmente, copre il periodo che va dal 1 Gennaio 2001 al 2 Gennaio 2018. I prodotti del progetto sono utilizzabili come input per tutti i modelli che necessitano di quelli che, nella logica DPSIR, vengono definiti determinanti meteorologici[4]. Parte integrante del progetto è rappresentata dall'analisi del database prodotto svolta tramite confronto statistico con le misure a disposizione, che consente di formulare una valutazione dettagliata della

qualità di questo. Inoltre, i risultati raggiunti in termini di metodi e strumenti, rappresentano un'importante acquisizione di consapevolezza nell'ambito della catena modellistica meteorologica.

Questa relazione ha lo scopo di illustrare lo stato di avanzamento del progetto descrivendo i prodotti ottenuti, di cui viene valutata la qualità, sottolineandone le applicazioni e proponendo possibili sviluppi futuri.

1.1 Obiettivi e motivazioni

Il principale obiettivo del progetto NAUSICA, come accennato nelle righe introduttive, era quello di generare un database di downscaling di analisi meteorologiche che, a partire da Gennaio 2001, riproducesse l'evoluzione dei principali campi meteorologici, utilizzando il modello meteorologico non idrostatico ad area limitata *Weather Research and Forecasting* (WRF) Model [3] le cui principali caratteristiche sono riassunte nel Paragrafo 2.1, e ne verificasse la qualità dei prodotti. Le **analisi** sono per definizione create guidando l'evoluzione temporale della simulazione tramite le misure disponibili di conseguenza il prodotto di tali run, costituito dall'evoluzione temporale dei principali campi tridimensionali, rappresenta la migliore riproduzione attualmente ottenibile delle condizioni meteorologiche passate. Il downscaling eseguito da NAUSICA raggiunge la risoluzione di 2 km che, essendo tipica della maggior parte dei fenomeni meteorologici al confine tra la mesoscala e la microscala riveste un elevato grado di interesse in quanto al limite della prevedibilità numerica operativa. In NAUSICA la simulazione è guidata da condizioni al contorno di tipo analisi, descritte nel Paragrafo 2.2, prodotte dall'*European Center for Medium-Range Weather Forecasts* (ECMWF) [5] e include l'aggiornamento delle temperature superficiali del mare descritta nel Paragrafo 3.3.3.

Il dominio coperto dai campi ad alta risoluzione, oltre al Friuli Venezia Giulia, include anche Veneto, Trentino Alto Adige e parte di Emilia Romagna, Slovenia, Austria, Germania e Croazia; in Figura 1.2 è evidenziato il bordo che lo delimita.

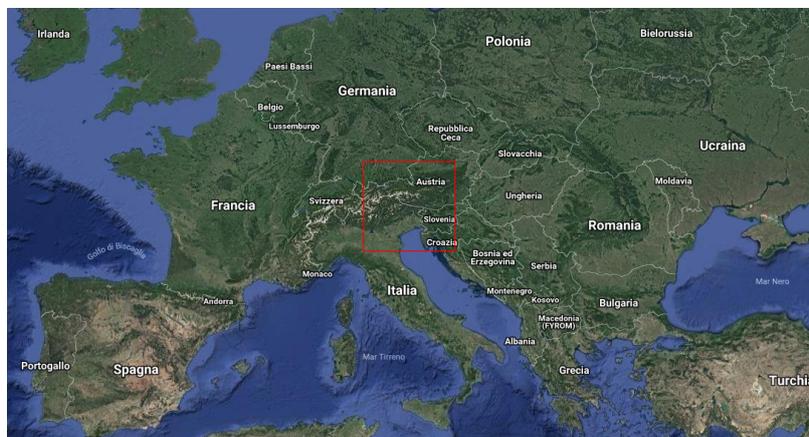


Figura 1.2: Dominio delle simulazioni ad alta risoluzione.

La categoria di appartenenza dello strumento scelto per la gestione dei flussi di lavoro, il *Workflow Manager ecFlow* descritto nel Paragrafo 2.4, è caratterizzata da un elevato grado di automazione e ripetibilità al variare dei parametri d'ambiente delle catene di processi che ha permesso di rendere operativa la catena implementata.

Tra i prodotti delle simulazioni vengono archiviati sia gli output nativi del WRF in formato *netCDF* [6] che i file in formato *GRIB* [7] risultato della post elaborazione svolta dall'*Unified Post Processor* (UPP) [8].

L'ambito di utilizzo dei prodotti di queste simulazioni è ad ampio spettro: questo database crea infatti gli input per qualsiasi modello che necessiti di condizioni meteorologiche come determinante, per esempio per modelli di dispersione di inquinanti in aria, modelli fotochimici per la qualità dell'aria, modelli oceanografici, modelli idrologici e modelli agronomici, oltre a rappresentare l'oggetto di qualsiasi richiesta di fornitura di dati meteorologici.

In particolare, gli output della post-elaborazione generati dalla suite *WRF ana post* descritta nel Paragrafo 4.2, forniscono gli input meteorologici alla catena modellistica *F-AIR* che è stata utilizzata per la valutazione della qualità dell'aria 2016 [9].

1.2 Outline del report

Nel Capitolo 2 sono elencati gli strumenti utilizzati mettendo in luce le caratteristiche che li rendono particolarmente adatti allo scopo, ma anche descrivendone eventuali problematiche. Il metodo di lavoro adottato nello svolgimento dell'intero progetto, strettamente correlato all'utilizzo di alcuni di essi, trova un'accurata descrizione in un paragrafo ad esso dedicato. A conclusione sono riportate le caratteristiche delle risorse di calcolo messe a disposizione da ARPA FVG [10] per lo svolgimento del progetto. La scelta di alcuni parametri fondamentali relativi alle simulazioni meteorologiche svolte, come ad esempio l'estensione del dominio di calcolo a più elevata risoluzione spaziale, è infatti strettamente connessa a questa disponibilità in quanto determinante dei tempi di calcolo.

All'interno della relazione viene successivamente analizzata la configurazione scelta per il run definitivo della simulazione, evidenziando in particolare le differenze rispetto al corrispondente utilizzo del modello numerico nell'attività di previsione [11]. L'ambito di appartenenza dei parametri analizzati spazia dalle parametrizzazioni fisiche all'interno del modello [12] descritte in Sezione 3.3, all'ottimizzazione nell'utilizzo delle risorse di calcolo descritta in Sezione 3.1, dalla scelta di copertura dei domini di calcolo descritta in Sezione 3.2, al settaggio della post elaborazione trattata in Sezione 3.3, sino alla valutazione della necessità di assimilare misure nel modello discussa in Sezione 3.4. Il processo di selezione di questi parametri include, oltre ad alcuni test preliminari, l'analisi di problemi riscontrati in produzioni propedeutiche e le relative soluzioni adottate.

Nel capitolo successivo vengono descritti scopi e caratteristiche dei tre flussi di lavoro costruiti per gestire e controllare pre-processamento,

produzione, post-elaborazione e analisi della qualità dei campi al suolo prodotti. Un ulteriore paragrafo tratta la problematica della valutazione dei campi in quota.

I risultati delle analisi di qualità del database prodotto costituiscono l'oggetto principale dell'ultimo capitolo.

2

Strumenti e Metodi

In questo Capitolo sono descritti i principali strumenti utilizzati nello svolgimento del progetto, incluse le risorse di calcolo a disposizione, e le caratteristiche delle condizioni al contorno con cui sono state guidate le simulazioni eseguite. Essendo connesso agli strumenti a disposizione viene descritto anche il metodo di lavoro utilizzato nella fase di sviluppo del progetto.

2.1 Modello WRF e pre-processore WPS

IL *Weather Research and Forecasting* (WRF) Model [3] è un modello di previsioni atmosferiche ad area limitata utilizzato sia per scopi di ricerca che per attività operative presso numerosi centri di varie nazionalità che ne contribuiscono costantemente allo sviluppo. È concepito per essere flessibile e portabile su piattaforme di calcolo parallele. Ha un'elevata versatilità in quanto adatto all'utilizzo su scale spaziali che vanno dal chilometro fino alle centinaia di chilometri, è infatti in grado di simulare i fenomeni appartenenti agli ordini di grandezza tipici della mesoscala, coprendo anche la transizione con la microscala α .

Il WRF può essere utilizzato sia per run ideali che per run reali guidati da opportune condizioni al contorno ed include la possibilità di assimilare misure.

Il pre-processamento dei dati è svolto dal *WRF Preprocessor System* (WPS) [13] che tramite tre eseguibili si occupa di

- Definire il dominio e calcolare i campi statici (*geogrid.exe*)
- Decodificare i file GRIB contenenti le condizioni al contorno (*ungrib.exe*)
- Interpolare i dati orizzontalmente sulla griglia creata (*metgrid.exe*)

Inoltre, nella catena del modello, prima del solutore numerico vero e proprio è presente il *real*: un altro eseguibile di pre-processamento che si occupa di creare i file contenenti le condizioni al contorno ed iniziali e di interpolare verticalmente i campi.

Il solutore numerico dinamico utilizzato in NAUSICA è l'*Advanced Research WRF* (ARW) [3] di cui si elencano le caratteristiche principali:

- È un solutore Euleriano comprimibile, non idrostatico

- Utilizza coordinate verticali *terrein-following* (le coordinate verticali sono staggerate: il geopotenziale è definito sulle superfici inferiore e superiore delle celle mentre gli altri campi nel centroide)
- Utilizza un'integrazione temporale di tipo Runge-Kutta al terzo ordine
- È caratterizzato da una discretizzazione spaziale di ordine elevato
- Contiene termini di Coriolis e curvatura completi
- Utilizza il nesting (1 o 2 way)
- Le condizioni al bordo possono essere di vario tipo (periodiche, specificate, ideali, reali,..)
- Include processi fisici di bordo, superficiali e radiativi. In particolare sono implementate diverse parametrizzazioni per la microfisica, la fisica del sottosuolo e del *boundary layer* mentre il mare è modellizzato come un bagno termico caratterizzato da una temperatura fissa o variabile (a meno di non accoppiare il WRF a modelli oceanografici che ne descrivano i processi fluido-termodinamici interni)
- Contiene la descrizione completa dei fenomeni meteorologici alla mesoscala

In Figura 2.1 è riportato lo schema a blocchi di WPS e WRF considerando unicamente gli elementi coinvolti nell'ambito del progetto NAUSICA. I file intermedi del processo e gli output del modello sono tutti, ad eccezione dell'output dell'*ungrib*, in formato *netCDF* [6].

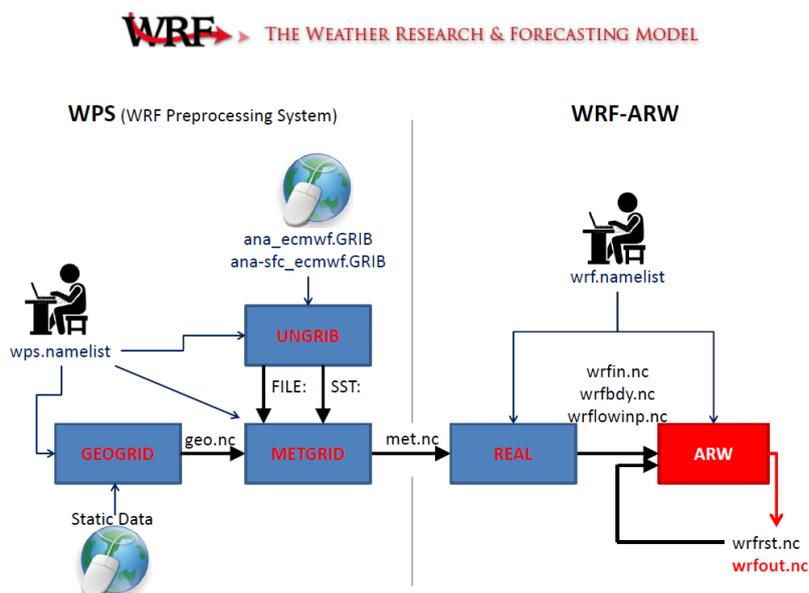


Figura 2.1: Schema a blocchi di WPS e WRF, sono riportati solo gli elementi coinvolti nell'ambito del progetto NAUSICA.

2.2 Condizioni al contorno: analisi ECMWF

Le simulazioni NAUSICA sono guidate da condizioni al contorno di tipo analisi prodotte da ECMWF [5]. Queste sono create guidando la simulazione del modello meteorologico tramite assimilazione di tutte le misure disponibili (stazioni, radar, satelliti, palloni sonda, boe, lidar, ecc).

Come indicato in Figura 2.1, i file GRIB di analisi ECMWF disponibili per il progetto NAUSICA sono di due tipi:

ana_ecmwf contenenti le condizioni al contorno su tutti i livelli (in atmosfera e nel suolo) con risoluzione 0.5°

ana-sfc_ecmwf contenenti le condizioni al contorno sulla superficie a risoluzione 0.1° . In particolare contengono i campi: temperatura a 2 metri, pressione a livello del suolo, pressione a livello del mare, copertura nuvolosa, componenti *u* e *v* dei venti a 10 metri, maschera *land/sea*, temperatura superficiale, temperatura della superficie marina (SST), temperatura del punto di rugiada a 2 metri. Questi ultimi sono stati introdotti come database per l'aggiornamento delle temperature superficiali del mare descritto nel Paragrafo 3.3.3.

Il database NAUSICA (2001-2017) è stato costruito utilizzando solo i file del primo tipo.

2.3 Altri strumenti utilizzati nella pre- e post-elaborazione

In Tabella 2.1 è riportato l'elenco di altri strumenti utilizzati.

Nell'ambito del progetto è stato scelto il GRIB (*GRIdded Binary*) [8], formato standard della *World Meteorological Organization* (WMO) [14] per la trasmissione e l'archiviazione di campi definiti su grigliato regolare in quanto estremamente compatto, veloce da codificare e decodificare, auto-descrittivo e flessibile. Sono stati archiviati anche gli output netCDF nativi, i *wrfout*, che però non sono destaggerati e non contengono alcuni dei campi calcolati in fase di post processamento ed archiviati nei GRIB. Per la creazione dei file GRIB è stato utilizzato l'*Unified Post Processor* (UPP) [8].

I linguaggi utilizzati nell'implementazione dei codici sono BASH [15] e FORTRAN 90 [16]. Per quanto riguarda quest'ultimo sono stati rispettati gli standard internazionali di programmazione [17] prendendo anche spunto dall'applicazione di questi al codice COSMO [18].

2.4 Workflow manager ecFlow

Nell'ambito degli strumenti ha rappresentato una scelta strategica l'utilizzo di un workflow manager [19] che coordina, ad un livello di gestione superiore a tutti gli altri software, la sequenza di operazioni da svolgere e le chiamate a questi. In particolare è stato scelto

Utilizzo	Software
Post elaborazione dati	UPP (<i>Unified Post Processor</i>)
Gestione file netCDF	ncdump , ncview , CDO
Visualizzazione campi da netCDF	ncview
Gestione file GRIB	GRIB API , CDO , WGRIB , acc2dif.exe, miss.exe, grib_api_extraction.exe, mask.exe
Visualizzazione campi da GRIB	GRADS
Analisi dati	R , eseguibili f90 point_in_polygon.exe, f_serie_temp.exe
Visualizzazione analisi dati	R , gnuplot
Estrazione serie temporali misure	estrai.exe, ET_comb_tab.sh, A_ET_comb_tab.sh
Assimilazione dati nel modello	GSI (<i>Gridpoint Statistical Interpolation System</i>)

Tabella 2.1: Elenco degli strumenti di pre e post elaborazione utilizzati nel progetto NAUSICA oltre a WPS e WRF. Gli eseguibili e gli script nominati sono stati prodotti specificamente per il progetto e saranno descritti nei capitoli successivi (Appendici E, G).

ecFlow[20], il workflow manager prodotto da ECMWF [5]. Questo strumento, oltre a consentire appunto una gestione ad un livello organizzativo superiore (*job* automaticamente generati da template e lanciati all'orario stabilito, sequenza tra blocchi logici regolata tramite *cron* e *trigger*) e di conseguenza permettere un'ottimizzazione totale di tempi e risorse (trigger dipendenti dallo stato di completamento delle operazioni precedenti e dipendenti dallo stato di esecuzione di queste) riduce al minimo gli eventuali errori umani nella gestione delle catene. Ha inoltre i vantaggi sia di sfruttare al massimo la modularità dello scripting, ovvero di riutilizzare gli stessi pezzi di codice in tutte le occasioni necessarie al variare dei parametri ambientali, che di fornire un elevato grado di controllo. È infatti pensato per creare e gestire anche il feedback dei processi ed è corredato da un visualizzatore, *l'ecflowview* [20] (o *ecflow_ui* nella nuova versione), caratterizzato da un'elevata leggibilità dello stato di esecuzione che consente di individuare istantaneamente lo stato dei processi.

Grazie a questo strumento, una volta implementata una struttura logica completa (*suite*) costituita da sequenze di blocchi logici che gestiscono parti dei processi di elaborazione in modo coerente, è automatico poterla riutilizzare al variare dei valori dei parametri iniziali senza modificare niente all'interno dei processi.

In Figura 2.2 è riportato un esempio dello stato di esecuzione, visualizzato tramite *ecflowview* [20], di alcune suite che saranno descritte in seguito (Capitolo 4). Anche senza entrare nei dettagli del funzionamento si nota immediatamente che in questo caso c'è stato un problema nel flusso all'interno del nodo di colore rosso. Si possono inoltre osservare indicatori di stato in percentuale di un processo o indicatori che si accendono al verificarsi di una determinata condizione (tecnicamente detti *meter* ed *event*).

2.5 Metodo di lavoro: TRAC, Jenkins e GIT

Il metodo di lavoro, definito come insieme delle operazioni di attribuzione dei compiti di sviluppo e produzione, controllo ed auto-controllo nello svolgimento di questi, verifica dello stato del progetto in ogni suo obiettivo, versionamento dei codici ed archiviazione dei

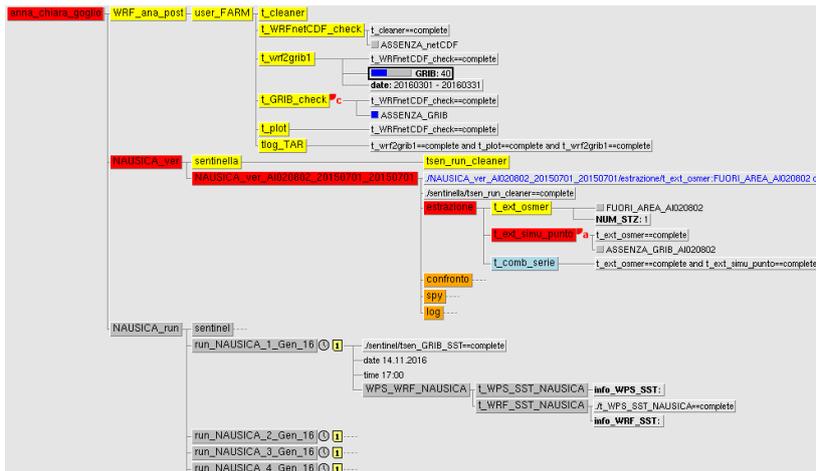


Figura 2.2: Esempio dello stato di esecuzione, visualizzato tramite eflowview, di alcune suite *ecFlow* relative al progetto. Anche senza entrare nei dettagli del funzionamento del software si nota che in questo run c'è stato un problema nel flusso all'interno del nodo di colore rosso, i nodi gialli sono quelli la cui esecuzione è terminata, quelli grigi sono stati caricati ma non mandati in esecuzione, quelli azzurri aspettano il segnale del trigger per attivarsi e quelli arancioni sono stati sospesi. Si possono inoltre osservare indicatori di stato in percentuale di un processo (GRIB in figura) o indicatori che si accendono al verificarsi di una determinata condizione (ASSENZA netCDF e ASSENZA GRIB in figura rispettivamente acceso e spento), definiti *meter* ed *event*.

dati, documentazione dell'attività e svolgimento di test di base sui codici prodotti, merita un paragrafo a sé in quanto, nell'ambito del progetto NAUSICA, è stato gestito tramite due software, TRAC [21] e JENKINS [22], che hanno contribuito a caratterizzare il progetto.

Il TRAC consente infatti di tracciare l'attività svolta da chiunque abbia contribuito al progetto tramite l'assegnazione di *ticket*: compiti specifici, completi di data di scadenza, che devono essere accettati dalla persona indicata per lo svolgimento, svolti ed infine chiusi al termine dell'attività specifica. Questo permette di ottimizzare la programmazione del lavoro e di poter ricostruire a posteriori la cronologia dell'intero sviluppo tramite diagrammi di GANTT che indicano la percentuale di completamento di ciascun obiettivo di questi. Nell'esempio di diagramma di GANTT riportato in Figura 2.3, corrispondente allo stato di avanzamento di NAUSICA a Dicembre 2016, si osserva ad esempio l'interruzione delle attività avvenuta ad Ottobre per problemi contrattuali e la successiva ripresa in Novembre solo di alcune sezioni del progetto, questo a causa della mancanza di risorse disponibili per il completamento di tutte le attività, in particolare della creazione del database.

Il repository GIT [23] associato al progetto traccia, parallelamente a ticket e documentazione, l'evoluzione dei codici implementati versionandoli consentendo facilmente il confronto tra versioni diverse, come mostrato nell'esempio di Figura 2.4.

Per la documentazione completa delle attività è inoltre presente una pagina WIKI-CRMA [1] sull'intranet dell'agenzia costantemente aggiornata durante lo svolgimento del progetto.

Il software JENKINS [22] consente invece di eseguire test di base automatici sui codici versionati per controllarne la sintassi essenziale

(per esempio la presenza del comando per la gestione dell'esecuzione in caso di errori "set -e" negli script in BASH [15]). Anche in questo caso lo stato dei test è reso graficamente evidente dall'interfaccia del software.

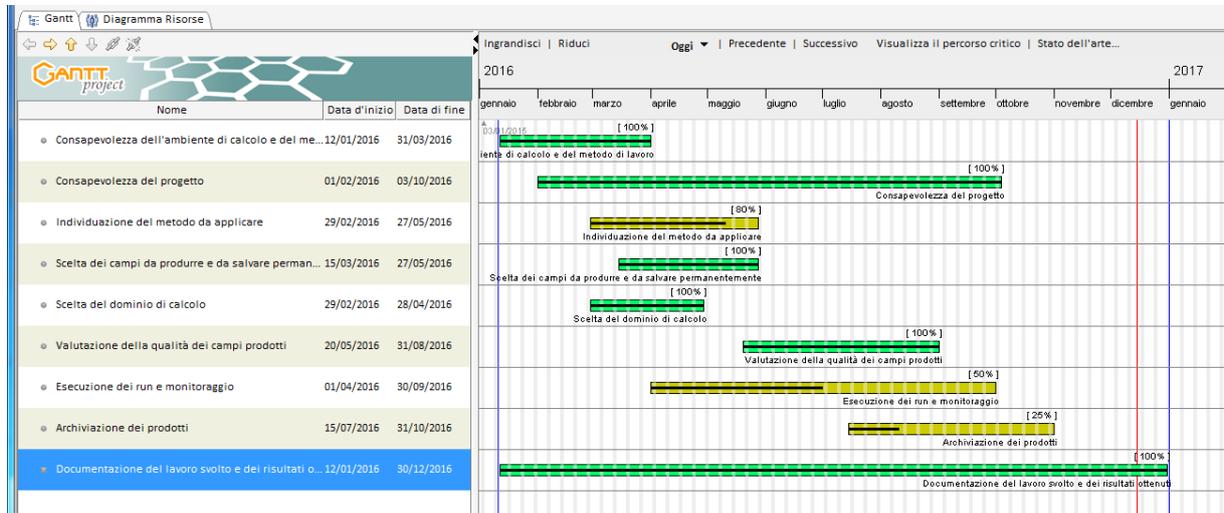


Figura 2.3: Diagramma di GANTT relativo allo stato di avanzamento di NAUSICA a Dicembre 2016. Si possono notare ad esempio l'interruzione delle attività avvenuta ad Ottobre per problemi contrattuali e la successiva ripresa a fine mese solo di alcune sezioni del progetto a causa della mancanza di risorse disponibili per il completamento di tutte le attività, prima tra queste la produzione del database.

	Non modificato	Aggiunto	Rimosso
637	696		
638	697		
639			
640	698		
640	699		

```

IF ( es /= 0 ) THEN
CALL grib_get_error_string(es,error_str)
PRINT*,error_str
WRITE(UNIT=*,FMT='(A)',error_str
END IF
    
```

Figura 2.4: Esempio di utilizzo del repository GIT per confrontare una parte di codice tra due versioni di questo. Come si può osservare questo strumento è caratterizzato da un'ottima leggibilità delle modifiche.

2.6 Risorse di calcolo disponibili

Le risorse di calcolo disponibili influiscono, tramite i tempi di esecuzione dei processi, sia sull'implementazione della struttura dei codici, richiedendo ad esempio un'analisi per la determinazione dei software più performanti, che sulla scelta delle parametrizzazioni della fisica descritta dal modello che devono rappresentare il miglior compromesso tra accuratezza del risultato e velocità di calcolo. Il cluster di calcolo regionale, la FENICE [10], dispone attualmente di tre code di calcolo: arpa, hp e Julia le cui caratteristiche tecniche fondamentali sono le seguenti:

ARPA 2 nodi di calcolo 64 core Intel 2.10 GHz 63Gb RAM connessi al master node con rete *Ethernet*

HP 4 nodi di calcolo 32 core Intel 2.30 GHz 126Gb RAM connessi al master node con rete *InfiniBand*

JULIA 8 nodi di calcolo 40 core Intel 2.10 GHz 126Gb RAM connessi al master node con rete *InfiniBand*

Per la prima parte di attività del progetto è stata utilizzata la coda HP (in alternanza con l'attività operativa) in quanto più veloce della coda ARPA grazie alla rete *Infiniband*.

Grazie all'attivazione della coda JULIA a completa disposizione del progetto NAUSICA, disponibile dal 15 Marzo 2017, i tempi di produzione dei database sono stati ridotti di circa un terzo.

3

Configurazione dei run

I flussi di lavoro relativi al progetto sono gestiti tramite il settaggio di vari parametri la cui configurazione caratterizza i run eseguiti. Questi parametri regolano il flusso sia dal punto di vista computazionale che della fisica riprodotta. Da questi dipendono infatti variabili come il tempo di esecuzione e le risorse utilizzate ma anche il tipo di parametrizzazione scelta tra quelle disponibili nel modello per descrivere i vari processi simulati.

La definizione della configurazione da adottare per il run definitivo ha rappresentato una delle attività fondamentali nello sviluppo del progetto ed ha condotto, oltre che alla generazione di simulazioni ad elevato grado di accuratezza, ad una maggiore consapevolezza del funzionamento dei codici.

Come settaggio iniziale è stato scelto quello attualmente utilizzato per le simulazioni di previsione operative. La definizione finale si discosta però dalla configurazione operativa, in alcuni casi rappresentandone una possibile evoluzione futura, per quanto riguarda l'estensione del dominio regionale, la scelta del modello di suolo e gestione delle temperature superficiali del mare da utilizzare nel WRF e per il settaggio di alcuni parametri relativi all'archiviazione delle grandezze cumulate in UPP e WRF. In Appendice A sono riportate le namelist di WPS e WRF utilizzate per la produzione attualmente disponibile. Il processo di configurazione è stato caratterizzato da una fase iniziale di test preliminari, atti appunto allo studio degli effetti ottenuti sui run al variare dei parametri utilizzati, ma anche da una seconda fase, non programmata, rappresentata dall'analisi di problemi riscontrati in primi tentativi di produzioni e dalla ricerca delle relative soluzioni. Per questa seconda fase del processo è stato fondamentale l'utilizzo della suite NAUSICA_ver, descritta nel Capitolo 4, costruita proprio al fine di valutare quantitativamente i prodotti generati.

Un ulteriore effetto ottenuto è stato quello di evidenziare alcuni limiti nell'utilizzo attuale del WRF indicando una possibile direzione per il miglioramento, in particolare ci si riferisce alla necessità di aggiornamento di alcuni campi, nel caso specifico relative alla temperatura del mare, messa in luce dall'analisi descritta nel Paragrafo 3.3.3.

3.1 Scalabilità del modello

Il primo test preliminare eseguito è l'analisi della scalabilità del modello WRF sulla coda *hp* allo scopo di ottimizzare l'utilizzo della coda di calcolo disponibile. Il risultato, riportato in Figura 3.1, mostra che i run del WRF risultano più performanti all'aumentare dei nodi di calcolo utilizzati. È stato quindi scelto, nella prima fase di lavoro, di utilizzare tre, e quando possibile anche quattro, nodi di calcolo della coda *hp* in alternanza con l'attività operativa, che ne usa altrettanti, programmando quindi le simulazioni tra le 17:00 UTC e le 01:00 UTC di ogni giorno.

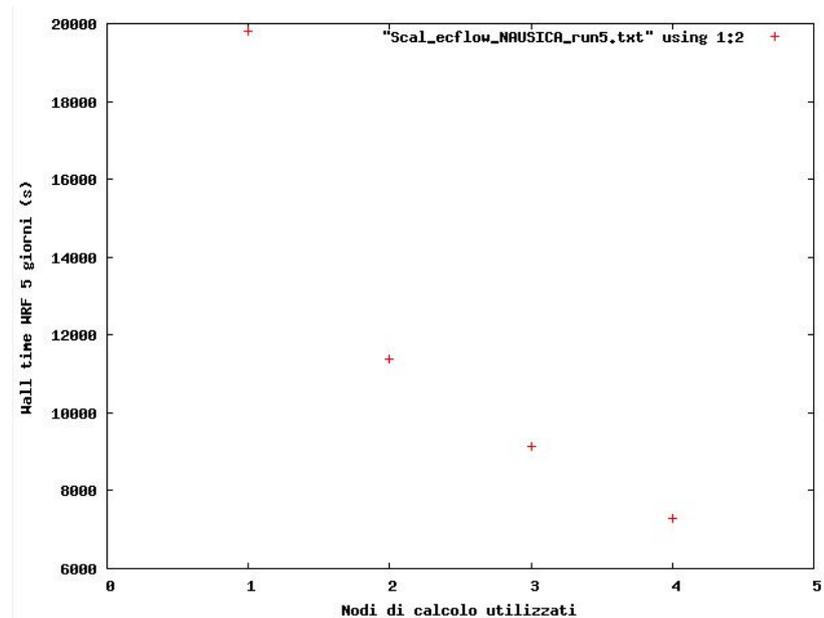


Figura 3.1: Scalabilità del modello WRF relativa alla coda *hp* eseguita con la configurazione operativa attuale, ovvero prima dell'estensione del dominio regionale descritta nel Paragrafo 3.2 e delle modifiche ai parametri descritte nel Paragrafo 3.3. Si può osservare che il modello scala bene all'aumentare delle risorse di calcolo utilizzate, ovvero che il *wall-time* è inversamente proporzionale al numero di nodi di calcolo utilizzati.

Per quanto riguarda la coda *Julia*, utilizzata per la produzione della maggior parte del database, non è ancora stato fatto uno studio sistematico della scalabilità. Sono stati eseguiti alcuni test sui tempi di simulazione, di cui si riportano i risultati in Tabella 3.1, che hanno portato alla decisione di riservare 180 processori (30 processori su 6 nodi) per le simulazioni del WRF.

Nodi	Processori	Δt [s] per simulazione di 10 giorni
6	20	≈ 9 h
6	30	≈ 6 h
5	30	> 6 h
6	40	CRASH

Tabella 3.1: Prove dei tempi di calcolo della coda *Julia* per simulazioni WRF di 10 giorni, da questi risultati si è deciso di riservare 180 processori (30 per sei nodi) per le simulazioni del WRF.

L'andamento della curva in Figura 3.2 che riporta i tempi di simulazione in funzione del periodo simulato, indica stabilità nei tempi di produzione.

Il wall-time del WRF su 30 processori per 6 nodi della coda *Julia* è quindi confrontabile con quello ottenuto utilizzando 4 nodi da 32 core sulla coda HP.

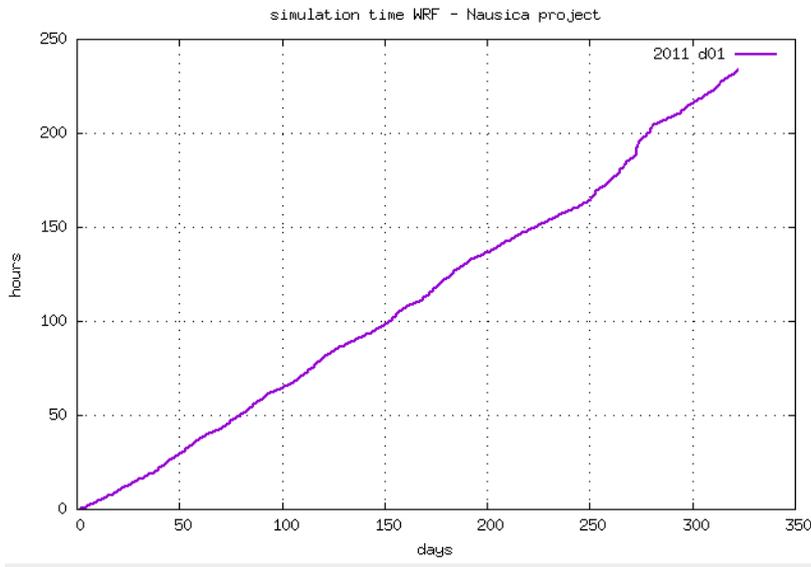


Figura 3.2: Ore di simulazione all'aumentare del numero di giorni simulato con il WRF utilizzando 30 processori per 6 nodi. Si nota stabilità nei tempi di produzione.

3.2 Estensione del dominio regionale

L'impostazione utilizzata per il modello WRF evolve temporalmente le condizioni meteorologiche su tre domini innestati: uno continentale (Europeo), uno nazionale (Italiano) e uno regionale ad alta risoluzione (2 km), centrato sul Friuli Venezia Giulia ma contenente anche Veneto, Trentino Alto Adige e parte di Slovenia, Austria, Germania e Croazia. I domini esterni, con rispettive risoluzioni di 50 km e 10 km, sono necessari per raccordare la sinottica descritta dalle condizioni al contorno alla scala del dominio regionale ad alta risoluzione.

In Tabella 3.2 sono riportate le risoluzioni spaziali e temporali e gli intervalli temporali di scrittura per i tre domini, in Figura 3.3 sono invece rappresentati i confini spaziali dei tre domini utilizzati nei flussi operativi, cioè come erano prima dell'estensione descritta.

Dominio	Risoluzione spaziale [km]	Risoluzione temporale [s]	Intervalli temporali di salvataggio [h]
Continentale	50	200	6
Nazionale	10	40	3
Regionale	2	8	1

Tabella 3.2: Risoluzione spaziale e temporale e intervallo temporale di salvataggio dei tre domini innestati utilizzati.

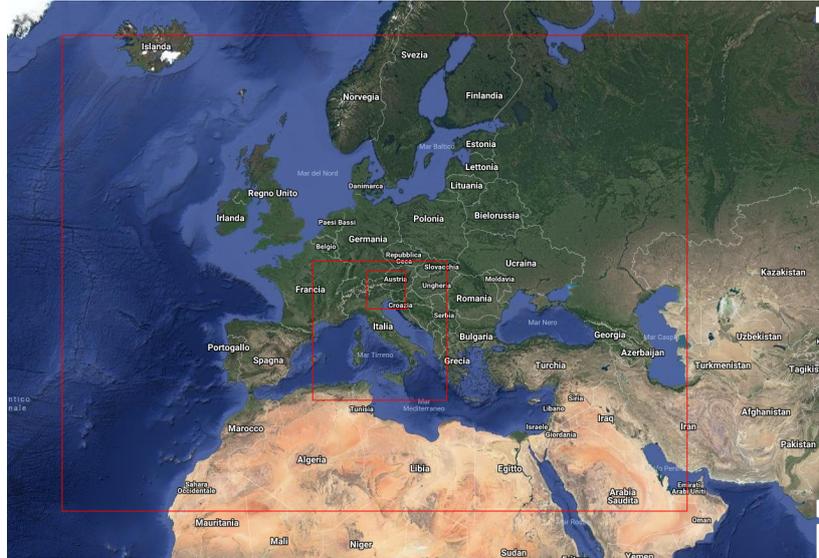


Figura 3.3: Confini spaziali dei tre domini utilizzati dai flussi operativi (prima dell'estensione del dominio regionale descritta nel Par. 3.2)

Il secondo test preliminare ha riguardato l'estensione spaziale del dominio regionale: si è cercato un compromesso accettabile tra area coperta da alta risoluzione e tempi di calcolo.

Di seguito sono elencate le varie fasi di estensione, le linee in Figura 3.4 ne rappresentano i confini e in Tabella 3.3 ne sono riassunte le caratteristiche:

1. Il primo run del test era caratterizzato da un dominio regionale di 91×91 punti di griglia, corrispondenti a $180 \text{ km} \times 180 \text{ km}$, centrato sul Friuli Venezia Giulia (quello attualmente operativo) che, per simulare 10 giorni (su 4 nodi da 32 core della coda *hp* o 6 nodi da 30 core sulla coda *Julia*) con la configurazione attuale dell'operative, impiega circa $\Delta t \approx 2 \text{ h } 30'$ (linea rossa in Figura 3.4).
2. La prima estensione, fatta per contenere anche Veneto, Trentino Alto Adige e la maggior parte della Slovenia, ha un'area che corrisponde a 201×141 punti di griglia, ovvero $400 \text{ km} \times 280 \text{ km}$. Questa estensione ha però il difetto di avere la chiusura superiore e parte di quella occidentale in corrispondenza dell'arco alpino, cosa che è preferibile evitare in quanto le zone montuose sono interessate da fenomeni locali molto più rilevanti rispetto a zone topograficamente omogenee e di conseguenza sconsigliata come bordo dove leggere le condizioni al contorno fornite dal dominio superiore (linea verde in Figura 3.4). In queste condizioni una simulazione di 10 giorni (su 4 nodi da 32 core della coda *hp* o 6 nodi da 30 core sulla coda *Julia*) con la configurazione dell'operative, impiega circa $\Delta t \approx 4 \text{ h}$.
3. L'estensione scelta, 201×201 punti di griglia corrispondenti a $400 \text{ km} \times 400 \text{ km}$, contiene anche l'Austria centro-meridionale, l'Istria ed una parte della Germania sud-orientale ed impiega circa $\Delta t \approx 5 \text{ h}$ per simulare 10 giorni (su 4 nodi da 32 core della coda *hp*

o 6 nodi da 30 core sulla coda *Julia*) con la configurazione attuale dell'operative (linea rossa in Figura 3.4). Resta, inevitabilmente, una parte di chiusura che attraversa l'arco alpino ma si è ritenuto che questa fosse sufficientemente distante dall'area di maggior interesse da renderne trascurabili gli effetti.



Figura 3.4: Fasi di estensione del dominio regionale. La linea blu corrisponde all'area attualmente utilizzata nelle simulazioni operative, la linea verde al primo tentativo di estensione e la linea rossa al dominio scelto. Le caratteristiche dei tre domini sono riassunte in Tabella 3.3.

Dominio	Punti di griglia	$\Delta X \times \Delta Y$ [km ²]	Area contenuta	Δt WRF
do3 operatività	91 x 91	180 x 180	Friuli Venezia Giulia	≈ 2h 30'
1° estensione do3	201 x 141	400 x 280	Friuli Venezia Giulia Istria (Croazia) Veneto Trentino Alto Adige Slovenia centro/occidentale Austria meridionale	≈ 5h
2° estensione do3	201 x 201	400 x 400	Friuli Venezia Giulia Istria (Croazia) Veneto Trentino Alto Adige Slovenia centro/occidentale Austria centro/meridionale Germania sud/orientale	≈ 6h

Tabella 3.3: Caratteristiche dei tre domini utilizzati nei test per la definizione del dominio regionale. I tempi riportati corrispondono a run del WRF di 10 giorni in configurazione operativa su 4 nodi da 32 core ciascuno.

3.2.1 Caratteristiche dei campi generati sul dominio nazionale

Oltre ai campi generati dal modello sul dominio regionale descritto nel Paragrafo 3.2, il database del progetto NAUSICA contiene anche

i campi relativi al dominio nazionale prodotti dal modello. È però da sottolinearsi che, a causa dell'attivazione del *feedback*, opzione del WRF che fa sì che le informazioni relative allo stato dei campi vengano comunicate anche tra livello inferiore e superiore e non solo viceversa consentendo di sfruttare al massimo le potenzialità del modello per quanto riguarda la riproduzione dei fenomeni alla mesoscala, alcuni campi relativi ai domini di raccordo archiviati presentano discontinuità in corrispondenza dell'area del dominio contenuto. Un esempio di questa discontinuità, che rende necessaria estrema cautela nell'utilizzo dei campi del dominio nazionale, è osservabile nella mappa della precipitazione riportata in Figura 3.5. La mappa dei venti riportata all'interno della stessa Figura dimostra che, per quanto riguarda questa osservabile, non sono invece presenti effetti così evidenti in corrispondenza dell'area del dominio interno.

3.3 Parametri di WRF e UPP

Il modello WRF è in grado di descrivere fenomeni fisici appartenenti ad un ampio range di ordini di grandezza, per la simulazione dei quali sono implementate al suo interno diverse parametrizzazioni [12]. Purtroppo la documentazione che descrive le caratteristiche di ciascuna di esse, e soprattutto ne analizza i limiti di applicabilità, è scarsa e frammentaria. Per questo motivo la scelta di un settaggio soddisfacente per il tipo di simulazione da svolgere è stato ottenuto osservando i risultati di run preliminari, modificandone di volta in volta i parametri in modo da minimizzare le differenze dei prodotti delle simulazioni rispetto alle misure a disposizione.

Se da un lato, con il processo fatto, si è ottenuto un accordo ritenuto soddisfacente ai fini del progetto, dall'altro restano sicuramente ancora da analizzare gli effetti del variare di molti dei parametri modificabili che, nei prodotti per ora disponibili, sono impostati secondo la configurazione attualmente operativa.

Le varie fasi del processo descritto sono trattate nei paragrafi seguenti descrivendo i test preliminari svolti e i vari problemi riscontrati nei run preliminari le cui relative soluzioni hanno portato alla ridefinizione di determinati parametri.

3.3.1 Grandezze cumulate e differenziali

Limite massimo nella scrittura dell'intervallo di integrazione (UPP)

Il primo problema riscontrato nel flusso è rappresentato da un errore restituito nella fase di post elaborazione generato nella lettura dei file GRIB prodotti per mezzo del software CDO [24]. Questo è risultato essere legato alle grandezze cumulate (pioggia, neve e radiazione) e causato dal fatto che il software di post processamento UPP [8] era settato in modo da destinare alla scrittura dell'intervallo di integrazione un unico byte, imponendo quindi un limite di 256 ore oltre il quale il valore di questo veniva azzerato.

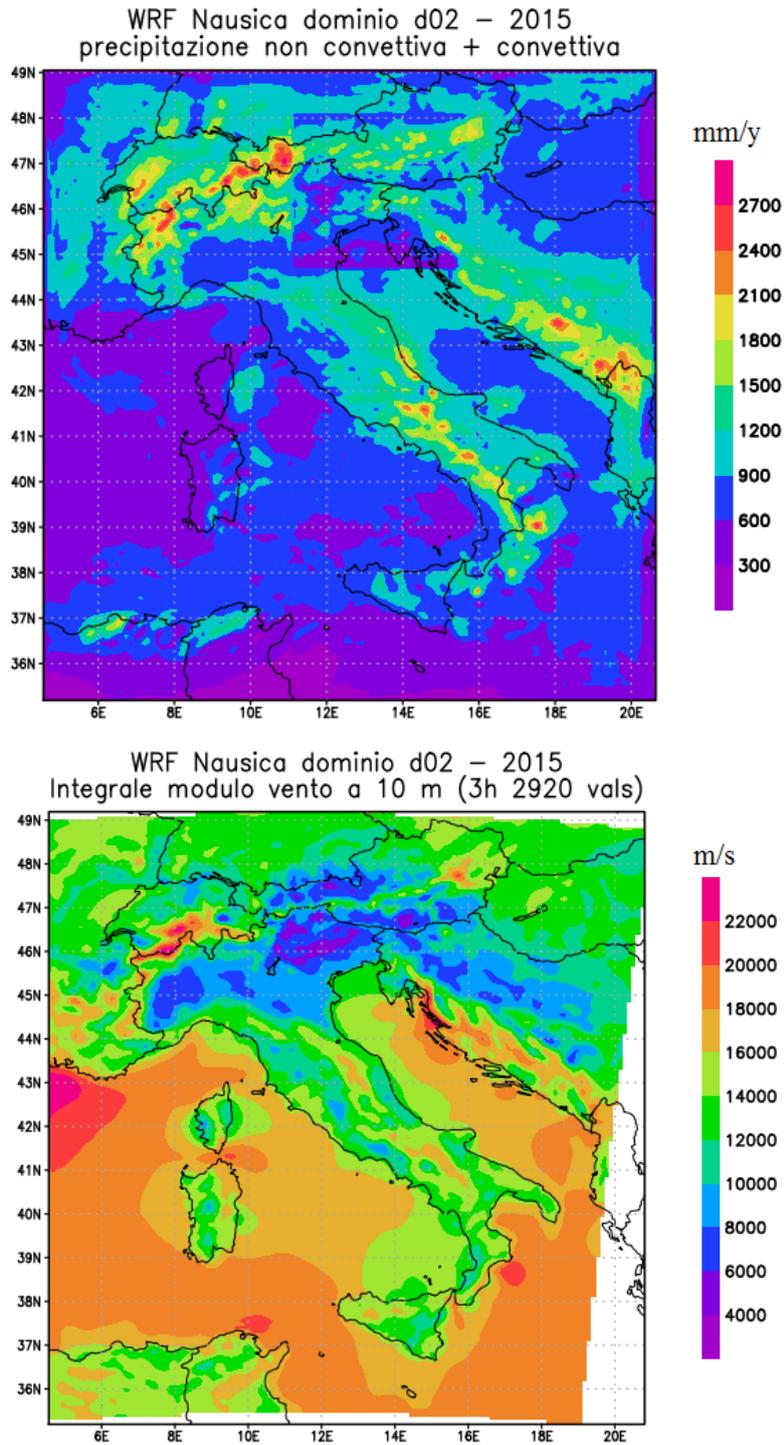


Figura 3.5: Campi di precipitazione cumulata (in alto) e vento a 10 metri (in basso) relativi al dominio nazionale d02. Si nota come, nel caso della precipitazione, i campi del database relativo al dominio d02 contengano un'evidente discontinuità in corrispondenza dell'area del dominio interno d03. Questo è dovuto all'opzione del WRF di *feedback* e rende necessaria estrema cautela nell'utilizzo dei campi generati dal modello per quanto riguarda i domini di raccordo.

In Figura 3.6 è riportata la descrizione del contenuto di alcuni record di un GRIB affetti da questo problema: si può appunto notare che l'intervallo temporale di integrazione attribuito alle grandezze cumulate è diverso dall'indicazione del tempo di forecast (cioè della distanza in ore dall'istante di inizializzazione della simulazione) di tutte le altre.

```

136:9276968:d=15010100:TMP:1000 mb:4812hr fcst:NAve=0
137:9336866:d=15010100:POT:1000 mb:4812hr fcst:NAve=0
138:9396764:d=15010100:RH:1000 mb:4812hr fcst:NAve=0
139:9466626:d=15010100:UGRD:1000 mb:4812hr fcst:NAve=0
140:9526524:d=15010100:VGRD:1000 mb:4812hr fcst:NAve=0
141:9586422:d=15010100:HGT:sfc:4812hr fcst:NAve=0
142:9666250:d=15010100:TMP:sfc:4812hr fcst:NAve=0
143:9751060:d=15010100:WEASD:sfc:4812hr fcst:NAve=0
144:9805974:d=15010100:TMP:2 m above gnd:4812hr fcst:NAve=0
145:9885802:d=15010100:RH:2 m above gnd:4812hr fcst:NAve=0
146:9955664:d=15010100:PRES:2 m above gnd:4812hr fcst:NAve=0
147:10040474:d=15010100:UGRD:10 m above gnd:4812hr fcst:NAve=0
148:10100372:d=15010100:VGRD:10 m above gnd:4812hr fcst:NAve=0
150:10165360:d=15010100:PRATE:sfc:4812hr fcst:NAve=0
151:10170450:d=15010100:CPRAT:sfc:0-204hr ave:NAve=0
152:10175540:d=15010100:PRATE:sfc:0-204hr ave:NAve=0
153:10180630:d=15010100:APCP:sfc:0-204hr acc:NAve=0
154:10260458:d=15010100:ACPCP:sfc:0-204hr acc:NAve=0
155:10265548:d=15010100:NCPCP:sfc:0-204hr acc:NAve=0

```

Figura 3.6: Descrizione del contenuto di alcuni record di un GRIB relativo ad un'unica scadenza (ottenuto tramite `wgrib`) affetto dal problema causato dall'insufficiente numero di ottetti destinati da UPP alla scrittura dell'intervallo temporale di integrazione. Si nota infatti che l'intervallo attribuito alle grandezze cumulate (2014 h) è diverso dall'indicazione del tempo di forecast (4812 h) di tutte le altre.

Il problema è stato risolto modificando il settaggio dell'UPP (in particolare la routine `UPPV2.2/src/unipost/GRIBIT.f`) in modo da destinare anche all'intervallo di integrazione delle grandezze cumulate due ottetti: il 19 e il 20 della sezione 1. L'informazione di destinazione di utilizzo di questi due byte è contenuta nell'ottetto successivo, il 21 della sezione 1, che deve essere settato a 10.

Si è osservato inoltre che mentre la versione utilizzata di UPP, la 2.2, consente questa modifica la successiva, V 3.0, non la permette.

Limite massimo del valore delle grandezze cumulate (WRF)

Un'altra modifica necessaria inerente le grandezze cumulate ha riguardato la presenza di un limite massimo, fissato di default, nel valore accumulato all'interno del modello WRF stesso, evidente in Figura 3.7.

Questo è stato risolto settando nella namelist del WRF il valore dei parametri `bucket_mm` e `bucket_J`, relativi rispettivamente a pioggia e radiazione, ad un valore di `-1` che corrisponde appunto all'assenza di un limite massimo di accumulazione. In Appendice I è riportata la descrizione dettagliata del processo di risoluzione relativa ai problemi sulle grandezze cumulate.

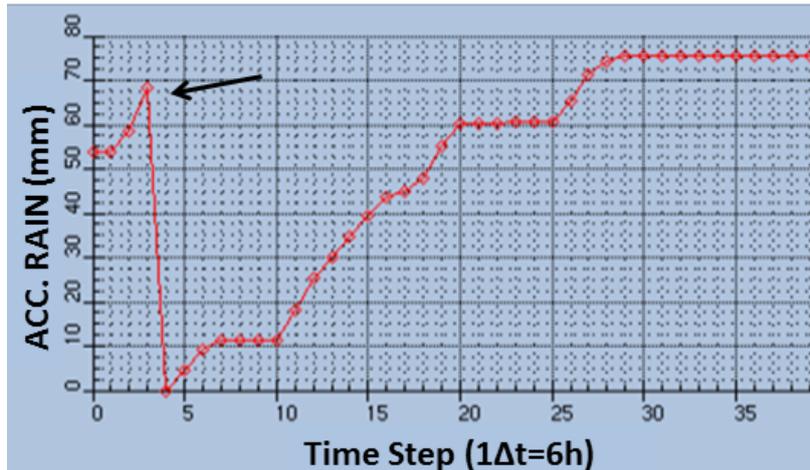


Figura 3.7: Serie temporale della pioggia cumulata [mm] in unità di *time-step* corrispondenti a 6 ore per unità letta dai file *wrfout.nc* nel caso in cui il limite massimo del valore di questa sia quello di default del WRF (100 mm per il *do1*). Si osserva infatti che in corrispondenza del superamento di questa soglia, indicato dalla freccia nera, il valore viene resettato a 0.

Produzione di grandezze differenziali tramite post-elaborazione

Essendo interessati per quanto riguarda le precipitazioni a confrontare le grandezze differenziali piuttosto che le cumulate, è stato implementato un eseguibile FORTRAN [25] (*acc2dif.exe*) che, utilizzando le GRIB API [26], calcola il campo differenziale scrivendolo in un GRIB con preambolo *PREC*. Al fine di valutarne il funzionamento sono stati confrontati i risultati ottenuti con le misure a disposizione, in Figura 3.8 è riportato un esempio di confronto che mostra un ottimo accordo.

Produzione diretta di GRANDEZZE ORARIE per le precipitazioni nel modello WRF

Il WRF dispone della possibilità di produrre direttamente campi di precipitazione differenziali aggiungendo nella namelist l'opzione *prec_acc_dt = 360,180,60* (dove i tre valori rappresentano gli intervalli di accumulazione per ciascun dominio). Questa modifica fa sì che vengano prodotti tre nuovi campi negli output: *prec_acc_c*, *prec_acc_nc* e *snow_acc_nc*. Per introdurre i rispettivi campi all'interno dei file GRIB è necessario sostituire nel file di parametri dell'UPP [8] i parametri *acm_total_precip* e *acm_total_precip* rispettivamente con *bucket_total_precip* e *bucket_conv_precip*. Si è deciso quindi di utilizzare questa opzione per i run definitivi del progetto. In Figura 3.9 si possono osservare la serie temporale della pioggia cumulata prodotta dal WRF e la descrizione del contenuto di alcuni record GRIB derivanti dal post-processamento dei *wrfout* mentre in Figura 3.8 si osserva che anche questo metodo produce risultati coerenti con le misure.

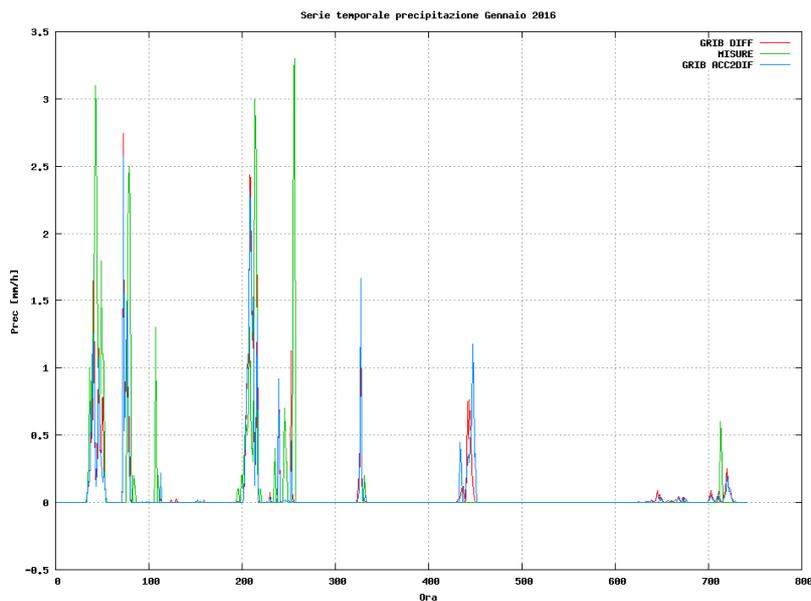


Figura 3.8: Confronto qualitativo tra le serie della precipitazione differenziale in Gennaio 2016 delle misure (linea verde), dei campi differenziali ottenuti tramite l'eseguibile *acc2diff.exe* a partire dai GRIB contenenti i valori cumulati (linea blu) e del campo differenziale prodotto direttamente dal WRF (linea rossa). Si osserva un accordo ritenuto soddisfacente.

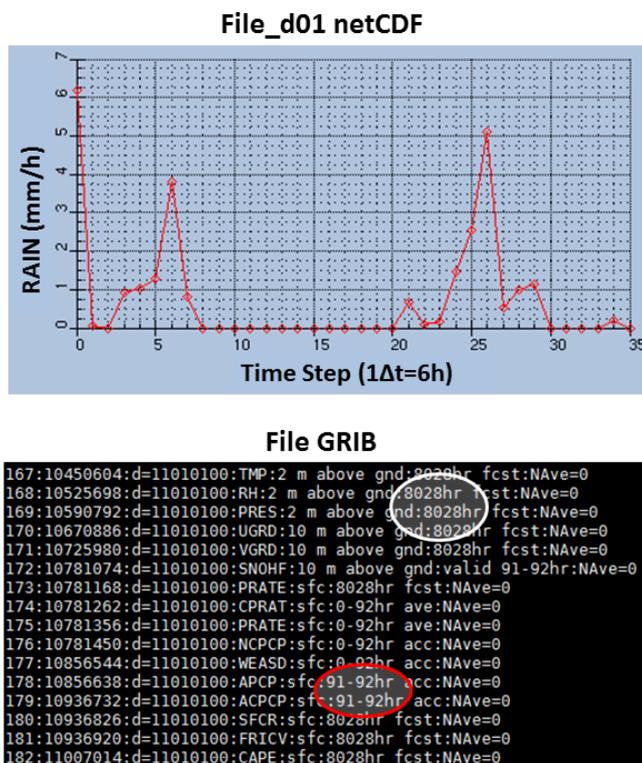


Figura 3.9: Serie temporale della pioggia cumulata prodotta dal WRF per il d01 e descrizione del contenuto di alcuni record GRIB derivanti dal post-processing dei wrfout per il d03. Si osserva che come atteso le grandezze relative alla precipitazione sono differenziali.

3.3.2 Coordinate dei campi nei file GRIB

In questo paragrafo è descritto un altro problema riscontrato legato all'UPP [8] per risolvere il quale non è però stato necessario modificare alcun parametro. L'esecuzione del *copygb* (eseguibile di UPP con il compito di destaggarare i campi negli output del modello ed eventualmente cambiare coordinate) restituisce un errore relativo al campo della *dew point temperature* (temperatura del punto di rugiada). Si è infatti osservato che questo, nei GRIB prodotti, presenta un valore divergente, sicuramente non verosimile, in un punto del dominio continentale. Essendo un campo non ritenuto necessario ai fini dell'utilizzo dei GRIB prodotti dal progetto è stato inizialmente eliminato dal post-processing senza ulteriori analisi sul problema riscontrato. In Appendice J è riportata la descrizione dettagliata del processo di indagine.

Dato che il problema si è successivamente ripresentato anche per altri campi è stato deciso di produrre i campi in proiezione *Lambert Conformal* e non *Lat/Lon* (il problema si presenta nella conversione).

3.3.3 Modello di suolo e aggiornamento delle SST

La principale divergenza di entità non ritenuta trascurabile nell'evoluzione delle condizioni meteorologiche simulate rispetto alle misure a disposizione, è stata riscontrata analizzando il campo delle temperature. Il modello infatti sottostimava questa variabile, in particolare la temperatura media a due metri dal suolo risultava inferiore fino a circa 20 gradi rispetto alle misure. Il problema riguardava tutto il dominio e risultava accentuato durante i mesi caldi e in corrispondenza delle zone di costa e pianura. Un esempio da cui si può dedurre l'entità della sottostima è riportato in Figura 3.10 (istogramma a sinistra), dove è mostrato un confronto tra le distribuzioni relative alla temperatura a due metri dal suolo a Trieste calcolata dal modello e misurata: è evidente che le temperature della simulazione non sono verosimili. Il problema descritto è stato risolto, come si vede confrontando le due distribuzioni in Figura 3.10 prodotte rispettivamente prima e dopo le modifiche, cambiando lo schema di suolo utilizzato dal modello ed introducendo l'aggiornamento delle temperatura superficiale del mare (*Sea Surface Temperature*). Queste modifiche vanno però a scapito dei tempi di calcolo che aumentano del 30% passando da 5 ore a 6.5 ore per la simulazione di 10 giorni (con 4 nodi da 32 core su coda *hp* o 6 nodi da 30 core su coda *Julia*).

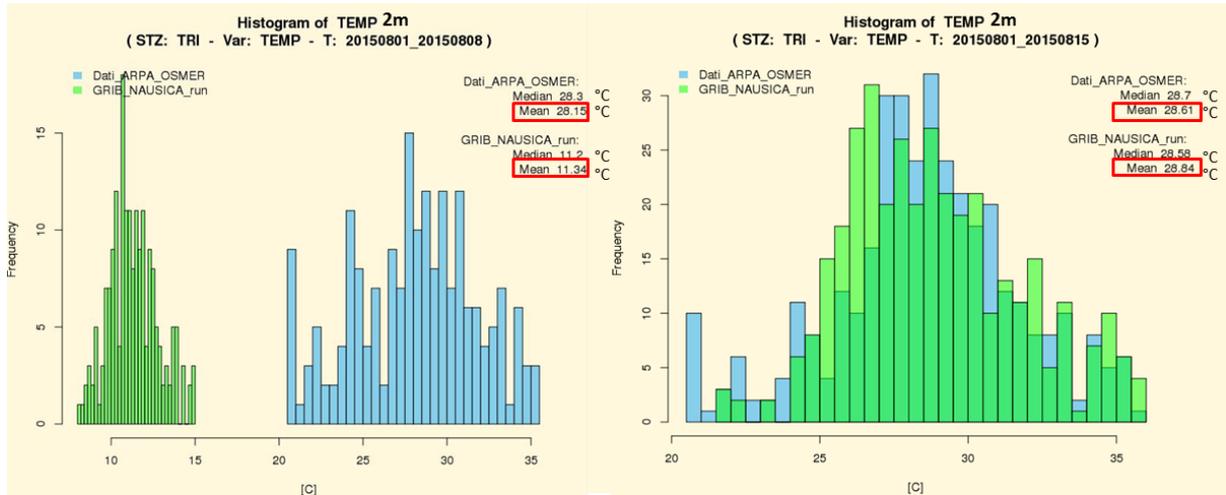


Figura 3.10: Confronto tra gli istogrammi relativi alla temperatura a due metri dal suolo calcolata dal modello e misurata in corrispondenza della stazione ARPA-OSMER di Trieste rispettivamente nella prima settimana di Agosto con la configurazione attualmente utilizzata in operatività: modello di suolo TDS in assenza di aggiornamento SST (sinistra) e nei primi 15 giorni di Agosto con la configurazione scelta per i run del modello NAUSICA: modello di suolo Noah LSM e aggiornamento SST ogni 6 ore (destra). Nel primo caso il modello sottostima la temperatura media di circa 17 °C, inoltre la temperatura massima raggiunta dal modello è inferiore di circa 5°C rispetto alla minima misurata mentre con la seconda parametrizzazione utilizzata si è ottenuta una simulazione con valori delle temperature verosimili. La frequenza che compare in ordinata corrisponde al numero di occorrenze di valori appartenenti al range della classe dell'istogramma tra le scadenze orarie disponibili sul periodo selezionato.

Modelli di suolo TDS e Noah LSM

Il suolo, all'interno del WRF, è modellizzato con diversi livelli sottostanti la superficie il cui strato inferiore, in assenza di flag aggiuntivi, è definito a temperatura costante.

Il primo schema di suolo testato nelle simulazioni (il quello attualmente adoperato nella configurazione operativa del WRF): il *Thermal Diffusion Scheme* (TDS) che include soltanto la conduzione di calore tra strati contigui. Questo schema considera quattro livelli con profondità massima di 23 cm al di sotto della superficie (indicata dalla linea rossa in Figura 3.11) mentre gli altri sono collocati a 2 cm, 5 cm e 11 cm da questa.

La parametrizzazione impostata per il suolo nel run definitivo, la cui scelta verrà giustificata nelle pagine seguenti, è indicata in letteratura come il *Noah Land Surface Model* (Noah LSM) [27] e introduce vari parametri come l'umidità del suolo e la presenza di vegetazione, oltre a descrivere la conduzione di calore tra i diversi strati. Il modello e i fenomeni fisici da esso presi in considerazione sono rappresentati nello schema riportato in Figura 3.11. Una caratteristica che lo differenzia dal TDS è la profondità dello strato inferiore che raggiunge i 2 metri. I livelli del sottosuolo in questo caso sono cinque.

Le condizioni iniziali del suolo sono anch'esse interpolate a partire dai GRIB di analisi ECMWF descritti nel Paragrafo 2.2, i livelli di suolo contenuti in queste sono quattro, corrispondenti alle fasce di profondità: 0-7 cm, 7-28 cm, 28-100 cm, 100-255 cm.

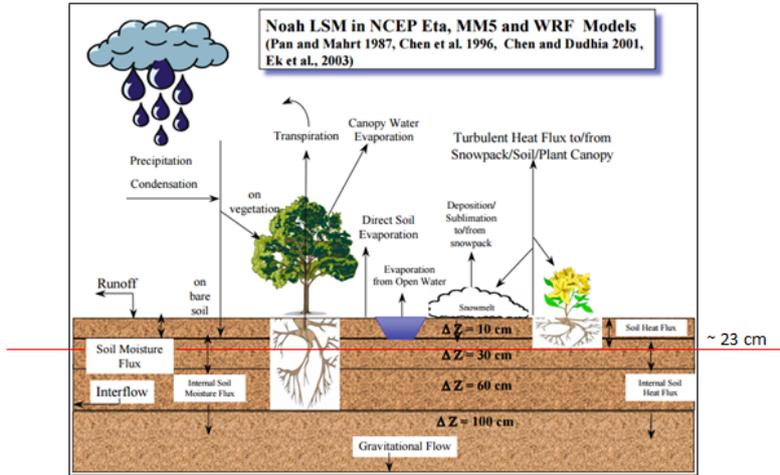


Figura 3.11: Schema del *Noah Land Surface Model*: il modello di suolo utilizzato per le simulazioni del progetto. La linea rossa corrisponde alla profondità massima raggiunta dall'analisi precedentemente utilizzata: il *Thermal Diffusion Scheme* che, a differenza del Noah LSM, considerava unicamente la conduzione di calore tra i diversi livelli.

In Figura 3.12 sono riportate le serie temporali estratte dai GRIB ECMWF della temperatura alle quattro profondità relative all'intero 2015, si può notare che, come atteso, l'ampiezza di oscillazione e la frequenza di variazione della temperatura aumentano avvicinandosi alla superficie.

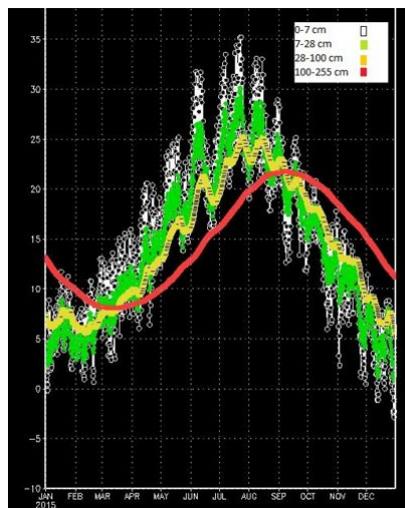


Figura 3.12: Serie temporali delle temperature in °C del suolo nei GRIB di analisi ECMWF (fonte delle condizioni iniziali ed al contorno della simulazione) alle quattro profondità descritte (fasce di profondità: 0-7 cm, 7-28 cm, 28-100cm, 100-255cm) relative all'intero 2015. Come atteso l'ampiezza di oscillazione e la frequenza di variazione della temperatura aumentano avvicinandosi alla superficie.

Aggiornamento delle (SST) e campi associati

A differenza del suolo, il mare è descritto nel modello WRF come una superficie a temperatura costante. È possibile settare il modello in modo che questa venga aggiornata con frequenza scelta dall'utente a partire dalle informazioni contenute in file di condizione al contorno superficiale come gli *ana-sfc* descritti nel Paragrafo 2.2.

Il processo di aggiornamento delle SST è illustrato in Figura 3.13, dove sono sottolineati in verde passaggi e prodotti intermedi coinvolti nel processo di aggiornamento SST. La prima differenza nella gestione del modello riguarda il WPS: *ungrib.exe* deve essere eseguito due volte, una per i GRIB *ana* (che genera file intermedi di prefisso *FILE:*) e una per gli *ana-sfc* (che genera file intermedi di prefisso *SST:*), il *metgrid* si occupa poi di leggere entrambi questi prodotti. Per ottenere queste operazioni le namelist del WPS e del WRF devono essere modificate [28], [29] come dettagliatamente descritto in Appendice C.

In caso di attivazione di questa opzione il WRF legge il valore del campo non dai file di restart *wrfrst.nc* ma dai file di input *wrflow.nc* generati dal *real*. La frequenza di lettura da file di input viene settata nella namelist del WRF come descritto in Appendice C.

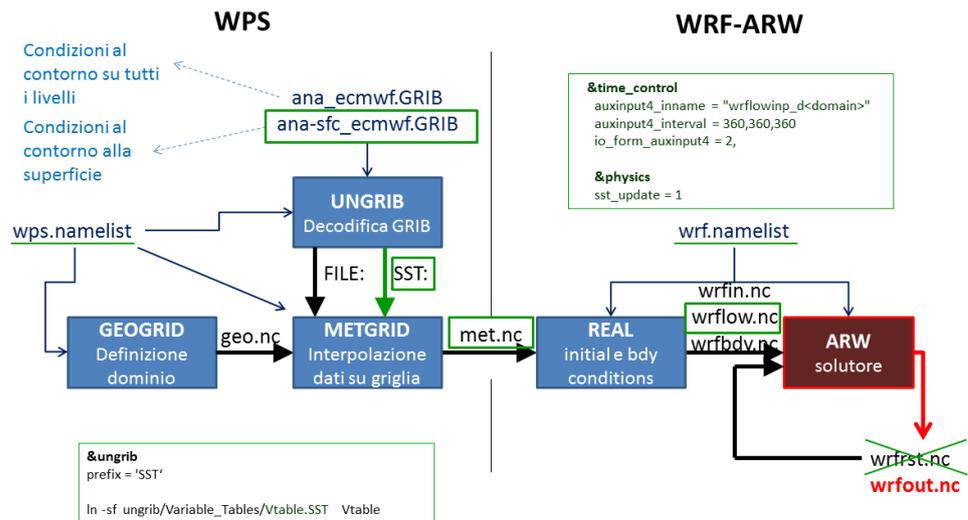


Figura 3.13: Illustrazione del processo di aggiornamento delle SST nello schema del WRF, gli elementi evidenziati in verde sono quelli coinvolti nel processo di aggiornamento.

Associando l'aggiornamento SST ad una parametrizzazione del suolo di tipo *Noah*, si ottiene anche l'aggiornamento di altri campi con la stessa cadenza temporale impostata per la lettura delle SST da file di input [12]. In particolare i campi introdotti ed aggiornati simultaneamente in presenza delle relative informazioni nei file di condizione al contorno sono l'umidità del suolo (i cui effetti sono rappresentati in Figura 3.14), la copertura nevosa, il ghiaccio marino e la frazione stagionale di vegetazione.

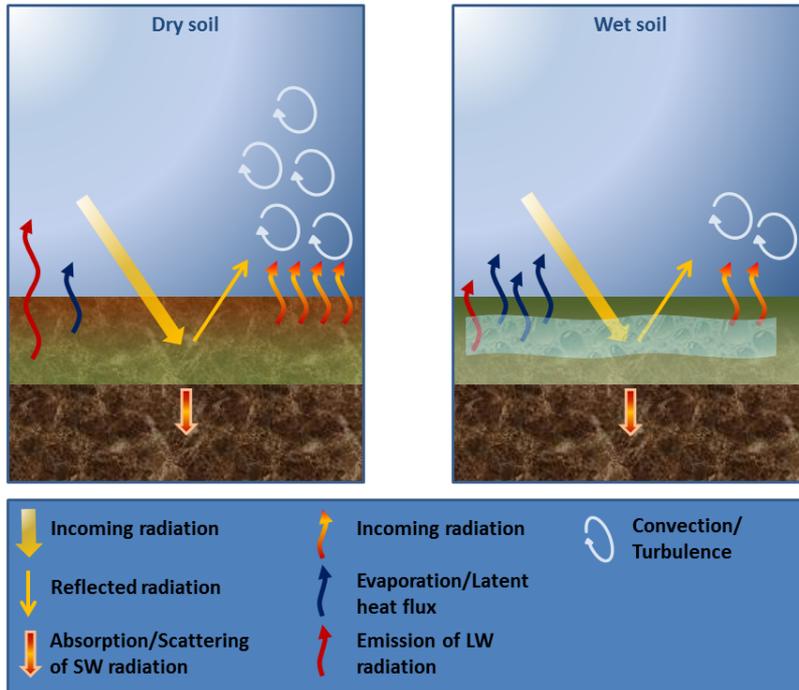


Figura 3.14: Effetto di un suolo caratterizzato da diversi gradi di umidità. Data la differenza tra i flussi di calore nelle diverse condizioni risulta comprensibile la necessità di utilizzare un modello di suolo, come il *Noah LSM model*, che introduca il campo relativo all'umidità di questo.

Aggiornamento della temperatura dello strato inferiore del suolo

Un'altra opzione del WRF relativa alla parametrizzazione del suolo è quella che attiva la modulazione della temperatura del livello più profondo (*TLAG*) che, in assenza dell'attivazione del flag *tmn_update* nella namelist del WRF, viene mantenuto a temperatura costante. La *TLAG* è modificata in base ai valori della *skin temperature* calcolando la modulazione su un periodo fissato nella namelist tramite la variabile *lagday*.

Come descritto nei prossimi paragrafi questa modifica al settaggio del WRF non è necessaria per simulazioni del tipo NAUSICA caratterizzate da una durata complessiva annuale.

Inoltre, per poter attivare questa opzione è necessario utilizzare la versioni di WRF precedenti alla 3.7 o correggere come descritto in rete [30] il codice per versioni successive a causa di un bug.

Parametrizzazioni consigliate per simulazioni annuali

Il WRF è un modello adatto per simulazioni caratterizzate da un periodo che varia da qualche ora alle durate necessarie per i run climatici. In Figura 3.15 è riportata la categorizzazione in classi e tipologia di appartenenza di queste in base alla durata del periodo simulato: previsioni ed analisi possono essere giornaliere, *medium-range* e stagionali mentre i run climatici variano da periodi stagionali ed annuali fino a periodi di decine di anni.

A seconda della classe di appartenenza è indicata una specifica configurazione parametrica che include l'aggiornamento dei campi necessari alla corretta descrizione dell'evoluzione temporale sul periodo scelto [28]. Concentrandosi sul caso di simulazione annuale che, come descritto nel Capitolo 3, è il periodo scelto per la produzione di NAUSICA attualmente disponibile, si legge dalla Figura 3.15 (linea blu) che è consigliato includere umidità del suolo, copertura nevosa, ghiaccio marino e frazione stagionale di vegetazione oltre all'aggiornamento delle SST. Questi aggiornamenti vengono attivati utilizzando il *Noah Land Surface model* e l'*SST update*. L'aggiornamento dello strato profondo del suolo risulta invece superfluo.

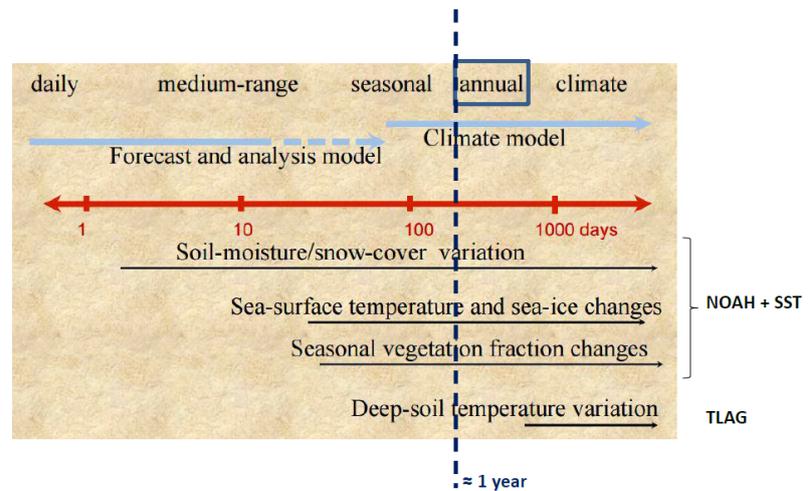


Figura 3.15: Categorizzazione in classi di run in base alla durata del periodo simulato. A seconda della classe di appartenenza è indicata una specifica configurazione parametrica che include l'aggiornamento dei campi necessari alla corretta descrizione dell'evoluzione temporale sul periodo scelto. Il periodo scelto per la produzione di NAUSICA attualmente disponibile è di un anno, linea blu, per il quale è consigliato includere umidità del suolo, copertura nevosa, ghiaccio marino e frazione stagionale di vegetazione oltre all'aggiornamento delle SST. La variazione della temperatura dello strato più profondo di suolo non risulta necessaria per questo tipo di simulazioni ed effettivamente, come verrà descritto nel paragrafo relativo, non si sono osservati effetti sui valori dei campi prodotti attivando questa opzione.

Analisi del problema relativo al modello di suolo

Nell'introduzione al Paragrafo 3.3.3 è descritta la problematica che ha portato alla selezione di un nuovo schema di parametrizzazione del suolo rispetto a quello utilizzato in operatività. L'effetto da correggere risulta evidente in Figura 3.10.

Parte fondamentale del processo risolutivo adottato è costituita dall'analisi svolta su tutte le variabili inerenti la temperatura del suolo che comparivano nell'intero flusso schematizzata in Figura 3.13. Sia il campo STL contenuto nei GRIB *ana-ECMWF* che le variabili SOIL-TEMP e TSLB contenute rispettivamente nei file *met.nc* e *wrfin.nc* risultavano verosimili mentre, nel caso di parametrizzazione del suolo TDS, il campo TSLB risultava affetto da problemi sia all'interno dei *wrfirst.nc* che dei *wrfout.nc*. Da queste osservazioni si è potuto concludere che il problema è generato all'interno del solutore.

Confrontando le distribuzioni empiriche relative alla temperatura a due metri della zona di costa del Friuli Venezia Giulia nella prima settimana di Agosto 2015 delle misure e nel primo caso, Figura 3.16 in alto a sinistra, di un run di restart inizializzato il primo Gennaio 2015 mentre nel secondo caso, Figura 3.16 in alto a destra, di un run inizializzato il primo Agosto 2015 (in entrambi i run si è utilizzato il TDS) si nota infatti che in entrambi la forma della distribuzione generata dal modello non è corretta ma solo nel secondo il range di temperature simulate è verosimile. È quindi confermato che il processo di restart (lettura dei wrfst affetti da errore dovuto al modello di suolo) causa la traslazione della distribuzione delle temperature verso il basso e che lo schema di suolo TDS non è sufficiente per run annuali.

L'effetto sulle temperature legato al TDS è dovuto al processo di restart come si può osservare anche negli output dei run operativi, Figura 3.16 in basso, nei quali si nota che in corrispondenza dell'operazione di restart quotidiana la temperatura minima viene inizializzata a valori molto inferiori rispetto a quelli calcolati dall'evoluzione. Si nota che l'effetto viene corretto in tempi brevi e non costituisce quindi un problema per simulazioni della durata di qualche giorno.

Effetti del cambiamento del modello di suolo

Introducendo la parametrizzazione *Noah LSM*, le temperature simulate per il suolo e di conseguenza anche in corrispondenza dei livelli superiori risultano verosimili a differenza di quelle ottenute con configurazione del suolo TDS, come si osserva in Figura 3.17.

Con l'introduzione del Noah LSM risulta quindi ridotta l'elevata differenza nelle temperature medie osservata nelle analisi degli output precedenti rispetto alle misure: si è passati da una differenza dell'ordine delle decine di gradi ad una differenza dell'ordine di qualche grado, come si nota in Figura 3.18. Anche la forma delle distribuzioni così ottenute è confrontabile con quella ricavata dalle misure.

Effetti dell'aggiornamento delle SST ogni 6 ore (da file ana)

In assenza di aggiornamento della temperatura superficiale del mare nel modello, essendo questa inizializzata il primo Gennaio alle ore 00 UTC, le temperature calcolate nelle zone in prossimità del mare risultano, come prevedibile, ancora inferiori a quelle misurate fino ad una decina di gradi. Questa tendenza è mostrata in Figura 3.19 (destra) nella quale sono riportati gli istogrammi risultanti dalle misure e dalle simulazioni per la stazione ARPA-OSMER Boa.

Attivando l'aggiornamento delle SST come descritto nei precedenti paragrafi si è ottenuta una configurazione del modello che descrive in modo soddisfacente i dati, sia relativamente alle temperature, come mostrato in Figura 3.19 (sinistra), che relativamente alle altre grandezze analizzate. Questo a scapito però, come già descritto nell'introduzione al Paragrafo corrente, di un aumento dei tempi di calcolo del 30%.

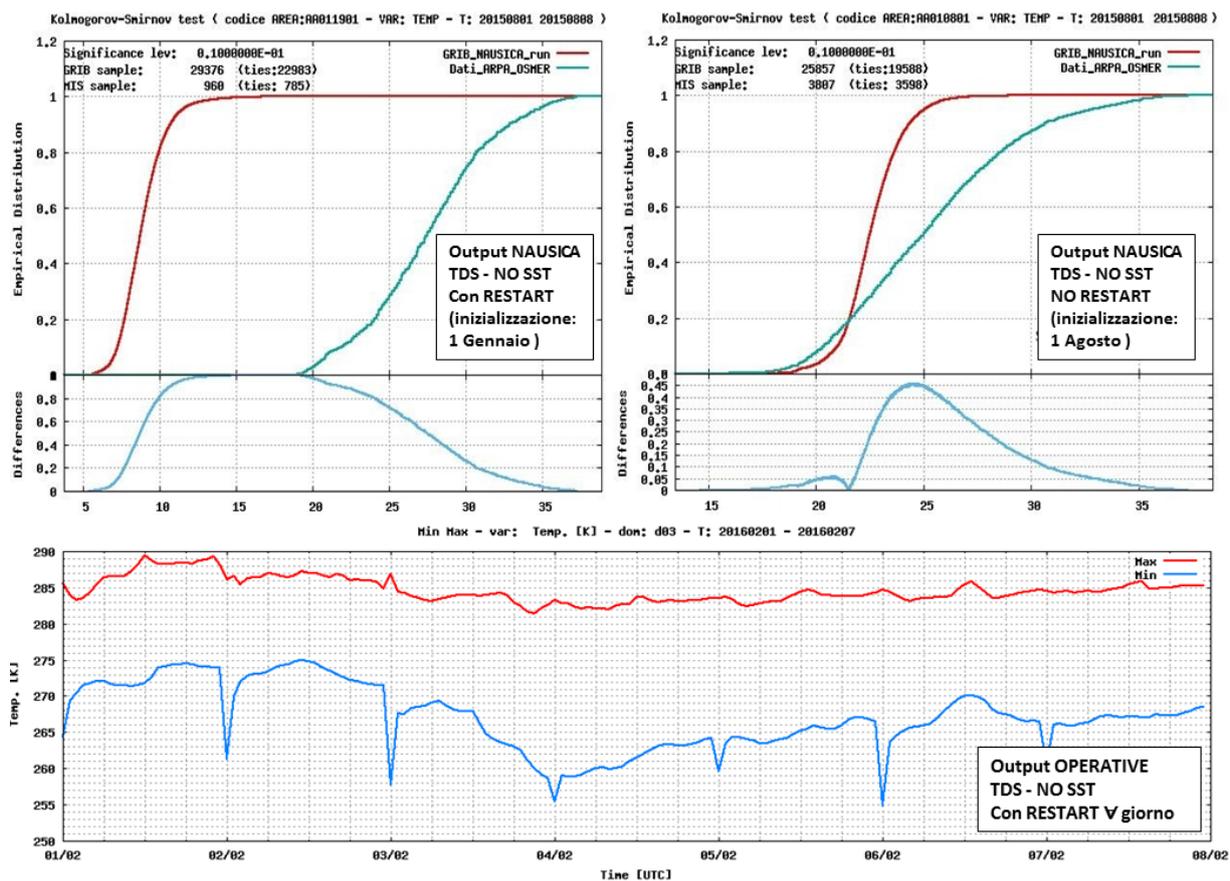


Figura 3.16: In alto: confronto tra le distribuzioni empiriche relative alla temperatura a due metri della zona di costa del Friuli Venezia Giulia nella prima settimana di Agosto 2015 costruiti considerando la misure e nel primo caso, grafico a sinistra, un run di restart inizializzato il primo Gennaio 2015 mentre nel secondo caso, grafico a destra, un run inizializzato il primo Agosto 2015 (nei run si è utilizzato il TDS). Si nota che in entrambi la forma della distribuzione generata dal modello non è corretta ma solo nel secondo il range di temperature simulate è verosimile. È quindi confermato che processo di restart (lettura dei *wrfrst* affetti da errore dovuto al modello di suolo) causa la traslazione della distribuzione delle temperature verso il basso. In basso: effetto sulle temperature legato al TDS e dovuto al processo di restart negli output operativi. Si nota che in corrispondenza del restart quotidiano la temperatura minima viene inizializzata a valori molto inferiori rispetto a quelli calcolati dall'evoluzione ma che l'effetto viene corretto in tempi brevi e non costituisce quindi un problema per simulazioni relative ad intervalli della durata di qualche giorno.

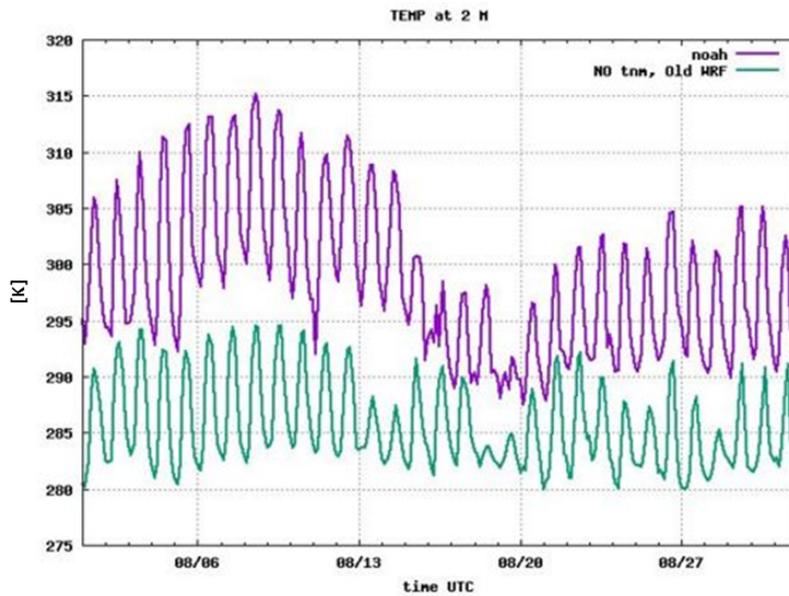


Figura 3.17: Serie temporale della temperatura a due metri nel mese di Agosto in corrispondenza delle coordinate di Udine con schemi di suolo TDS (verde) e Noah LSM (viola). Si nota che nel secondo caso le temperature riprodotte dalla simulazioni sono verosimili.

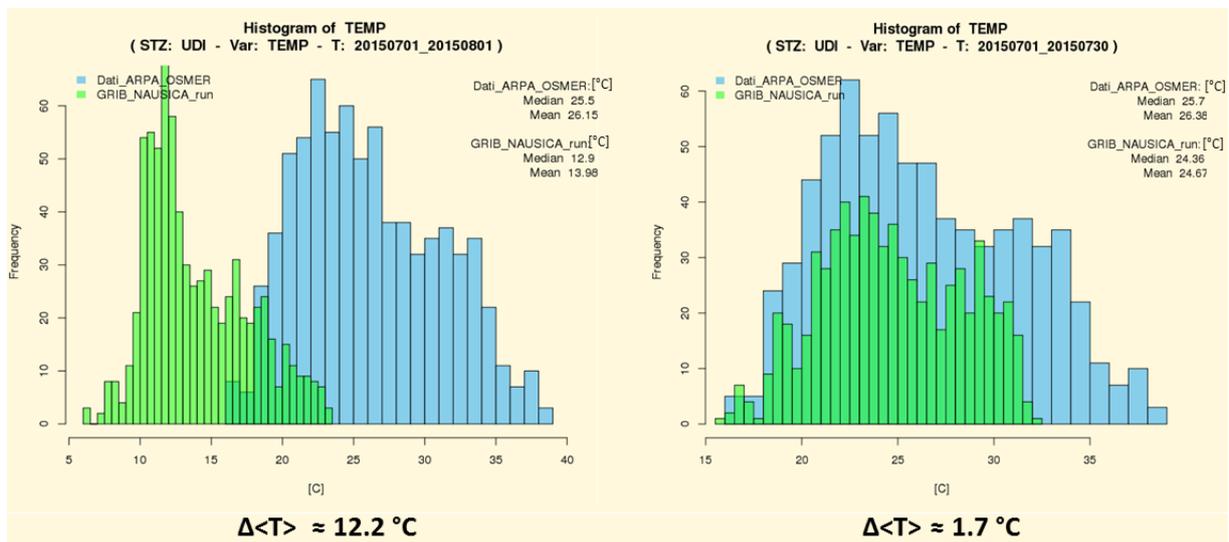


Figura 3.18: Confronto tra gli istogrammi relativi alla temperatura a due metri dal suolo calcolata dal modello e misurata in corrispondenza della stazione ARPA-OSMER di Udine nel mese di Luglio in assenza di aggiornamento SST rispettivamente con modello di suolo TDS (sinistra) e con modello di suolo Noah LSM (destra). Si nota che la parametrizzazione di suolo Noah LSM risolve l'elevata differenza nelle temperature osservata nel caso di schema TDS. La frequenza che compare in ordinata corrisponde al numero di occorrenze di valori appartenenti al range della classe dell'istogramma tra le scadenze orarie disponibili sul periodo selezionato.

È da sottolinearsi il fatto che nei file *ana-ECMWF* dai quali viene letto il valore del campo SST ogni 6 h questo varia solo ogni 24 h (in corrispondenza delle ore 18:00 UTC).

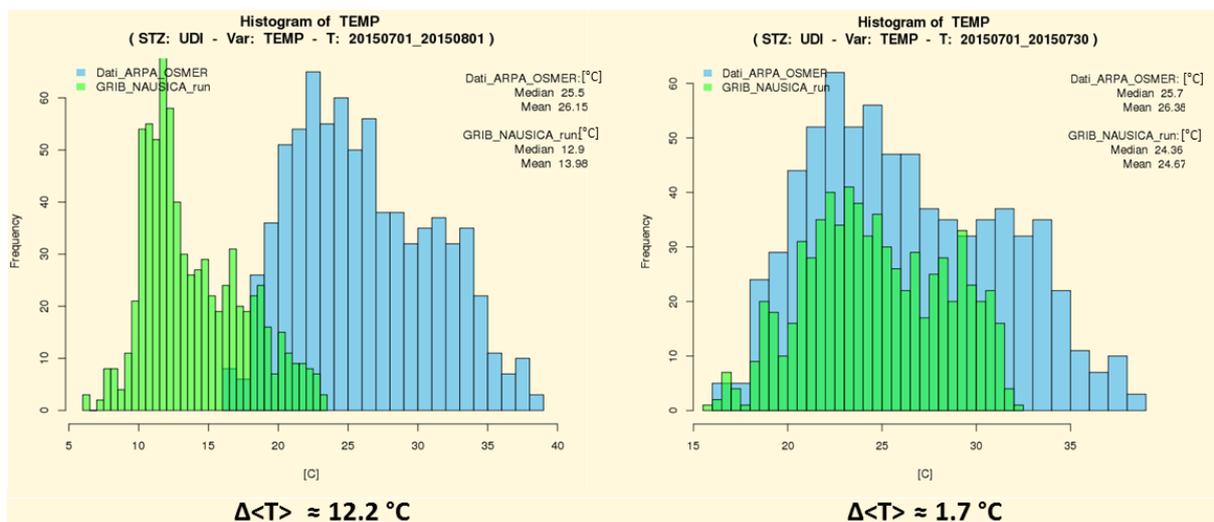


Figura 3.19: Confronto tra gli istogrammi relativi alla temperatura a due metri dal suolo calcolata dal modello con schema di suolo Noah LSM e misurata in corrispondenza della stazione ARPA-OSMER BOA nel mese di Maggio rispettivamente in assenza di aggiornamento SST (destra) e con aggiornamento delle SST ogni 6 ore (sinistra). Si nota che attivando l'aggiornamento delle SST come descritto nei precedenti paragrafi si è ottenuta una configurazione del modello che descrive in modo soddisfacente i dati. La frequenza che compare in ordinata corrisponde al numero di occorrenze di valori appartenenti al range della classe dell'istogramma tra le scadenze orarie disponibili sul periodo selezionato.

Effetti dell'aggiornamento delle SST ogni 6 ore (da file ana-sfc e settaggio maschere land/sea) per WPS V3.9.1

Come già detto il database NAUSICA è costruito con aggiornamento SST ogni 6 ore a partire da file *ana* (cioè con la stessa risoluzione delle altre condizioni al contorno, ovvero di 0.5°). Inoltre, dato che buona parte del mare rimane scoperta da punti di griglia, i valori delle temperature sono interpolati a partire dal campo di SKINTEMP che è non mascherato sul mare. Questo settaggio causa la presenza di un "cuneo" nelle TSK osservabile in Figura 3.20, il campo di SST resta mal definito negli output ed inoltre, dove definiti entrambi, sea prevarica land. Questi problemi sono stati risolti nel lavoro di tesi di Pierluigi Masai [31] tramite lettura delle SST da file *ana-sfc* (che hanno risoluzione pari a 0.1°) e modifica della definizione delle maschere nel file METGRID.TBL che si occupa di dare le direttive di interpolazione e mascheramento al *metgrid.exe*. Nei run relativi a questo lavoro, eseguiti però con WRF 3.9 e WPS 3.9.1 (in quanto WPS V3.9 presenta dei problemi), è stata mascherata la terra per il campo SST scegliendo la maschera corretta (LANDSEA) e mascherato il mare per il campo TSK scegliendo sempre la stessa maschera. Si ottiene così la definizione di entrambi i campi nei *wrfout*, come si osserva in Figura 3.21.

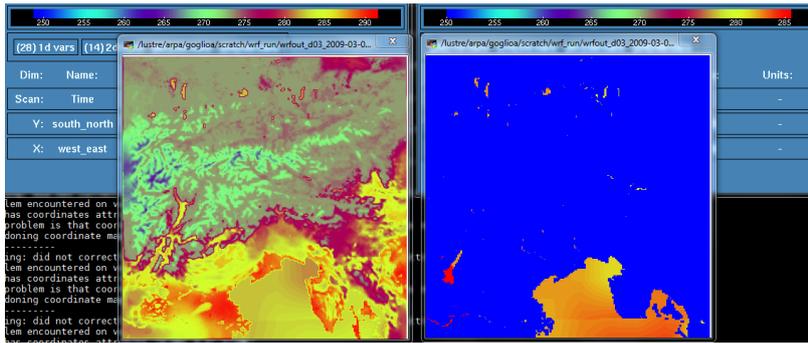


Figura 3.20: Campi di SKINTEMP (sinistra) ed SST (destra) nelle simulazioni NAUSICA. Il primo è ben definito anche se si nota che i valori sono interpolati a partire da condizioni al contorno a bassa risoluzione (“cuneo” in zona golfo di Trieste) e si intuisce la presenza di una discontinuità in corrispondenza della superficie su cui sono lette le SST (rettangolo a sud del dominio). Per il campo SST, entrambi i difetti sono molto marcati ed inoltre si osserva che non è stata settata la prevaricazione della terra sul mare dove presenti entrambi (parte dell’Istria scompare). In entrambe le figure la temperatura è espressa in Kelvin.

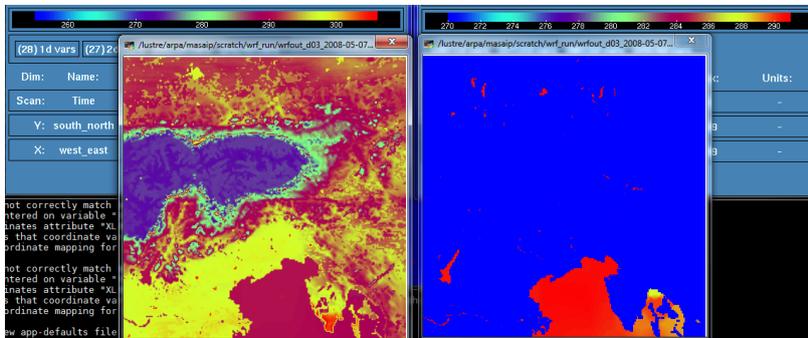


Figura 3.21: Campi di SKINTEMP (sinistra) ed SST (destra) nelle simulazioni eseguite per la tesi di Pierluigi Masai (WRF 3.9, WPS 3.9.1). In queste, oltre ad essere stata usata la risoluzione più alta (0.1° invece di 0.5°) per le condizioni al contorno della SST, è stato mascherato il mare per la definizione della SKINTEMP e la terra per la definizione del campo SST. Inoltre è stata attivata l’opzione per far prevaricare land rispetto a sea dove entrambi sono definiti.

Esempio di settaggio corretto dei campi del file METGRIG.TBL per WPS V3.9.1 (da associare all’aggiunta nella namelist del WPS nella sezione

```
&geogrid
```

```
di
```

```
geog\_data\_res = 'gtopo_5m+usgs_5m+nesdis_greenfrac+5m',
```

```
'gtopo_2m+usgs_2m+nesdis_greenfrac+2m',
```

```
'gtopo_30s+usgs_30s+nesdis_greenfrac+30s'
```

che indica il metodo di interpolazione corretto):

```

=====
name=SKINTEMP
  mpas_name=skintemp
  interp_option=sixteen_pt+four_pt+wt_average_4pt+
                +wt_average_16pt+search
  masked=water
  interp_water_mask = LANDSEA(0)
  fill_missing=0.
=====
name=SST
  interp_option=sixteen_pt+four_pt+wt_average_4pt+
                +wt_average_16pt+search
  masked=land
  interp_land_mask=LANDSEA(1)
  fill_missing=0.
  missing_value=-1.E30
  flag_in_output=FLAG_SST
=====

```

Effetti dell'aggiornamento della temperatura dello strato inferiore di suolo (TLAG)

L'ultima analisi svolta nell'ambito della descrizione del suolo riguarda l'attivazione della modulazione della temperatura profonda. Per poter svolgere questo test è stato necessario correggere un bug presente nel codice del WRF V 3.7.1 [30]. Si è constatato che, per simulazioni di un anno, l'utilizzo di questo parametro non influenza i campi al di sopra del livello del suolo, Figura 3.22. Per le simulazioni svolte non è quindi stata utilizzata questa modulazione.

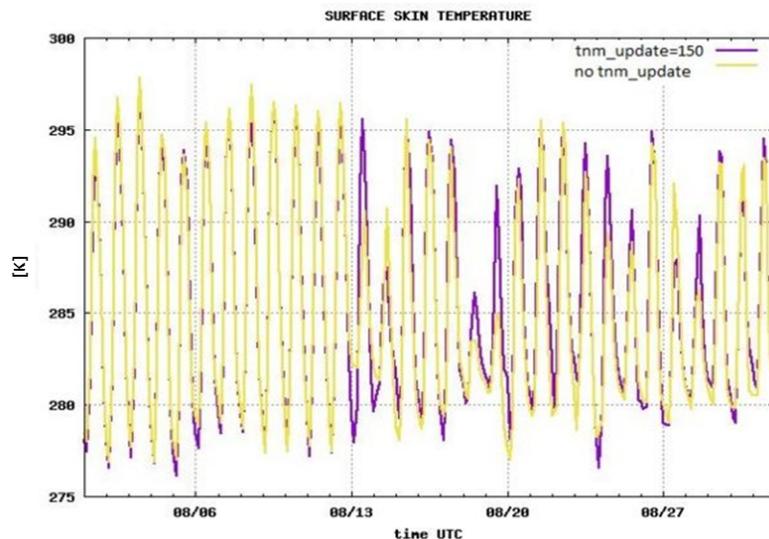


Figura 3.22: Confronto tra le serie temporali relative alla temperatura superficiale in presenza (linea viola) e assenza (linea gialla) dell'aggiornamento della temperatura profonda del suolo (TLAG) in configurazione TDS senza aggiornamento SST. Si nota che l'utilizzo di questo parametro non influenza i campi al di sopra del livello del suolo.

In Figura 3.23 si può osservare l'evoluzione della temperatura dello strato di suolo inferiore al variare di più parametri: le linee viola e azzurra sovrapposte descrivono il valore della variabile in configurazione di suolo TDS rispettivamente prima e dopo la correzione del bug, la quale doveva riguardare infatti soltanto l'utilizzo della TLAG che in questi casi era disattivata. La linea gialla descrive nella prima parte il valore in configurazione Noah LSM con TLAG disattivata: si osserva che, rispetto a tutti gli altri casi in TDS, la temperatura dello strato profondo è di circa dieci gradi più elevata. La seconda e la terza fase descritte dalla linea gialla rappresentano l'evoluzione temporale in caso di modello TDS con attivazione di `tmn_update` ma calcolo della modulazione della temperatura allo strato inferiore effettuato in base al valore della temperatura superficiale su periodi (`lagday` [giorni]) settati diversamente.

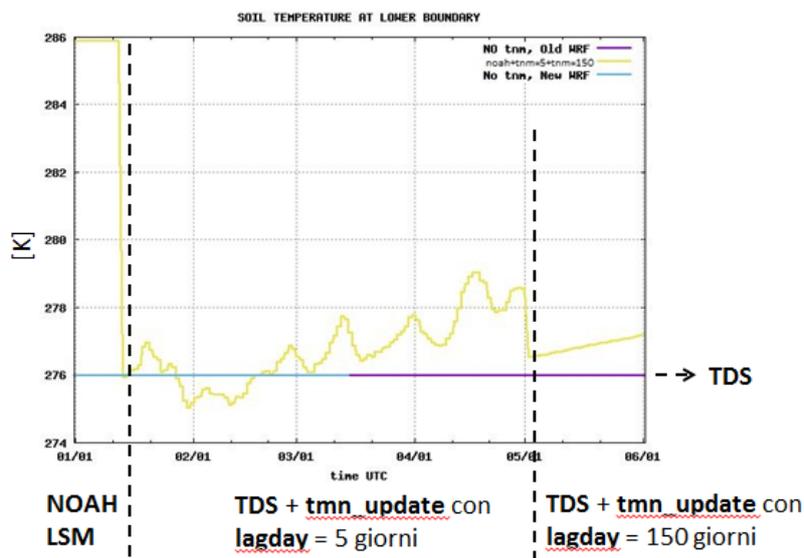


Figura 3.23: Evoluzione della temperatura dello strato di suolo inferiore al variare di più parametri: le linee viola e azzurra sovrapposte descrivono il valore della variabile in configurazione di suolo TDS rispettivamente prima (Old WRF) e dopo (New WRF) la correzione del bug al WRF, la quale doveva riguardare infatti soltanto l'utilizzo del `tmn_update` che in questi casi era disattivato. La linea gialla descrive nella prima parte il valore in configurazione Noah LSM con `tmn_update` disattivata: si osserva che, rispetto a tutti gli altri casi in TDS, la temperatura dello strato profondo è di circa dieci gradi più elevata. La seconda e la terza fase descritte dalla linea gialla rappresentano l'evoluzione temporale in caso di modello TDS con attivazione di `tmn_update` ma tempi di modulazione calcolati sul valore della skin temperature su periodi differenti (`lagday` [giorni]) rispettivamente di 5 e 150 giorni) ottenute dopo la correzione del bug.

3.4 Valutazione della necessità di assimilare misure

Dalla valutazione presentata in questo paragrafo, relativa alla capacità descrittiva delle condizioni al contorno fornite, che sono descritte nel Paragrafo 2.2, per quanto riguarda i fenomeni fisici che interessano il progetto sul dominio di calcolo definito, si può concludere che

sarebbe necessario assimilare dati almeno nell'ambito delle temperature marine superficiali. La scelta di questa variabile come oggetto dell'analisi è derivata dall'esperienza di utilizzo del codice WRF da parte dei colleghi. È però auspicabile che vengano valutate anche altre variabili.

Per verificare la predittività del modello per quanto concerne le temperature superficiali dell'Adriatico nella zona interessata dal dominio regionale e quantificarne la differenza rispetto ai dati misurati, sono stati confrontati i valori relativi alle serie temporali estratte dal campo SST contenuto nei GRIB ECMWF, come descritto nel Paragrafo 2.2, con misurazioni effettuate presso le stazioni OSMER di Trieste e Lignano nell'arco del 2015. Dall'analisi, descritta in dettaglio nei paragrafi successivi, si può concludere che i GRIB ECMWF, introdotti nel Paragrafo 2.2, descrivono i dati analizzati correttamente, come valutabile tramite i diagrammi di Taylor di Figura 3.31 A e 3.31 B, ad eccezione di

1. Episodi di Bora nella zona orientale del dominio regionale che, in periodi di stagione calda, causano *upwelling* (fenomeno schematizzato in Figura 3.24) non descritto dal modello. Questi episodi corrispondono ai punti allineati che si allontanano dalla bisettrice nel plot di regressione lineare relativo a Trieste (Figura 3.29 A).
2. Sovrastima della SST in zona lagunare nei mesi freddi e sottostima di questa nei mesi caldi, entrambe sistematiche e di non elevata rilevanza (effetti dell'ordine di circa 2°C rispetto a temperature che variano tra i 5°C e i 30°C). Questi corrispondono alle zone di lontananza dei punti dalla bisettrice nel plot di regressione lineare relativo a Lignano di Figura 3.29 B).
3. Episodi periodici di flusso e deflusso dalle acque lagunari non riprodotti nei GRIB osservabili in Figura 3.28 .

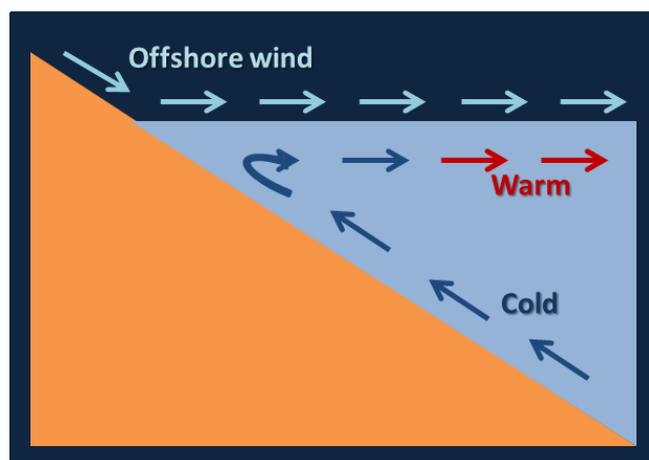


Figura 3.24: Schematizzazione del fenomeno di *upwelling* descrivibile come il rimescolamento delle acque causato dall'interazione della superficie marina con venti da terra molto intensi. Nel periodo estivo, essendo maggiore il gradiente di temperatura dell'acqua, il fenomeno ha un effetto più importante sulla diminuzione delle temperature superficiali e di conseguenza dell'intera zona di costa.

Risulta quindi necessario un ulteriore sviluppo del progetto NAUSICA per l'introduzione di una catena di assimilazione dei dati relativi alla temperatura superficiale del mare. A questo scopo è stato individuato il software GSI [32], elencato tra gli strumenti in Tabella 2.1, il cui utilizzo sarà oggetto della prossima attività inerente il progetto NAUSICA.

Descrizione dell'analisi eseguita

Come si osserva da Figura 3.25, nella quale è riportata la localizzazione delle due stazioni, quella di Trieste è situata in prossimità del molo Fratelli Bandiera, zona particolarmente soggetta alla Bora e conseguente effetto di *upwelling* di cui si vogliono valutare gli effetti, mentre la centralina di Lignano si trova in uno dei canali di accesso alla laguna di Marano. Come sarà descritto questa localizzazione comporta alcune peculiarità nell'andamento delle temperature rispetto al caso di mare aperto.

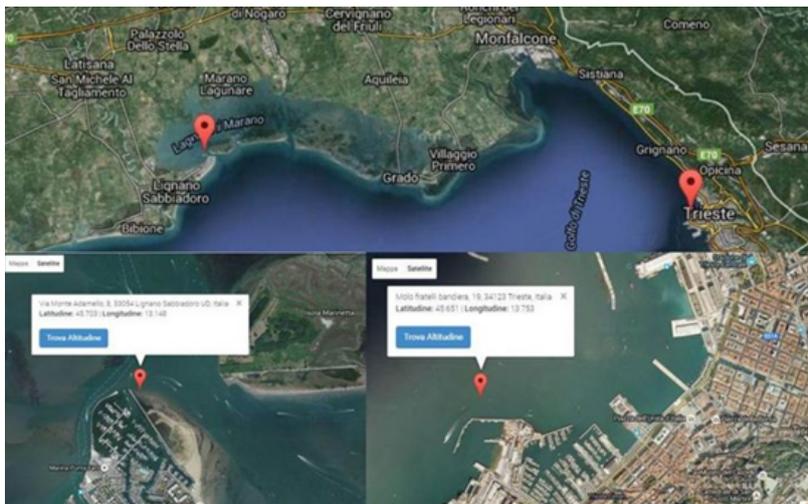


Figura 3.25: Localizzazione stazioni ARPA OSMER di Trieste (lat/lon 45.651;13.753) e Lignano (lat/lon 45.703;13.148)

Analisi relativa alla stazione di Trieste

Dal confronto tra il valore della temperatura ricavato dall'interpolazione del campo SST nelle coordinate relative alla stazione e i valori misurati presso la centralina di Trieste, riportato in Figura 3.4, si nota che

- L'accordo tra GRIB e misure è generalmente buono. Questa affermazione è supportata dalla concentrazione dei punti sulla bisettrice nel plot di regressione lineare riportato in Figura 3.30 A.
- In corrispondenza di pochi punti ben distinguibili in figura (per esempio quelli indicati dalle frecce) c'è un netto discostamento dei dati rispetto all'andamento del campo SST dei GRIB ECMWF

caratterizzato da una brusca diminuzione della temperatura dell'acqua e da una successiva curva di recupero. Questo effetto di brusco raffreddamento della temperatura superficiale del mare, caratterizzato da diminuzioni dell'ordine di cinque gradi e tempi caratteristici dell'ordine di qualche ora, è attribuibile all'effetto di *upwelling* causato da raffiche di Bora e alla successiva ritermalizzazione dell'acqua. I tempi di recupero sono dell'ordine di 48 ore. Figura 3.27 A mostra l'evidente correlazione tra il raffreddamento della SST e le raffiche di Bora, queste ultime sono facilmente distinguibili in quanto caratterizzate da venti provenienti da una direzione di circa 60° e intensità superiori ai 15 m/s. È stato scelto di riportare l'analisi relativa al mese di Luglio che presenta diversi episodi, due dei quali sono evidenziati in rosso e corrispondono ai picchi analogamente evidenziati in Figura 3.27. I picchi osservati corrispondono nella regressione lineare di Figura 3.30 A ai punti allineati che si discostano dalla bisettrice. Appare evidente che l'effetto descritto si ha solo in condizioni di gradiente verticale della temperatura iniziale dell'acqua superiore ad una certa soglia, infatti questi episodi sono osservabili nei dati analizzati tra inizio Maggio e fine Agosto. A conferma dell'ultima osservazione è riportato in Figura 3.27 B un estratto del mese di Novembre nel quale si osserva un episodio di Bora molto forte non correlato all'effetto di raffreddamento istantaneo descritto (episodio evidenziato in blu in Figura 3.27 B).

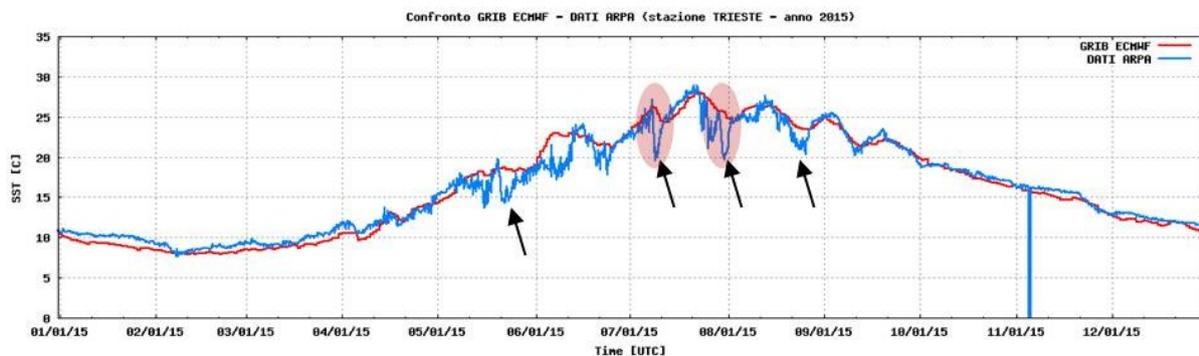


Figura 3.26: Confronto tra il valore della temperatura superficiale del mare ricavato dall'interpolazione del campo SST dei GRIB ECMWF nelle coordinate relative alla stazione di Trieste nei GRIB ECMWF e i valori misurati presso la centralina ARPA di Trieste nell'anno 2015. Le frecce indicano alcuni episodi di discostamento GRIB/dati riconducibili al fenomeno di *upwelling* (Fig. 3.24) causato da raffiche di Bora. I due episodi evidenziati in rosso sono analizzati in dettaglio in Fig. 3.27 (sinistra)

Analisi relativa alla stazione di Lignano

Osservando Figura 3.28, che riporta il confronto tra il valore della temperatura ricavato dall'interpolazione del campo SST nelle coordinate relative alla centralina di Lignano nei GRIB ECMWF e i valori misurati presso questa

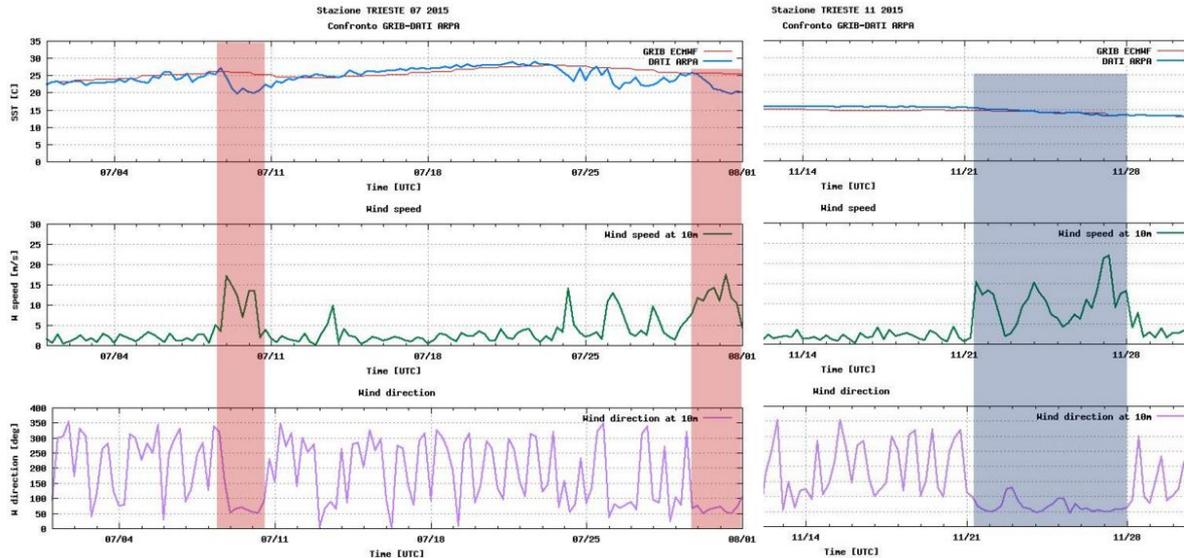


Figura 3.27: Confronto tra il valore della temperatura superficiale del mare ricavato dall'interpolazione del campo SST nelle coordinate relative al punto di griglia più vicino alla stazione di Trieste nei GRIB ECMWF e i valori misurati presso la centralina affiancati da velocità e direzione del vento misurati in Luglio 2015 (sinistra) e Novembre 2015 (destra). Le zone evidenziate corrispondono ad episodi di Bora, in particolare quelle evidenziate in rosso corrispondono agli stessi episodi evidenziati in Fig. 3.4

- Non si nota l'effetto di *upwelling*, schematizzato in Figura 3.24, che caratterizza invece l'andamento della curva relativa alla centralina di Trieste nei periodi di elevato gradiente verticale nella temperatura delle acque. Questo è dovuto sia al fatto che la laguna di Grado e Marano ha una profondità media dell'ordine di mezzo metro, mentre l'Adriatico in zona Trieste raggiunge fino a 40 m, che all'inferiore intensità della Bora in questa area. Come si nota da Figura 3.28 relativa alla centralina di Lignano, che riporta evidenziati in verde gli stessi episodi di Bora osservati in Luglio 2015 a Trieste ed evidenziati in rosso nei grafici di Figura 3.4 e Figura 3.27 A, la velocità del vento a Lignano è inferiore di circa 5 m/s e caratterizzata da un angolo maggiore ai 60° tipici di Trieste.
- Si osserva una sovrastima della SST da parte dei GRIB nei mesi freddi e una sottostima nei mesi caldi, entrambe sistematiche, identificabili in Figura 3.28 A ed evidenziate dalla distribuzione dei punti fuori bisettrice nella regressione lineare di Figura 3.30 B. Questa tendenza potrebbe essere riconducibile all'inferiore profondità delle acque lagunari, e alla sottostima dell'effetto di questa rispetto al caso di mare aperto, oppure alla torbidità dell'acqua lagunare non contemplata dagli algoritmi del modello.
- L'andamento oscillatorio che si nota in Figura 3.28 ben riprodotto nei GRIB è dovuto ai fenomeni di accoppiamento tra condizioni del mare e condizioni meteorologiche: raffreddamento e riscaldamento delle acque.
- Le oscillazioni con periodo dell'ordine di circa 24 ore, che si possono notare nella curva relativa ai dati in Figura 3.28, si può

ipotizzare che siano invece legate agli effetti di marea tramite le correnti di ingresso e uscita dell'acqua all'interno della laguna di Marano, particolarmente intense in prossimità dei canali di accesso come quello in cui è situata la centralina, oppure al movimento verticale della sonda stessa.

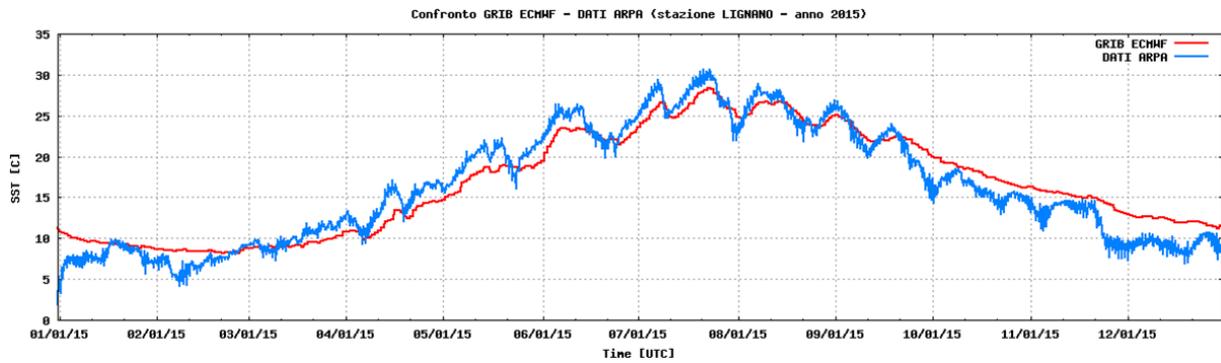


Figura 3.28: Confronto tra il valore della temperatura superficiale del mare ricavato dall'interpolazione del campo SST nelle coordinate relative alla stazione di Lignano nei GRIB ECMWF e i valori misurati presso la centralina nell'anno 2015.

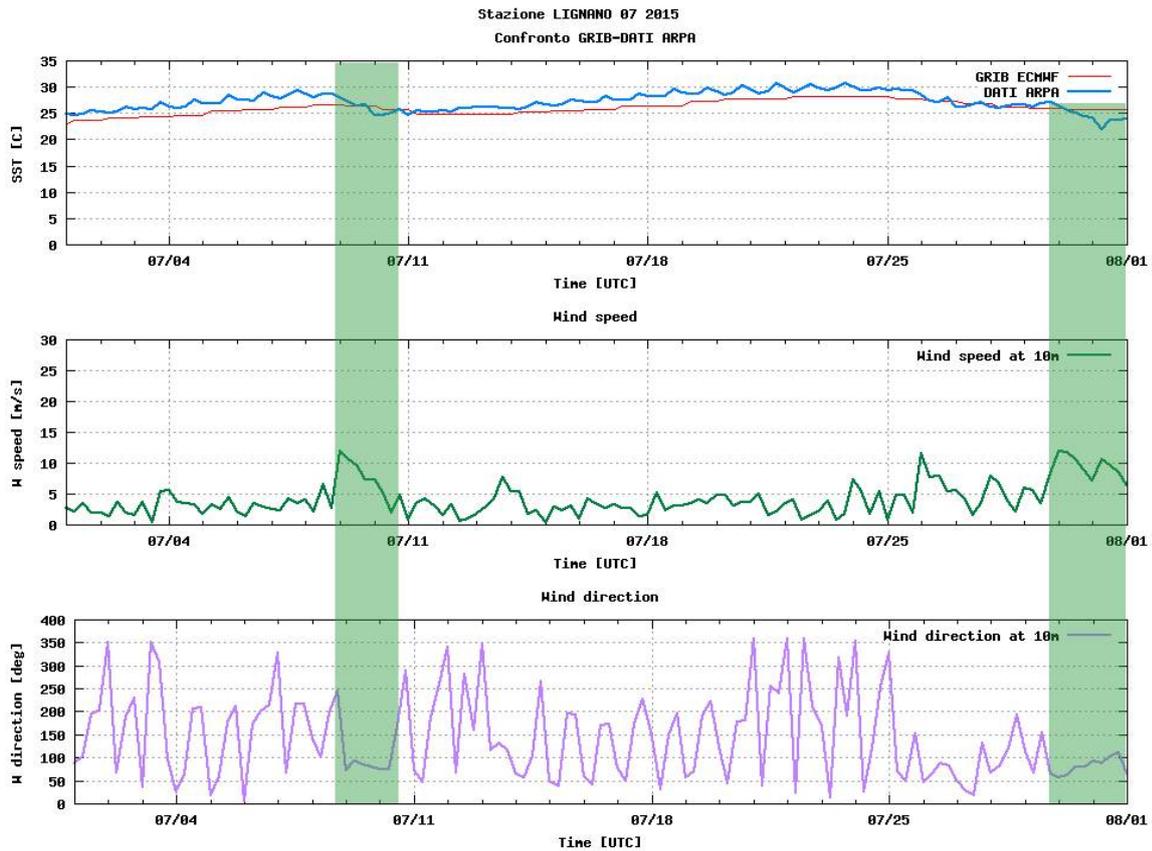


Figura 3.29: Confronto tra il valore della temperatura superficiale del mare ricavato dall'interpolazione del campo SST nelle coordinate relative alla stazione di Lignano nei GRIB ECMWF e i valori misurati presso la centralina affiancati da velocità e direzione del vento misurati relativi a Luglio 2015. Le zone evidenziate sono episodi di Bora corrispondenti agli stessi episodi evidenziati in Fig. 3.4 e Fig. 3.27 A relativi alla stazione di Trieste.

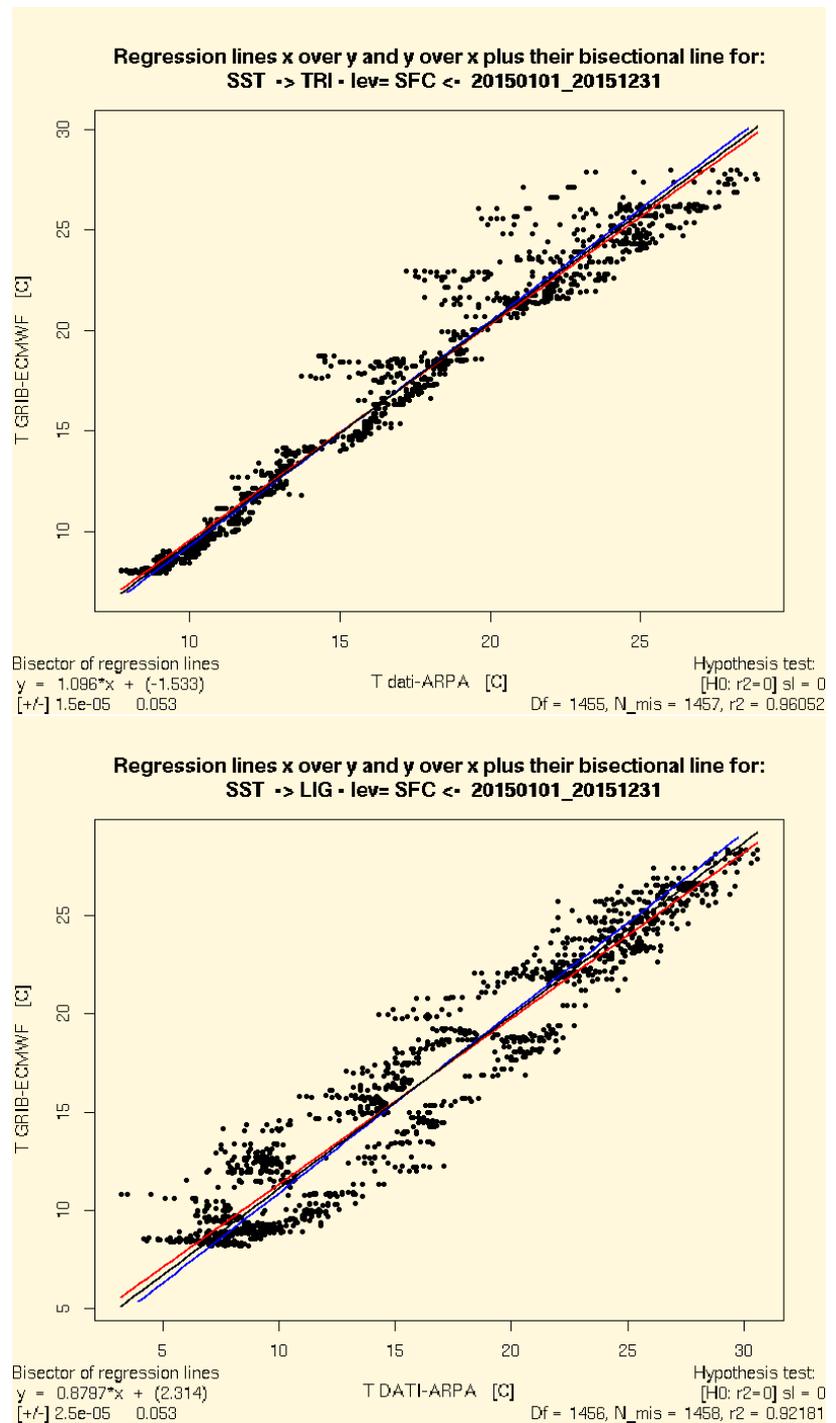


Figura 3.30: Plot di regressione lineare relativo al campo SST in GRIB ECMWF e misure ARPA per le stazioni di Trieste (in alto) e Lignano (in basso)

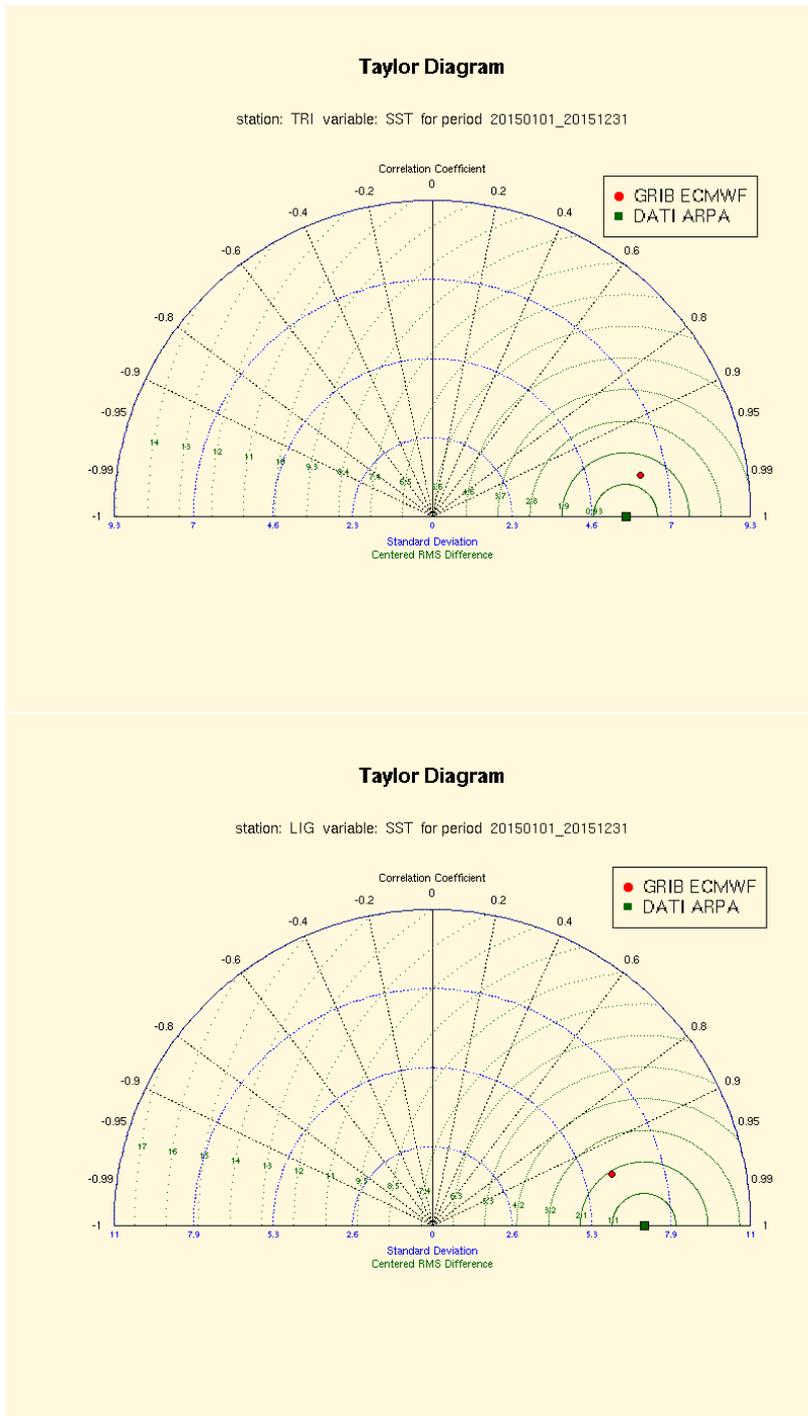


Figura 3.31: Diagramma di Taylor per la valutazione della stima delle misure ARPA da parte dei GRIB ECMWF relativamente alla SST presso le stazioni di Trieste (in alto) e Lignano (in basso)

4

Flussi di lavoro

In questo capitolo sono descritti i flussi di lavoro implementati nell'ambito del progetto:

NAUSICA_run e *NAUSICA_pp* oggetto del Paragrafo 4.1, contenenti il pre-processamento, la simulazione meteorologica vera e propria e una prima catena di post processamento

WRF_ana_post descritta nel Paragrafo 4.2, per la conversione degli output in formato GRIB [7]

NAUSICA_ver Paragrafo 4.3, con lo scopo di quantificare statisticamente la qualità dei campi al suolo prodotti dalle simulazioni tramite confronto con i dati misurati dalle stazioni ARPA.

È inoltre descritta la problematica riscontrata nella definizione di un metodo per la valutazione della qualità dei campi lungo la verticale, attualmente implementata nella suite *NAUSICA_rds* descritta nel Paragrafo 4.4. All'interno del capitolo sono presenti anche alcuni cenni all'organizzazione logica delle principali strutture computazionali implementate ed alle modalità di utilizzo di questi, descritte esaustivamente in Appendice E, G ed F.

4.1 Produzione del database e post-processamento degli output

La suite *ecflow NAUSICA_run* gestisce il flusso della simulazione meteorologica vera e propria. Include una fase iniziale che esegue il pre-processamento necessario tramite WPS [13] e la simulazione meteorologica tramite modello WRF [3] a partire dalle condizioni iniziali ed al contorno ECMWF, descritte nel Paragrafo 2.2. È inoltre implementata la possibilità di eseguire la conversione degli output in file GRIB [7] sia sui livelli del modello che su quelli isobarici.

La suite *NAUSICA_pp* è una copia di *NAUSICA_run* da utilizzare nel caso si voglia eseguire il post-processing di output del WRF in modo indipendente da un run di questo in atto contemporaneamente. La configurazione dei run che hanno generato i prodotti attualmente disponibili, descritti nel Capitolo 5, è quella definita tramite l'analisi descritta nel Capitolo 3.

In Figura 4.2 è riportato lo schema della struttura logico/computazionale della suite i cui dettagli sulla struttura e l'utilizzo sono esaustivamente riportati in Appendice E.

È da sottolinearsi il fatto che l'attivazione dell'intero processo si ottiene tramite l'attribuzione di un valore a pochi parametri in un unico *shell script* che crea il file di definizione della suite. Un esempio di settaggio di questo è riportato in Figura 4.1, per i dettagli sui singoli parametri fare riferimento alla descrizione in Appendice E.

```
#####
#
#          Parametri
#
#          NOME_SUITE="NAUSICA_run"
#
#-----#
#
#          DA MODIFICARE :
#
#          NUMERO_FAM_run_NAUSICA=2
#
#          NOMI_FAM_run_NAUSICA_Ar=( "1_Gen_16" "2_Gen_16" )
#
#          DATE_START_FAM_run_NAUSICA_Ar=( 20160101 20160108 )
#
#          DATE_END_FAM_run_NAUSICA_Ar=( 20160108 20160115 )
#
#          RESTART_FLAG_FAM_run_NAUSICA_Ar=( 0 1 )
#
#          DAYS_PER_RUN_FAM_run_NAUSICA_Ar=( 7 7 )
#
#          WRF_RUN=1
#
#          SST_FLAG=1
#
#          GRIB_Ar=( "CRMA" "MOD_LEV" )
#
#          SCAN_TYPE="CRMA"
#
#          NAUSICA_RUN_TIME=17:00
#
#          NAUSICA_RUN_DATE=20161114
#
#          NAUSICA_RUN_DT=1
#
#          CLEAN=1
#
#          ARCHIV=1
#
#####
```

Figura 4.1: Esempio di settaggio dei parametri di inizializzazione delle suite NAUSICA_run e NAUSICA_pp (per i dettagli sui singoli parametri fare riferimento alla descrizione in Appendice).

4.1.1 Prodotti delle suite NAUSICA_run e NAUSICA_pp

Di ogni run delle suite NAUSICA_run e NAUSICA_pp vengono archiviati sia gli output nativi: *wrfout.nc*, che i file GRIB prodotti dal post-processing di questi se attivato. È possibile inoltre salvare in un file compresso gli output del flusso ecFlow, gli *output* ed *error* dei job. Per quanto riguarda il controllo della produzione è possibile fare in modo che i risultati di questo vengano riportati su un'apposita pagina HTML. Le caratteristiche dei prodotti sono riassunte in Tabella 4.1.

wrfout.nc

Gli output nativi del modello sono archiviati in cartelle giornaliere con il codice, attribuito secondo la convenzione CRMA [33], 1390FA01D0_YYYY (dove YYYY è l'anno a cui si riferiscono).

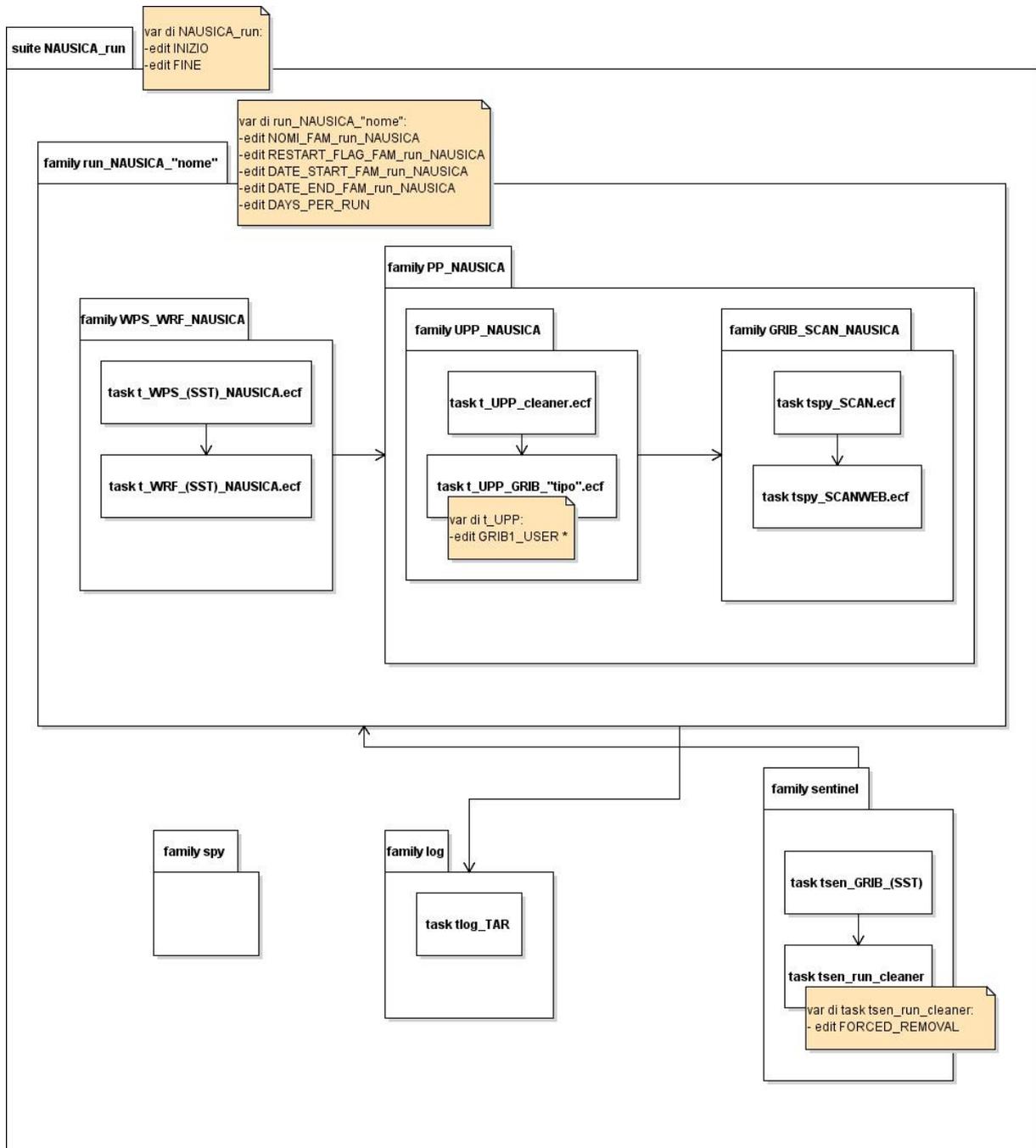


Figura 4.2: Schema della struttura logico/computazionale della suite NAUSICA_run (e NAUSICA_pp) i cui dettagli sulla struttura e l'utilizzo sono esaurientemente riportati in Appendice.

Ogni cartella giornaliera contiene, oltre ad un file *netCDF* per il dominio europeo, uno per il dominio nazionale e quattro per quello regionale, i tre file di restart *wrfrst.nc*: uno per ogni dominio.

Un anno di simulazioni con la configurazione descritta nel Capitolo 3 occupa uno spazio su disco pari a **1.5 Tb**. Quindi per l'archiviazione dei netCDF relativi all'intero database dei 17 anni programmati sono necessari **25.5 Tb**.

Per quanto riguarda i tempi, sempre nella configurazione scelta, si ha che per ogni anno simulato sono necessari circa 10 giorni di run su 4 nodi da 32 core. Di conseguenza il *wall-time* totale per i **17 anni** previsti dal progetto è di **170 giorni**.

Prodotti e formato	Suite di produzione	Tempi di produzione	Dimensioni	Caratteristiche
wrfout.nc (wrfrst.nc)	NAUSICA_run NAUSICA_pp	10 gg per anno con 4 nodi da 32ppn	1.5 Tb per anno	Codice di archiviazione: 1390FA01Do_YYYY
CRMA.GRIB	NAUSICA_pp NAUSICA_run	30 h per anno	0.11 Tb per anno	Produzione parallela PRS_LEV e MOD_LEV
FARM.GRIB	WRF_ana_post	27 h per anno	0.18 Tb per anno	Sui livelli isobarici
Info_stat.txt Lin_reg .ps .png Taylor .ps .png Kol-smf .ps .png Spettro .ps .png	NAUSICA_ver NAUSICA_nc	7 gg per anno in trimestri	1.4 Gb per anno	Su tutte le stazioni, tutte le aree, tutte le variabili, tutte la analisi, tutti i formati di output
Info_stat.txt Kol-smf .ps .png bias.png vert.png grad.png	NAUSICA_rds			

Tabella 4.1: Elenco delle caratteristiche dei prodotti delle suite utilizzate nel progetto NAUSICA. A causa delle problematiche riscontrate nella suite NAUSICA_rds descritte non è ancora stata completata la struttura della suite. Per questo motivo non è attualmente possibile fornire informazioni complete su dimensioni e tempi di produzione.

crma.GRIB

I prodotti del post-processing eseguito dalle suite NAUSICA_run e NAUSICA_pp sono file GRIB con prefisso *CRMA* archiviati in due alberature parallele contenenti rispettivamente i campi sui livelli isobarici (PRS_LEV) e sui livelli del modello (MOD_LEV). Questi sono organizzati in cartelle giornaliere contenenti ciascuna le scadenze del dominio continentale con cadenza di 6 ore per file, quelle del dominio nazionale con 3 ore per file e quelle regionali con 1 ora per file.

Un anno di cartelle giornaliere occupa 0.11 Tb, quindi 2 Tb di file GRIB per l'intero periodo del progetto. Come si osserva il formato GRIB [7] è molto compatto, cosa che rappresenta una delle ragioni per le quali è stato scelto come formato di archiviazione.

Il controllo sulla presenza delle scadenze attese per i file GRIB prodotti viene riportato insieme alla statistica (massimi, minimi e valori medi) di tutti i campi contenuti in questi in un'apposita pagina HTML. Attualmente questo è implementato solo per i GRIB sui livelli isobarici e, dato che utilizza le CDO [24] e non le GRIB_API [26] con cui queste andrebbero sostituite, non è ottimizzato per l'utilizzo su periodi più lunghi di qualche giorno. Un esempio di output è riportato in Figura 4.3.

Stato avanzamento acquisizione scadenze				
FILE NAME	2015-01-05T00:00:00	2015-01-05T06:00:00	2015-01-05T12:00:00	2015-01-05T18:00:00
CRMA_WRF-4-CRMA_001_20150105_00 GRIB	0	0	0	0
FILE NAME	2015-01-06T00:00:00	2015-01-06T06:00:00	2015-01-06T12:00:00	2015-01-06T18:00:00
CRMA_WRF-4-CRMA_001_20150106_00 GRIB	0	0	0	0

Statistiche variabili								
#FILE	TIME	ACRONIMO	VAR	LEVEL	LTYPE	MIN	MEAN	MAX
CRMA_WRF-4-CRMA_001_20150105_00 GRIB	2015-01-05 00:00:00	Height300_[gpm]	var7	30000	100	8454.5	9068.3	9628.1
CRMA_WRF-4-CRMA_001_20150105_00 GRIB	2015-01-05 00:00:00	Height500_[gpm]	var7	50000	100	5073.5	5540.5	5886.7
CRMA_WRF-4-CRMA_001_20150105_00 GRIB	2015-01-05 00:00:00	Height700_[gpm]	var7	70000	100	2564.0	2935.6	3223.5
CRMA_WRF-4-CRMA_001_20150105_00 GRIB	2015-01-05 00:00:00	Height800_[gpm]	var7	80000	100	1622.1	1948.5	2142.7
CRMA_WRF-4-CRMA_001_20150105_00 GRIB	2015-01-05 00:00:00	Height925_[gpm]	var7	92500	100	482.89	779.61	947.11
CRMA_WRF-4-CRMA_001_20150105_00 GRIB	2015-01-05 00:00:00	T300_[K]	var11	30000	100	215.79	225.28	239.10
CRMA_WRF-4-CRMA_001_20150105_00 GRIB	2015-01-05 00:00:00	T500_[K]	var11	50000	100	234.04	247.78	265.53
CRMA_WRF-4-CRMA_001_20150105_00 GRIB	2015-01-05 00:00:00	T700_[K]	var11	70000	100	249.66	264.45	281.01
CRMA_WRF-4-CRMA_001_20150105_00 GRIB	2015-01-05 00:00:00	T850_[K]	var11	85000	100	254.58	272.36	287.77
CRMA_WRF-4-CRMA_001_20150105_00 GRIB	2015-01-05 00:00:00	T925_[K]	var11	92500	100	249.54	275.13	287.73

Figura 4.3: Esempio di pagina html per il controllo della produzione dei CRMA GRIB tramite suite NAUSICA_run o NAUSICA_pp sui livelli isobarici. Nella tabella superiore si legge lo stato delle scadenze attese (presenza indicata dal colore verde, assenza dal colore rosso). Nella tabella inferiore vengono riportati data, acronimo, nome, livello, tipo di livello, massimo, minimo e media di ciascun campo contenuto nei file analizzati.

4.2 Post-elaborazione del database

Il compito della suite WRF_ana_post è quello di eseguire il post-processing tramite UPP [8] degli output netCDF del modello WRF generando i file GRIB necessari come input per la catena modellistica F-AIR. Come input possono essere utilizzati i wrfout di tipo analisi prodotti dalla catena NAUSICA o alternativamente i wrfout di tipo forecast prodotti dalla catena operativa. È implementata anche la possibilità di creare GRIB con le caratteristiche di tipo previsione (nome dei file, tipologia, descrizione delle scadenze) a partire dai netCDF di analisi prodotti da NAUSICA. Questa opzione è stata utilizzata per colmare alcune lacune negli output della catena operativa.

La suite WRF_ana_post esegue inoltre il controllo dei GRIB prodotti creando dei plot di controllo su variabili scelte da file di inizializzazione. I dettagli tecnici di struttura e funzionamento sono riportati in Appendice F mentre in Figura 4.4 ne è rappresentata la struttura logica.

4.2.1 Prodotti della suite WRF_ana_post

I GRIB prodotti da questa suite hanno prefisso “FARM” e contengono i campi sui livelli isobarici necessari alla catena F-AIR. Sono organizzati in cartelle giornaliere contenenti ciascuna le scadenze del dominio continentale con cadenza di 6 ore per file, quelle del dominio nazionale con 3 ore per file e quelle regionali caratterizzate da scadenza oraria.

Come riassunto in Tabella 4.1 un anno di cartelle giornaliere occupa 0.18 Tb, quindi 3.1 Tb di file GRIB per l'intero periodo del progetto

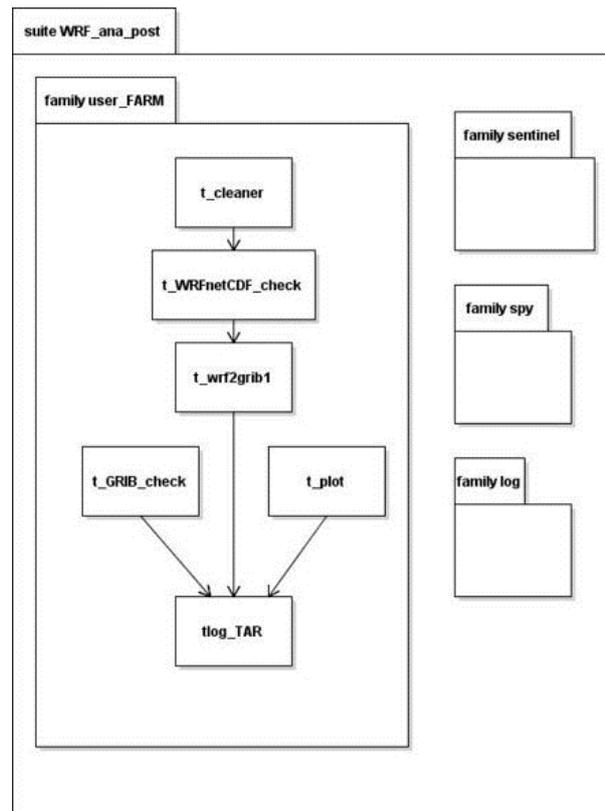


Figura 4.4: Schema della struttura logico-computazionale della suite WRF_ana_post i cui dettagli sulla struttura e l'utilizzo sono esaustivamente riportati in Appendice.

mentre il tempo di simulazione di un anno è di circa 27 ore che corrispondono a 19 giorni per l'intero periodo di 16 anni.

Il controllo della produzione avviene tramite plot che riportano massimi e minimi dei GRIB su intervalli scelti dall'utente, in Figura 4.5 è riportato un esempio di plot di controllo.

Vengono inoltre salvati in un file compresso gli output del flusso ecFlow e gli output ed error dei job.

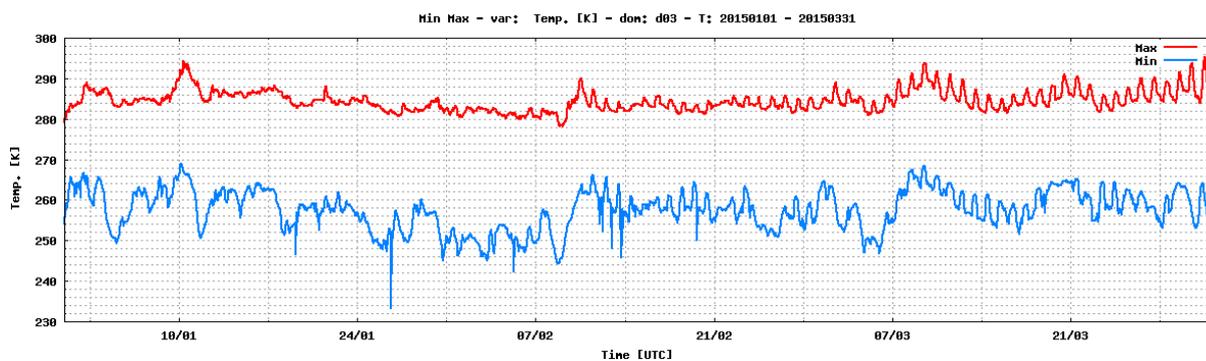


Figura 4.5: Esempio di plot di controllo della produzione di GRIB della suite WRF_ana_post. Per alcune grandezze scelte da file di inizializzazione sono graficate i valori massimo (linea rossa) e minimo (linea blu) letti nelle varie scadenze dei file GRIB.

4.3 Controllo qualità dei campi al suolo

Le suite NAUSICA_ver e NAUSICA_nc sono state create al fine di valutare la qualità delle simulazioni meteorologiche prodotte per mezzo di confronti statistici tra grandezze misurate e i rispettivi campi simulati contenuti nei file GRIB (suite NAUSICA_ver, Appendice G) o nei file netCDF (suite NAUSICA_nc, Appendice G).

Viene eseguito il confronto statistico, i test disponibili sono elencati nel Paragrafo 4.3.1, tra le serie temporali estratte in corrispondenza delle coordinate delle stazioni per le diverse variabili scelte da file di inizializzazione, le variabili disponibili sono elencate nel Paragrafo 4.3.2. Questo confronto, pur rappresentando un ottimo metodo di valutazione, risulta però affetto da discrepanze locali tra dati e simulazioni non necessariamente riscontrabili nell'evoluzione temporale delle condizioni meteorologiche che sono calcolate su scale più ampie. Queste possono, per esempio, essere dovute all'operazione di interpolazione del valore dei campi a partire dai punti di griglia più vicini alle coordinate fornite in particolari condizioni: casi di orografia complessa o vicinanza di confini terra/mare, entrambe presenti sul territorio regionale o possono essere dovute a fenomeni locali riprodotti traslati rispetto alle coordinate reali dell'evento di una distanza inferiore alla risoluzione utilizzata. Di conseguenza è stato scelto di introdurre nell'analisi anche il confronto tra le distribuzioni estratte su aree opportunamente definite, si veda il Paragrafo 4.3.3. Se da un lato con questo tipo di analisi si perde l'informazione temporale, viene infatti necessariamente considerata l'informazione integrata sul periodo scelto, dall'altro si ottiene una valutazione dell'accuratezza delle simulazioni ad una scala spaziale più significativa per i processi simulati. Un altro limite di questa seconda analisi è costituita dalla rappresentatività delle stazioni scelte rispetto alle aree e viceversa: è necessario che, per le grandezze analizzate, le stazioni coprano tutti i principali andamenti presenti nell'area corrispondente.

La struttura logico-computazionale di queste suite, Figura 4.7, risulta più complessa di quelle delle suite descritte nei paragrafi precedenti, ad esempio nei Paragrafo 4.1 e 4.2, questo è dovuto alla maggiore quantità di compiti distinti da eseguire: dalle estrazioni delle misure alle interpolazioni dei campi estratti dai GRIB, dall'esecuzione dei test statistici oggetto del Paragrafo 4.3.1, alla creazione dei plot relativi a questi. Utilizzo e struttura dettagliata di NAUSICA_ver sono descritte in Appendice G. Sono necessari inoltre diversi script in R [34] che eseguono i vari test statistici, script gnuplot [35] ed R per la visualizzazione dei risultati di alcuni di questi, eseguibili in FORTRAN [16] per l'estrazione delle serie temporali dalle misure, per la selezione spaziale tramite maschere e per la conversione e il calcolo di alcune variabili [25].

È interessante osservare che la gestione da parte dell'utente di tutta la suite avviene tramite un unico file *csv* contenente le informazioni essenziali all'attivazione del flusso in forma di codice. Un esempio di inizializzazione tramite questo file è riportato in Figura 4.6.

Anche per questa suite, come per le altre precedentemente descritte, vengono archiviati in un file TAR compresso gli *output* del flusso *ecFlow* e gli *output* ed *error* dei job.

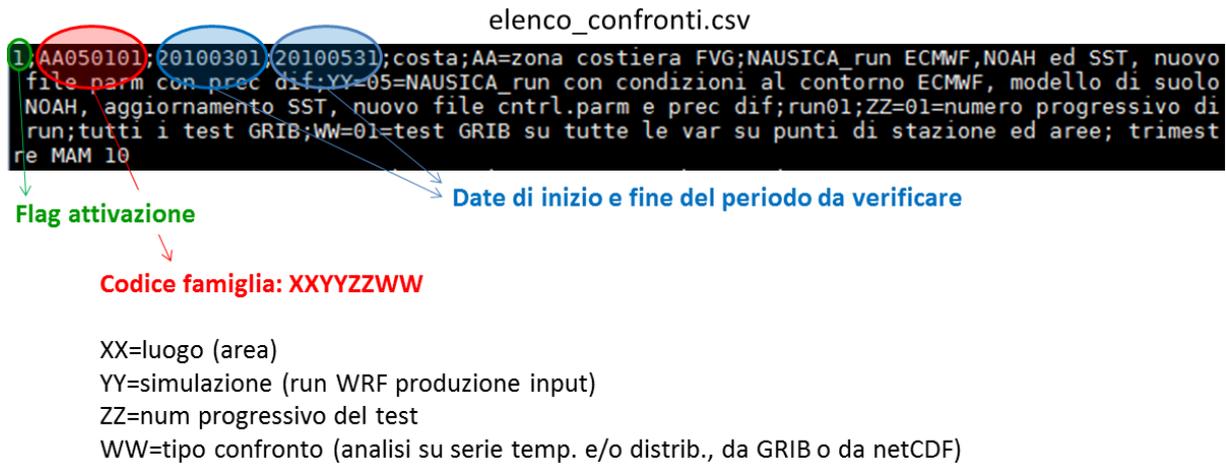


Figura 4.6: Esempio di attivazione della suite NAUSICA_ver tramite file csv. Ogni riga del file contiene le informazioni necessarie ad un run della suite codificate secondo la convenzione descritta in Appendice.

4.3.1 Test statistici eseguibili

Le due famiglie principali di analisi sono quella relativa ai confronti tra le serie temporali simulata e misurata nei punti di stazione e quella relativa ai confronti tra le distribuzioni sulle aree definite nel Paragrafo 4.3.3. La struttura della suite permette anche di scegliere se eseguire tutti i test o solo quelli di una delle due famiglie.

In Figura 4.8 sono riportati vari esempi dei prodotti dei test statistici descritti nelle prossime pagine. I risultati delle analisi sono invece discussi nel Capitolo 5.

Confronto tra misure e simulazione per serie temporali

Le prime informazioni che vengono ricavate dalle serie temporali e salvate in un file *txt*, a cui vengono poi progressivamente aggiunti i risultati di tutti i test eseguiti, sono medie, mediane, percentili (in particolare i 1°, 25°, 75°, 99°) e valori massimi e minimi per entrambe le fonti di dati.

Il primo test eseguito è la regressione lineare, da cui si può fare una prima valutazione della discrepanza tra campi misurati e simulati, corredato dal calcolo tramite R del bias e dei coefficienti di Pearson, Kendall e Spearman che restituiscono il grado di correlazione tra le due serie. Vengono poi creati gli istogrammi, sovrapposti in un ulteriore plot, da cui si può fare un primo confronto tra le forme delle distribuzioni.

Per leggere in modo chiaro il grado di correlazione tra le serie temporali e la deviazione standard che le caratterizza si utilizzano i

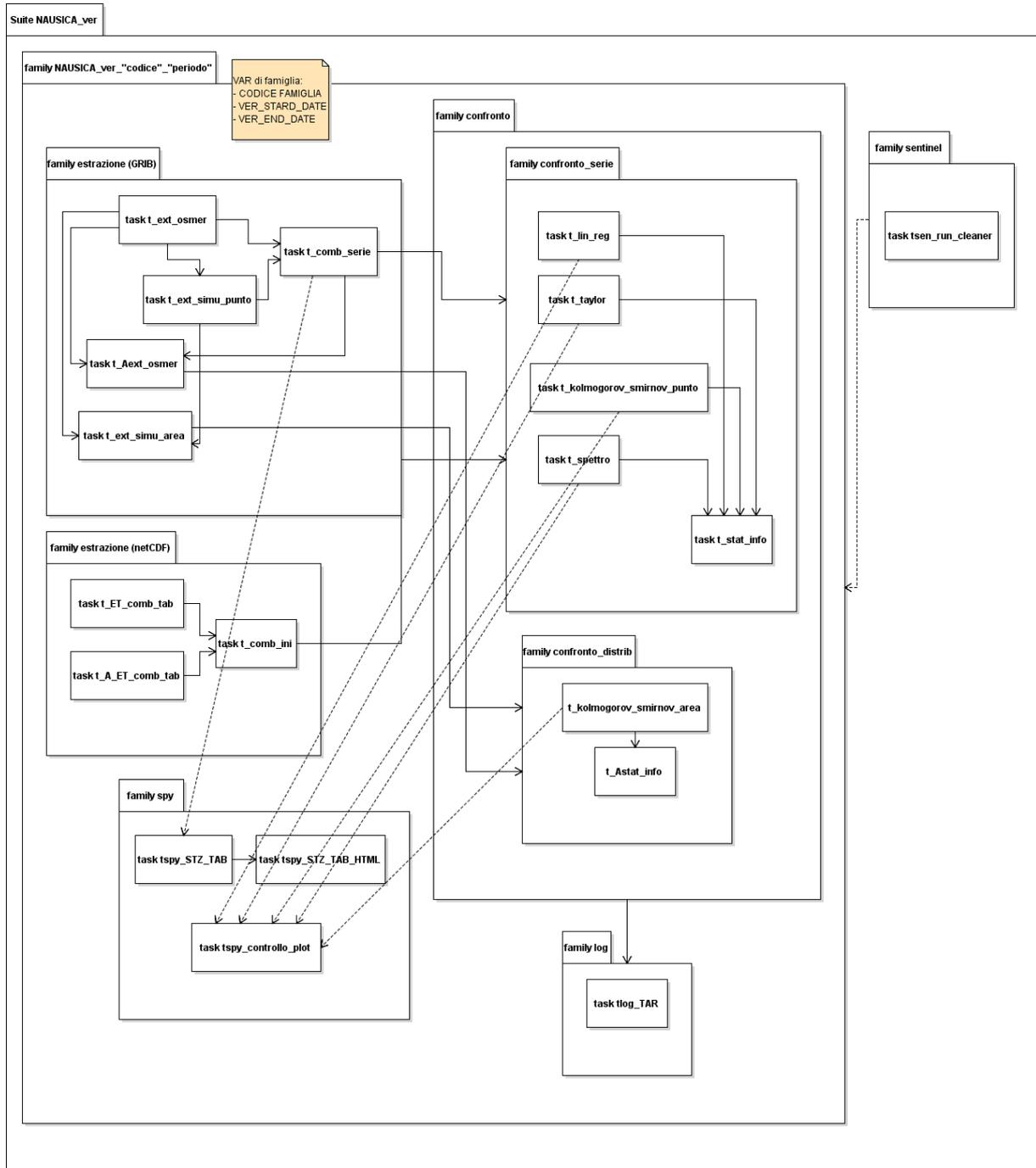


Figura 4.7: Schema della struttura logico/computazionale della suite NAUSICA_ver i cui dettagli sulla struttura e l'utilizzo sono esaustivamente riportati in Appendice.

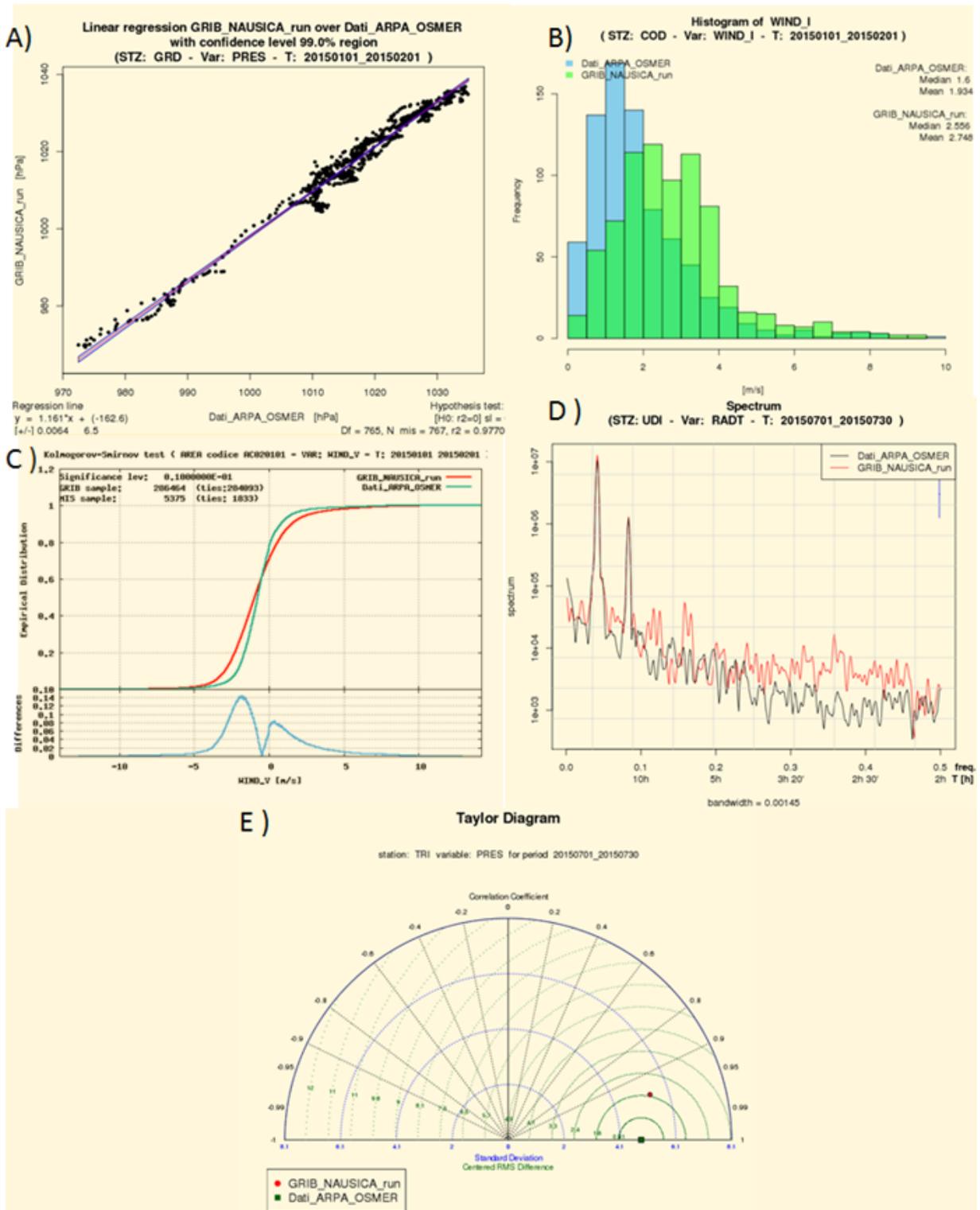


Figura 4.8: Esempi di alcuni plot prodotti dei test statistici di confronto tra misure e simulazioni eseguiti. In particolare in (A) è riportato il plot di regressione lineare tra le serie temporali relative alla pressione a livello di stazione nelle coordinate corrispondenti alla stazione ARPA di Grado nel mese di Gennaio 2015. In (B) sono mostrati gli istogrammi relativi alle serie temporali della componente u (EW) del vento presso la stazione ARPA di Codroipo nel mese di Gennaio 2015. La frequenza che compare in ordinata corrisponde al numero di occorrenze di valori appartenenti al range della classe dell'istogramma tra le scadenze orarie disponibili sul periodo selezionato. La (C) è il risultato del test di Kolmogorov-Smirnov applicato alle distribuzioni empiriche relative all'area di media pianura, definita nel Paragrafo 4.3.3, per la variabile wind v (NS) in gennaio 2015. Il plot in (D) contiene il confronto tra gli spettri della radiazione visibile relativi alle serie temporali nel punto di stazione di Udine in Luglio 2015. In (E) c'è invece un esempio di diagramma di Taylor (stazione di Trieste, pressione a livello di stazione, Luglio 2015).

diagrammi di Taylor. Dato che questi non contengono però informazione su un'eventuale differenza e sua localizzazione nella forma delle distribuzioni, viene eseguito il test di Kolmogorov-Smirnov che quantifica appunto la differenza delle distribuzioni empiriche all'interno del range di appartenenza dei dati.

L'ultimo confronto eseguito è quello tra gli spettri in potenza. Da questo, oltre che ottenere un'ulteriore valutazione della qualità della simulazione, si ricavano informazioni interessanti sulle caratteristiche di alcuni campi descritti nel Capitolo 5.

Confronto tra misure e simulazioni per distribuzioni su aree

Le distribuzioni su area vengono create, per quanto riguarda le misure, estraendo tutti i valori misurati disponibili per le stazioni contenute nell'area scelta nel periodo definito, si ottengono invece dalle simulazioni estraendo i valori del campo su tutti i punti di griglia interni per tutte le scadenze disponibili.

Anche in questo caso viene creato un file di testo contenente medie, mediane, percentili (in particolare il 1°, 5°, 75°, 99°) e valori massimi e minimi a cui vengono aggiunti il bias calcolato tramite R ed il risultato del test di Wilcoxon (test di ipotesi che valuta la popolazione di provenienza dei campioni confrontati). Ultimo test eseguito è il Kolmogorov-Smirnov per il confronto tra le distribuzioni empiriche.

4.3.2 Variabili confrontabili

Per effettuare i confronti descritti nel paragrafo precedente, Paragrafo 4.3.1, sono state utilizzate alcune delle osservabili che vengono misurate dalle stazioni ARPA OSMER anche se, per alcune stazioni, non sono disponibili alcune di esse o non lo sono in alcuni periodi.

Per confrontare il campo estratto dalle simulazioni con quello estratto dalle misure è stato necessario elaborare i risultati delle estrazioni prima di poterli mettere a confronto. Questa operazione è dovuta sia al fatto che le unità di misura non sempre sono le stesse (temperatura dell'aria a 180 cm che deve essere convertita perché misurata in °C ed estratta in K dagli output del modello, pressione istantanea dell'aria a livello di stazione che deve essere convertita perché misurata in hPa ed estratta in Pa) sia al fatto che alcune grandezze devono essere ricavate da altre (radiazione solare visibile misurata in kJ/h*mq ed estratta dai file GRIB in W/mq, componenti u e v della velocità istantanea del vento a 10 m di altezza estratte dagli output del WRF ma da calcolare a partire da intensità e direzione per i dati misurati e velocità istantanea del vento a 10 m di altezza misurata dalle stazioni ma da calcolare a partire dalle sue componenti nei dati estratti dalle simulazioni).

In alcuni casi è inoltre stato deciso di fare dei tagli sui valori estratti in modo da rendere più leggibili le analisi statistiche: per precipitazione oraria e radiazione sono stati considerati solo i valori non nulli, si veda Figura 4.9 che mostra il risultato del taglio sul confronto tra distribuzioni empiriche, mentre per l'umidità, a causa di problemi

sulle misure (strumenti che rilevano saturazione non realistica) si è deciso di considerare solo le umidità inferiori al 95%. L'effetto di questo taglio è rappresentato in Figura 4.9. Se da un lato questa scelta rende i confronti più significativi, dall'altro fa perdere di significato i valori assoluti delle statistiche. Sono state analizzate anche le precipitazioni orarie ad alta intensità: maggiori di 10 e 25 mm/h, per le quali è però necessario eseguire i test su periodi di tempo sufficientemente lunghi da avere una statistica significativa (nel caso delle stazioni scelte deve essere superiore o uguale all'anno. In Tabella 4.2 sono riassunte le variabili utilizzabili per il confronto, le operazioni necessarie per poterle confrontare e le unità di misura disponibili per l'analisi.

Variabile	Operazioni necessarie al confronto	Unità di misura
Precipitazione oraria $\neq 0$	Taglio valori = 0	[mm/h]
Precipitazione oraria > 10 mm/h	Taglio valori < 10 mm/h	[mm/h]
Precipitazione oraria > 25 mm/h	Taglio valori < 25 mm/h	[mm/h]
Umidità relativa < 95%	Taglio valori > 95%	[%]
Temperatura a 2 metri	Conversione unità di misura (misurata in C° ed estratta in K)	[K] e [C°] (consigliato)
Pressione a livello di stazione	Conversione unità di misura (misurata in hPa ed estratta in Pa)	[hPa]
Radiazione solare visibile $\neq 0$	Conversione e Taglio valori = 0 (misurata in kJ/h*mq ed estratta in W/mq)	[W/mq]
Componenti u e v della velocità del vento a 10 m	Estrate dai GRIB e da calcolare a partire da intensità e direzione per i dati misurati	[m/s]
Intensità del vento a 10 m	misurata e da calcolare a partire dalle sue componenti nei dati estratti dai GRIB	[m/s]
Evapotraspirazione giornaliera	Da calcolare a partire dal flusso orario	[mm/mq]

Tabella 4.2: Variabili utilizzabili per il confronto eseguito dalla suite NAUSICA_ver, operazioni necessarie per poterle confrontare e unità di misura disponibili per l'analisi.

4.3.3 Definizione delle aree di analisi

La definizione delle aree di analisi è una questione delicata in quanto, per poter ricavare delle distribuzioni il cui andamento abbia rilevanza fisica leggibile, devono essere scelte in modo da contenere zone meteorologicamente omogenee. Anche l'utilizzo delle distribuzioni per il confronto tra misure e simulazioni risente negativamente della presenza di zone disomogenee all'interno della stessa area perché, avendo al momento a disposizione un numero limitato di stazioni, può essere che le misure non siano rappresentative di tutti i possibili andamenti. Un esempio del problema appena descritto è rappresentato dall'effetto che si osserva nelle aree montane ed alpine che sono inevitabilmente disomogenee a causa della topografia molto varia. In alcuni casi anche nei risultati relativi alla media e alta pianura, come osservabile per esempio in Figura 5.6, sono evidenti due comportamenti distinti all'interno della stessa distribuzione.

I risultati attualmente disponibili relativi alle zone che sono state definite *Montagna* ed *Alpi*, presentati nel Capitolo 5, andranno quindi considerati meno significativi di quelli afferenti a tutte le altre aree.

Di seguito è riportato l'elenco delle aree definite, i codici ad esse attribuiti, il numero di stazioni contenute e il range assunto all'interno dell'area dei valori di altezza sul livello del mare dei punti di griglia dei campi simulati, il quale può essere considerato un indicatore grezzo dell'omogeneità meteorologica della zona.

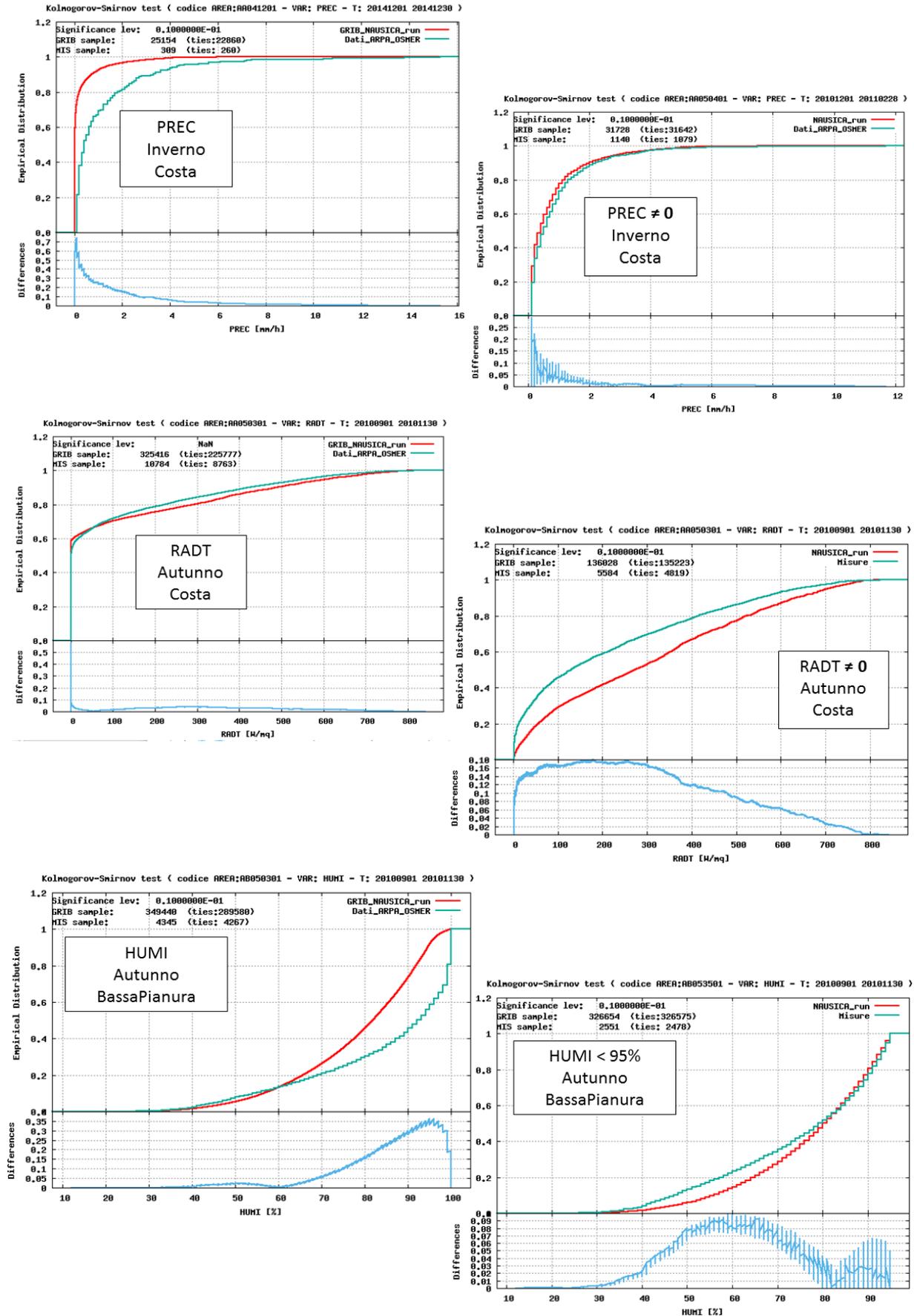


Figura 4.9: Effetto del taglio dei valori nulli sulle distribuzioni empiriche per precipitazione oraria e radiazione e del taglio al di sopra della soglia del 95% per l'umidità relativa.

In Figura 4.10 sono rappresentate le zone descritte e le stazioni considerate nell'analisi.

Elenco aree di analisi della suite NAUSICA_ver:

Zona di COSTA (cod AA):

- Range altezza sul livello del mare 0-5.4 m
- Stazioni contenute 5

Zona di BASSA PIANURA (cod AB):

- Range altezza sul livello del mare 0-148 m
- Stazioni contenute 3

Zona di MEDIA PIANURA (cod AC):

- Range altezza sul livello del mare 5.1-214 m
- Stazioni contenute 7

Zona di ALTA PIANURA (cod AD):

- Range altezza sul livello del mare 8.3-327 m
- Stazioni contenute 9

Zona di MONTAGNA (cod AE):

- Range altezza sul livello del mare 33-2121 m
- Stazioni contenute 13

Zona ALPI (cod AF):

- Range altezza sul livello del mare 610-2145 m
- Stazioni contenute 5

Zona CARSO (cod. AI)

- Altezza sul livello del mare 268 m
- Stazioni contenute 1

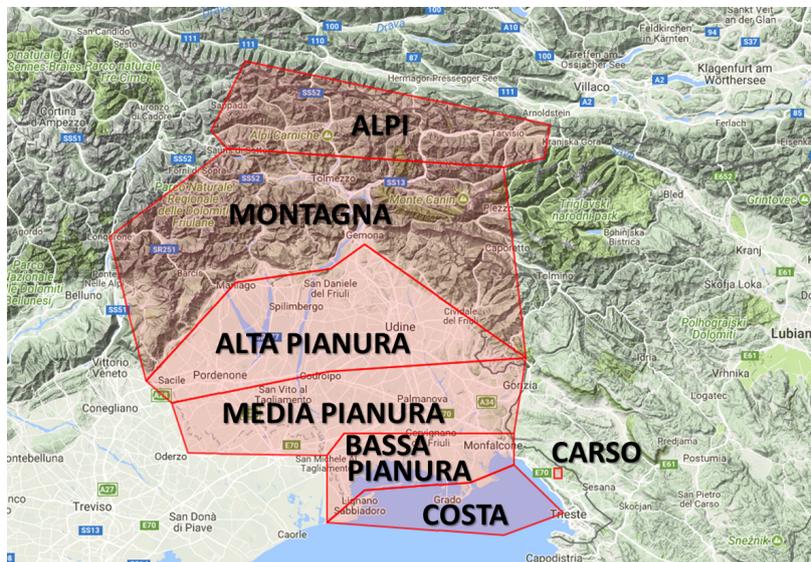


Figura 4.10: Rappresentazione delle aree meteorologicamente omogenee definite per l'analisi delle distribuzioni.

4.3.4 Prodotti delle suite NAUSICA_ver e NAUSICA_nc

I prodotti di questa suite sono rappresentati da file di testo contenenti le informazioni statistiche relative a serie temporali (file con prefisso *info_stat*) e distribuzioni su area (file con prefisso *info_stat_A*) ed i risultati dei test statistici per il confronto tra queste. Di questi file ne vengono prodotti uno per ogni combinazione variabile-stazione e uno per ogni combinazione variabile-area e sono salvati in cartelle con nome relativo al codice attribuito alla famiglia di run, del tipo XXYYZZWW-data-inizio_data-fine.

Vengono inoltre prodotti ed archiviati vari grafici in formato *png* e *ps*. In particolare vengono generati i plot relativi alle regressioni lineari e gli istogrammi di confronto tra le distribuzioni (file con prefisso *Lin_reg*), quelli contenenti il test di Taylor (*Taylor*), quelli che rappresentano il confronto tra le distribuzioni empiriche (prefisso *Kol_smn*) e i confronti tra gli spettri (prefisso *Spettro*).

Per quanto riguarda il confronto tra le serie temporali la produzione, nel caso di utilizzo di tutte le stazioni e variabili disponibili, è di 12 plot per ogni coppia variabile-stazione mentre per le analisi delle distribuzioni viene prodotto un unico plot per variabile per ogni area: il confronto tra le distribuzioni empiriche (prefisso *Kol_smn_A*).

Le dimensioni totali ottenute dalla produzione di tutti i file *txt*, *png* e *ps* relativi a tutti i test implementati (Paragrafo 4.3.1), per tutte le stazioni ARPA e per tutte le aree definite (Paragrafo 4.3.3), confrontando tutte le variabili disponibili (Paragrafo 4.3.2) risulta di circa 5 Gb per anno di simulazione se si considera di impostare confronti stagionali. Di conseguenza, con queste impostazioni, per la durata dell'intero progetto si otterrebbe un'occupazione del disco di 87 Gb. Se si decidesse di salvare solo i file *txt* ovviamente si ridurrebbe drasticamente il volume occupato fino a 3.4 Mb per anno, che corrispondono a circa 57 Mb per l'intero periodo. Le caratteristiche dei prodotti sono riassunte in Tabella 4.1, mentre in Tabella 4.3 sono riportate le dimensioni dei prodotti per area per mese.

Area	Dimensioni [Mb] (1 trimestre, 9 variabili, tutti i test)
COSTA e stazioni contenute	90
BASSA PIANURA e stazioni contenute	33
MEDIA PIANURA e stazioni contenute	124
ALTA PIANURA e stazioni contenute	156
MONTAGNA e stazioni contenute	179
ALPI e stazioni contenute	57
Stazione di Sgonico	18

Tabella 4.3: Dimensioni dei prodotti di NAUSICA_ver per area per trimestre considerando tutte le stazioni, tutte le aree, tutte le variabili, tutte le analisi disponibili e tutti i formati di output disponibili per i plot (*ps* e *png*).

I tempi di analisi sono strettamente legati al software utilizzato per l'estrazione dei campi dai GRIB: utilizzando le GRIB API [26] sono necessari circa 3 giorni per l'analisi mensile della simulazione relativa ad un anno considerando tutte le stazioni, tutte le aree, tutte le variabili, tutte le analisi disponibili e tutti i formati disponibili per i plot prodotti. Di conseguenza per tutti i 17 anni con questa configurazione

servono circa 50 giorni di run della suite. Per la versione della suite che utilizza le CDO [24] i tempi sono circa il doppio di questi.

A conclusione dell'analisi relativa ad ogni area viene aggiornata dalla suite un'apposita pagina html di controllo che mostra in modo leggibile se sono presenti tutte le scadenze attese per ogni variabile nei punti di stazione (Figura 4.11).

#NOME_STZ	NOME_VAR	DATA_I	DATA_F	FILE_COMB_OUT	RIGHE_LETTE	RIGHE_ATESE
Boa	HUMI	20150701	20150730	lustre_arpa_goglioa_scratch_ver_run_AA020701_20150701_20150730\tabella_comb_BOA_HUMI.tab	719	719
Boa	PREC	20150701	20150730	lustre_arpa_goglioa_scratch_ver_run_AA020701_20150701_20150730\tabella_comb_BOA_PREC.tab	0	719
Boa	PRES	20150701	20150730	lustre_arpa_goglioa_scratch_ver_run_AA020701_20150701_20150730\tabella_comb_BOA_PRES.tab	719	719

Figura 4.11: Esempio di tabella html per il controllo della presenza del numero di scadenze attese per l'estrazione di misure e campi dai GRIB nei punti di stazione, il verde corrisponde alla presenza di tutte le scadenze per variabile per stazione mentre il rosso indica che il numero di scadenze è inferiore a quello atteso. La tabella viene aggiornata a conclusione dell'analisi relativa a ciascuna area.

4.3.5 Visualizzazione dei risultati di controllo della qualità

Il numero di file che costituiscono tutti prodotti della valutazione della qualità delle simulazioni WRF realizzate nell'ambito del progetto NAUSICA è elevato e la ricerca delle informazioni in essi contenute necessitava di uno strumento che agevolasse l'utente in tale compito. Al fine di rendere consultabile il database di analisi svolte è stata progettata e costruita una pagina HTML che fosse utilizzabile anche da non esperti di computazione e portabile, assieme a tutti i file dei prodotti, su computer con sistemi operativi anche diversi dal Linux. In Figura 4.12 è riportata l'immagine della pagina HTML preparata. Il menu realizzato è dinamico: a seconda dell'area selezionata vengono mostrati i luoghi ad essa correlati (stazioni contenute o area intera). è inoltre presente un bottone per la visualizzazione delle aree su mappa Google Maps.

The screenshot shows the 'VERIFICHE NAUSICA (Campi al suolo)' web interface. It features a search panel with the following elements:

- Logo for 'arpa FVG' (Agenzia Regionale per la Protezione dell'Ambiente del Friuli Venezia Giulia) and 'CRMA' (Centro Regionale Meteo).
- A dropdown menu for 'CAMPO'.
- A dropdown menu for 'ANALISI'.
- Dropdown menus for 'AREA' and 'LUOGO'.
- Dropdown menus for 'PERIODO' and 'ANNO'.
- Buttons for 'VAI' and 'Reset'.
- A button labeled 'AREE FVG' at the bottom.

Figura 4.12: Pagina HTML realizzata per la visualizzazione dei prodotti della valutazione della qualità delle simulazioni WRF realizzate nell'ambito del progetto NAUSICA.

4.4 Controllo qualità dei campi in quota

Le suite NAUSICA_rds descritta nel Paragrafo 4.4.1 e la suite NAUSICA_srds, Paragrafo 4.4.3, sono state create al fine di valutare lungo la verticale i campi prodotti dalle simulazioni tramite confronto con i radiosondaggi disponibili. Nel Paragrafo 4.4.1 verrà descritta la metodologia adottata che si è però rivelata utilizzabile soltanto per il confronto tra campi che variano linearmente con la quota e per i gradienti. Il paragrafo successivo, Paragrafo 4.4.2, sarà quindi dedicato alla spiegazione del problema. Nel Capitolo 5 saranno di conseguenza riportati i risultati delle sole analisi svolte con metodologie corrette. Inoltre, nel Capitolo 6, verrà indicata una possibile modifica nell’approccio da implementare nel caso si decida di voler proseguire con questo tipo di indagine.

Lo schema della suite NAUSICA_rds segue quello della suite NAUSICA_ver come si può osservare da Figura 4.13 in quanto il flusso è analogo a quello che gestisce il confronto tra aree di quest’ultima.

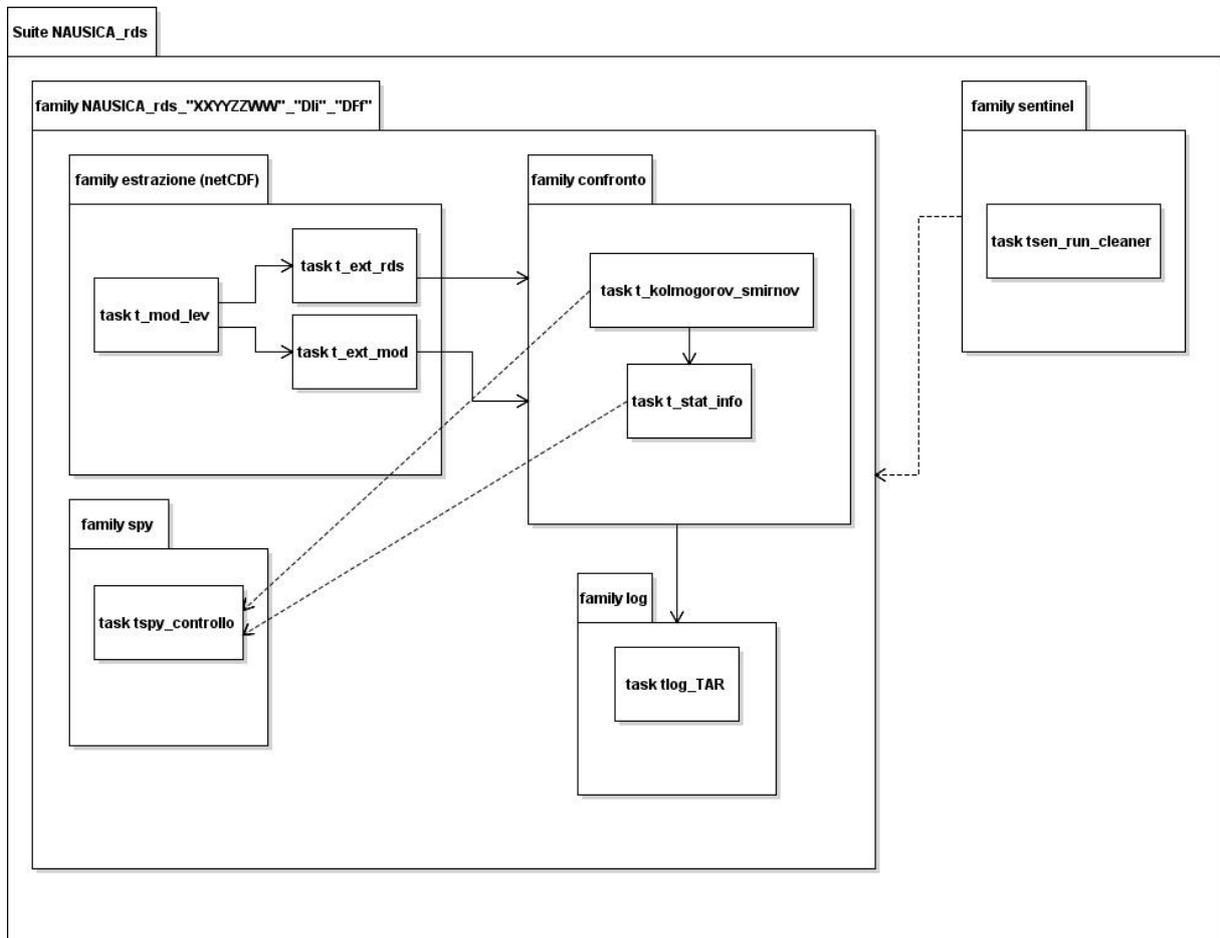


Figura 4.13: Flusso di lavoro della suite NAUSICA_rds.

4.4.1 Test eseguibili e variabili confrontabili in NAUSICA_rds

La suite NAUSICA_rds, descritta in Appendice G, esegue un confronto lungo la verticale tra misure dei radiosondaggi e campi estratti

dalla simulazione in analogia con quanto fatto per le aree superficiali dalla suite NAUSICA_ver.

Il confronto viene eseguito tra il campione di dati estratto da un'area appartenente alla superficie del modello centrato sul punto di griglia più vicino alle coordinate di lancio del radiosondaggio e il campione di dati corrispondente a tutte le misure del radiosondaggio relative ad una banda verticale centrata sull'altezza del livello del modello nel punto di griglia più vicino alla stazione di lancio, come schematizzato in Figura 4.14.

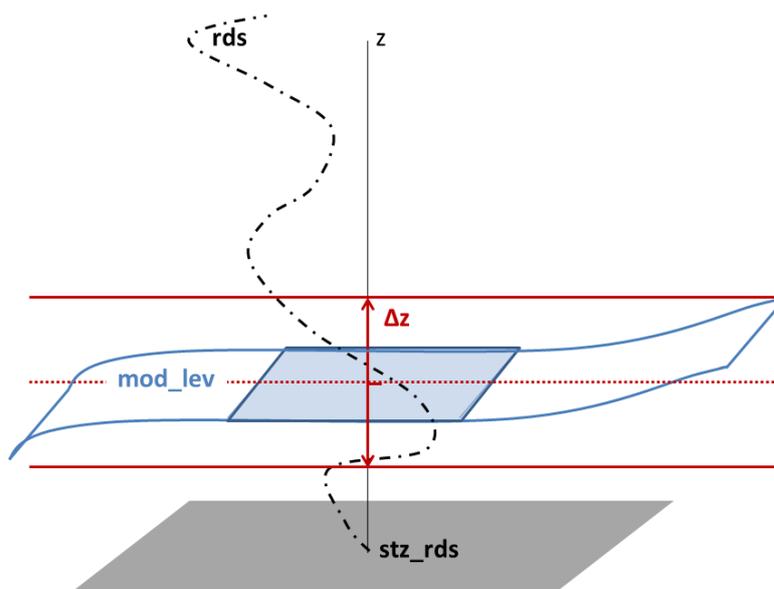


Figura 4.14: Schematizzazione dei due campioni confrontati in NAUSICA_rds per ogni livello estratti rispettivamente da modello (tutti i punti di griglia interni all'area selezionata, in blu) e radiosondaggio (tutte le misure effettuate nella parte di traiettoria contenuta all'interno della banda rossa). Come descritto nel Paragrafo 4.4.2 questo tipo di analisi è valido solo per campi che variano linearmente spostandosi lungo la verticale.

I confronti eseguiti dalla suite sono gli stessi di quelli eseguiti tra aree nella suite NAUSICA_ver, si veda il Paragrafo 4.3.1, ovvero viene creato un file di testo contenente medie, mediane, percentili (in particolare il 1°, 25°, 75°, 99°) e valori massimi e minimi a cui vengono aggiunti il bias calcolato tramite R ed il risultato del test di Wilcoxon (test di ipotesi che valuta la popolazione di provenienza dei campioni confrontati). Viene poi eseguito il test di Kolmogorov-Smirnov per confrontare le distribuzioni empiriche.

La suite è inoltre corredata da script accessori che calcolano e graficano il confronto tra i gradienti medi dei campi costruiti associando il valore all'altezza media delle fasce verticali, il bias e la differenza dei 25° e 75° percentili tra modello ed osservazioni per i singoli livelli del modello. In Figura 4.16 è riportato un esempio per ciascun tipo di analisi. Inoltre è possibile ottenere il profilo verticale medio alle 12:00 UTC e 24:00 UTC per radiosondaggi e simulazioni, Figura 4.15.

Attualmente è implementato il confronto per la sola stazione di lancio di Udine (Rivolto dal 2016, prima Campoformido) ma l'analisi è facilmente estendibile ad altre stazioni.

Le variabili confrontabili sono

- Temperatura
- Temperatura potenziale
- Pressione (affetta dalla problematica descritta nel Paragrafo 4.4.2)
- Intensità del vento (anch'esso affetto dalla problematica descritta nel Paragrafo 4.4.2)

Tramite gli script accessori descritti in Appendice H è possibile ottenere un confronto grafico tra i profili verticali dei gradienti. Queste ultime analisi sono da ritenersi non affette dalla problematica descritta nel Paragrafo 4.4.2.

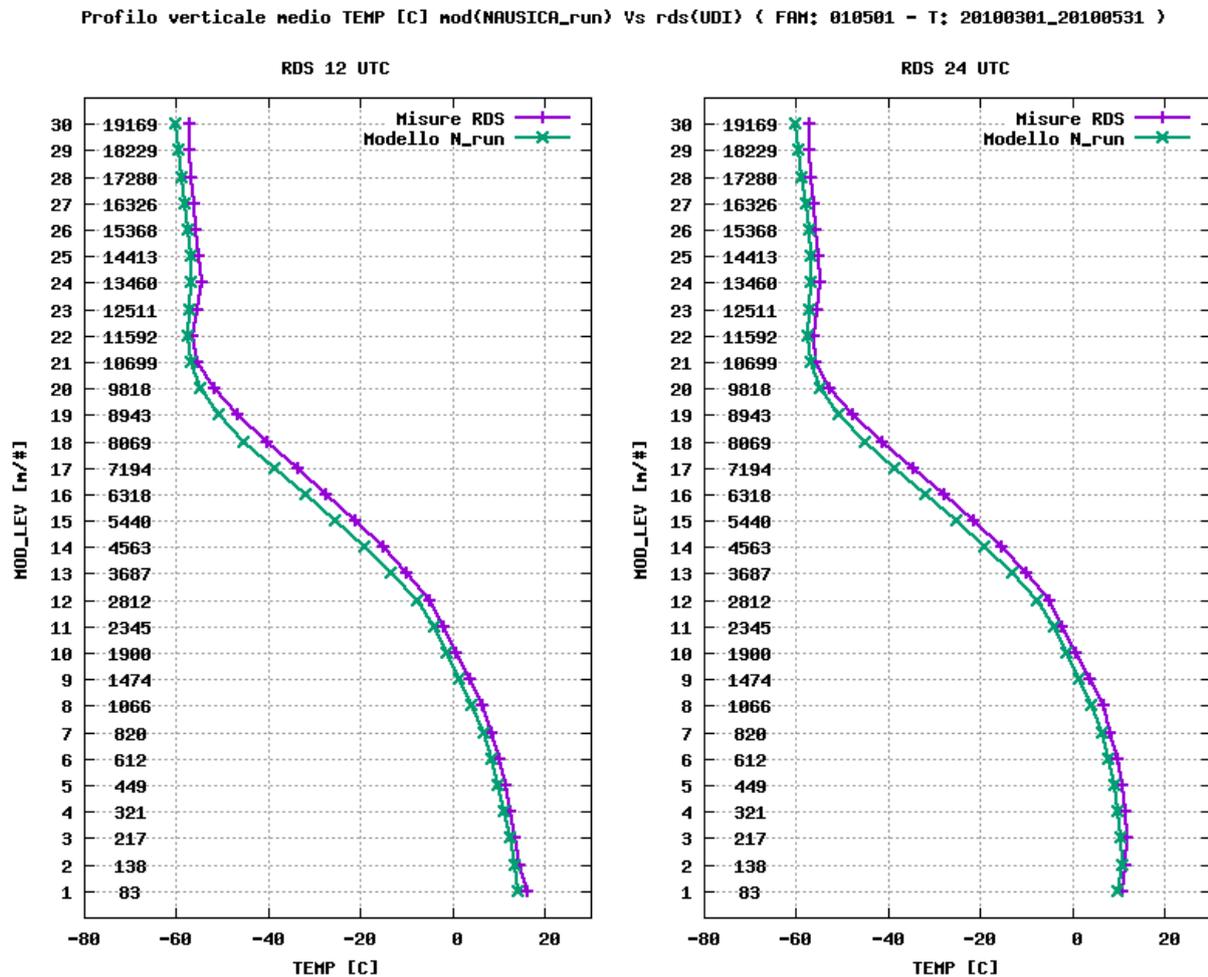


Figura 4.15: Profilo verticale medio della temperatura alle 12:00 UTC e 24:00 UTC nella primavera del 2010 a Rivolto per modello e misure dei radiosondaggi.

4.4.2 Problematiche riscontrate nella metodologia utilizzata

Osservando i risultati del confronto tra campi lungo la verticale ottenuti tramite la suite NAUSICA_rds, implementata come descritto nel

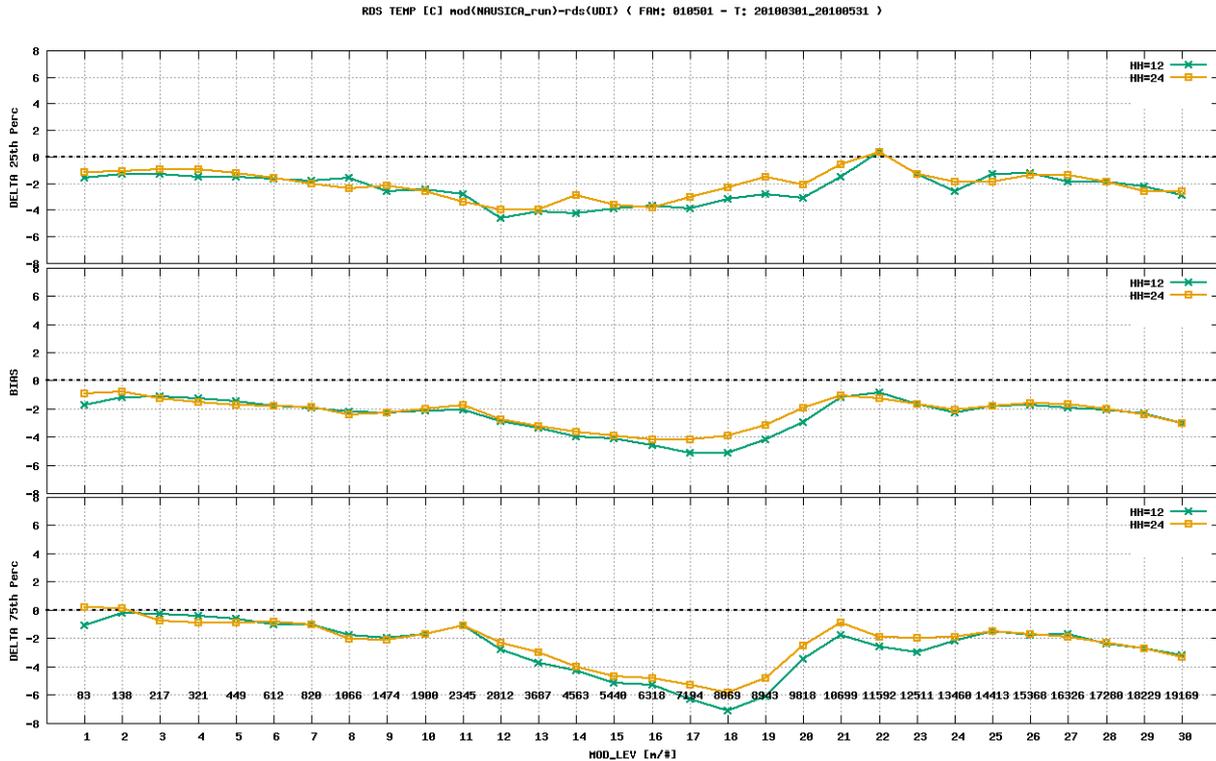


Figura 4.16: Bias (al centro), differenza tra venticinquesimo (in alto) e settantacinquesimo (in basso) percentile tra distribuzione relativa al modello e distribuzione relativa alle misure per la temperatura in corrispondenza di ogni livello del modello, per la primavera del 2010 a Rivolto.

Paragrafo 4.4.1 e in Appendice H, è parso evidente che la metodologia utilizzata è applicabile solo a variabili la cui variazione verticale è approssimabile ad una variazione lineare nel range Δz considerato; nel prossimo paragrafo verranno spiegati i motivi. Il risultato è che per la temperatura potenziale l'analisi resta valida, per la pressione può essere considerata valida per Δz più piccoli di quelli scelti nei run eseguiti (cioè dell'ordine di 10-20 m e non 100 m) mentre per il vento non ha senso. Di conseguenza anche l'analisi della temperatura risulta invalidata in quanto questa è ricavata per il modello a partire dalla temperatura potenziale utilizzando la pressione. In Figura 4.18 sono riportati alcuni esempi di output dell'analisi, le discontinuità nel caso della pressione sono da attribuirsi alla problematica appena descritta.

Nell'ipotesi di un'osservabile che vari linearmente lungo la verticale, per es. nel caso della temperatura per la quale $T(z) = T_0 + \Gamma * z$, si ottiene che il valore di aspettazione sull'intervallo $0 - \Delta z$ è pari al valore atteso nel punto $\Delta z/2$, infatti

$$T_{mod}(\Delta z/2) - E[T(\Delta z)] = (T_0 + \Gamma * \Delta z/2) - \int_0^{\Delta z} \Delta z T(z) dz = 0$$

Ci si aspetta quindi che, se estratti da campioni appartenenti alla stessa popolazione, i bias relativi a queste osservabili siano nulli. Si possono di conseguenza attribuire i valori ottenuti per il bias alla qualità della simulazione.

Viceversa, per osservabili che variano con la quota in modo non approssimabile con una variazione lineare nella banda verticale scelta, ad esempio per la pressione la cui dipendenza dalla quota è esprimibile come $P(z) = P_0 + e^{-az}$, dove a è dell'ordine di $10^{-4}[\text{m}^{-1}]$, si ottiene una differenza tra valore alla quota $\Delta z/2$ e valore di aspettazione che dipende da Δz e non più nullo. In particolare, questo risulta nullo solo per Δz tendente a 0 e ad infinito, come atteso dato che per Δz tendente a zero la variazione esponenziale è approssimabile con una variazione lineare e per Δz infinito stiamo considerando tutto il range di altezze. In particolare, nel caso della pressione, si trova che

$$\frac{P_{mod}(\Delta z/2) - E[T(\Delta z)]}{P_0}$$

ha l'andamento riportato in Figura 4.17.

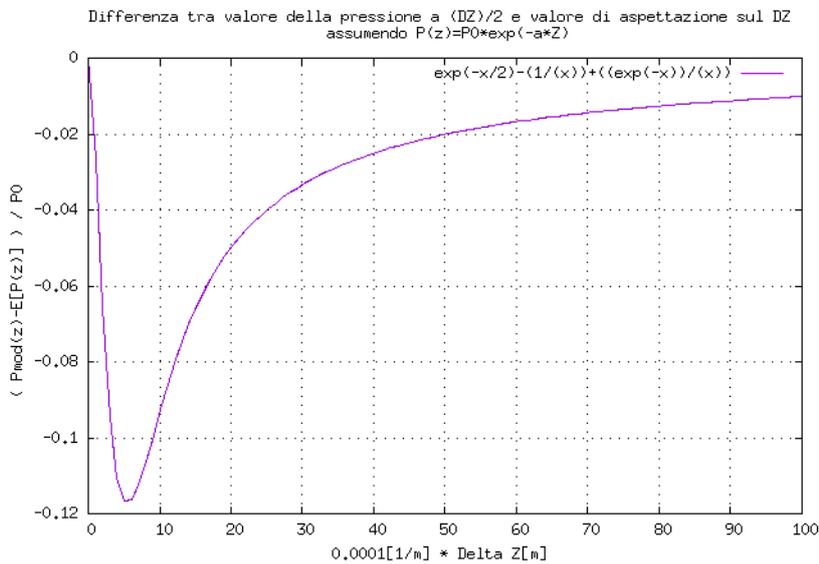


Figura 4.17: Andamento della differenza tra valore del campo atteso a $\Delta z/2$ e valore di aspettazione del campo sull'intero Δz in funzione del Δz , nell'ipotesi che la pressione abbia una dipendenza dalla quota del tipo $P(z) = P_0 + e^{-az}$, dove a è dell'ordine di $10^{-4}[\text{m}^{-1}]$. Le discontinuità che si osservano in Figura 4.18 sono dovute al fatto che il valore di Δz scelto per l'analisi (in particolare per i livelli dall'ottavo in poi) corrisponde ad un valore non nullo di questa funzione e di conseguenza aggiunge un contributo dominante al bias rendendo trascurabile il contributo dovuto alle differenze tra campi simulati e misurati a cui siamo interessati.

Sarebbe interessante trovare il valore di Δz al di sotto del quale l'analisi può considerarsi valida per tutte le variabili ma questo è funzione della dipendenza specifica del campo con la quota (non sempre nota) e inoltre c'è il problema che per Δz troppo piccoli il campione di misure risulta non rappresentativo. Si ritiene quindi che non abbia senso lavorare in questa direzione ma piuttosto nel cambiare la metodologia di estrazione dei campioni come verrà descritto nel Capitolo 5.

A conferma dei risultati della discussione presentata nelle righe precedenti si riporta in Figura 4.19 l'analisi della pressione ottenuta analogamente a quella mostrata in Figura 4.18) ottenuta diminuendo il Δz a 50 m per tutti i livelli. Si osserva la diminuzione dell'effetto dovuto alla problematica sul bias.

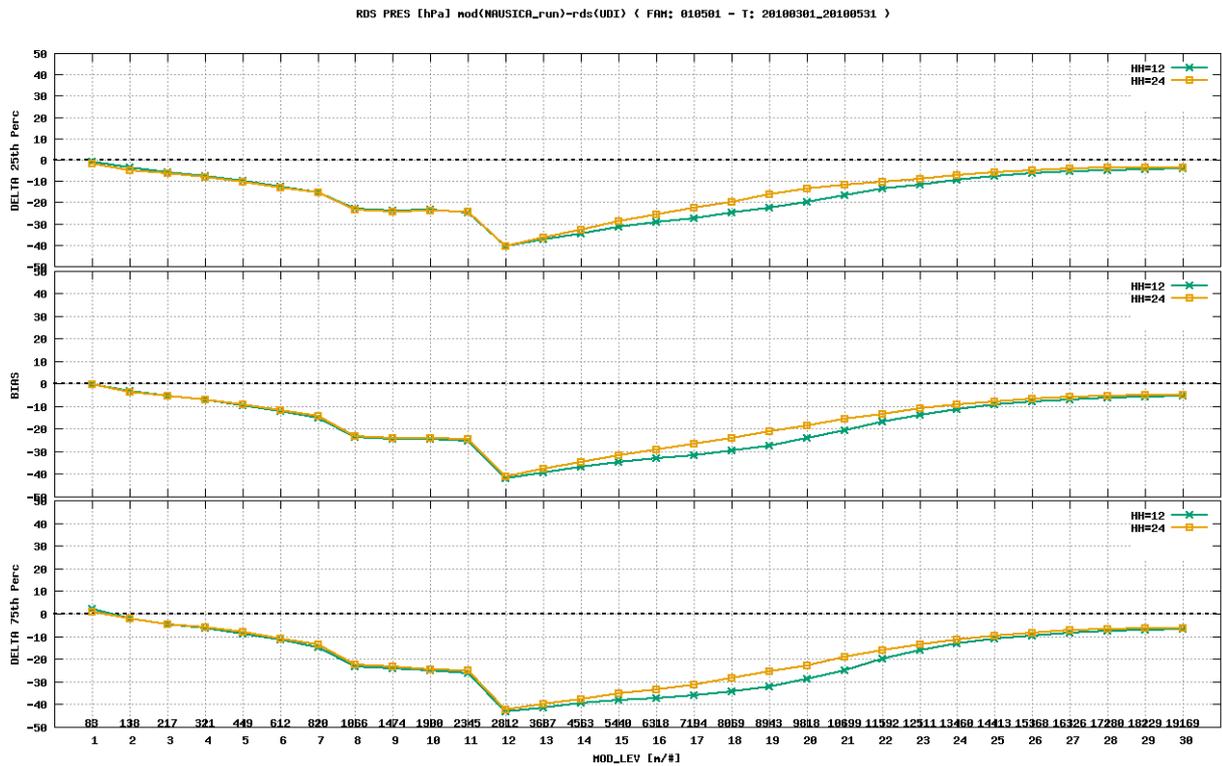
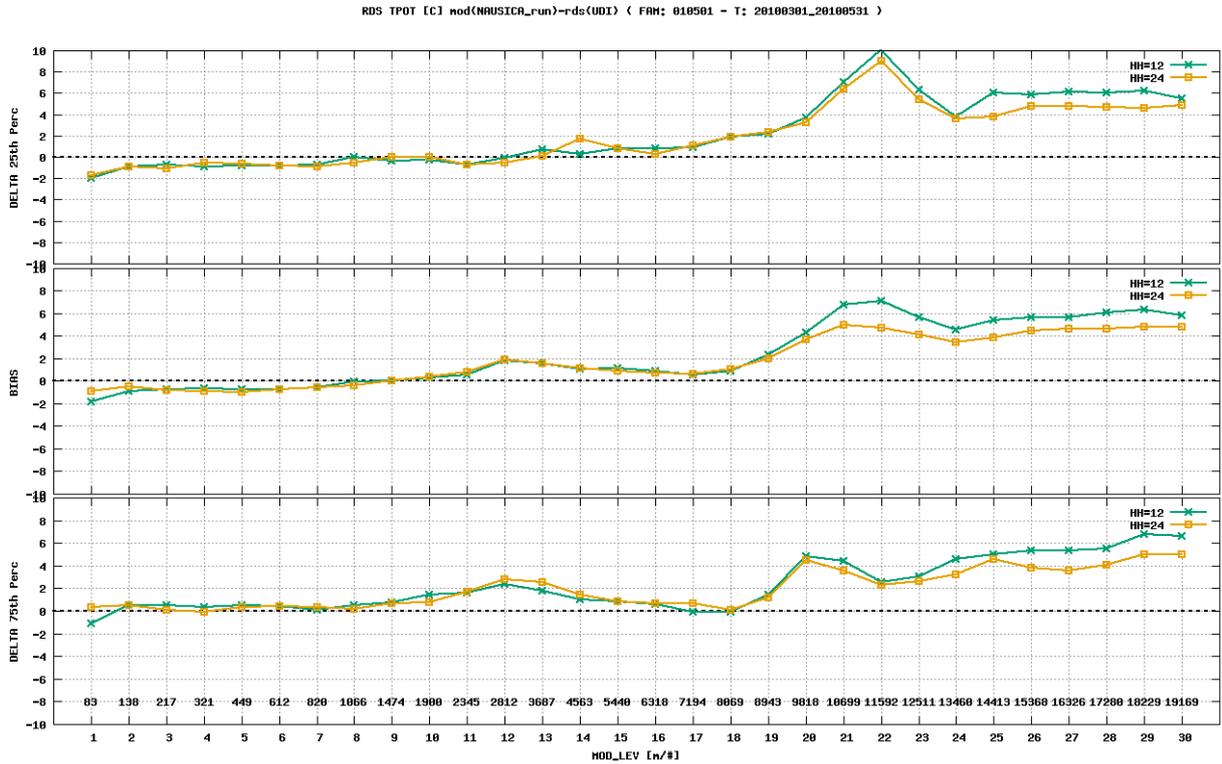


Figura 4.18: Alcuni risultati delle analisi svolte con la suite NAUSICA_rds da cui emerge la problematica descritta, nel grafico superiore è riportata l'analisi valida del bias e differenza tra 25° e 75° percentile delle distribuzioni di modello e dati raccolti da radiosondaggi del campo di temperatura potenziale nella primavera del 2010 a Rivolto utilizzando bande verticali di 50 m per i primi 8 livelli e di 200 m per i livelli successivi, mentre in quello inferiore è riportata la stessa analisi relativa alla pressione. In questo secondo caso l'analisi è da ritenersi NON attendibile per i livelli successivi all'ottavo come spiegato nel Paragrafo

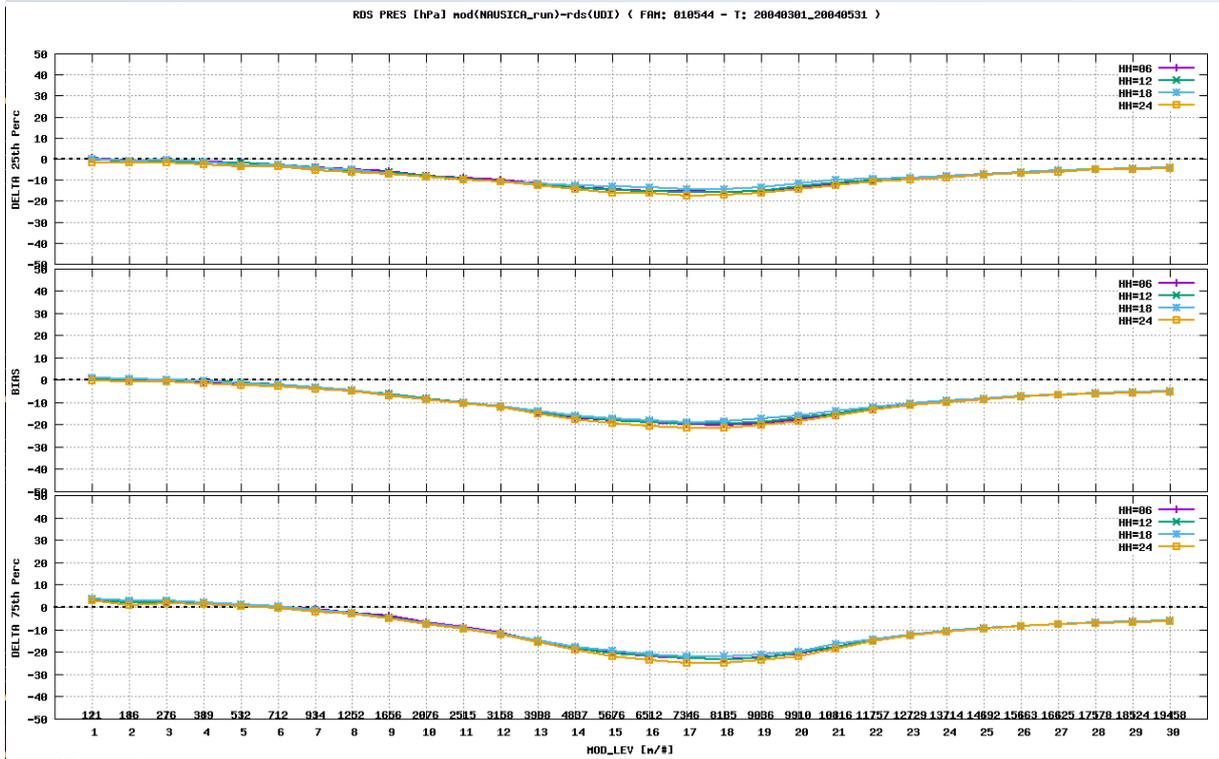


Figura 4.19: Analisi analoga a quella riportata in Figura 4.18) (bias e differenza tra 25° e 75° percentile delle distribuzioni di modello e dati raccolti da radiosondaggi del campo di temperatura potenziale nella primavera del 2004 a Rivolto) ottenuta utilizzando bande verticali di 50 m per tutti i livelli. Si può notare la scomparsa delle discontinuità presenti nel caso precedente.

4.4.3 Suite NAUSICA_srds

La suite NAUSICA_srds, anch'essa descritta in Appendice H, è stata costruita per sviluppare un ulteriore tipo di analisi verticale ma anche questa è affetta da problematiche (sempre descritte in Appendice H) che la rendono non attendibile. In pratica lo scopo era quello di effettuare le stesse analisi statistiche della suite NAUSICA_rds giorno per giorno e non sul periodo di tempo selezionato. Ciò che rende questa analisi non attendibile è l'intervallo di tempo di salvataggio del valore dei campi nelle simulazioni (un salvataggio ogni ora) che rende il confronto uno a uno di queste con le misure dei singoli lanci di radiosondaggi (ovviamente non istantanee) troppo stringente.

5

Risultati del progetto disponibili

All'interno di questo capitolo sono elencati i prodotti del progetto disponibili. Questi sono rappresentati dagli output nativi della simulazione meteorologica generati dalla suite NAUSICA_run, dai risultati delle varie post-elaborazioni eseguite dalle suite NAUSICA_pp e WRF_ana_post, descritte rispettivamente nei Paragrafi 4.1 e 4.2, e dalle analisi svolte da NAUSICA_ver, elencate nel Paragrafo 4.3, le quali includono i file contenenti la statistica, i plot relativi ai test statistici e le conclusioni tratte dall'analisi di alcuni di questi.

Lo stato di avanzamento dell'intero progetto a Gennaio 2017 è indicato dal diagramma di Gantt riportato in Figura 2.3 mentre quello relativo a Giugno 2017 è indicato dal *diagramma di Gantt* riportato in Figura 5.1. Da questo si può notare che la totalità dei prodotti è stata generata a partire da metà marzo 2017, cioè da quando sono state disponibili le risorse di calcolo della coda *Julia*. A causa delle modifiche apportate alla configurazione è stato necessario ripetere anche i run svolti nel periodo precedente. In particolare, rispetto alle produzioni precedenti:

- Negli output del WRF è stata sostituita la precipitazione oraria a quella cumulata, si veda il Paragrafo 3.3.1
- E' stato scelto di produrre GRIB in Lambert conformal e non Lat Lon, vedere il Paragrafo 3.3.2
- E' stata aggiunto il test sull'evapotraspirazione descritto nel Paragrafo 4.3.2
- Sono stati tagliati dai campioni analizzati nei test statistici i valori nulli di precipitazione e radiazione e le umidità relative superiori al 95% come descritto nel Paragrafo 4.3.2
- Sono state introdotte analisi annuali delle precipitazioni superiori a 10 mm/h e 25 mm/h, si veda il Paragrafo 4.3.2

Nel lavoro svolto durante gli ultimi mesi di progetto ci si è principalmente concentrati sul completamento di simulazioni ed analisi. Inoltre sono stati interpretati alcuni spettri di vento e alcune delle analisi dei profili verticali valide nonostante la problematica descritta nel Paragrafo 4.4.2.

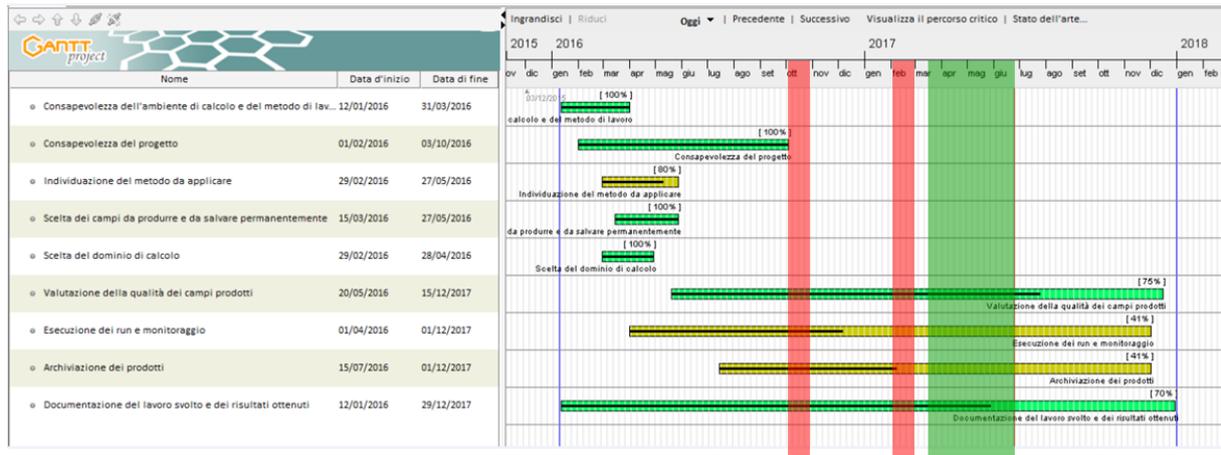


Figura 5.1: Diagramma di GANTT relativo allo stato di avanzamento del progetto NAUSICA a Giugno 2017. Si possono notare le interruzioni delle attività avvenute ad Ottobre 2016 e Febbraio 2017 per problemi contrattuali (periodi evidenziati in rosso) e l'attesa delle risorse di calcolo necessarie per la produzione, disponibili solo a partire dal 15 marzo 2017 (periodo evidenziato in verde), che ha causato ritardi nella produzione del database.

5.1 File archiviati

In Tabella 5.1 sono riassunti i file e i relativi periodi di copertura e dimensioni disponibili. Nei prossimi paragrafi saranno elencate le caratteristiche di ciascuno di essi in dettaglio.

File disponibili	Periodo coperto	Descrizione	Dimensione
wrfout.nc	Dal 2001 al 2017 compresi	Output run annuali WRF	≈ 24Tb
FARM.GRIB	2015 completo (da nc analisi) 2016 primo trimestre (da netCDF analisi) 2016 primo semestre (da netCDF forecast) 2017 completo (da nc analisi)	Post-elaborazione UPP	≈ 1.06Tb
CRMA.GRIB	Dal 2001 al 2017 compresi	Post-elaborazione UPP	≈ 2.0Tb
info_stat.txt	Trimestri MAM, JJA, SON e DJF per Dal 2001 al 2017 compresi Precipitazioni > 10 mm/h e 25 mm/h per Dal 2001 al 2017 compresi EVTR trimestri MAM, JJA, SON e DJF per Dal 2001 al 2017 compresi	Informazioni statistiche su serie temporali e distribuzioni spaziali	≈ 30Gb
Lin_reg .ps .png Taylor .ps .png Kol-smf .ps .png Spettro .ps .png	Dal 2001 al 2017 compresi Trimestri MAM, JJA, SON e DJF per Precipitazioni > 10 mm/h e 25 mm/h per 2010 EVTR trimestri MAM, JJA, SON e DJF per Dal 2001 al 2017 compresi	Plot test statistici	≈ 34Gb
Out (plot e txt) NAUSICA_rds	Trimestri MAM, JJA, SON e DJF Dal 2001 al 2017 compresi	Analisi e plot relativi	≈ 2 Gb

Tabella 5.1: Elenco dei prodotti del progetto disponibili

wrfout.nc

Il primo prodotto disponibile è rappresentato dagli output nativi di run annuali del modello WRF [3], non destaggerati, in formato netCDF [6], ottenuti da run guidati da condizioni al contorno ECMWF, Paragrafo 2.2, e caratterizzati dalla configurazione descritta nel Capitolo 3, ovvero su tre domini innestati definiti nel Paragrafo 3.2: continentale, nazionale e regionale, con risoluzioni spaziali ed intervalli temporali di salvataggio rispettivamente di 50 km, 10 km, 2

km e 6 ore, 3 ore, 1 ora, con modello di suolo *Noah LSM*, con aggiornamento delle temperature superficiali del mare ogni 6 ore, si vedano rispettivamente i Paragrafi 3.10 e 3.10, e senza limite per le grandezze cumulate, Paragrafo 3.3.1.

E' da sottolinearsi il fatto che l'inizializzazione delle simulazioni è stata fatta ogni primo gennaio tramite condizioni iniziali lette dai file GRIB di analisi prodotti da ECMWF. Sono archiviati insieme ai *wrfout* anche i file di restart *wrfst* prodotti per ogni giorno simulato.

Il periodo attualmente simulato copre il periodo che va dal 1 Gennaio 2001 al 31 Dicembre 2017.

Le dimensioni occupate dai file descritti sono di circa 24 Tb.

Il codice di archiviazione di questi file, stabilito in base alla convenzione per la classificazione delle simulazioni del CRMA [33], è 1390FA01D0.

FARM.GRIB

I periodi attualmente post-elaborati per mezzo di run dell'UPP [8] gestiti della suite *WRF_ana_post* sono: gli anni 2015 e 2017 completi e il primo trimestre dell'anno 2016 ottenuti dai *netCDF* prodotti da *NAUSICA_run* e il primo semestre dell'anno 2016 creato dai *netCDF* prodotti dalla catena operativa. Per quest'ultima produzione, a causa dell'assenza di alcune scadenze nei *netCDF* operativi, sono stati coperti alcuni intervalli di date mancanti utilizzando gli output derivanti da *NAUSICA_run* e convertendo questi in un formato analogo ai corrispondenti risultati da previsioni. In particolare i periodi coinvolti da questa operazione sono dal 5 al 10 Gennaio 2016 e dal 13 al 16 Maggio 2016.

Le cartelle giornaliere prodotte da file di tipo analisi contengono un intervallo corrispondente a 24 ore mentre, quelle generate dalle previsioni, ne contengono 121 ciascuna. Di conseguenza lo spazio totale occupato dai GRIB prodotti è di circa 1 Tb di cui una parte relativa alle scadenze derivanti dal post-processing delle previsioni.

CRMA.GRIB

Questi sono il risultato del flusso di post elaborazione tramite UPP [8] degli output della suite *NAUSICA_run* per mezzo di *NAUSICA_pp* e *NAUSICA_run* e sono stati utilizzati come input per la suite *NAUSICA_ver*. In questa produzione sono stati generati solo i GRIB relativi ai campi sui livelli isobarici in quando, essendo disponibili anche gli output nativi, è sempre possibile utilizzare la suite per produrre anche quelli sui livelli del modello.

Questa post-elaborazione copre il periodo 2001-2017 compresi gli estremi. La dimensione totale occupata su disco è pari a 2 Tb.

info_stat.txt

Nei file testo di prefisso *info_stat* sono contenute le informazioni statistiche relative alle serie temporali e alle distribuzioni su area

(*info_stat_A*) descritte nel Paragrafo 4.3.3 e i risultati di tutti i test statistici eseguiti, di questi file ce ne sono per ogni intervallo, uno per ogni combinazione variabile-stazione e uno per ogni combinazione variabile-area e sono salvati in cartelle con nome relativo al codice attribuito alla famiglia di run, del tipo *XXYYZZWW-data-inizio_data-fine*, da NAUSICA_ver.

Sono stati analizzati tutti i trimestri stagionali (MAM=Marzo, Aprile, Maggio; JJA=Giugno, Luglio, Agosto; SON=Settembre, Ottobre, Novembre; DJF=Dicembre, Gennaio, Febbraio) del periodo 2001-2017 e i run annuali del periodo 2001-2017 per le precipitazioni intense.

plot.ps e plot.png (Lin_reg, Taylor, Kol_smn, Spettro)

I test statistici associati al periodo descritto nel paragrafo precedente, cioè i trimestri stagionali 2010-2014 e precipitazioni intense annuali per il 2010, sono corredati dai corrispondenti plot: regressioni lineari e istogrammi (prefisso *Lin_reg*), test di Taylor (prefisso *Taylor*), confronti tra distribuzioni empiriche per i punti di stazione e per le aree (prefissi *Kol_smn* e *Kol_smn_A_*) e confronti tra gli spettri (prefisso *Spettro*).

La dimensione totale dei file prodotti è di circa 34 Gb.

5.2 Valutazione della qualità delle simulazioni

La valutazione della qualità delle simulazioni tramite confronto dei risultati di queste con le misure a disposizione è lo scopo del flusso di lavoro delle suite NAUSICA_ver, NAUSICA_nc e NAUSICA_rds, la cui struttura è descritta nel Capitolo 4.

I test di confronto sono stati svolti per periodi stagionali, ovvero trimestrali, mentre le precipitazioni intense sono state valutate annualmente in modo da avere un campione sufficientemente numeroso. Le analisi dei campi in quota non inficcate dalla problematica descritta nel Paragrafo 4.4.2 sono state svolte su periodi stagionali.

In generale si nota un accordo molto soddisfacente tra modello e misure per tutte le variabili analizzate, a dimostrazione di questa osservazione si valutino gli esempi riportati nelle Figure 5.2, 5.9, 5.10, 5.14, 5.16, 5.17, 5.19, 5.20, 5.29, 5.24, 5.24. In particolare per quanto riguarda i valori medi di temperatura, umidità e precipitazioni intense, superiori cioè a 10 mm/h, la qualità delle simulazioni è ottima. Le discrepanze principali si osservano per i venti, che risultano sovrastimati dal modello di circa 2 m/s sulla maggior parte delle aree e per la radiazione che risulta invece sovrastimata sistematicamente dal modello di circa 100 W/mq. Discrepanze di ordine inferiore si notano nelle code delle distribuzioni di temperatura ed umidità per le quali il modello produce un range di valori più stretto, per i valori bassi della precipitazione e l'evapotraspirazione che risultano sottostimate dalle simulazioni e nella distribuzione della pressione che risulta più larga rispetto a quella misurata.

Sono riportati di seguito i risultati dei test analizzati per singola variabile. Si ricorda che, come spiegato nel Paragrafo 4.3.3, i test

per le aree di Montagna e Alpi sono da considerarsi meno significativi in quanto le stazioni contenute non ne rappresentano tutta la fenomenologia e che l'area Carso contiene una sola stazione.

5.2.1 *Temperatura*

I test relativi alla temperatura mostrano un'elevata accuratezza nella riproduzione dei valori medi da parte del modello. In Figura 5.2 sono riportati due esempi di confronto tra le distribuzioni empiriche di misure e modello, il primo concernente uno dei casi riprodotti meglio dalla simulazione e il secondo relativo ad uno dei casi con maggiore discrepanza. I bias risultano inferiori a circa 0.8°C in estate ed autunno, inferiori a circa 2°C in primavera e a circa 2.5°C in inverno. Come si può leggere in Tabella 5.3 i valori medi risultano tendenzialmente sottostimati nelle stagioni calde e sovrastimati in quelle fredde.

Le code delle distribuzioni, la cui tendenza è evidenziata in Figura 5.4, sono state valutate tramite l'analisi della differenza tra primi e novantanovesimi percentili delle distribuzioni. I valori ricavati sono riportati in Tabella 5.5. Dall'analisi emerge che la distribuzione simulata è più stretta rispetto a quella dei valori misurati. In particolare le massime, specialmente quelle estive, risultano sottostimate e le minime, specialmente le invernali, sovrastimate. Il 99° percentile (coda della distribuzione relativa ai valori alti) risulta infatti sottostimato in tutte le aree e in tutte le stagioni mentre il 1° percentile (coda della distribuzione relativa ai valori bassi) tendenzialmente sovrastimato. È interessante notare l'andamento discontinuo nella derivata della distribuzione empirica, soprattutto nella zona di media pianura riportata in Figura 5.6, che indica che si sta considerando un'area contenente zone con comportamenti meteorologici differenti.

Analizzando invece il confronto tra gli spettri di temperatura si osserva un ottimo accordo misure-modello, si veda Figura 5.7.

È interessante notare che, considerando un periodo trimestrale, nello spettro della temperatura compare, oltre ai picchi dovuti all'oscillazione diurna ed armoniche successive, anche il segnale relativo all'oscillazione stagionale chiaramente visibile anche nella serie temporale.

5.2.2 *Intensità del vento*

Questa è la variabile caratterizzata dalla peggior accuratezza nella riproduzione delle misure da parte del modello tra quelle analizzate. È infatti presente un bias simulazioni-misure di circa 1 o 2 m/s su quasi tutte le aree, un esempio è riportato in Figura 5.8. L'area con accordo migliore è la costa, mentre Carso, bassa e media pianura presentano il bias maggiore. In Figura 5.9 è riportato un esempio di confronto sulla costa.

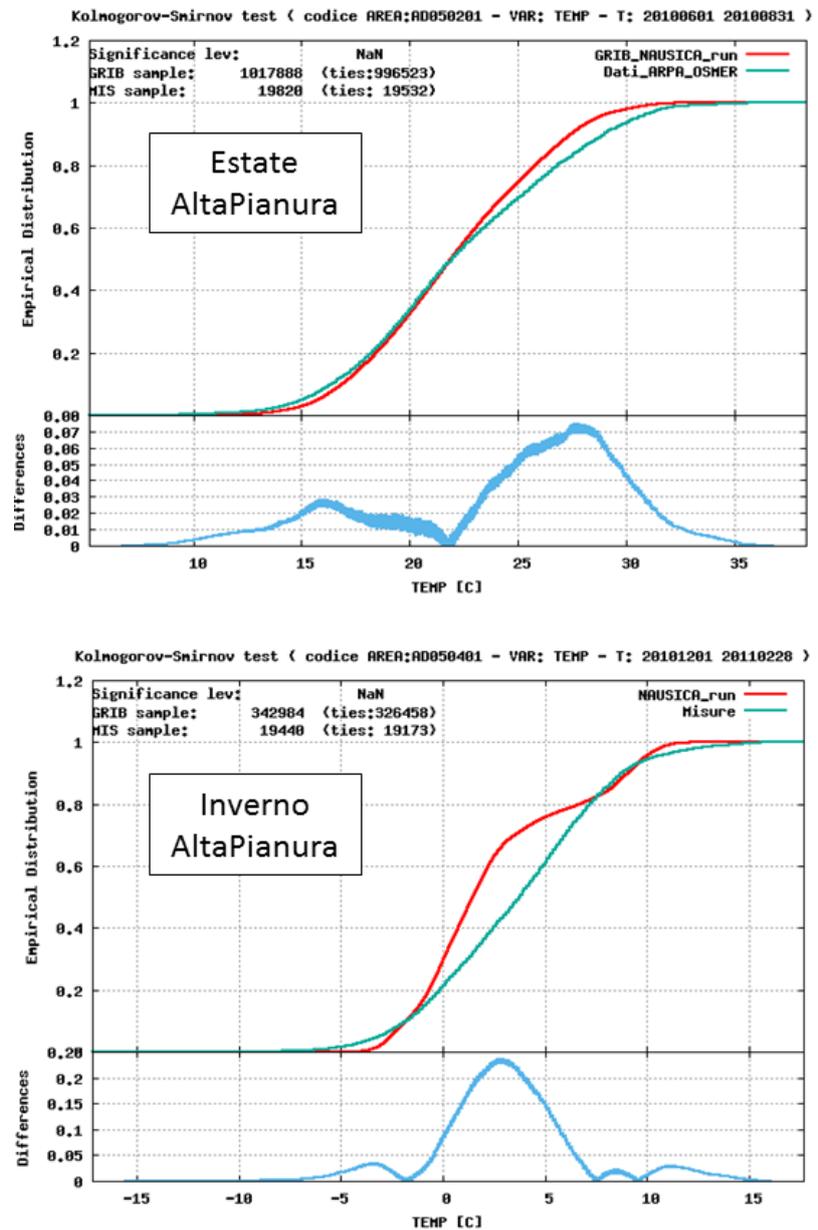


Figura 5.2: Confronto tra distribuzioni empiriche di modello e misure della temperatura sull'area di alta pianura in estate (Giugno, Luglio e Agosto 2010) e in inverno (Dicembre 2010, Gennaio e Febbraio 2011). Questi due confronti rappresentano uno dei casi di miglior riproduzione della distribuzione dei valori misurati e uno dei casi con bias misure/modello maggiore. Si può inoltre notare che le distribuzioni prodotte dal modello sono leggermente più strette di quelle relative alle misure (si veda anche Figura 5.3).

BIAS TEMPERATURE 2010: media(simulazioni) – media(misure)

STAGIONE AREA	Primavera	Estate	Autunno	Inverno
Costa	-1.5 °C	-0.8 °C	-0.6 °C	-0.1 °C
Bassa Pianura	-0.8 °C	-0.1 °C	0.7 °C	0.2 °C
Media Pianura	-0.8 °C	-0.3 °C	0.1 °C	-1.0 °C
Alta Pianura	-0.8 °C	-0.3 °C	0.0 °C	-1.1 °C
Montagna	-1.4 °C	-0.3 °C	-0.2 °C	-2.1 °C
Alpi	0.0 °C	0.9 °C	1.5 °C	-0.9 °C
Carso	-1.5 °C	-0.7 °C	-0.7 °C	-1.0 °C

BIAS TEMPERATURE 2011: media(simulazioni) – media(misure)

STAGIONE AREA	Primavera	Estate	Autunno	Inverno
Costa	-1.6 °C	-0.8 °C	-0.5 °C	2.1 °C
Bassa Pianura	-1.4 °C	0.0 °C	1.0 °C	2.6 °C
Media Pianura	-1.0 °C	0.2 °C	1.0 °C	1.5 °C
Alta Pianura	-1.0 °C	0.1 °C	0.9 °C	1.5 °C
Montagna	-1.1 °C	0.2 °C	0.7 °C	1.8 °C
Alpi	-1.6 °C	-0.8 °C	-0.5 °C	2.1 °C
Carso	-1.8 °C	-0.4 °C	0.0 °C	-1.7 °C

■ media(simulazioni) < media(misure)
 ■ media(simulazioni) > media(misure)

Figura 5.3: Bias (differenza tra i valori medi delle simulazioni e quelli delle misure) per le temperature stagionali degli anni 2010 e 2011. Il bias è sempre inferiore ai 2°C. Si può notare che tendenzialmente i valori medi risultano sottostimati nelle stagioni calde e sovrastimati in quelle fredde.

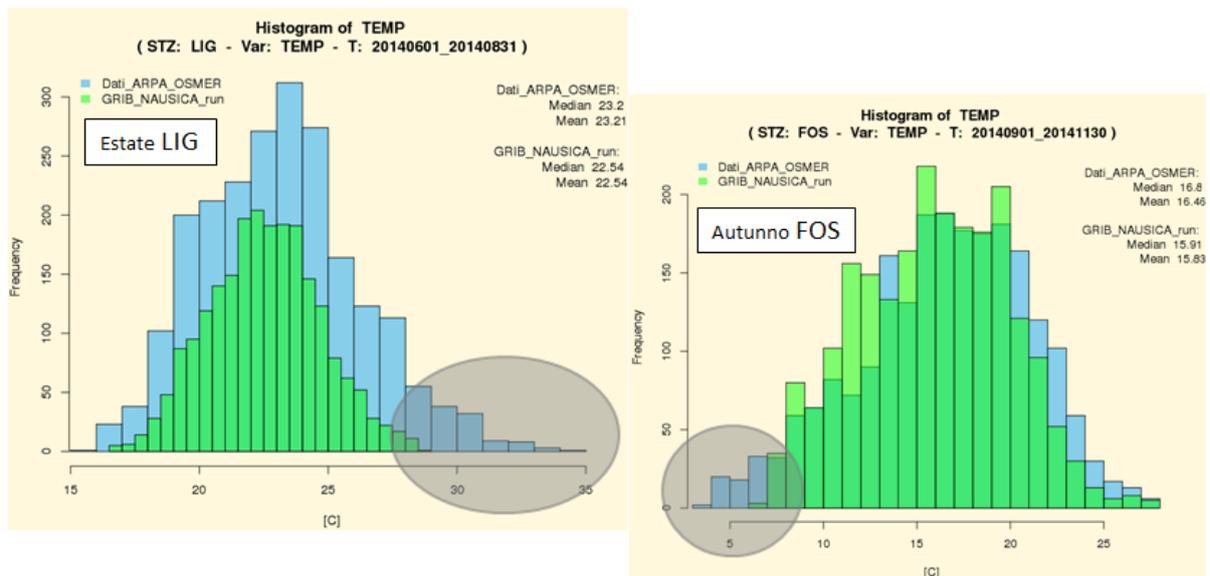


Figura 5.4: Istogrammi della temperatura di simulazioni e misure relativi all'estate 2014 a Lignano e all'autunno 2014 a Fossalon. E' evidenziata la tendenza delle simulazioni a sottostimare le temperature massime, specialmente in estate, e a sovrastimare le minime, specialmente invernali. Questa tendenza è analizzata in generale tramite lo studio dei 1° e 99° percentile riportati in Tabella 5.5.

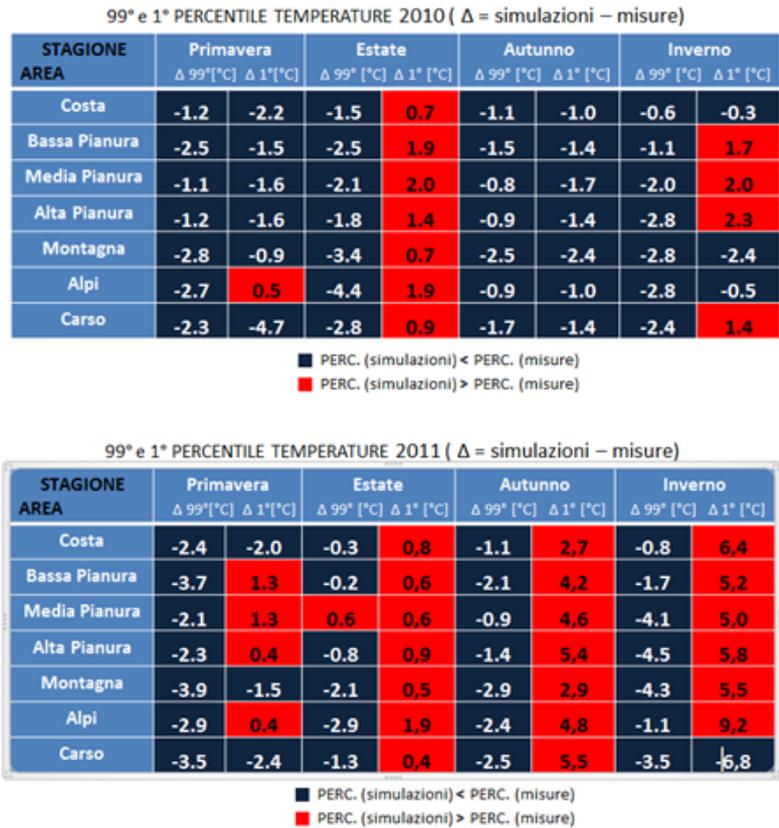


Figura 5.5: Differenza tra 99° e 1° percentili di simulazioni e dati per le temperature stagionali per gli anni 2010 e 2011. Il 99° percentile risulta sottostimato in tutte le aree e in tutte le stagioni mentre il 1° percentile tendenzialmente sovrastimato.

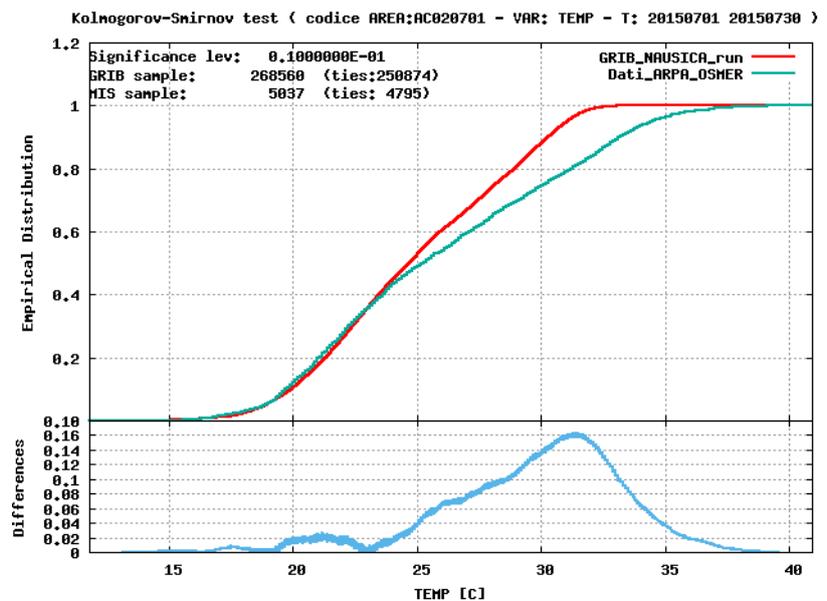


Figura 5.6: Confronto tra distribuzioni empiriche di simulazioni e modello della temperatura sull'area di media pianura per il mese di Luglio 2015. Si nota una sottostima delle alte temperature e l'andamento di crescita discontinuo nella che indica che si sta considerando un'area contenente zone con comportamenti meteorologici differenti.

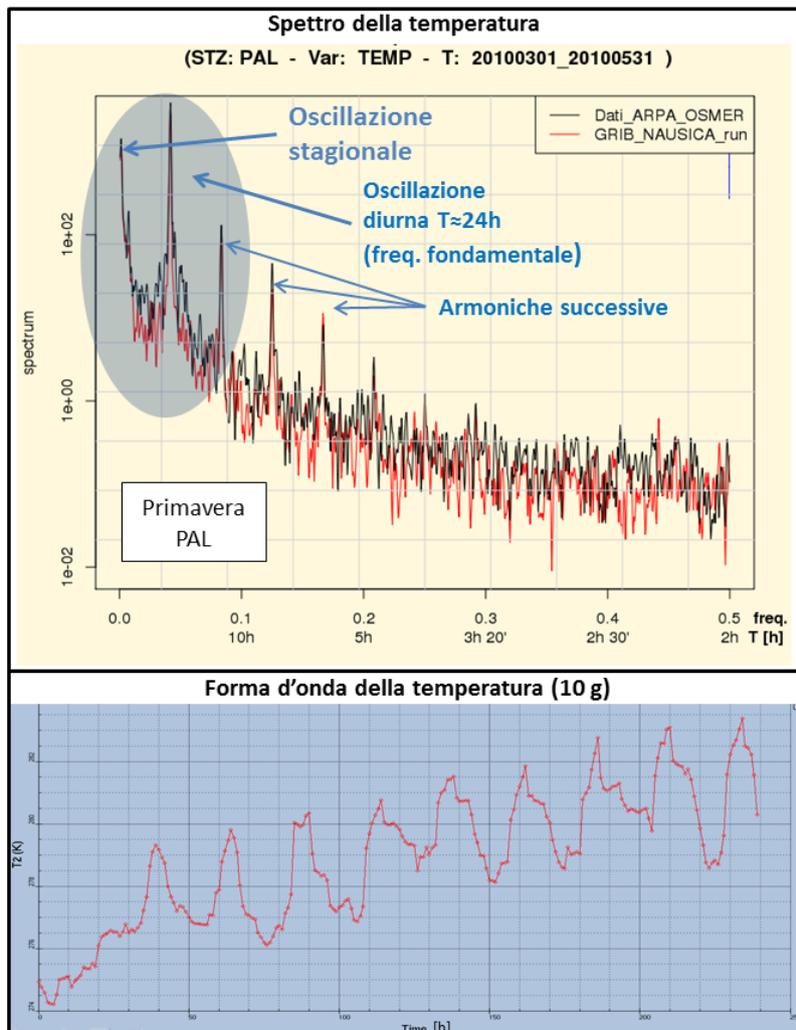


Figura 5.7: Spettro in potenza (sopra) e forma d'onda su 10 giorni (sotto) della temperatura presso la stazione di Palazzolo dello Stella in Primavera (Marzo, Aprile e Maggio 2010). Si nota un ottimo accordo tra simulazione e dati. Nello spettro compare, oltre ai picchi dovuti all'oscillazione diurna ed armoniche successive, anche il picco relativo all'oscillazione stagionale.

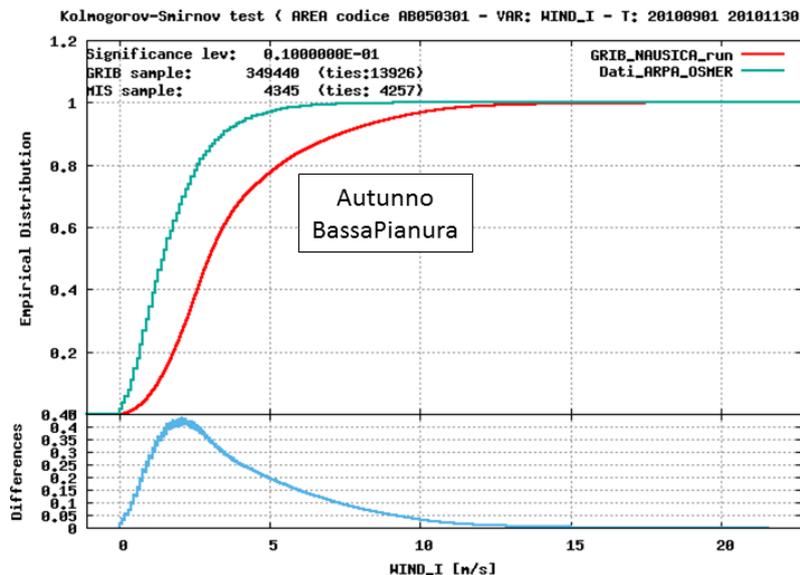


Figura 5.8: Confronto delle distribuzioni di misure e modello dell'intensità del vento sull'area di bassa pianura in autunno (Settembre, Ottobre e Novembre 2010). Rappresenta un buon esempio della generale tendenza del modello a sovrastimare i venti.

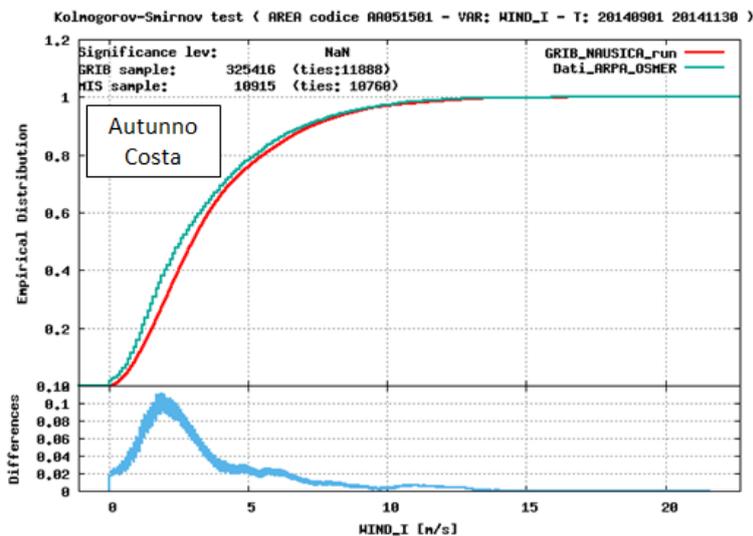


Figura 5.9: Confronto tra distribuzioni simulata e misurata dell'intensità del vento sull'area di costa in autunno (Settembre, Ottobre e Novembre 2010). Come si può anche osservare confrontandola con il caso di bassa pianura di Figura 5.8, in quest'area l'accuratezza nella riproduzione dei dati da parte del modello risulta maggiore.

La qualità nella riproduzione dell'intensità del vento risulta però molto variabile localmente: l'andamento presso alcuni punti di stazione, ad esempio Cividale di Figura 5.10 (grafico in alto), è ben rappresentato in tutte le stagioni anche nei valori di bassa intensità, inferiori cioè a 5 m/s, che caratterizzano la scala tipica di effetti locali come le brezze di valle, mentre in altri casi, per esempio Faedis, plot in basso in Figura 5.10, il bias è sistematicamente molto elevato.

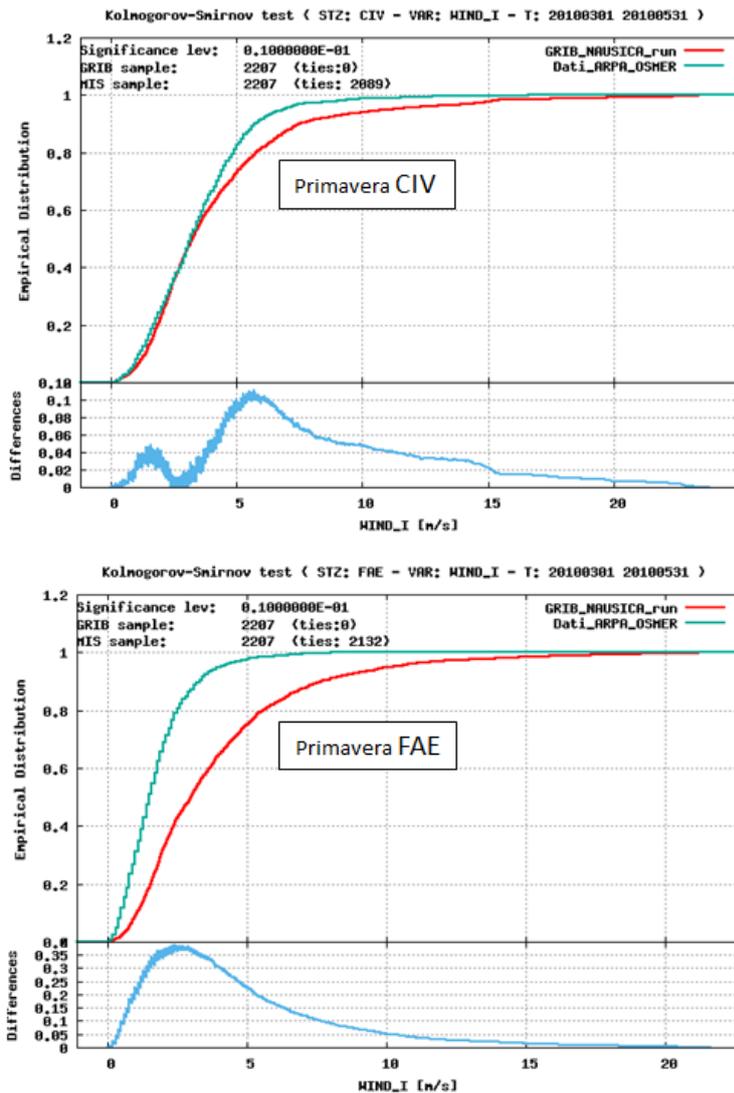
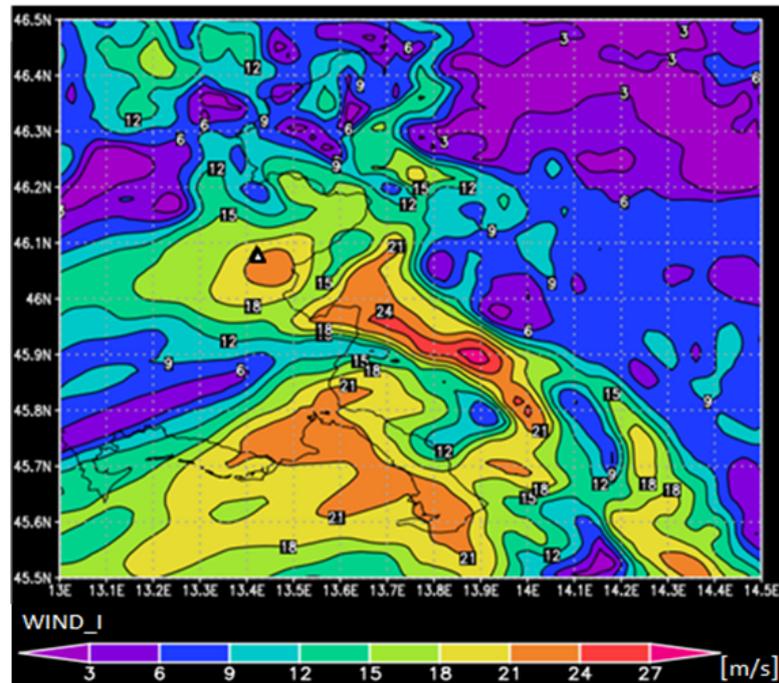


Figura 5.10: Confronto delle distribuzioni misurata e simulata dell'intensità del vento presso le stazioni di Cividale (sopra) e Faedis (sotto) in primavera (Marzo, Aprile e Maggio 2010). L'accuratezza nella riproduzione dell'intensità del vento risulta molto variabile localmente.

Nei prodotti delle simulazioni si osservano episodi di venti forti non presenti nelle misure. Attraverso l'analisi di alcune delle scadenze in cui si presentavano i massimi di questa discrepanza questi episodi sono stati ricondotti a fenomeni locali che non hanno interessato le stazioni. Un esempio è riportato in Figura 5.11: nella scadenza analizzata il modello indica un valore di intensità del vento presso la stazione di Cividale di circa 24 m/s, mentre le misure ne mostrano

uno dell'ordine di 18 m/s. Come si può notare la discrepanza è dovuta al fatto che la stazione di Cividale si trova nella zona periferica del temporale chiaramente visibile in figura, il valore ricavato dalle simulazioni è quindi influenzato anche dai valori in punti di griglia interessati dal temporale mentre la stazione non è stata interessata dal fenomeno.



Δ = Stazione ARPA-OSMER CIVIDALE

MAX WIND_I 2010 stz CIV [m/s]:

AD050101_20100301_20100531_CIV_WIND_I.txt:Max. Mis	18.3
AD050101_20100301_20100531_CIV_WIND_I.txt:Max. Grib	23.78
AD050201_20100601_20100831_CIV_WIND_I.txt:Max. Mis	13.1
AD050201_20100601_20100831_CIV_WIND_I.txt:Max. Grib	12.14
AD050301_20100901_20101130_CIV_WIND_I.txt:Max. Mis	12.6
AD050301_20100901_20101130_CIV_WIND_I.txt:Max. Grib	20.29
AD050401_20101201_20110228_CIV_WIND_I.txt:Max. Mis	14.9
AD050401_20101201_20110228_CIV_WIND_I.txt:Max. Grib	21.18
AD050501_20110301_20110531_CIV_WIND_I.txt:Max. Mis	13.2
AD050501_20110301_20110531_CIV_WIND_I.txt:Max. Grib	20.44

Figura 5.11: Intensità del vento presso la stazione di Cividale, indicata dal Δ, in una scadenza per la quale la simulazione prevede un massimo molto superiore a quello misurato (i massimi di misure e simulazioni sono riportati nella tabella di destra). Si può osservare che la discrepanza è dovuta al fatto che la stazione si trova nella zona periferica del temporale chiaramente visibile in figura, il valore ricavato dalle simulazioni è quindi influenzato anche dai valori in punti di griglia interessati dal temporale mentre la stazione non è stata interessata dal fenomeno.

Dal confronto tra i test a disposizione si osserva che non c'è dipendenza del bias dalla stagione ma si notano differenze a seconda dell'anno considerato, come evidente dalla Tabella 5.12.

Osservando lo spettro dell'intensità del vento nella stagione estiva a Cividale in Figura 5.13, si nota che il modello riproduce bene le caratteristiche brezze di valle che corrispondono alle frequenze carat-

BIAS WIND_I: media(NAUSICA) – media(misure)

STAGIONE AREA	Primavera		
	2010	2011	2014
Costa	1.4 m/s	0.7 m/s	0.8 m/s
Bassa Pianura	1.9 m/s	0.1m/s	1.6 m/s
Media Pianura	1.2 m/s	0.7 m/s	1.2 m/s
Alta Pianura	1.1 m/s	0.7 m/s	1.1 m/s
Montagna	1.4 m/s	0.7 m/s	1.1 m/s
Alpi	1.5 m/s	0.5 m/s	1.4 m/s
Carso	2.5 m/s	1.4 m/s	1.8 m/s

Figura 5.12: Bias (differenza tra la media delle simulazioni e quella relativa ai dati) dell'intensità del vento per la primavera 2010, 2011 e 2014. Si osserva una dipendenza dall'anno considerato.

terizzate da un picco nello spettro. C'è però un bias sulle basse frequenze che risultano sottostimate dal modello, questo lo si nota anche per gli altri spettri, nelle altre stazioni e nelle altre stagioni.

5.2.3 Umidità relativa

L'umidità relativa inferiore al 95% è riprodotta con un elevato grado di accuratezza dalle simulazioni: il bias è inferiore al 5% in tutti i casi analizzati ad esclusione dell'area costa, come si può leggere in Tabella 5.14. Anche l'area Alpina è molto ben riprodotta come si può vedere in Figura 5.15.

In zona costa, a causa della presenza di molti punti di griglia di tipo mare, per i quali la parametrizzazione dell'umidità non è sufficientemente accurata, è presente un bias sistematico di circa il 10%.

La distribuzione simulata risulta più stretta di quella relativa alle misure, come evidenziato in Figura 5.16, in particolare si notano in pianura sovrastima valori bassi e in montagna sottostima dei valori alti.

5.2.4 Precipitazioni

In generale il bias nelle distribuzioni relative alla precipitazione è inferiore ad 1 mm orario in tutte le stagioni e su tutte le aree, il 75° percentile è anch'esso inferiore a circa 1.5 mm orari, come si può osservare in Figura 5.17. Si nota però una sottostima del range 2mm/h - 10 mm/h del modello nella maggior parte dei casi analizzati, come si nota ad esempio in Figura 5.18.

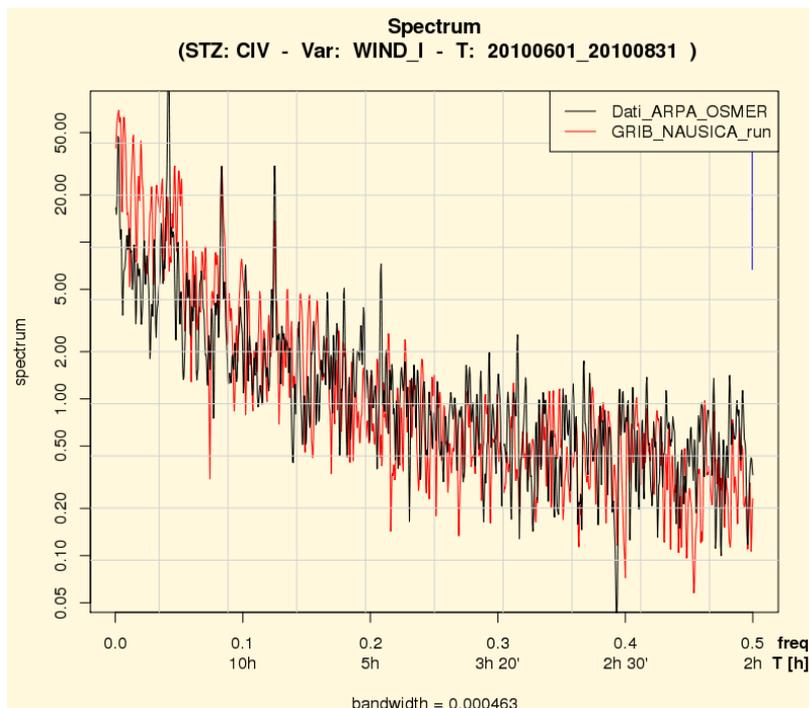


Figura 5.13: Spettro di potenza, ricavato da misure e modello, dell'intensità del vento a Cividale nella stagione estiva del 2010. I picchi corrispondono alle frequenze delle tipiche brezze di valle, fenomeni locali che risultano ben riprodotti dal WRF. Di contro si osserva una sottostima del modello per quanto riguarda le basse frequenze.

BIAS HUMI: media(NAUSICA) – media(misure)

STAGIONE AREA (XX)	Primavera	Estate	Autunno	Inverno
Costa	9.45%	9.76%	10.31%	14.78%
Bassa P.	3.44%	3.47%	2.47%	3.97%
Media P.	0.34%	1.17%	-1.18%	0.87%
Alta P.	-0.28%	-0.75%	-2.78%	-2.02%
Montagna	1.73%	-1.86%	-2.01%	-0.19%
Alpi	0.90%	0.51%	-5.60%	-4.01%

Figura 5.14: Bias (differenza tra la media delle simulazioni e quella relativa ai dati) dell'umidità relativa stagionale del 2010. Come si può notare il bias è inferiore al 5% in tutte le aree e stagioni considerate ad eccezione della costa che risulta affetta da un bias sistematico di circa il 10% dovuto alla presenza di numerosi punti di griglia di tipo mare.

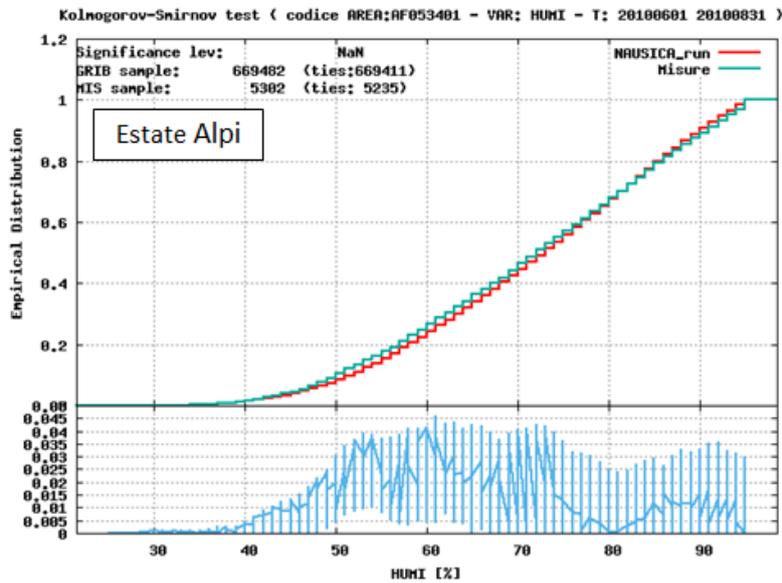


Figura 5.15: Confronto tra distribuzione empirica delle misure e simulata relativo all'umidità relativa sull'area alpina in estate (Giugno, Luglio e Agosto 2010).

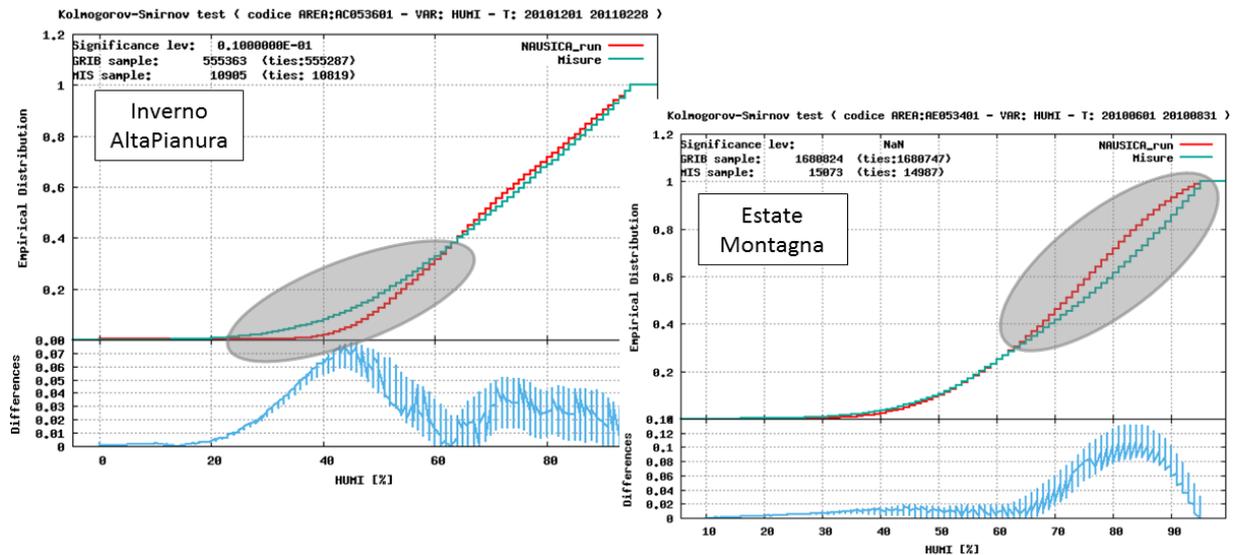


Figura 5.16: Confronto delle distribuzioni empiriche ricavate rispettivamente da modello e misure dell'umidità relativa sulle aree di alta pianura e montagna rispettivamente in inverno (Dicembre 2010, Gennaio e Febbraio 2011) ed estate (Giugno, Luglio e Agosto 2010). La distribuzione simulata risulta più stretta di quella relativa alle misure, in particolare si notano in pianura sovrastima valori bassi e in montagna sottostima dei valori alti.

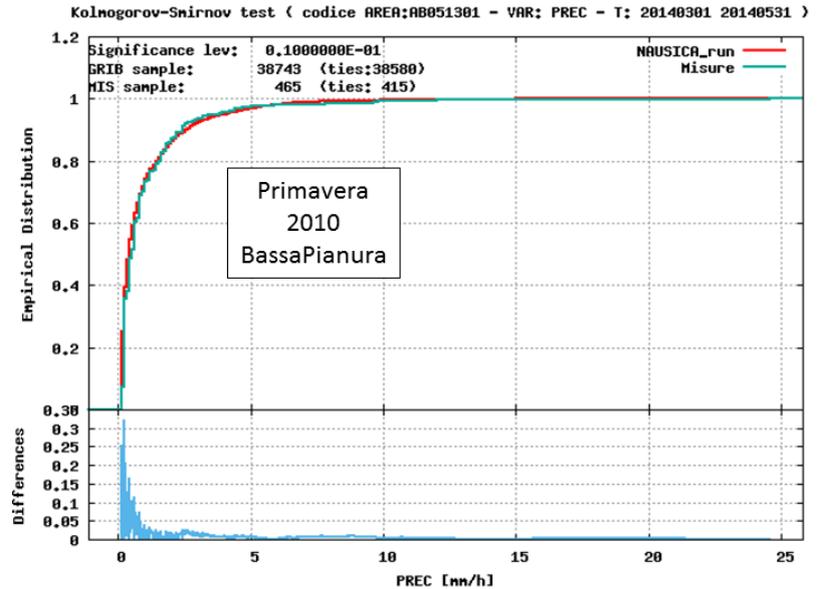


Figura 5.17: Confronto delle distribuzioni di misure e modello della precipitazione oraria sull'area di bassa pianura in primavera (Marzo, Aprile, Maggio 2010). In generale il bias nelle distribuzioni relative alla precipitazione è inferiore ad 1 mm orario in tutte le stagioni e su tutte le aree.

5.2.5 Precipitazioni > 10 mm/h

Come già descritto nel Paragrafo 4.3.2, per avere una statistica sufficiente a questa analisi è necessario considerare un intervallo temporale annuale.

A differenza del caso delle precipitazioni deboli, in questo range la distribuzione riprodotta dal modello presenta un bias inferiore al mm orario su tutte le aree. Due esempi sono riportati in Figura 5.19.

5.2.6 Precipitazioni > 25 mm/h

Anche per questa analisi è stato considerato un intervallo annuale ma non c'è comunque statistica sufficiente sulle aree contenenti poche stazioni, cioè Costa, Bassa pianura e Alpi.

L'accordo è ottimo anche in questo caso, come si nota in Figura 5.20. Il bias modello-misure è infatti di circa -1.5 mm/h per la media pianura, di 0.45 mm/h per l'alta pianura e di 3.06 mm/h per l'area di montagna. Si osservano sottostima su media pianura e sovrastima su area di montagna di entità contenuta.

5.2.7 Pressione

Per la pressione non è significativa l'analisi sulle aree a causa della complessità orografica della regione analizzata. Questo problema è evidente anche dall'analisi in alcuni punti di stazione come osservabile in Figura 5.21. Per le altre stazioni la distribuzione calcolata dal modello è più larga e ha bias simulazione-misure variabile tra 0.5 e 7.0 hPa. Si notano inoltre una sottostima nella simulazione dei valori bassi nelle stagioni fredde e una sovrastima dei valori alti nelle stagioni calde. In Figura 5.22 sono riportati due casi in cui sono evidenti le tendenze descritte.

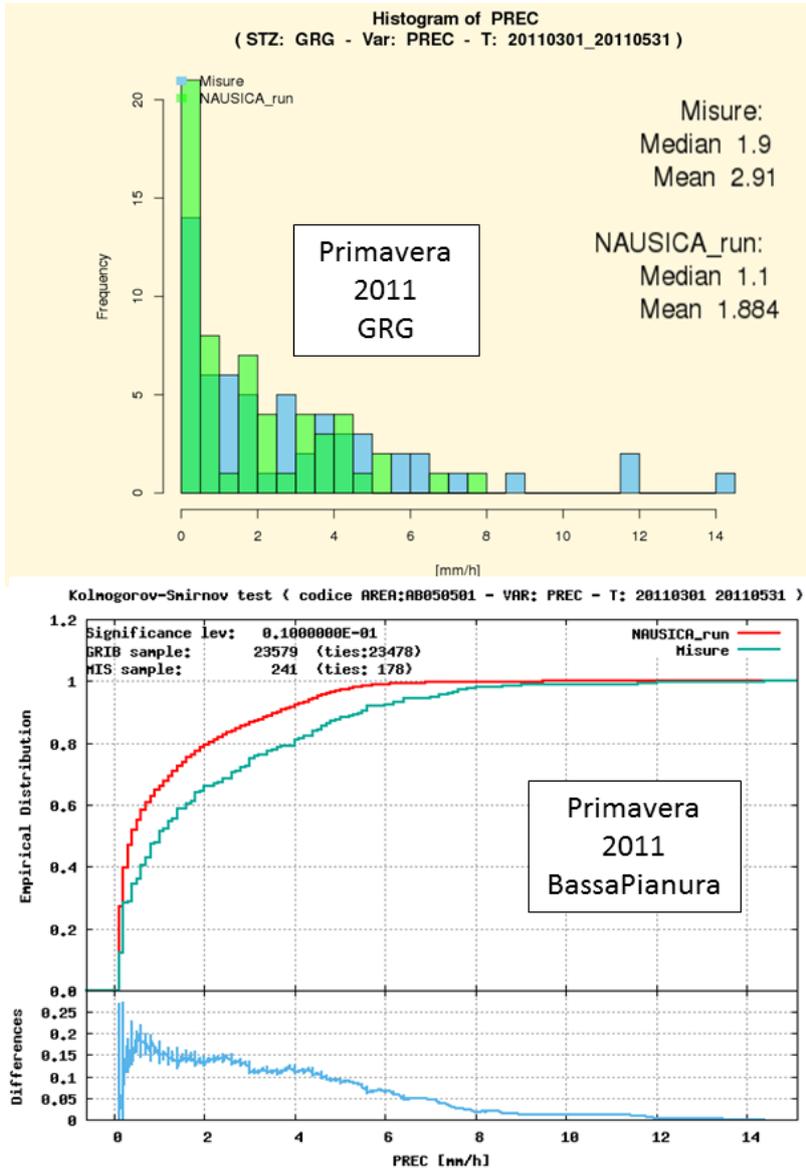


Figura 5.18: Confronto tra distribuzione simulata e misurata della precipitazione oraria sull'area di bassa pianura in primavera (Marzo, Aprile, Maggio 2011) e, in alto, istogramma di una delle stazioni contenute in questa area (Gorgo) nel medesimo periodo. Si nota una sottostima del range 2mm/h -10 mm/h comune alla maggior parte delle aree e stazioni analizzate.

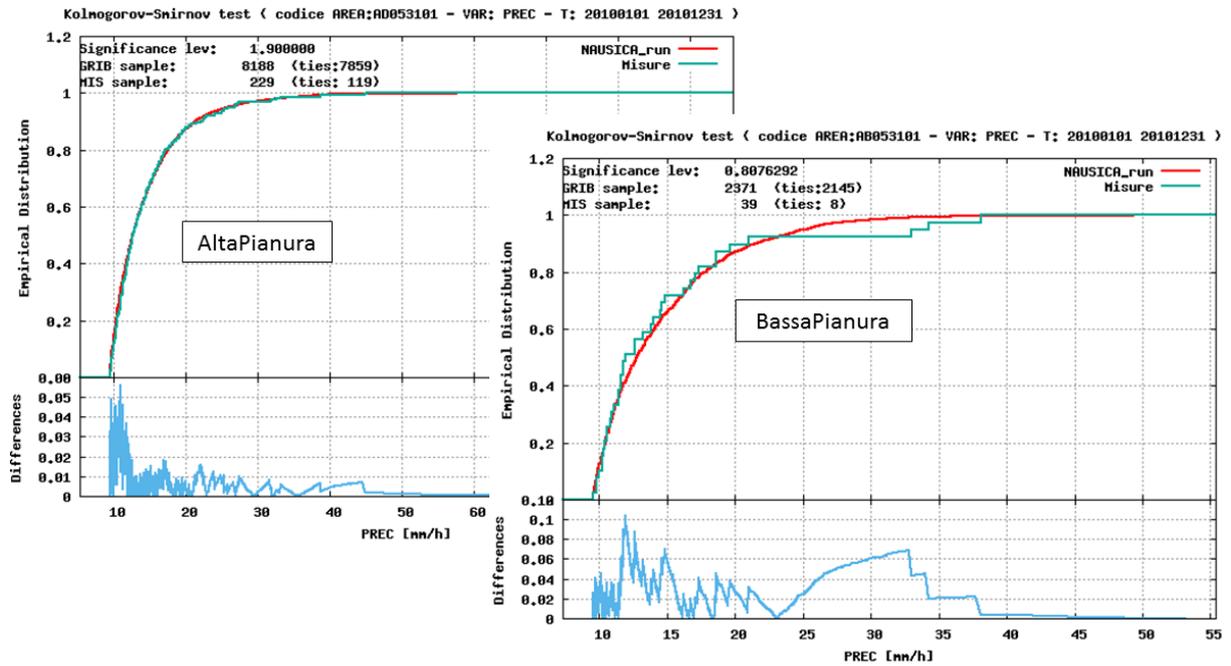


Figura 5.19: Confronto tra distribuzioni estratte dal modello e misurate della precipitazione oraria di intensità superiore ai 10 mm orari sulle aree di alta e bassa pianura per l'anno 2010. Il bias come si può notare sempre inferiore ad 1 mm/h.

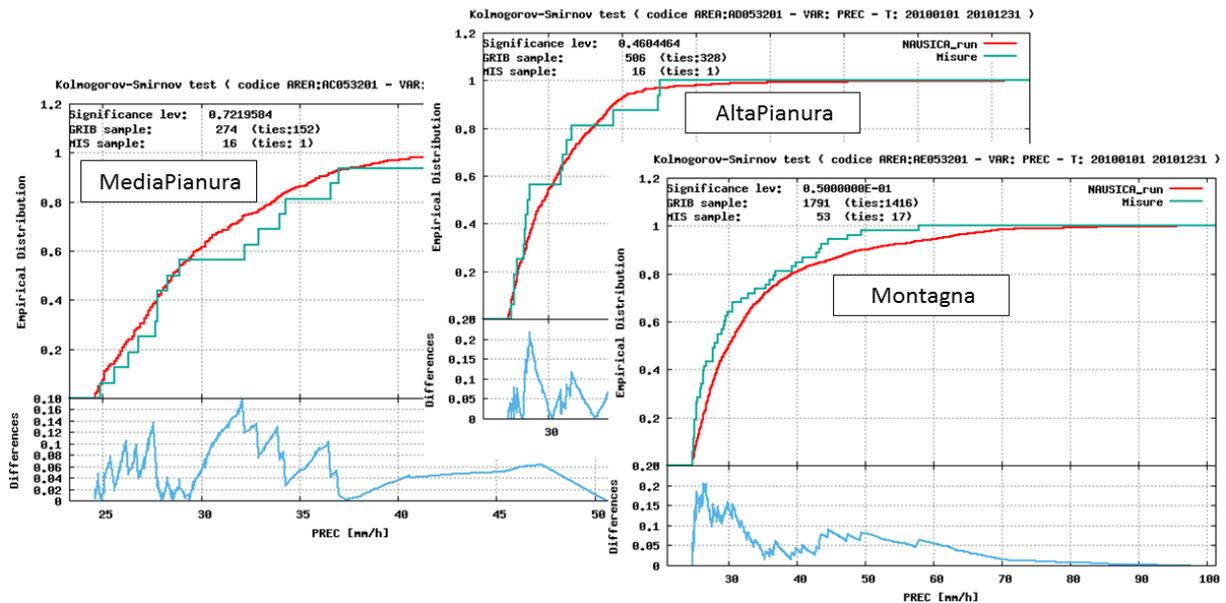


Figura 5.20: Confronto tra distribuzioni simulate e misurate della precipitazione oraria di intensità superiore ai 25 mm orari sulle aree di montagna, media e alta pianura per l'anno 2010. Si osservano nelle simulazioni una sottostima su media pianura e una sovrastima su area di montagna di entità contenuta.

Dato che ciò che conta è il gradiente della pressione e che, come si osserva in Figura 5.23, la forma della distribuzione risulta ben riprodotta il risultato della simulazione può considerarsi comunque soddisfacente.

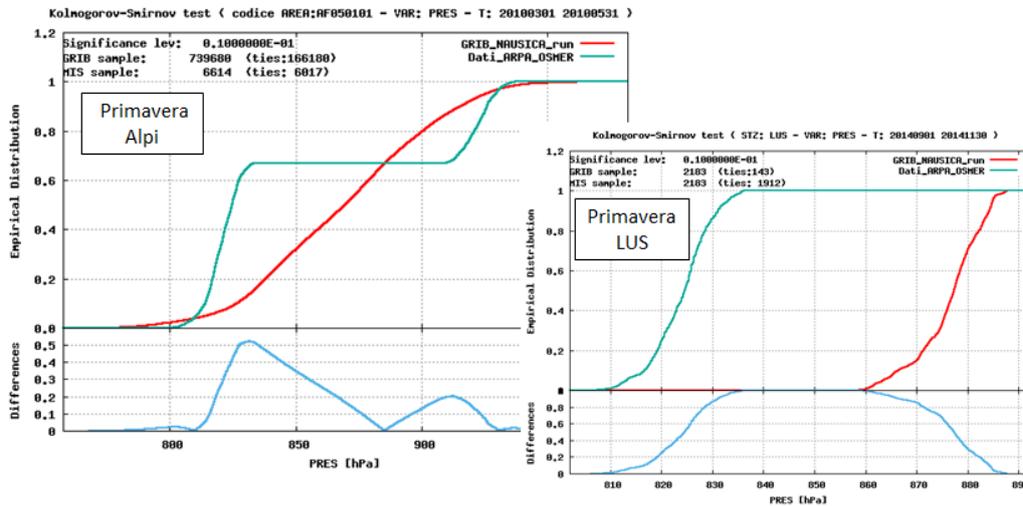


Figura 5.21: Confronto delle distribuzioni modello/misure della pressione sull'area Alpi in primavera (Marzo, Aprile, Maggio 2010) e relativo al punto di stazione del monte Lussari nello stesso periodo. Si nota la sovrapposizione di tre zone caratterizzate da andamenti differenti nell'area considerata che rendono poco significativa tutta l'analisi.

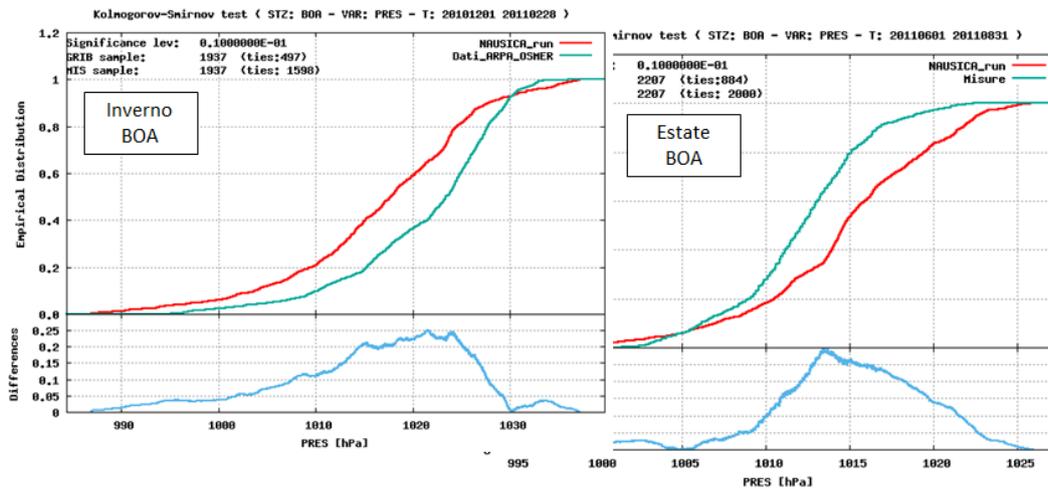


Figura 5.22: Confronto delle distribuzioni modello/misure della pressione nella stazione Boa in inverno (Dicembre 2010, Gennaio e Febbraio 2011) ed estate (Giugno, Luglio ed Agosto 2010). La distribuzione calcolata dal modello è più larga e ha bias variabile tra 0.5 e 7.0 hPa. Si notano inoltre una sottostima nella simulazione dei valori bassi nelle stagioni fredde e una sovrastima dei valori alti nelle stagioni calde.

Analizzando invece il confronto tra gli spettri di pressione si osserva un ottimo accordo misure-modello, è interessante notare che lo spettro è dominato dal picco corrispondente alle modulazioni su scala sinottica, vedere Figura 5.24.

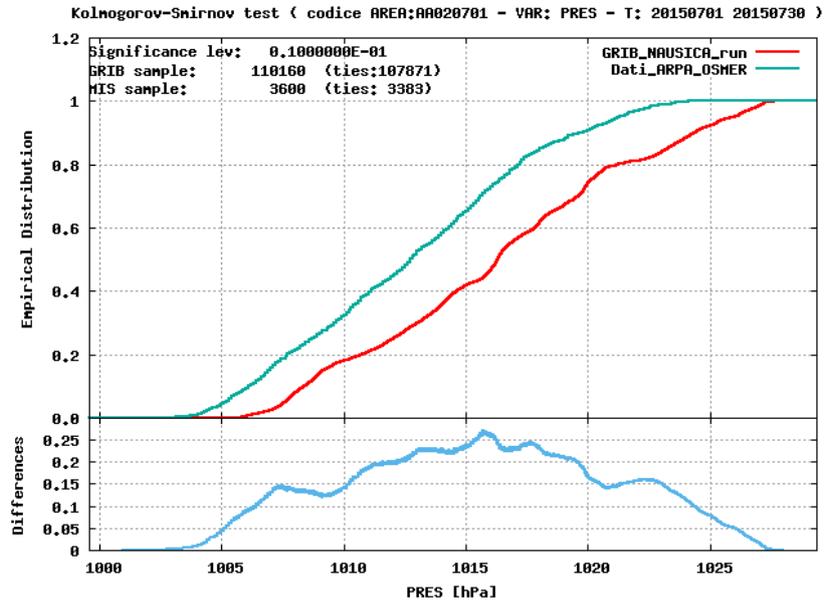


Figura 5.23: Confronto tra distribuzioni simulata e misurata della pressione sull'area di costa per il mese di Luglio 2015. Si nota una sovrastima da parte del modello dell'ordine del 5%.

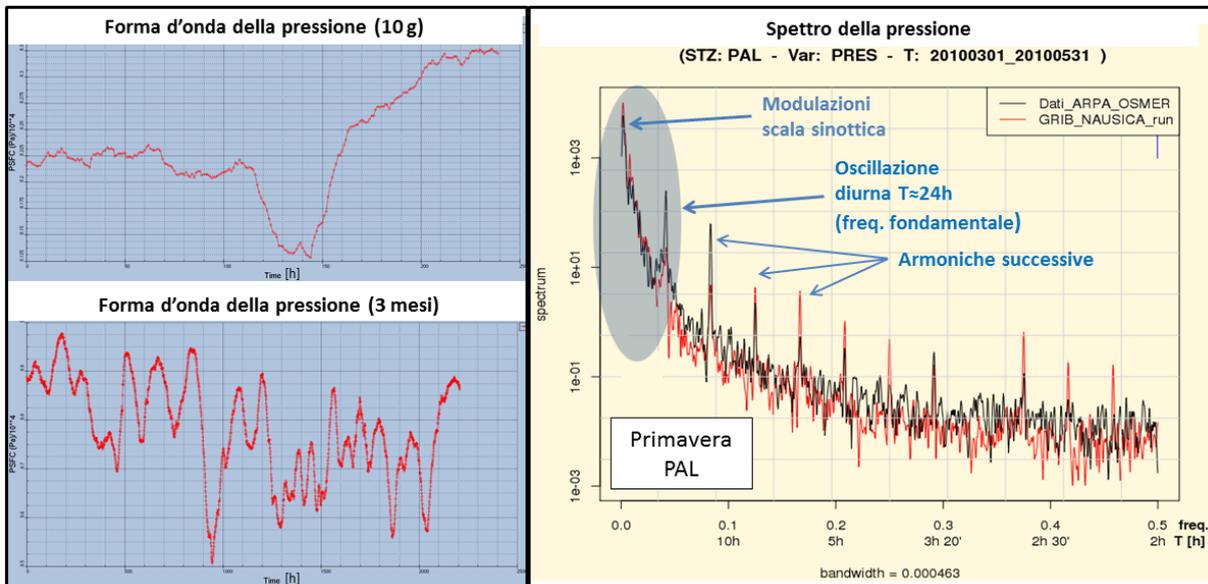


Figura 5.24: Forma d'onda su 10 e 90 giorni (sinistra) e spettro in potenza (destra) della pressione presso la stazione di Palazzolo dello Stella in Primavera (Marzo, Aprile e Maggio 2010). Si nota un ottimo accordo tra simulazione e dati. Lo spettro è caratterizzato dal picco relativo alle modulazioni su scala sinottica.

5.2.8 Radiazione

Dalle analisi relative a questa osservabile si riscontra una sovrastima sistematica da parte del modello. Il bias varia da circa 60 W/mq per le zone di pianura e montagna a circa 140 W/mq per le aree di costa ed Alpi, come si nota in Figura 5.25. Nella forma della distribuzione si osserva una sovrastima dei valori bassi nelle stagioni fredde mentre le stagioni calde sono caratterizzate da una distribuzione di forma analoga a quella delle misure ma più larga, come si osserva in Figura 5.25 e 5.26.

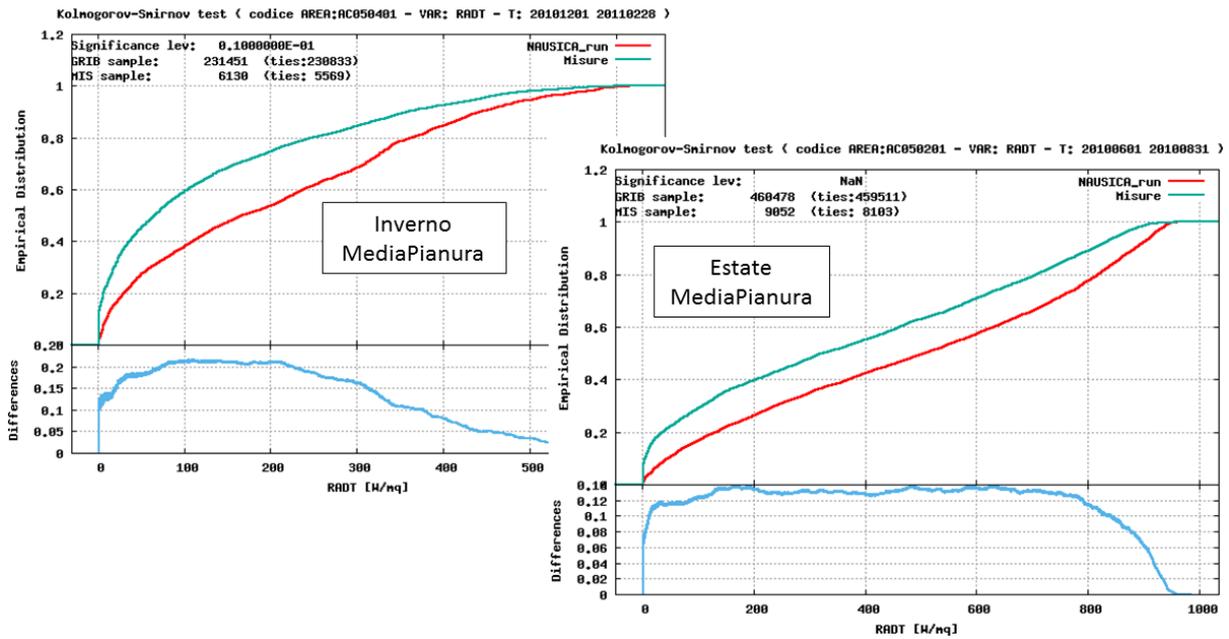


Figura 5.25: Confronto tra distribuzioni ricavata dal modello e misurata della radiazione nell'area di media pianura in inverno (Dicembre 2010, Gennaio e Febbraio 2011) ed estate (Giugno, Luglio ed Agosto 2010). Si vede il bias dovuto alla sovrastima del modello rispetto alle misure. Nella forma della distribuzione si osserva una sovrastima dei valori bassi nelle stagioni fredde mentre le stagioni calde sono caratterizzate da una distribuzione di forma analoga a quella delle misure ma più larga, come evidenziato anche in Figura 5.26.

Per quanto riguarda lo spettro in potenza della radiazione visibile, come si può osservare dal confronto tra spettri misurati e simulati riportato in Figura 5.27, l'accordo tra dati e simulazioni è buono. Sono infatti ben riprodotte negli spettri previsti dal modello le frequenze caratteristiche del segnale, corrispondenti ai picchi di maggiore ampiezza. Tra questi quello di maggiore ampiezza corrisponde alla frequenza caratteristica di periodo 24 ore attribuibile alla variazione diurna della radiazione. Infatti come si nota in Figura 5.27 (grafico in alto), dove è rappresentata la serie temporale della radiazione ad onda corta estratta dai *netCDF* relativi alle prime settimane del mese di Luglio 2015, il segnale in assenza di copertura nuvolosa è periodico con periodo di 24 ore. Le ultime ore dell'intervallo riportato nel grafico sono invece caratterizzate da un segnale di ampiezza inferiore e irregolare dovuto alla presenza di nuvolosità che è anche causa dell'elevata densità di picchi, tra l'altro più difficilmente riprodotti

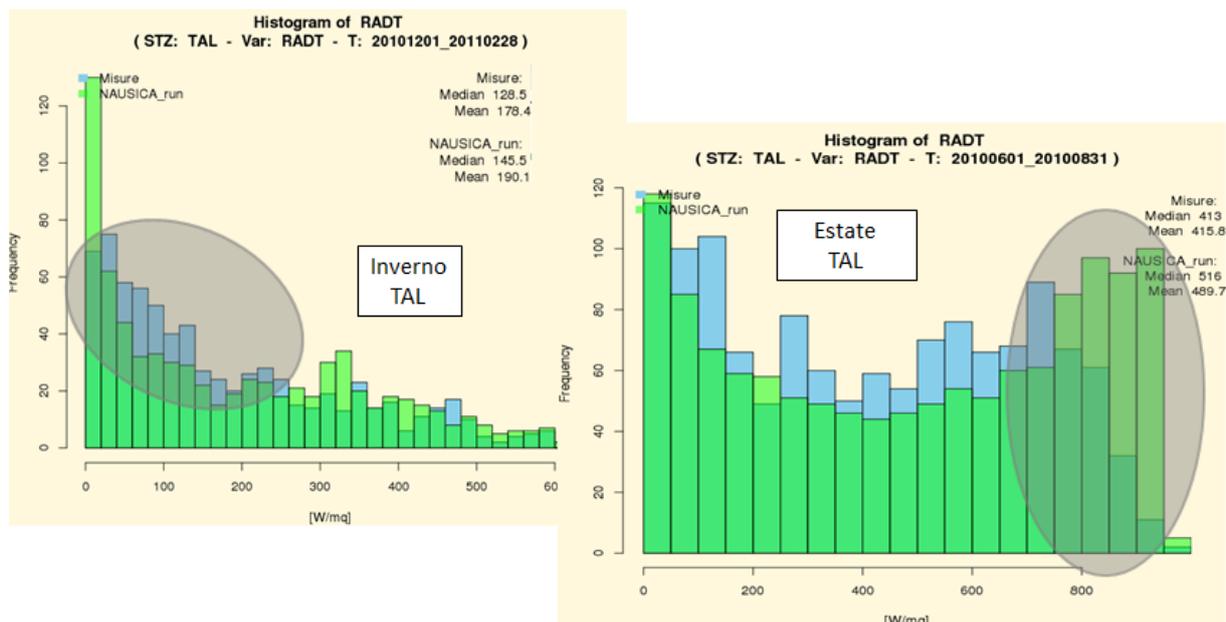


Figura 5.26: Istogrammi delle distribuzioni della radiazione in una delle stazioni appartenenti all'area di media pianura (Talmasson) in inverno (Dicembre 2010, Gennaio e Febbraio 2011) ed estate (Giugno, Luglio ed Agosto 2010). Nella forma della distribuzione è evidenziata la sovrastima dei valori bassi nelle stagioni fredde e la tendenza nelle stagioni calde ad avere una distribuzione di forma analoga a quella delle misure ma più larga.

dal modello, che caratterizzano le alte frequenze dello spettro. In Figura 5.27 (grafico in alto a destra) è anche riportato un singolo ciclo diurno relativo ad un giorno caratterizzato da bassa copertura nuvolosa estratto dalla serie temporale: come atteso è approssimabile, almeno per quanto riguarda le prime armoniche, ad un segnale periodico quadro. La forma d'onda della radiazione osservata può essere approssimata almeno per i primi ordini dello sviluppo con un'onda quadra, la cui forma e il cui spettro sono rappresentati in Figura 5.27 (immagine in basso). È di conseguenza ragionevole attendersi che nello spettro della radiazione siano presenti le prime armoniche tipiche dello sviluppo di un'onda quadra mentre si attende uno spettro più denso di picchi per le alte frequenze. Questo si osserva chiaramente nello spettro riportato in Figura 5.27 (figura di mezzo) dove sono indicati i picchi corrispondenti a queste componenti. Considerando intervalli temporali maggiori, vedere Figura 5.28, compare anche il picco relativo alla oscillazione stagionale della radiazione.

5.2.9 Evapotraspirazione giornaliera

L'evapotraspirazione giornaliera riprodotta dal modello è in ottimo accordo con le misure: il bias è sempre inferiore a 0.7 mm/mq (Figura 5.29) ad esclusione dell'area di costa, di cui un esempio è riportato in Figura 5.31, che a causa della presenza di molti punti di griglia di tipo mare presenta una notevole discrepanza soprattutto nelle stagioni calde. Questo problema è stato ridotto considerando nell'analisi solo i punti di tipo *land*.

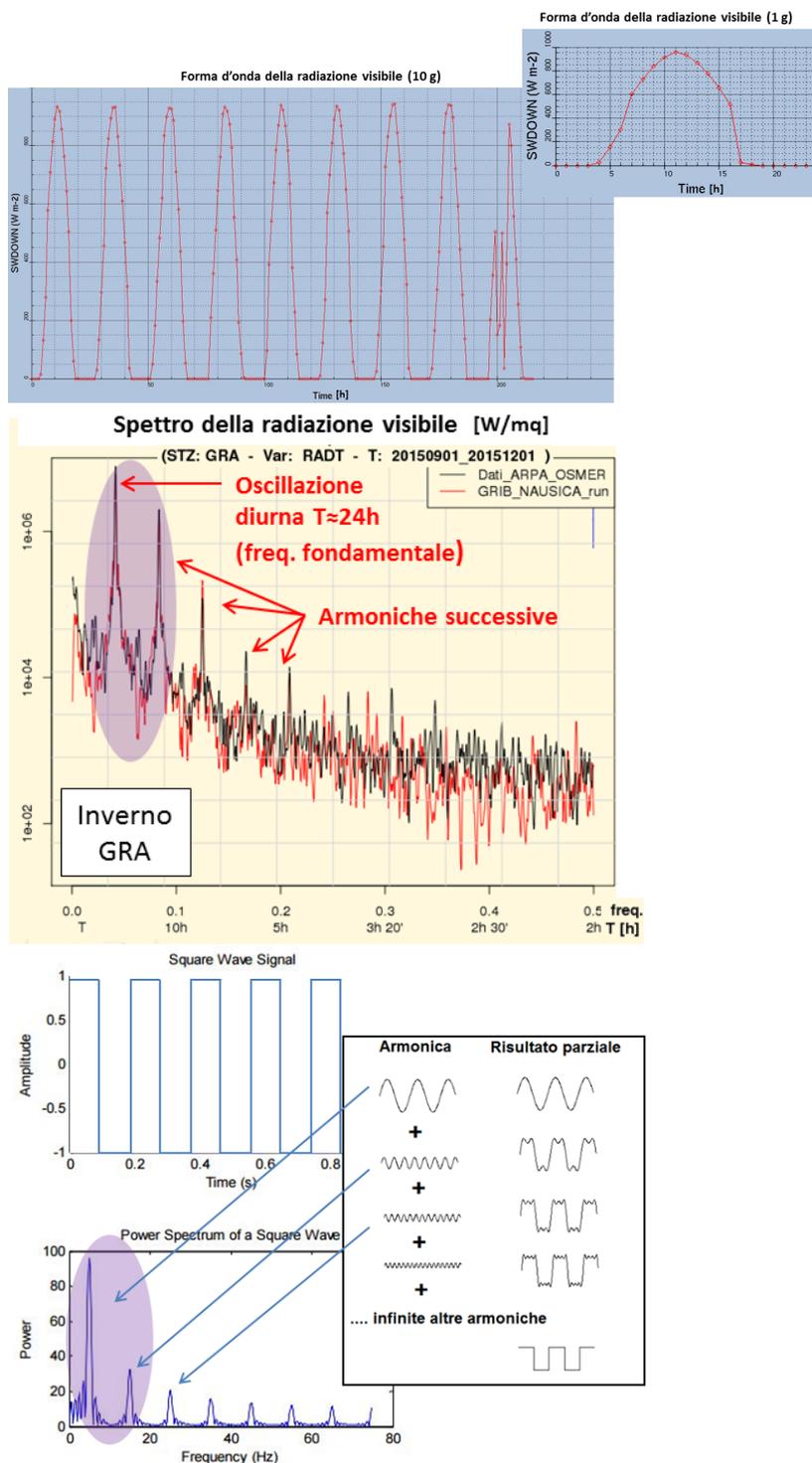


Figura 5.27: Analisi dei risultati ricavati riguardo lo spettro in potenza della radiazione visibile. In Fig. in alto a destra è riportato un singolo ciclo diurna relativo ad un giorno caratterizzato da bassa copertura nuvolosa estratto dalla serie temporale dei netCDF (1 Luglio 2015, Udine): si può osservare che è caratterizzato da un valore nullo durante le ore notturne. In alto a sinistra è rappresentata la serie temporale della radiazione ad onda corta estratta dai netCDF relativi alle prime 200 ore del mese di Luglio 2015, la parte finale dell'intervallo temporale riportato è interessata dagli effetti di copertura nuvolosa che modificano la forma del segnale. Nella Fig. centrale sono sovrapposti gli spettri delle serie temporali della radiazione visibile ricavati da misure e simulazioni nella stazione di Gradisca nel trimestre Settembre, Ottobre e Novembre 2015. L'accordo tra dati e simulazioni è ritenuto soddisfacente soprattutto per quanto riguarda la riproduzione dei picchi relativi alla frequenza fondamentale, dovuta al ciclo diurna e di conseguenza di periodo 24 ore, e alle prime armoniche indicate dalle frecce in figura. La discordanza che si osserva per le alte frequenze è dovuta al rumore introdotto dalla copertura nuvolosa e al fatto che nemmeno in assenza di nuvolosità l'onda è quadra, vedere anche La Figura in alto a destra di destra. Nella Fig. in basso sono invece riportati un esempio di onda quadra e il relativo spettro la cui costruzione è illustrata nello schema a destra.

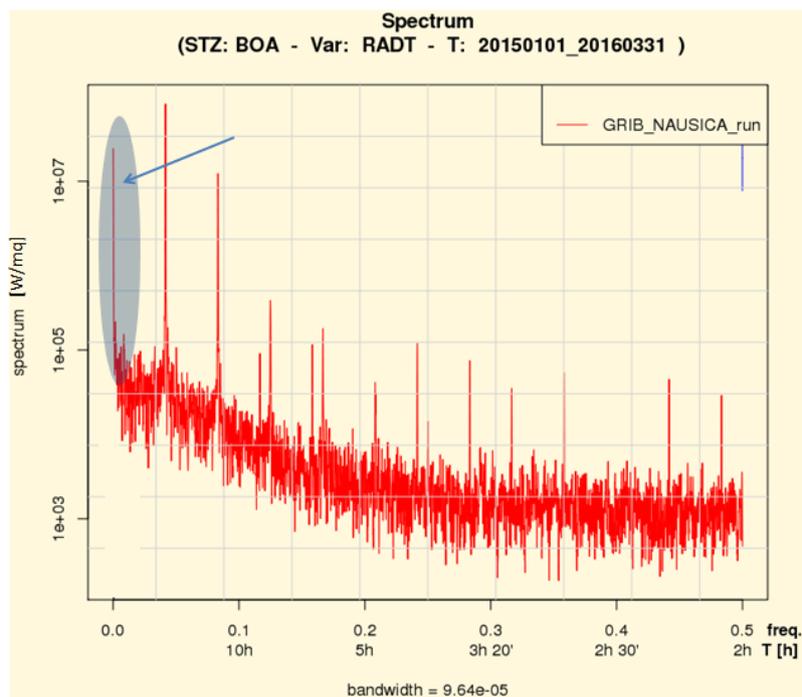


Figura 5.28: Spettro in potenza della radiazione visibile relativo a 15 mesi (Gennaio 2015-Marzo 2016) ricavato dalla serie temporale dei netCDF nelle coordinate della stazione BOA: si nota la frequenza caratteristica dell'oscillazione stagionale della radiazione indicata dalla freccia azzurra.

L'estate è la stagione la cui forma della distribuzione è riprodotta con accuratezza più elevata, l'inverno è invece caratterizzato da una sottostima da parte del modello. Questo si vede chiaramente nelle Figure 5.29 e 5.31.

Anche per questa variabile si osserva che la qualità della riproduzione locale è molto variabile: presso alcune stazioni si ha un accordo misure-simulazioni molto buono mentre in altre si osserva un bias sistematico di entità non trascurabile chiaramente visibile in Figura 5.32.

5.3 Analisi dei campi in quota

Attualmente non sono ancora state analizzate nel dettaglio le analisi corrette che, come spiegato nel capitolo 4.4, si limitano all'andamento con la quota della temperatura potenziale e dei gradienti di temperatura, temperatura potenziale, intensità del vento e pressione.

Ad una prima analisi si può osservare che la temperatura potenziale è molto ben riprodotta dal modello: bias sempre inferiore a 2°C, nella maggior parte dei casi anche a 1°C fino ad una quota di circa 10 km, vedere Figura 5.33. Per quote superiori si osservano divergenze maggiori spesso caratterizzate da distribuzioni simulate più strette rispetto alle rispettive misurate. È sempre presente una divergenza in corrispondenza dei livelli più alti attribuibile ad effetti di bordo.

Alla stessa conclusione, cioè rappresentatività del modello qualitativamente buona, si può giungere osservando il confronto tra i profili medi stagionali della temperatura potenziale.

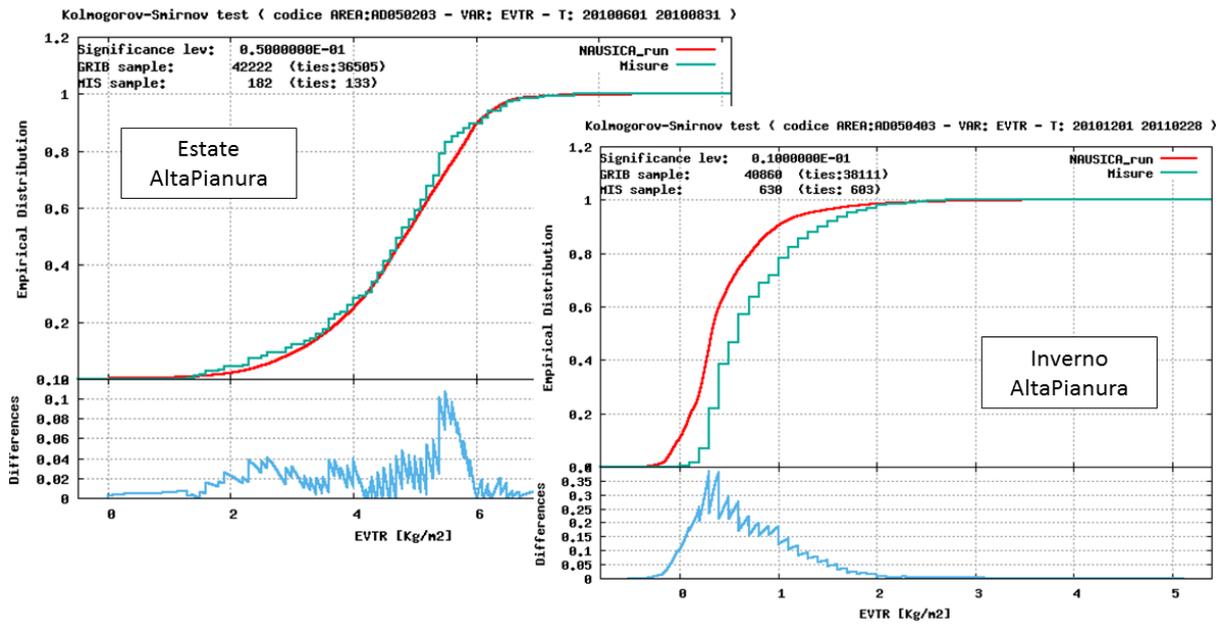


Figura 5.29: Confronto delle distribuzioni misurata/simulata dell’evapotraspirazione nell’area di alta pianura in inverno (Dicembre 2010, Gennaio e Febbraio 2011) ed estate (Giugno, Luglio ed Agosto 2010). L’evapotraspirazione giornaliera riprodotta dal modello è in ottimo accordo con le misure.

L’estate è la stagione con accordo migliore mentre l’inverno è caratterizzato da una sottostima da parte del modello.

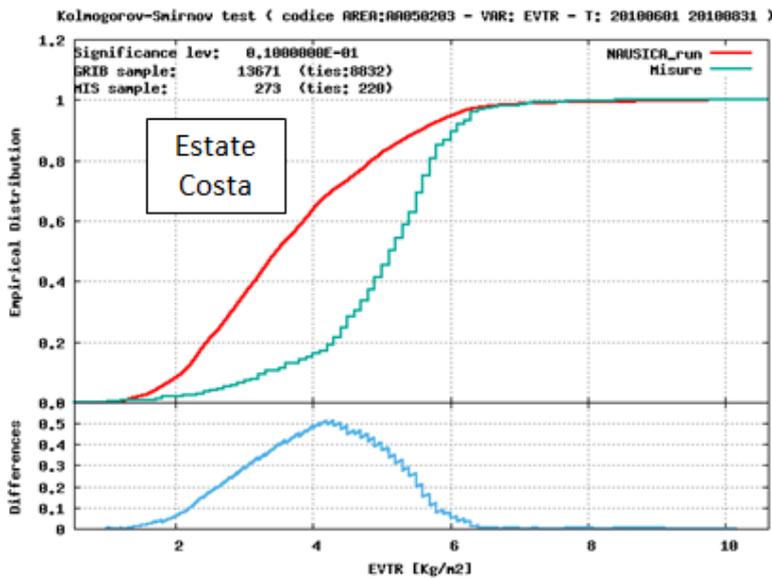


Figura 5.30: Confronto delle distribuzioni simulata/misurata dell’evapotraspirazione nell’area di costa in estate (Giugno, Luglio ed Agosto 2010). Si nota il bias dovuto alla presenza all’interno dell’area di molti punti di griglia di tipo mare. Il problema è stato ridotto considerando solo i punti di tipo land.

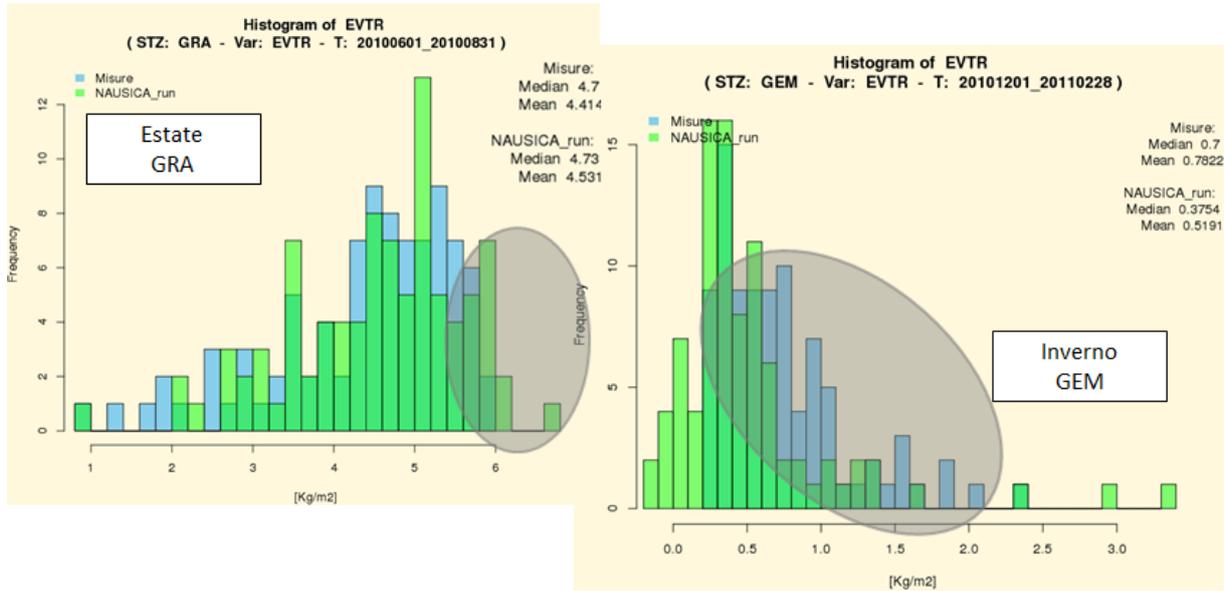


Figura 5.31: Istogrammi delle distribuzioni dell'evapotraspirazione in una delle stazioni appartenenti all'area di media pianura (Gradisca) in inverno (Dicembre 2010, Gennaio e Febbraio 2011) e in zona montagna (Gemona) in estate (Giugno, Luglio ed Agosto 2010). Si notano una leggera sovrastima dei valori alti da parte del modello in estate e la sottostima dei valori centrali della distribuzione in inverno.

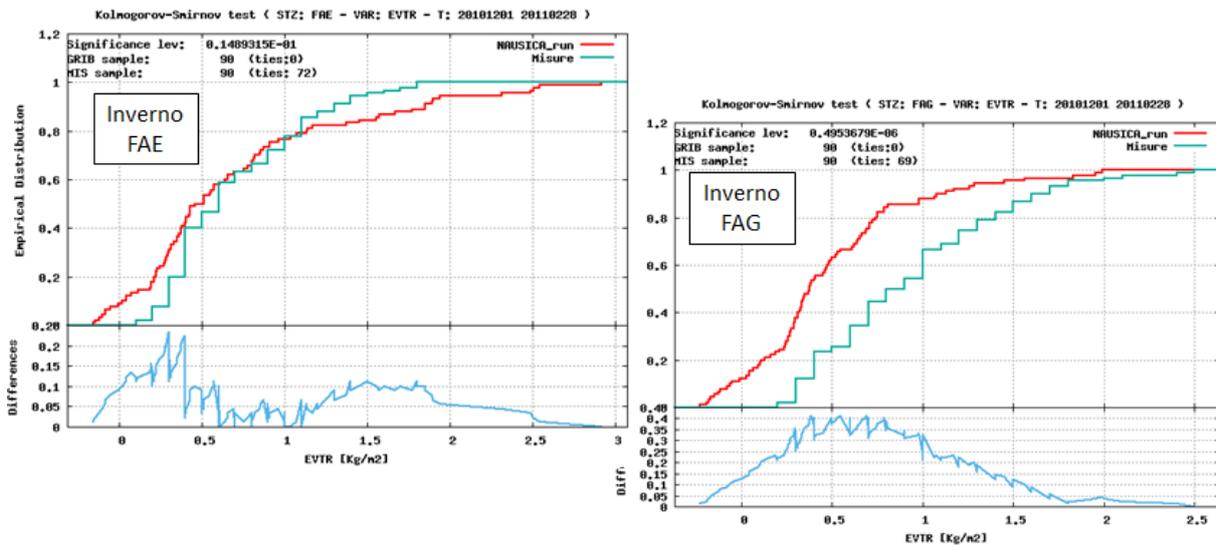


Figura 5.32: Confronto tra distribuzioni empiriche misurata e simulata dell'evapotraspirazione nelle stazioni di Faedis e Fagagna, entrambe appartenenti all'area di alta pianura, in inverno (Dicembre 2010, Gennaio e Febbraio 2011). Si osserva che la qualità della riproduzione locale è molto variabile.

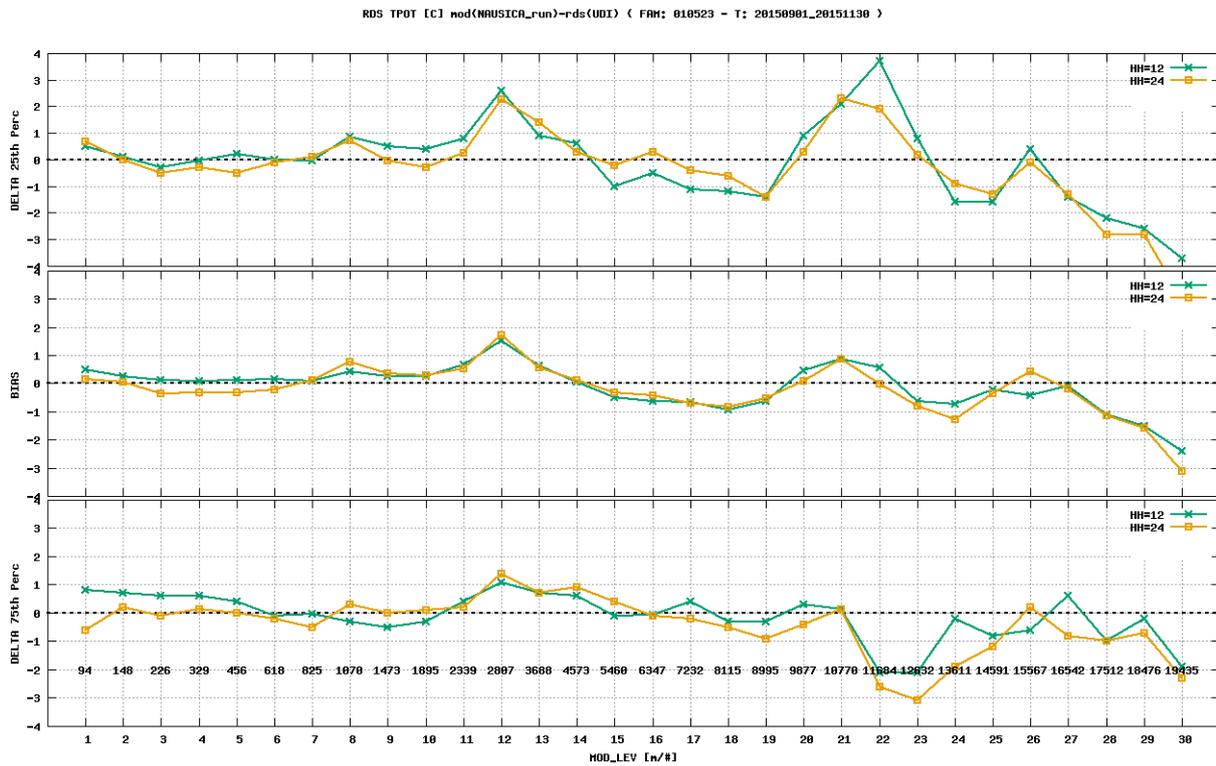


Figura 5.33: Bias e differenza tra 25° e 75° percentile delle distribuzioni di modello e dati misurati da radiosondaggi del campo di temperatura potenziale nell'ottobre del 2015 a Rivolto utilizzando bande verticali di 50 m per i primi 8 livelli e di 200 m per i livelli successivi. Si può notare che il modello riproduce molto bene le misure almeno fino ad una quota di circa 10 km oltre la quale la distribuzione dei dati simulati risulta un po' più stretta di quella reale anche se il bias risulta sempre inferiore ai 2°C. La divergenza per i livelli più alti può essere attribuita ad effetti di bordo.

Nonostante l'ampio range della scala necessario alla visualizzazione dell'intero profilo che non rende quantificabile la differenza, l'accordo appare qualitativamente buono, si veda Figura 5.34.

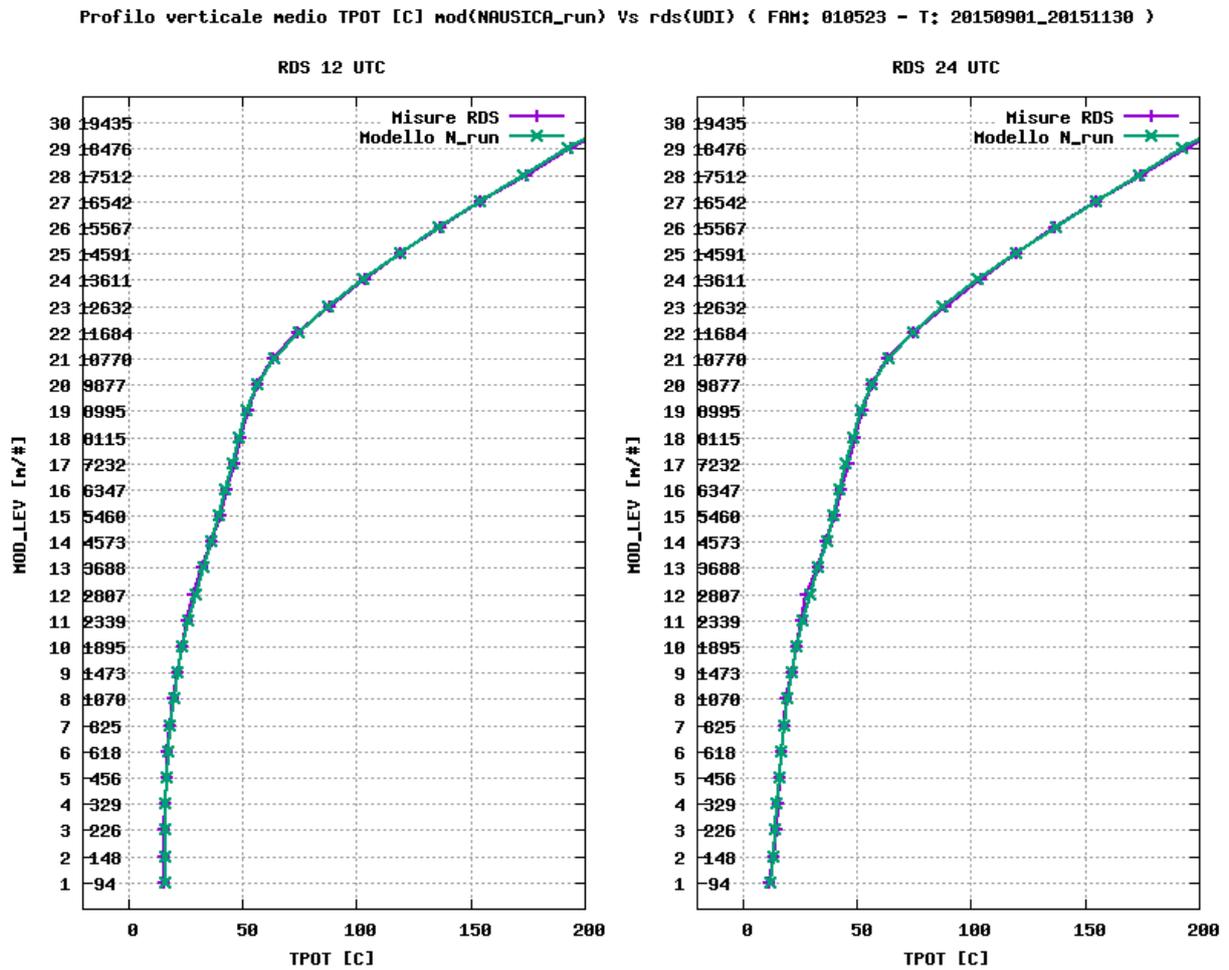


Figura 5.34: Profilo verticale medio stagionale della temperatura alle 12:00 UTC e alle 24:00 UTC nell'ottobre del 2015 a Rivolto per modello e misure dei radiosondaggi. Nonostante l'ampio range della scala (necessario alla visualizzazione dell'intero profilo) che non rende quantificabile la differenza, l'accordo appare qualitativamente buono.

Dal profilo verticale dei gradienti di temperatura, intensità del vento, singole componenti del vento, pressione e temperatura potenziale, di cui alcuni esempi sono riportati nelle Figure 5.35 e 5.36 , si può notare qualitativamente che l'accuratezza delle simulazioni NAUSICA è molto buona. I livelli con maggiore bias sul gradiente risultano essere i due livelli inferiori. Il segno del gradiente resta comunque sempre ben riprodotto.

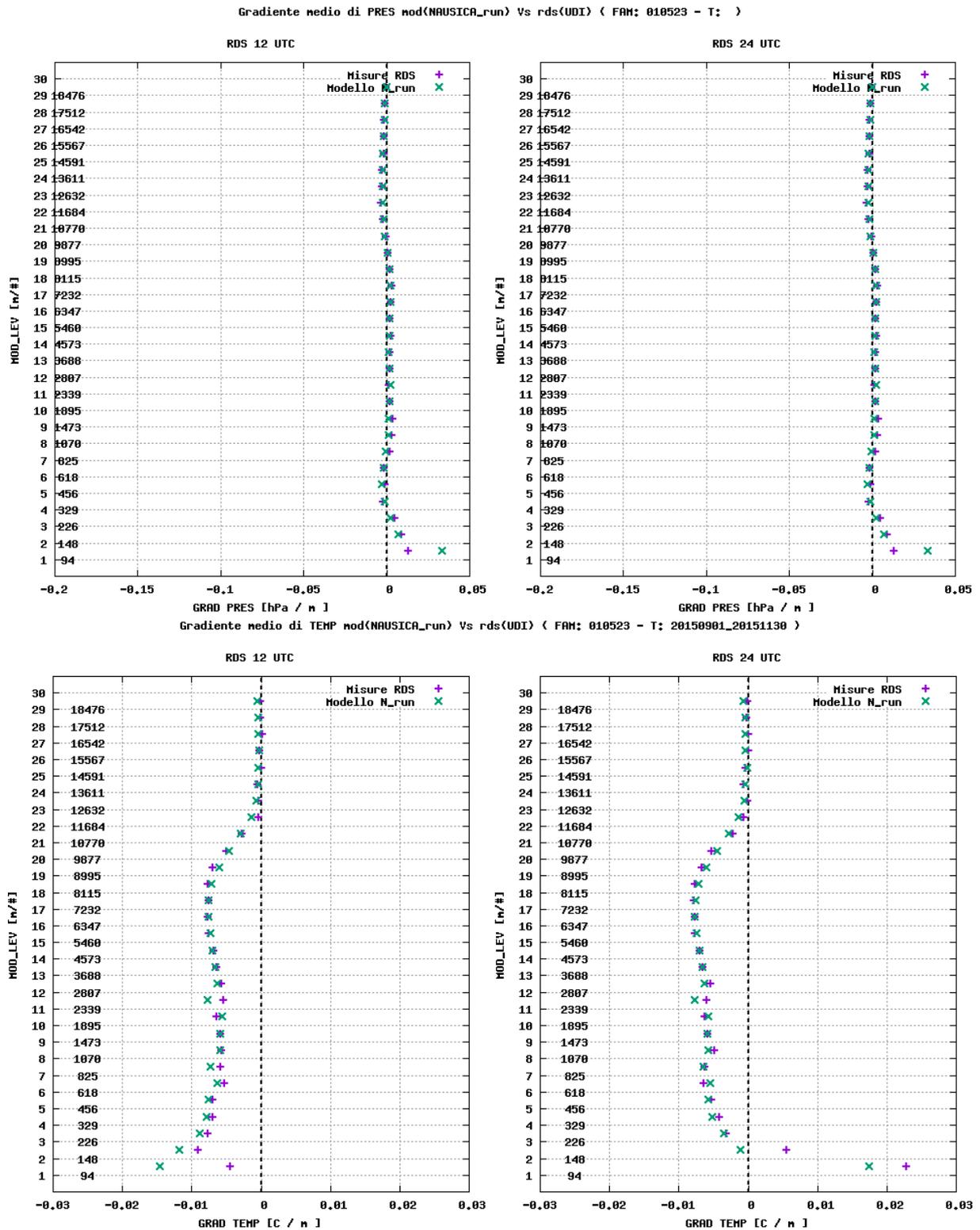
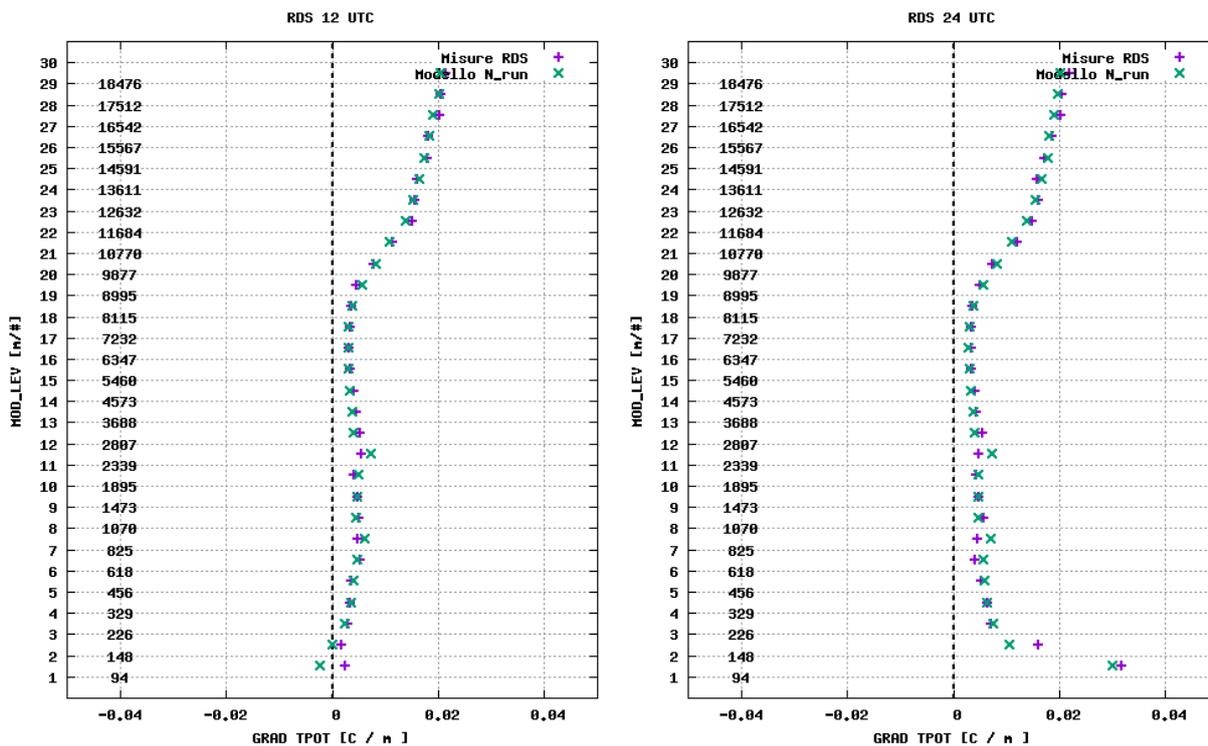


Figura 5.35: Confronto tra i gradienti medi di pressione e temperatura misurati e calcolati per l'autunno del 2015 alle 12 e 24 UTC per ciascun livello del modello sulla verticale nel punto di lancio del radiosondaggio di Rivolto. Qualitativamente l'accuratezza delle simulazioni NAUSICA è molto buona. I livelli con maggiore bias sul gradiente risultano essere i due livelli inferiori. Il segno del gradiente resta comunque sempre ben riprodotto.

Gradiente medio di TPOT mod(NAUSICA_run) Vs rds(UDI) (FAM: 010523 - T: 20150901_20151130)



Gradiente medio di WIND_I mod(NAUSICA_run) Vs rds(UDI) (FAM: 010523 - T: 20150901_20151130)

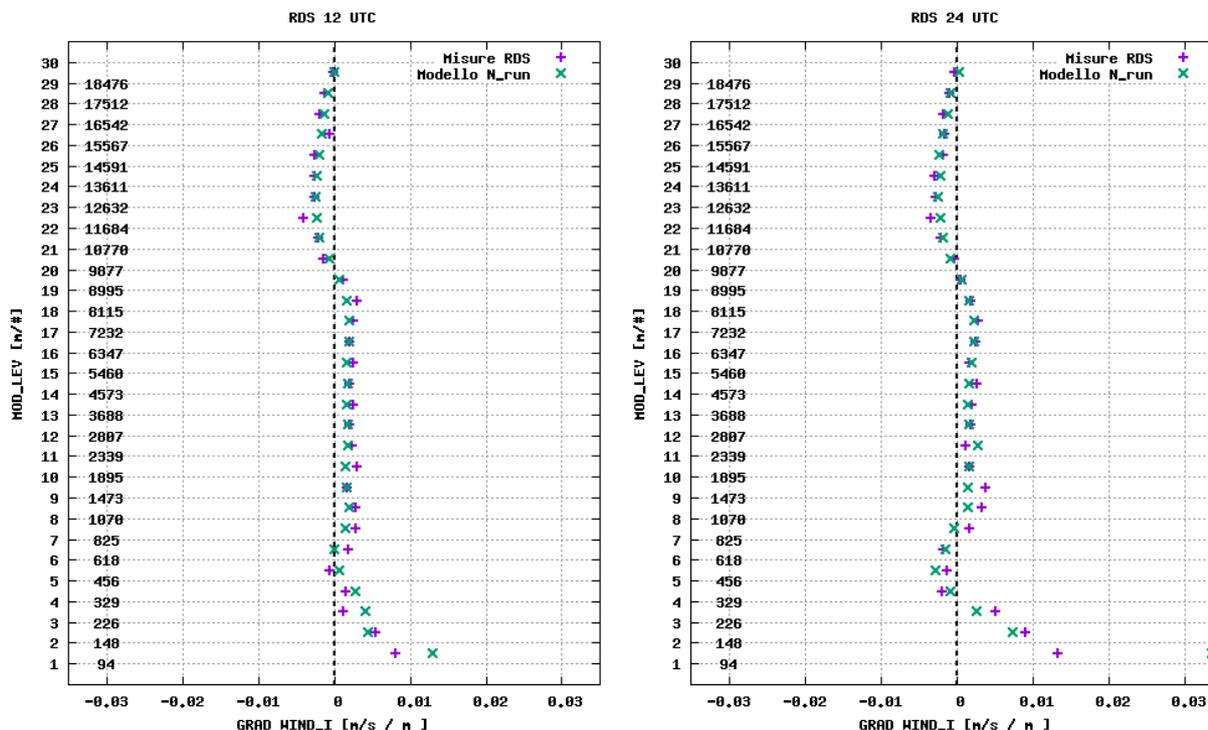


Figura 5.36: Confronto tra i gradienti medi di temperatura potenziale ed intensità del vento misurati e calcolati per l'autunno del 2015 alle 12 e 24 UTC per ciascun livello del modello sulla verticale nel punto di lancio del radiosondaggio di Rivolto. Qualitativamente l'accuratezza delle simulazioni NAUSICA è molto buona. I livelli con maggiore bias sul gradiente risultano essere i due livelli inferiori. Il segno del gradiente resta comunque sempre ben riprodotto.

6

Sviluppi futuri del progetto

In questo capitolo sono discussi vari sviluppi del progetto NAUSICA. Su alcuni di essi si è lavorato nel corso 2018, verranno quindi descritti nel report del progetto successivo a questo.

Principali sviluppi affrontati nel 2018:

- Messa in operatività di una suite NAUSICA_gen che si occupa di aggiornare automaticamente il database Nausica
- Produzione, tramite la nuova suite NAUSICA_stat, dei campi statistici derivati dai principali campi NAUSICA (medie, mediane, percentili, minimi, massimi, somme, ecc.) su decenni, mesi ed anni e confronto dei risultati con i corrispondenti valori misurati

Alcune proposte di sviluppo del progetto nate nel corso del lavoro descritto in questa relazione sono:

- L'assimilazione di dati nel modello WRF tramite software GSI [32] all'interno della suite NAUSICA_run (Paragrafo 4.1)
- Miglioramento nei criteri di selezione delle aree di analisi utilizzate dalla suite NAUSICA_ver (Paragrafo 4.3) in modo da aumentare l'omogeneità delle caratteristiche meteorologiche all'interno di ciascuna di esse. Ci si riferisce principalmente alla separazione tra "zone di vetta" e "zone di valle" nelle aree di montagna e Alpi. Questa operazione è stata inclusa tra le opzioni dell'eseguibile *extraction.exe*
- Ottimizzazione di altri parametri WRF per run annuali in NAUSICA_run. Il WRF dispone infatti di tante opzioni e parametrizzazioni che non sono state analizzate.
- Prova di esecuzione di run più lunghi di un anno in vista di future simulazioni climatiche
- Rivisitazione della metodologia di analisi della qualità dei profili verticali dei campi che, come descritto nel Paragrafo 4.4.2, è affetta da una problematica che ne invalida gran parte dei risultati. In particolare un'idea è quella di eseguire i confronti tra volumi di spazio considerando nel campione di dati estratto dalle simulazioni anche l'estrapolazione (fatta con le cdo) del valore dei campi ogni

10 m (come si ha per le misure) su un intervallo verticale pari a quello scelto per il corrispondente campione del radiosondaggio. In questo modo si dovrebbe eliminare il problema anche se, ovviamente, potrebbe essere introdotto un bias causato dal processo di estrapolazione stesso

A

Namelist di WRF e WPS

A.1 Namelist WPS

```
&share
wrf_core = 'ARW',
max_dom = 3,
start_date = '%SYYYY%-%SMM%-%SDD%_00:00:00', '%SYYYY%-%SMM%-%SDD%_00:00:00', '%SYYYY%-%SMM%-%SDD%_00:00:00'
end_date   = '%EYYYY%-%EMM%-%EDD%_00:00:00', '%EYYYY%-%EMM%-%EDD%_00:00:00', '%EYYYY%-%EMM%-%EDD%_00:00:00'
interval_seconds = 21600
opt_output_from_geogrid_path = '../Domains'
io_form_geogrid = 2,
/

&geogrid
parent_id      = 0, 1, 2,
parent_grid_ratio = 1, 5, 5,
i_parent_start = 1, 36, 55,
j_parent_start = 1, 27, 106,
e_we          = 96, 136, 201,
e_sn          = 96, 156, 201,
geog_data_res = '5m', '2m', '30s',
dx = 50000,
dy = 50000,
map_proj = 'lambert',
ref_lat  = 45.0,
ref_lon  = 11.3,
truelat1 = 45.0,
truelat2 = 45.0,
stand_lon = 11.3,
geog_data_path = '/lustre/arpa/operative/data/wrfop/geog'
/

&ungrib
out_format = 'WPS',
prefix = 'SST'
/

&metgrid
fg_name = 'FILE', 'SST',
io_form_metgrid = 2,
opt_output_from_metgrid_path = '../Domains',
/
```

```

&mod_levs
  press_pa = 201300 , 200100 , 100000 ,
            95000 , 90000 ,
            85000 , 80000 ,
            75000 , 70000 ,
            65000 , 60000 ,
            55000 , 50000 ,
            45000 , 40000 ,
            35000 , 30000 ,
            25000 , 20000 ,
            15000 , 10000 ,
            5000 , 1000
/

```

A.2 Namelist WRF

```

&time_control
  run_days           = %DAYS_PER_RUN%,
  run_hours          = 0,
  run_minutes        = 0,
  run_seconds        = 0,
  start_year         = %SYYYY%, %SYYYY%, %SYYYY%,
  start_month        = %SMM%, %SMM%, %SMM%,
  start_day          = %SDD%, %SDD%, %SDD%,
  start_hour         = 00, 00, 00,
  start_minute       = 00, 00, 00,
  start_second       = 00, 00, 00,
  end_year           = %EYYYY%, %EYYYY%, %EYYYY%,
  end_month          = %EMM%, %EMM%, %EMM%,
  end_day            = %EDD%, %EDD%, %EDD%,
  end_hour           = 00, 00, 00,
  end_minute         = 00, 00, 00,
  end_second         = 00, 00, 00,
  interval_seconds   = 21600
  input_from_file    = .true., .true., .true.,
  history_interval   = 360, 180, 60,
  frames_per_outfile = 4, 8, 6,
  adjust_output_times = .true.,
  restart            = %RESTART%,
  restart_interval   = 1440,
  io_form_history    = 2
  io_form_restart    = 2
  io_form_input      = 2
  io_form_boundary   = 2
  debug_level        = 100
  write_hist_at_0h_rst = %RESTARTda0%
  auxinput4_inname   = "wrfinput_d<domain>",
  auxinput4_interval = 360, 360, 360,
  io_form_auxinput4  = 2
/

&domains
  time_step          = 200,
  time_step_fract_num = 0,
  time_step_fract_den = 1,
  max_dom            = 3,
  e_we               = 96, 136, 201,

```

```

e_sn                = 96,   156,  201,
e_vert              = 31,   31,   31,
num_metgrid_levels = %NUM_METGRID_LEVELS%,
num_metgrid_soil_levels = 4,
p_top_requested     = 5000,
dx                  = 50000, 10000, 2000,
dy                  = 50000, 10000, 2000,
grid_id             = 1,    2,    3,
parent_id           = 0,    1,    2,
parent_grid_ratio   = 1,    5,    5,
i_parent_start     = 1,   36,  55,
j_parent_start     = 1,   27, 106,
parent_time_step_ratio = 1,    5,    5,
use_adaptive_time_step = .false.,
step_to_output_time = .true.,
min_time_step      = 100,  20,  4,
max_time_step      = 400,  80, 16,
target_cfl         = 1.2, 1.2, 1.2,
max_step_increase_pct = 5,   51,  51,
starting_time_step = 100,  20,  4,
nproc_x            = %NPROC_X%,
nproc_y            = %NPROC_Y%,
feedback           = 1,
smooth_option      = 0
/
nproc_x            = 8,
nproc_y            = 16,

&physics
mp_physics         = 8,    8,    8,
ra_lw_physics     = 1,    1,    1,
ra_sw_physics     = 1,    1,    1,
radt              = 50,   50,   50,
sf_sfclay_physics = 2,    2,    2,
sf_surface_physics = 2,    2,    2,
bl_pbl_physics    = 2,    2,    2,
bldt              = 0,    0,    0,
cu_physics        = 1,    1,    0,
cudt              = 5,    5,    0,
isfflx            = 1,
ifsnow            = 0,
icloud            = 1,
surface_input_source = 1,
num_soil_layers   = 4,
sf_urban_physics = 0,
maxiens           = 1,
maxens            = 3,
maxens2           = 3,
maxens3           = 16,
ensdim            = 144,
sst_update        = 1,
bucket_mm         = -1,
bucket_J          = -1,
prec_acc_dt       = 360,  180,  60,
/

```

```

&fdda

```

```
/

&dynamics
w_damping           = 1,
diff_opt            = 1,
km_opt              = 4,
diff_6th_opt        = 0,      0,      0,
diff_6th_factor     = 0.12,  0.12,  0.12,
base_temp           = 290.
damp_opt            = 0,
zdamp               = 5000., 5000., 5000.,
dampcoef            = 0.2,   0.2,   0.2
khdif               = 0,     0,     0,
kvdif               = 0,     0,     0,
non_hydrostatic     = .true., .true., .true.,
moist_adv_opt        = 1,     1,     1,
scalar_adv_opt       = 1,     1,     1,
/

&bdy_control
spec_bdy_width      = 5,
spec_zone            = 1,
relax_zone           = 4,
specified             = .true., .false., .false.,
nested               = .false., .true., .true.,
/

&grib2
/

&namelist_quilt
nio_tasks_per_group = 0,
nio_groups = 1,
/
```

B

Codici dei prodotti delle suite

B.1 Codici NAUSICA_ver

CODICI (XXYYZZWW):

- XX(AREA):
 - AA=Costa
 - AB= Bassa Pianura
 - AC=Media Pianura
 - AD=Alta Pianura
 - AE=Montagna
 - AF=Alpi
 - AI=Carso (solo Sgonico)

- YY(run del WRF):
 - 01=WRF con condizioni ECMWF
 - 02=WRF con condizioni ECMWF + NOAH LSM
 - 03=WRF con condizioni ECMWF + NOAH LSM + aggiornamento SST
 - 04=WRF con condizioni ECMWF + NOAH LSM + aggiornamento SST + prec cum
 - 05=WRF con condizioni ECMWF + NOAH LSM + aggiornamento SST + prec diff

- ZZ (n° di run)
 - Si veda la tabella successiva

- WW(tipo di run)
 - 01=Test confronto GRIB/misure arpa-osmer su area e stazioni
 - 02=Test confronto GRIB/misure arpa-osmer solo su stazioni
 - 03=Test da tabelle già pronte (senza estrazioni) su area e stazioni
 - 04=Test da tabelle già pronte (senza estrazioni) solo su stazioni
 - 05=Test EVTR (estrazione da netCDF+estrazione dati ET) su area e stazioni
 - 06=Test EVTR (estrazione da netCDF+estrazione dati ET) solo su stazioni

Variabile	PREC ≠ 0	PREC > 10 mm/h	PREC > 25 mm/h	TEMP	HUMI <95%	WIND I,U,V	PRES	RADT	EVTR
Periodo									
MAM 10	01	No	No	01	01	01	01	01	01
JJA 10	02	No	No	02	02	02	02	02	02
SON 10	03	No	No	03	03	03	03	03	03
DJF 10_11	04	No	No	04	04	04	04	04	04
MAM 11	05	No	No	05	05	05	05	05	05
JJA 11	06	No	No	06	06	06	06	06	06
SON 11	07	No	No	07	07	07	07	07	07
DJF 11_12	08	No	No	08	08	08	08	08	08
MAM 12	09	No	No	09	09	09	09	09	09
JJA 12	10	No	No	10	10	10	10	10	10
SON 12	11	No	No	11	11	11	11	11	11
DJF 12_13	12	No	No	12	12	12	12	12	12
MAM 13	16	No	No	16	16	16	16	16	16
JJA 13	17	No	No	17	17	17	17	17	17
SON 13	18	No	No	18	18	18	18	18	18
DJF 13_14	19	No	No	19	19	19	19	19	19
MAM 14	13	No	No	13	13	13	13	13	13
JJA 14	14	No	No	14	14	14	14	14	14
SON 14	15	No	No	15	15	15	15	15	15
DJF 14_15	20	No	No	20	20	20	20	20	20
Anno 2010	No	31	32	No	No	No	No	No	No
Anno 2011	No	41	42	No	No	No	No	No	No
Anno 2012	No	43	44	No	No	No	No	No	No
Anno 2013	No	37	38	No	No	No	No	No	No
Anno 2014	No	39	40	No	No	No	No	No	No
MAM 15	48	No	No	48	48	48	48	48	48
JJA 15	49	No	No	49	49	49	49	49	49
SON 15	50	No	No	50	50	50	50	50	50
DJF 15_16	51	No	No	52	51	51	51	51	51
MAM 16	52	No	No	52	52	52	52	52	52
JJA 16	53	No	No	53	53	53	53	53	53
SON 16	54	No	No	54	54	54	54	54	54
DJF 16_17	55	No	No	55	55	55	55	55	55
Anno 2015	No	56	57	No	No	No	No	No	No
Anno 2016	No	58	59	No	No	No	No	No	No
MAM 01	60	No	No	60	60	60	60	60	60
JJA 01	61	No	No	61	61	61	61	61	61
SON 01	62	No	No	62	62	62	62	62	62
DJF 01_02	63	No	No	63	63	63	63	63	63
MAM 02	64	No	No	64	64	64	64	64	64
JJA 02	65	No	No	65	65	65	65	65	65
SON 02	66	No	No	66	66	66	66	66	66
DJF 02_03	67	No	No	67	67	67	67	67	67
MAM 03	68	No	No	68	68	68	68	68	68
JJA 03	69	No	No	69	69	69	69	69	69
SON 03	70	No	No	70	70	70	70	70	70
DJF 03_04	71	No	No	71	71	71	71	71	71
MAM 04	72	No	No	72	72	72	72	72	72
JJA 04	73	No	No	73	73	73	73	73	73
SON 04	74	No	No	74	74	74	74	74	74
DJF 04_05	75	No	No	75	75	75	75	75	75
MAM 05	76	No	No	76	76	76	76	76	76
JJA 05	77	No	No	77	77	77	77	77	77
SON 05	78	No	No	78	78	78	78	78	78
DJF 05_06	79	No	No	79	79	79	79	79	79
MAM 06	80	No	No	80	80	80	80	80	80
JJA 06	81	No	No	81	81	81	81	81	81
SON 06	82	No	No	82	82	82	82	82	82
DJF 06_07	83	No	No	83	83	83	83	83	83
MAM 07	84	No	No	84	84	84	84	84	84
JJA 07	85	No	No	85	85	85	85	85	85
SON 07	86	No	No	86	86	86	86	86	86
DJF 07_08	87	No	No	87	87	87	87	87	87
MAM 08	88	No	No	88	88	88	88	88	88
JJA 08	89	No	No	89	89	89	89	89	89
SON 08	90	No	No	90	90	90	90	90	90
DJF 08_09	91	No	No	91	91	91	91	91	91
MAM 09	92	No	No	92	92	92	92	92	92
JJA 09	93	No	No	93	93	93	93	93	93
SON 09	94	No	No	94	94	94	94	94	94
DJF 09_10	95	No	No	95	95	95	95	95	95
Anno 2001	No	21	22	No	No	No	No	No	No
Anno 2002	No	23	24	No	No	No	No	No	No
Anno 2003	No	25	26	No	No	No	No	No	No
Anno 2004	No	27	28	No	No	No	No	No	No
Anno 2005	No	29	30	No	No	No	No	No	No
Anno 2006	No	33	34	No	No	No	No	No	No
Anno 2007	No	35	36	No	No	No	No	No	No
Anno 2008	No	96	97	No	No	No	No	No	No
Anno 2009	No	98	99	No	No	No	No	No	No

Tabella B.1: Codici ZZ dei prodotti della suite NAUSICA_ver

B.2 Codici NAUSICA_rds

CODICI (XXYYZZWW):

- XX(STAZIONE DI LANCIO DEL RADIOSONDAGGIO):
01=UDINE (Campoformido/Rivolto)
- YY(run del WRF):
05=WRF con condizioni ECMWF + NOAH LSM + aggiornamento SST +
prec diff
- ZZ (n° di run; mod lev analizzati Δh [m]; Δq [deg²])
Si veda la tabella successiva.
- WW(orario di lancio del radiosondaggio)
06
12
18
24

C

Settaggio del WRF per l'aggiornamento delle SST

Il WRF è predisposto per la lettura di specifici file contenenti un campo relativo alla temperatura superficiale del mare (SST), per i dettagli sul file necessari si veda il Par. 3.3.3.

Come riportato nella documentazione [29] l'aggiunta di queste informazioni può avvenire esclusivamente all'istante iniziale oppure ad ogni intervallo temporale della simulazione. Per l'implementazione della lettura e aggiornamento delle SST nella suite NAUSICA_run [29] è stata scelta una scadenza di 6 ore, in parallelo con la lettura delle condizioni al bordo, anche se all'interno delle condizioni al contorno il valore del campo varia solo una volta al giorno in corrispondenza delle 18:00 UTC.

Per utilizzare le SST è necessario (procedura descritta nella documentazione e nei tutorial Climate Run Example del WRF [29],[1]):

1. Fare run separati dell'*ungrib*:

- per i GRIB ECMWF (che producono output con prefisso *FILE*;) utilizzando *Vtable* e *namelist* ordinarie del WPS
- per i GRIB SST (che producono output con prefisso *SST*;) utilizzando le *Vtable.SST* (si trovano nella directory di WPS in *ungrib/Variable_Tables/Vtable.SST*) e facendo le seguenti modifiche alla *namelist*:

```
&share
start_date = \%START\_DATE\%,
end_date = \%END\_DATE\%,
interval_seconds = 21600,
&ungrib
prefix = 'SST',
```

2. Fare un run unico per il metgrid modificando la *namelist* del WPS come segue:

```
&metgrid
fg_name = 'FILE', 'SST'
```

3. Aggiungere alla *namelist* del WRF le seguenti righe:

```
&time_control
auxinput4_inname = "wrflowinp_d<domain>"
auxinput4_interval = 360,360,360
io_form_auxinput4 = 2,
&physics
sst_update = 1,
```


D

Misure a disposizione

D.1 Misure stazioni OSMER

Database di misure effettuate presso le stazioni dell'OSMER []

- Periodo: dal 1990 al 2017 (ad oggi)
- Dati giornalieri
 - pioggia totale giornaliera mm
 - temperatura giornaliera minima a 180 cm di altezza °C
 - ora della temperatura giornaliera minima a 180 cm
 - temperatura giornaliera media a 180 cm °C
 - temperatura giornaliera massima a 180 cm di altezza °C
 - ora della temperatura giornaliera massima a 180 cm
 - umidità giornaliera minima a 180 cm di altezza %
 - ora dell'umidità giornaliera minima a 180 cm
 - umidità giornaliera media a 180 cm di altezza %
 - umidità giornaliera massima a 180 cm di altezza %
 - ora dell'umidità giornaliera massima a 180 cm
 - tempo totale giornaliero di bagnatura fogliare min
 - vento da nord - velocità media a 10 m di altezza m/s
 - vento da nord - numero totale di minuti nel giorno min
 - vento da nordest - velocità media a 10 m di altezza m/s
 - vento da nordest - numero totale di minuti nel giorno min
 - vento da est - velocità media a 10 m di altezza m/s
 - vento da est - numero totale di minuti nel giorno min
 - vento da sudest - velocità media a 10 m di altezza m/s
 - vento da sudest - numero totale di minuti nel giorno min
 - vento da sud - velocità media a 10 m di altezza m/s
 - vento da sud - numero totale di minuti nel giorno min
 - vento da sudovest - velocità media a 10 m di altezza m/s
 - vento da sudovest - numero totale di minuti nel giorno min
 - vento da ovest - velocità media a 10 m di altezza m/s
 - vento da ovest - numero totale di minuti nel giorno min
 - vento da nordovest - velocità media a 10 m di altezza m/s
 - vento da nordovest - numero totale di minuti nel giorno min

- calma di vento - numero totale di minuti nel giorno min
- velocità media giornaliera del vento a 10 m di altezza m/s
- velocità massima giornaliera del vento a 10 m di altezza m/s
- direzione della velocità massima giornaliera a 10 m gr. nord
- ora della velocità massima giornaliera a 10 m
- velocità media giornaliera del vento a 2 m di altezza m/s
- velocità massima giornaliera del vento a 2 m di altezza m/s
- dir vento a 10 m in corrisp. della velocità massima giornaliera a 2 m gradi nord
- ora della velocità massima giornaliera a 2 m
- pressione giornaliera minima al livello di stazione hPa
- ora della pressione giornaliera minima
- pressione giornaliera media al livello di stazione hPa
- pressione giornaliera massima al livello di stazione hPa
- ora della pressione giornaliera massima
- radiazione globale giornaliera totale kJ/mq
- tempo totale giornaliero di insolazione min
- temperatura gior min a 0 cm °C caso H3 (temperatura gior min a 50 cm °C)
- ora della temp giorn min a 0 cm caso H3 (ora della temp giorn min a 50 cm)
- temperatura giornaliera media a 0 cm di altezza °C
- temperatura giorn mas a 0 cm °C caso H3 (temperatura giorn mas a 50 cm °C)
- ora della temp giorn mas a 0 cm caso H3 (ora della temp giorn mas a 50 cm)
- temperatura giorn min a 50 cm °C caso H3 (temperatura giorn min a 20 cm °C)
- ora della temp giorn min a 50 cm caso H3 (ora della temp giorn min a 20 cm)
- temperatura giornaliera media a 50 cm di altezza °C
- temperatura giorn mas a 50 cm °C caso H3 (temperatura giorn mas a 20 cm °C)
- ora della temp giorn mas a 50 cm caso H3 (ora della temp giorn mas a 20 cm)
- temperatura giorn min a 20 cm °C caso H3 (temperatura giorn min a 0 cm °C)
- ora della temp giorn min a 20 cm caso H3 (ora della temp giorn min a 0 cm)
- temperatura giornaliera media a 20 cm di altezza °C
- temperatura giorn mas a 20 cm °C caso H3 (temperatura giorn mas a 0 cm °C)
- ora della temp giorn mas a 20 cm caso H3 (ora della temp giorn mas a 0 cm)
- temperatura giorn minima a -10 cm nel terreno °C
- ora della temp giorn minima a -10 cm
- temperatura giornaliera media a -10 cm nel terreno °C

- temperatura giornaliera massima a -10 cm nel terreno °C
- ora della temperatura giornaliera massima a -10 cm
- temperatura giornaliera minima a -20 cm nel terreno °C
- ora della temperatura giornaliera minima a -20 cm
- temperatura giornaliera media a -20 cm nel terreno °C
- temperatura giornaliera massima a -20 cm nel terreno °C
- ora della temperatura giornaliera massima a -20 cm
- Dati orari
 - pioggia totale oraria mm
 - temperatura istantanea dell'aria a 180 cm di altezza °C
 - umidità relativa istantanea dell'aria %
 - tempo totale orario di bagnatura fogliare min
 - direzione istantanea di provenienza del vento a 10 m gr. nord
 - velocità istantanea del vento a 10 m di altezza m/s
 - velocità media oraria del vento a 10 m di altezza m/s
 - velocità media oraria del vento a 2 m di altezza m/s
 - pressione istantanea dell'aria al livello di stazione hPa
 - radiazione solare globale oraria totale kJ/mq
 - tempo totale orario di insolazione min
 - temp ist aria a 0 cm °C caso H2 (temp ist aria a 50 cm °C)
 - temp ist aria a 50 cm °C caso H2 (temp ist aria a 20 cm °C)
 - temp ist aria a 20 cm °C caso H2 (temp ist aria a 0 cm °C)
 - temperatura istantanea del terreno a 10 cm di profondità °C
 - temperatura istantanea del terreno a 20 cm di profondità °C
- Stazioni disponibili (e relativo codice)
 - Vivaro VIV
 - San_Vito_al_Tagliamento SAN
 - Brugnera BRU
 - Faedis FAE
 - Fagagna FAG
 - Udine UDI
 - Talmassons TAL
 - Palazzolo_dello_Stella PAL
 - Capriva_del_Friuli CAP
 - Gradisca_d'Isonzo GRA
 - Fossalon_di_Grado FOS
 - Cervignano_del_Friuli CER
 - Trieste TRI
 - Sgonico SGO
 - Gemona GEM
 - Codroipo COD
 - Pordenone POR
 - Lignano LIG

- Cividale_del_Friuli CIV
- Monte_Zoncolan ZON
- Monte_Lussari LUS
- Pala_d'Altei PDA
- Monte_Matajur MAT
- Monte_S._Simeone SIM
- Tarvisio TAR
- Enemonzo ENE
- Forni_di_Sopra FSP
- Monfalcone MNF
- Grado GRD
- N_Adriatic_ITT_Boa BOA
- Pradamano PRD
- Lauzacco LAU
- Bicinicco BIC
- Gorgo GRG
- Pontebba PNT
- Cave_del_Predil CDP
- Tolmezzo TOL
- Musi MUS
- Chievolis CHI
- Bordano BOR
- Pian_Cavallo PIA
- Barcis BAR
- Coritis COR

D.2 Misure Evapotraspirazione (stazioni OSMER)

Database di misure effettuate presso le stazioni dell'OSMER

- Periodo dal 2000 a Marzo 2017
- Stazioni disponibili (e codice relativo)
 - Vivaro VIV
 - San_Vito_al_Tagliamento SAN
 - Brugnera BRU
 - Faedis FAE
 - Fagagna FAG
 - Udine UDI
 - Talmassons TAL
 - Palazzolo_dello_Stella PAL
 - Capriva_del_friuli CAP
 - Gradisca_d_Isonzo GRA
 - Fossalon_di_Grado FOS
 - Cervignano_del_Friuli CER
 - Trieste TRI

- Sgonico SGO
- Gemona GEM
- Codroipo COD
- Pordenone POR
- Lignano LIG
- Cividale_del_Friuli CIV
- Tarvisio TAR
- Enemonzo ENE
- Pradamano PRD
- Pian_Cavallo PIA
- Monte_Zoncolan ZON
- Monte_Lussari LUS

D.3 Radiosondaggi

- Periodo: 1991-2017
- Stazione di lancio (codice e coordinate)
 - UDI Campoformido 16055 (con lat/lon che varia tra 46.03/13.18 e 46.04 13.19 a seconda dei periodi) fino al 17/01/2016
 - UDI Rivolto 16045 (lat/lon = 45.97/13.05) dal 18/01/2016
- Altre osservazioni
 - Fino al 2004 4 lanci al giorno: alle 00,06,12,18 UTC
 - Per il periodo successivo 3 lanci al giorno: 00,06,12 UTC
 - Dal 2010 2 lanci al giorno: 00,12 UTC
 - Il RDS 01071818.UDI ha vento non attendibile (commento all'interno del file)
 - Il RDS 02101118.UDI contiene un commento NON preceduto da "#"
- Campi misurati
 - Secondi dal lancio
 - Altezza raggiunta dalla radiosonda in m sul livello del mare
 - Pressione misurata dalla radiosonda in decimi di hPa
 - Temperatura misurata dalla radiosonda in decimi di grado centigrado
 - Umidità relativa misurata dalla radiosonda in % del mare
 - Temperatura di rugiada misurata dalla radiosonda in decimi di grado centigrado
 - Direzione di provenienza del vento in gradi Nord
 - Intensità del vento in decimi di m/s

E

Struttura ed utilizzo delle suite

NAUSICA_run e NAUSICA_pp

E.1 Utilizzo della suite

Per utilizzare questa suite si deve:

1. Settare le variabili nel file di File di inizializzazione della suite
2. Settare le *variabili da settare* nello Shell Script di creazione del file di definizione della suite
3. Se si vuole utilizzare il controllo della produzione dei GRIB modificare il file *f_found* che contiene l'ultima scadenza controllata)
4. Lanciare lo Shell Script di creazione della *suite.def* reindirizzandone l'output nel file *NAUSICA_run.def*.

E.2 Struttura della Suite

La struttura di questa suite (Fig. 4.2) sfrutta il metodo di localizzazione dei file.ecf di ECFLOW [20]:

- I task sono tutti posizionati in un'unica cartella "nome_suite" allo stesso livello
- La cartella "nome_suite" contiene anche la cartella "include" al cui interno sono definiti i file.h da includere
- Il file di definizione della suite contenente le direttive per la suite, la definizione dell'alberatura e le dipendenze viene creato tramite uno Shell Script di creazione della suite.def

Sono definite due variabili di suite che contengono data di inizio e fine dell'intera simulazione:

- INIZIO = data di inizio dell'intera simulazione
- FINE = data di fine dell'intera simulazione (ATTENZIONE: Del giorno corrispondente alla data di fine della simulazione verranno soltanto scritti wrfout e wrfrst relativi alle ore 00)

Queste vengono settate dallo Shell Script di creazione della suite.def in base alla data di start del primo blocco famiglia e alla data di fine dell'ultimo blocco famiglia.

I parametri della simulazione sono definiti in un apposito file di inizializzazione (prima versione *NAUSICA_run.ini*, versione attuale aggiornata e con aggiunta di possibilità di utilizzo di SST, creazione di GRIB e controllo di questi *NAUSICA_run_Suite_def_TOT.sh*, *NAUSICA_run_HP.ini*).

Famiglie run_NAUSICA_*

La suite NAUSICA_run è composta da famiglie run_NAUSICA_“nome” che

contengono una famiglia WPS_WRF_NAUSICA che esegue il run di WPS e WRF per il periodo definito in apposite variabili di famiglia. Ogni famiglia esegue

1. un run del WPS su tutto il periodo indicato (task “t_WPS_NAUSICA”)
2. più run del WRF a seconda di sotto-periodi indicati in un’apposita variabile di famiglia (task “t_WRF_NAUSICA”)
3. archivia output e restart del WRF (task “t_WRF_NAUSICA”)

ATTENZIONE: la “label ecfLOW” indica il periodo dell’ultimo run eseguito per WPS e WRF. Quindi nel caso di WPS tutto il periodo mentre per il WRF l’ultimo sotto-periodo.

Queste famiglie sono costruite dallo script di creazione della suite a partire da un blocco di base modificando le variabili di famiglia e sono triggerate dal completamento della famiglia precedente.

Nella versione “NAUSICA_run_HP” l’esecuzione di ogni famiglia è controllata anche da un trigger di tipo “date” e da un trigger di tipo “time” in modo da poter decidere a che ora di che giorno far partire il run della famiglia.

Variabili di famiglia

- NOMI_FAM_run_NAUSICA = nome di ciascuna famiglia (per es. Gen, Feb, Anno_2015, Anno_2014, Lun, Mar, ecc.)
- RUN_START_DATE = data inizio blocco famiglia (formato YYYYMMDD)
- RUN_END_DATE = data fine blocco famiglia (formato YYYYMMDD)
- RESTART_FLAG = flag 0 o 1 che indica se la famiglia deve iniziare da restart la simulazione (RESTART_FLAG = 1) o meno (RESTART_FLAG = 0)
- DAYS_PER_RUN = numero di giorni per ogni singolo job del WRF sottomesso

E.2.1 Famiglie di controllo

“Famiglia SENTINEL:”

1. Controlla all’inizio dell’intera simulazione l’esistenza di tutti i GRIB necessari per le condizioni al contorno (task “tsen_GRIB”)
2. Pulisce la cartella log alla fine dell’esecuzione di ogni famiglia di run (task *tsen_run_cleaner*)

“Famiglia LOG:”

1. Copia output del flusso, namelist, output ed error dei job in un file.tar alla fine dell’esecuzione di tutte le famiglie di run (task “tlog_TAR”)

ATTENZIONE: la cartella relativa all’output del task tlog_TAR stesso non può essere né copiata né rimossa, ricordarsi quindi di rimuoverla manualmente prima di ogni nuova simulazione.

ATTENZIONE: si può scegliere di archiviare solo l’output del flusso delle famiglie contenute nell’ultimo run oppure di tutte le famiglie di run presenti dall’ultimo run del task “tsen_run_cleaner” (per modificare l’opzione cambiare la parte di codice commentata all’interno del task “tlog_TAR”)

"Famiglia GRIB_SCAN:"

Questa famiglia contenuta all'interno delle famiglie di post-processing
PP_NAUSICA

1. Controlla la presenza dei file GRIB attesi per scadenza
2. Controlla la presenza all'interno di questi di alcune tra le variabili attese (elencate in "wrf_model/etc/NAUSICA_run_lista_fields.csv")
3. Riporta i valori minimo massimo e medio assunti da ciascuna di queste variabili

ATTENZIONE: è sconsigliato l'utilizzo di questo task di controllo per periodi di durata superiore ad una settimana per questioni di tempi.

OSSERVAZIONE: Nella prima versione (versione "SPY") questa famiglia è contenuta all'interno della famiglia di controllo SPY, la frequenza del controllo viene scelta tramite una variabile: "SPY_SCAN_DT" presente nello script di creazione della suite. Anche il tipo di GRIB da esaminare è determinato tramite una variabile: "SCAN_TYPE" (per ora è possibile analizzare solo GRIB di tipo CRMA). Nell'ultima versione, dato che è logico che il controllo avvenga alla fine di ogni produzione di GRIB, questa è stata spostata all'interno della famiglia di post-processing ed è stata eliminata la variabile relativa alla frequenza di controllo che viene invece triggerata dal completamento del run di UPP.

L'analisi riprende ogni volta dall'ultimo file controllato (i nomi dei file controllati vengono scritti in "/scratch/log/f_found.out", se si vuole far ripartire il controllo da capo, o da una scadenza specifica basta modificare questo file.

ATTENZIONE: ricordarsi di rimuovere sempre il file f_found prima di iniziare una nuova simulazione. Queste operazioni sono gestite da due task `spy:t_SCAN.ecf` et `SCANWEB.ecf` corredati da due job template ("wrf_model/src/NAUSICA_run_fields_control_PBSjob.tpl", "wrf_model/src/NAUSICA_run_makehtml_fieldscontrol_PBSjob.tpl") e una tabella di elenco variabili da cercare ("wrf_model/etc/NAUSICA_run_lista_fields.csv").

I risultati di queste tre analisi vengono riportati in modo leggibile su una pagina html: <http://grid0.mercuriofvog.it/nomeutente/>

E.2.2 Shell Script di creazione della suite.def

Scrivere su standard output il file di definizione della suite, basta quindi redirigerlo in un file "NAUSICA_run.def" da posizionare nella cartella ECF_HOME. Struttura:

1. Intestazione e definizione delle date di inizio e fine dell'intera simulazione come descritto nel paragrafo Suite NAUSICA_run
ATTENZIONE: Del giorno corrispondente alla data di fine della simulazione verranno soltanto scritti wrfout e wrfrst relativi alle ore 00
2. Definizione delle direttive ECF tramite funzione
ATTENZIONE: La variabile \$SUITE non viene riconosciuta da ecf se utilizzata all'interno della definizione della variabile ECF_FILES. E' stata quindi definita nello script un'altra variabile NOME_SUITE
3. Definizione della famiglia "Sentinel" tramite funzione
4. Definizione della famiglia "log" tramite funzione
5. Definizione delle Famiglie run_NAUSICA_*. Le variabili di famiglia elencate nella sezione famiglie run_NAUSICA_* devono essere inizializzate per ogni famiglia in appositi array definiti all'inizio dello script.

“Trigger:” I trigger (sia per le famiglie di run che per i task di controllo) sono definiti per sostituzione dallo script, l’esecuzione di ogni famiglia diversa dalla prima è triggerata dal completamento dell’esecuzione della famiglia precedente, nella versione HP sono presenti anche trigger di tipo date e time in modo da settare orario e data dell’esecuzione di ognuna.

Variabili da settare (ultima versione = NAUSICA_run_Suite_def.sh)

- NOME_SUITE="NAUSICA_run"
- NUMERO_FAM_run_NAUSICA = n Numero famiglie di run (ESCLUSE le famiglie di controllo log, sentinel e spy) Da questo dipende il numero di elementi letti negli array successivi
- NOMI_FAM_run_NAUSICA_Ar = ("nome1" "nome2" ...) Nomi delle famiglie di run (per es. Gen, Feb, Anno_2015, Anno_2014, Lun, Mar, ecc.)
- DATE_START_FAM_run_NAUSICA_Ar = (20150101 20150105 ...) Formato YYYYMMGG Per utilizzare il restart creato dalla famiglia precedente si deve avere data_start fam_(N) = data_end fam_(N-1)

- DATE_END_FAM_run_NAUSICA_Ar = (201505 20150110 ...) Formato YYYYMMGG Di questo giorno viene scritto wrfout e wrfst solo delle ore 00

ATTENZIONE: il wall-time dell’UPP va aumentato in modo proporzionale all’intervallo di tempo simulato da ogni famiglia: modificare la variabile DIS_WALL_TIME nel file di inizializzazione (circa 12 giorni processati per ogni ora di wall-time)

ATTENZIONE: il wall-time dei job SCAN e SCANWEB va aumentato in modo proporzionale agli intervalli di tempo intercorrenti tra la fine di una famiglia e la fine della famiglia successiva (vanno infatti calcolati anche eventuali intervalli temporali non simulati contenuti in questi nel caso di famiglie disgiunte): modificare la variabile DIS_WALL_TIME nel file di inizializzazione (circa 60 giorni processati per ogni ora di wall-time)

- RESTART_FLAG_FAM_run_NAUSICA_Ar = (0 1 ...) Se = 1 la famiglia comincia il run da file di restart Se = 1 deve esistere il wrfst nell’archivio corrispondente al primo giorno
- DAYS_PER_RUN_FAM_run_NAUSICA_Ar = (5 5 ...) Numero di giorni per ogni singolo job del WRF sottomesso (ESCLUSO il giorno finale di cui vengono scritte solo le ore 00) Deve essere un divisore del num di giorni dell’intervallo da simulare associato alla famiglia (ESCLUSO il giorno finale di cui vengono scritte solo le ore 00)

ATTENZIONE: se si aumentano i valori è necessario aumentare il wall-time del job modificando la variabile WRF_WALL_TIME nel file di inizializzazione (per 5 giorni servono piu’ di 2 ore e 30 min di wall-time)

- WRF_RUN = 1 (presente solo nella versione NAUSICA_run_Suite_def.sh e successive)

Flag che mi dice se eseguire il run del wrf o solo il post processing se = 0 viene eseguito SOLO il POST-PROCESSING (devono essere presenti i netCDF nell’archivio) se = 1 viene eseguita anche la produzione dei netCDF

- SST_FLAG = 0 (presente solo nella versione con SST e successive) Per utilizzo o meno delle temp. sup. del mare (se = 1 devono esistere i file GRIB SST)
- GRIB_Ar = ("CRMA" "MOD_LEV") (presente solo nella versione con produzione di GRIB e successive) Per utilizzare UPP e scegliere il tipo di GRIB (sono disponibili CRMA: livelli isobarici e MOD_LEV: livelli del modello). Per ogni tipologia scelta viene creato il task UPP corrispondente. Se GRIB="" si ferma alla produzione dei wrfout

- SPY_SCAN_DT="00:45" (Presente SOLO nella versione con SPY:
NAUSICA_run_Suite_def_SPY.sh)
Frequenza di controllo dei GRIB prodotti da parte della SPY
Formato hh:mm (per es. 00:45)
- SPY_SCAN_TYPE="CRMA" (Presente solo nelle ultime versioni:
NAUSICA_run_Suite_def_TOT.sh, NAUSICA_run_Suite_def_HP.sh)
Tipo di output da analizzare con la spy
CRMA = GRIB livelli isobarici
ATTENZIONE: per ora è l'unica possibilità implementata
ATTENZIONE: è un'analisi pesante, quindi utilizzarla solo per run su pochi GRIB.
ATTENZIONE: per il caso di GRIB con il problema CDO togliere i commenti nel job.tpl
- NAUSICA_RUN_TIME (Presente solo nella versione
NAUSICA_run_Suite_def_HP.sh)
Orario di inizio di ogni blocco di simulazione (ogni NAUSICA_RUN_DT giorni)
Formato HH:MM
ATTENZIONE: è l'ora UTC
- NAUSICA_RUN_DATE (Presente solo nella versione
NAUSICA_run_Suite_def_HP.sh) Data di inizio simulazione Formato YYYYMM-MDD
- NAUSICA_RUN_DT=1 (Presente solo nella versione
NAUSICA_run_Suite_def_HP.sh) Intervallo tra i giorni di simulazione Per es. 1=tutti i giorni, 2=ogni 2 giorni
- CLEAN=0 (Presente solo nella versione
NAUSICA_run_Suite_def_HP.sh) Flag per controllare la pulizia iniziale dei log: Se = 0 NON viene pulita la cartella dei log (verranno salvati nell'archivio dal task tlog_TAR alla fine della simulazione se ARCHIV=1) Se = 1 vengono eliminati tutti i log presenti creati in precedenza all'inizio del run della suite
- ARCHIV=1 (Presente solo nella versione
NAUSICA_run_Suite_def_HP.sh) Flag per attivare la copia di output flusso, jobout e joberror in un file compresso Se = 0 non viene creato l'archivio Se = 1 viene creato l'archivio Per esempio se si vuole archiviare un'intera simulazione divisa in piu' run della suite in un unico archivio attivare l'opzione solo all'ultimo run

E.2.3 File di inizializzazione della suite .ini

E' stato creato un file di inizializzazione della suite per ogni versione di *shell script* di creazione della suite: *NAUSICA_run_nome_sh_script.ini* Utilizzo di SST nella suite NAUSICA_run Per includere la possibilità di aggiungere i dati sulle temperature superficiali del sono state apportate le seguenti modifiche alla struttura della suite NAUSICA_run:

1. è stato preparato un nuovo script di creazione della suite (prime versioni: *NAUSICA_run_Suite_def_SST.sh, NAUSICA_run_Suite_def_TOT.sh, NAUSICA_run_Suite_def_HP.sh* ultima versione: *NAUSICA_run_Suite_def.sh*) contenente una variabile flag aggiuntiva SST_FLAG che indica se utilizzare (SST_FLAG=1) o meno (SST_FLAG=0) le SST
2. è stato preparato un nuovo task per il WPS la cui esecuzione nel caso di utilizzo di SST va necessariamente divisa in due fasi come descritto nella sezione di utilizzo SST ed un nuovo task per il WRF:
 - *t_WPS_SST_NAUSICA.ecf* che esegue geogrid, ungrib per file ECMWF, ungrib per file SST e metgrid

- *t_WRF_SST_NAUSICA.ecf* che esegue il WRF
3. Sono stati creati nuovi job template per WPS e WRF per il caso SST_FLAG=1
 4. Sono state preparate nuove namelist template per WPS e WRF per il caso SST_FLAG=1
 5. è stato modificato il file di inizializzazione (prime versioni: *NAUSICA_run_SST.ini*, *NAUSICA_run_TOT.ini* ultima versione: *NAUSICA_run_Suite_def_HP.ini*) in modo da contenere l'indicazione dei path dei nuovi template e dei GRIB SST necessari

L'unica differenza nella struttura della suite è che, per WPS, WRF e sentinella, vengono utilizzati i task con opzione SST.

E.3 Post-processing nella suite NAUSICA_run

Ad ogni famiglia di run può essere aggiunta una famiglia PP_NAUSICA dedicata al post-processing degli output del wrf implementata nella struttura della suite nel modo seguente:

1. Nello script di creazione della suite (versioni *NAUSICA_run_Suite_def_TOT.sh*, *NAUSICA_run_Suite_def_HP.sh*, *NAUSICA_run_Suite_def.sh*) è presente una variabile aggiuntiva GRIB che indica se si vogliono produrre i GRIB (tramite UPP) a partire dai wrfout e che tipo di GRIB (leggere il prossimo paragrafo per conoscere le tipologie di GRIB disponibili).
2. I task *t_UPP_GRIB_*.ecf* contenuti nella famiglia UPP_NAUSICA gestiscono il run dell'UPP.
3. File di inizializzazione, template job e namelist e file parametri necessari all'UPP sono indicati nel file di inizializzazione (versioni: *NAUSICA_run_TOT.ini*, *NAUSICA_run_Suite_def_HP.sh*)

Si possono scegliere due tipi di GRIB: CRMA (livelli isobarici) e MOD_LEV (livelli del modello).

Se si scelgono entrambi verranno creati due task all'interno della famiglia UPP_NAUSICA eseguiti contemporaneamente.

Per poter fare in modo che possano essere eseguiti o uno o l'altro o entrambi è stato aggiunto un task *t_UPP_cleaner* che prepara i run dell'UPP pulendo i link ai wrfout nella cartella di lavoro e creando i nuovi. Inoltre per permettere l'esecuzione contemporanea sono stati creati file di inizializzazione e template analoghi per i due casi invece di utilizzare un template con sostituzione (questo potrebbe essere implementato nel task *t_UPP_cleaner*).

E.3.1 Controllo dei GRIB prodotti e delle variabili in essi contenute

All'interno della famiglia di post-processin PP_NAUSICA è stata aggiunta una famiglia GRIB_SCAN descritta tra le famiglie di controllo che genera una pagina html nella quale sono riportati i risultati della ricerca di file GRIB per scadenza, variabili e breve statistica sui valori di queste (massimi, minimi e valori medi).

ATTENZIONE: ricordarsi di rimuovere sempre il file *f_found* prima di iniziare una nuova simulazione

F

Struttura ed utilizzo della suite WRF_ana_post

F.1 Utilizzo della suite

Per utilizzare questa suite si deve:

1. Scegliere quale produzione attivare cambiando il valore della variabile GRIB1_USER nel file di definizione della suite. In particolare la famiglia FARM produce GRIB di tipo analisi a partire dai wrfout di NAUSICA_run, la famiglia FARM_fc produce GRIB di tipo forecast a partire dagli output della catena operativa e la famiglia FARM_ana2fc produce GRIB con la forma di quelli generati da FARM_bc ma a partire dai wrfout di analisi di NAUSICA.
2. Settare il "file di inizializzazione di famiglia" ("../wrf_model/etc/WRF_ap_user-famiglia.ini") contenente le informazioni necessarie al run delle famiglie di post produzione.
In particolare settare le variabili:

- START_DATE: data di inizio dell'intervallo di tempo post processato
- END_DATE: data di fine dell'intervallo di tempo post processato (compresa nei giorni post processati)
- DT_GRIB_CHECK: quanti giorni di GRIB prodotti si vogliono controllare per ogni job lanciato
- SLEEP_MIN ogni quanti minuti si vuole eseguire il controllo dei GRIB prodotti

F.2 Struttura della Suite

Questa suite è pensata per un eventuale utilizzo nella catena operativa, le variabili che caratterizzano i run sono quindi definite in file di inizializzazione di famiglia modificabili senza ricaricare la suite.

I parametri generali della simulazione sono definiti in un apposito file di inizializzazione della suite "WRF_ana_post.ini" mentre i parametri relativi alle singole famiglie sono definiti in file di inizializzazione di famiglia del tipo: "../wrf_model/etc/WRF_ap_user-famiglia.ini"

F.2.1 Famiglie di produzione user_%user-famiglia%

Ognuna di queste famiglie è caratterizzata da:

- START_DATE data di inizio dell'intervallo di tempo post processato
- END_DATE data di fine dell'intervallo di tempo post processato (compresa nei giorni post processati)

- file `"../wrf_model/etc/WRF_ap_wrf_4_"user-famiglia"_cntrl.parm"` che definisce campi e livelli da includere nei GRIB prodotti

F.2.2 Task di produzione

"task t_wrf2grib1":

1. Crea i GRIB1 dai wrfout netCDF tramite job template
`"../wrf_model/src/WRF_ap_"user-fam"-run_unipost.sh
 ../wrf_model/src/WRF_ap_"user-fam"_unipost_frames.template
 ../wrf_model/src/WRF_ap_upp_GRIB Extr_sub-hourly_PBS.tpl.job "`
 La percentuale di GRIB creati è indicata tramite il meter "GRIB"

F.2.3 Task di controllo

"task t_cleaner":

1. Pulizia log e work directory di famiglia all'inizio del run

"task t_WRFnetCDF_check":

1. Creazione work e log dir
2. Controllo presenza wrfout netCDF relativi alle date scelte definite nel file di inizializzazione di famiglia
3. Creazione dei link a questi nella work dir Se mancano dei netCDF si accende l'event "ASSENZA_netCDF"

"task t_GRIB_check":

1. Controllo presenza GRIB1 creati per ogni scadenza di ogni dominio relativi al periodo indicato nel file di inizializzazione di famiglia Se mancano dei GRIB si accende l'event "ASSENZA_GRIB" Questo task è triggerato tramite le variabili
 - DT_GRIB_CHECK: quanti giorni di GRIB prodotti si vogliono controllare ad ogni controllo
 - SLEEP_MIN: ogni quanti minuti si vuole eseguire il controllo dei GRIB prodotti

"task t_plot"

1. Crea plot di controllo su tutto il periodo definito nel file di inizializzazione di famiglia su tutti i domini per le variabili indicate nel file di inizializzazione di famiglia tramite due job template i quali creano progressivamente una tabella con i valori massimo e minimo per i campi scelti leggendoli dai GRIB tramite WGRIB (`"../wrf_model/src/WRF_ap_plot.tpl.job"`) e generano i plot con gnuplot (`"../wrf_model/src/WRF_ap_tab.tpl.job"`).
2. Alla fine dei run archivia i plot in `"data/WRF_ap_plot/"user-famiglia"_"variabile"_"dominio"_"START-DATE"_"END-DATE".png"` Questo task è triggerato tramite le variabili
 - DT_GRIB_CHECK: quanti giorni di GRIB prodotti si vogliono controllare ad ogni controllo
 - SLEEP_MIN: ogni quanto si vuole eseguire il controllo dei GRIB prodotti
 - SLEEP_PLOT_D ritardo in minuti della partenza del task "t_plot" rispetto al task "t_GRIB_check" (sono i due task di controllo sulla produzione dei GRIB)

Aggiorna ad ogni run i plot prodotti fino a coprire tutto il periodo indicato nel file di inizializzazione di famiglia.

OSSERVAZIONE: Questo task contiene anche il compito del task

t_GRIB_check ad esclusione dell'attivazione dell'event di visualizzazione per l'assenza di file GRIB che potrebbe però essere aggiunto. Viene mantenuto anche il task t_GRIB_check per questioni di versatilità della suite. OSSERVAZIONE: Se mancano delle scadenze nei file GRIB il plot viene creato comunque sul periodo indicato.

ATTENZIONE: Nel caso si voglia fare un run di questo task senza rieseguire tutto il flusso (per esempio su un intervallo temporale post-elaborato in precedenza) è necessario fare prima il run del task di pulizia: t_cleaner.

"task tlog_TAR:

1. Copia output flusso, jobout e joberror in un file compresso: "/data/tar_log/"

ATTENZIONE: la cartella relativa a tlog output va rimossa a mano alla fine delle simulazioni

F.2.4 Famiglia FARM_ana2fc

La famiglia ana2fc produce file con nomi e date/ore da correggere per convertirle utilizzare, con opportune modifiche, i seguenti comandi e poi spostarli a mano nelle cartelle giornaliere corrispondenti:

```
for F in $( ls /scratch/WRF_ap_run/FARM_ana2fc/CRMA_WRF-4-FARM_d03_2016010100+000*.GRIB ); do
    H=$( echo $F | cut -f 2 -d"+" | cut -f 1 -d"h" )
    N=$( printf "\%i\n" $((10#{$H})) )
    if [[ $N -ge 120 ]] && [[ $N -le 143 ]]; then
        N_OK=$(( $N - 120 ))
        F_OK=CRMA_WRF-4-FARM_d03_2016010600+`printf "\%06i" $((10#{$N_OK}))`h.GRIB
        echo $F_OK
        grib_set -s dataDate=20160106,stepRange=$N_OK $F /scratch/WRF_ap_run/FARM_ana2fc/$F_OK
    fi
done
```

F.2.5 Aggiunta del taglio del dominio

La suite WRF_ana_post descritta è stata modificata perché il FARM ha attualmente problemi a leggere il dominio do3 NAUSICA (é troppo grande). Per poterlo ridurre alle dimensioni del do3 operativo è stato introdotto il flag Do3_CUT nel file di inizializzazione. Il taglio viene eseguito del job: /scratch/WRF_ap_run/FARM_box/ritaglio.job che

- Taglia i netCDF per mezzo dell'utility WRF **iowrf** (che serve per tagliare sulle tre dimensioni i campi dai wrfout, destaggerare i campi nei netcdf e comprimerli). Per compilare questo eseguibile è necessario utilizzare la versione 3.6 delle netCDF e non versioni più recenti altrimenti restituisce un errore. Il taglio impiega circa 10 – 15 secondi per file wrfout_do3 (quindi circa 5 ore per anno).
- Post-processa i netCDF tagliati con UPP che legge questi file e li processa senza generare errori

Le griglie dei GRIB FARM prodotti passano, tramite quest'ultima modifica nella catena, da/a:

Prima:

```
1 : lonlat > size : dim = 8372 nx = 92 ny = 91
lon : first = 11.913 last = 14.225 inc = 0.0254065934 degrees_east
lat : first = 45.3 last = 46.917 inc = 0.0179666667 degrees_north
```

Dopo:

```
1 : lcc > size : dim = 13200 nx = 110 ny = 120 North Pole
originLon = 11.479 originLat = 45.306 lonParY = 11.3
lat1 = 45 lat2 = 45 xinc = 2000 m yinc = 2000 m
```


G

Struttura ed utilizzo delle suite NAUSICA_ver e NAUSICA_nc

Entrambe le suite svolgono gli stessi compiti e utilizzano gli stessi job template ma NAUSICA_ver è pensata per l'estrazione di campi dai GRIB risultanti dalla post elaborazione degli output delle simulazioni di NAUSICA_run e il confronto di questi con le misure delle stazioni ARPA-OSMER. La suite NAUSICA_nc esegue invece i confronti direttamente a partire da tabelle create dall'utente. E' stata pensata ed utilizzata per confrontare campi estratti direttamente dai wrfout.nc (evapotraspirazione). Dato che questa analisi è stata introdotta tra quelle routinarie è stata implementata anche la possibilità di estrazione di questa variabile direttamente dai wrfout.nc del modello e contiene anche le operazioni necessarie al confronto.

G.1 Utilizzo della suite NAUSICA_ver e NAUSICA_nc

Per utilizzare questa suite si deve:

1. Settare il file csv "elenco confronti" contenente le informazioni necessarie alla creazione delle famiglie di estrazione (entrambe le suite leggono dallo stesso file).
2. Settare il template del file di inizializzazione di famiglia cambiando SOLO nomi, livelli, unità di misura delle variabili da estrarre, numero plot attesi, path e template del nome dei file GRIB da utilizzare per il confronto. Se si vogliono settare questi parametri diversamente per ogni famiglia è necessario seguire le operazioni del punto 6 (solo per estrazione da GRIB, nel caso di estrazione EVTR o di run a partire da tabelle già create è settata automaticamente ad EVTR).
3. (Creare, se non è già presente, per ogni famiglia di estrazione il file con l'elenco dei vertici della poligonale delimitante l'area da estrarre senza commenti.)
4. (Settare i nomi da associare a GRIB e misure nei plot nel file di inizializzazione della suite.)
5. Lanciare lo shell script di creazione della suite indirizzandone l'output in "NAUSICA_ver.def" o "NAUSICA_nc.def".
6. (Se è necessario settare i parametri di famiglia diversamente per ognuna di queste, editare i file ini di famiglia
"/scratch/ini_ver/ NAUSICA_ver_"codice-famiglia"_"data-inizio"_"data-fine".ini." creati dallo shell script di creazione della suite).

G.2 Struttura della suite NAUSICA_ver

La struttura di questa suite sfrutta il metodo di localizzazione dei file.ecf di ECFLOW:

- I task sono tutti posizionati in un'unica cartella "nome_suite" allo stesso livello
- La cartella "nome_suite" contiene anche la cartella "include" al cui interno sono definiti i file.h da includere
- Il file di definizione della suite contenente le direttive per questa, la definizione dell'alberatura e le dipendenze viene creato tramite uno shell script che legge un file contenente tutti i confronti che si intendono eseguire e quelli già svolti tra misure e simulazioni individuati tramite una codifica univoca: il file elenco confronti.

A partire da quest'ultimo file lo script definisce tre "variabili di suite" che contengono i codici delle famiglie di confronto da svolgere, data di inizio e fine dell'estrazione di ogni singola famiglia:

- CODICE_FAMIGLIA = codice univoco della famiglia di estrazione definito nel file elenco confronti
- VER_START_DATE = data di inizio estrazione di ogni famiglia definita nel file elenco confronti
- VER_END_DATE = data di fine estrazione di ogni famiglia definita nel file elenco confronti

I parametri generali della simulazione sono definiti in un apposito file di inizializzazione "NAUSICA_ver.ini" mentre i parametri relativi alle singole famiglie sono definiti in file di inizializzazione di famiglia individuati tramite il relativo codice "/scratch/ini_ver/ NAUSICA_ver_"codice-famiglia"_"data-inizio"_"data-fine".ini" creati dallo shell script a partire da due template da settare:

"/src/NAUSICA/wrf_model/etc/ NAUSICA_ver_TPL_A.ini" (per le analisi su punti ed aree: codice WW=01) e

"/src/NAUSICA/wrf_model/etc/ NAUSICA_ver_TPL_P.ini" (per le analisi solo su punti di stazione: codice WW=02).

G.2.1 File elenco confronti

File csv contenente le informazioni necessarie alla creazione delle famiglie di estrazione, letto dallo shell script di creazione della suite. Elenco di campi separati da ";" contenenti:

1. codice svolgimento (0 = da fare, 1 = fatto);
2. codice identificativo famiglia (formato XXYZZWW)
3. data inizio confronto (formato YYYYMMDD);
4. data fine confronto (formato YYYYMMDD);
5. descrizione sintetica XX (informazioni sull'AREA)
6. descrizione estesa XX
7. descrizione sintetica YY (informazioni sulle condizioni al contorno della simulazione)
8. descrizione estesa YY
9. descrizione sintetica ZZ (numero di run)
10. descrizione estesa ZZ
11. descrizione sintetica WW (informazioni su variabili e test da eseguire)
12. descrizione estesa WW
13. commento al confronto

G.2.2 Variabili confrontabili

Le variabili confrontabili con la suite sono (le operazioni necessarie sono svolte dall'eseguibile `f_serie_temp` se richiedono una conversione di unità di misura o devono essere ricavate da altre):

- Estribili dai GRIB:
- `PREC` \neq 0: pioggia totale oraria [mm/h]
- `HUMI` < 95%: umidità relativa istantanea dell'aria [%]
- `TEMP`: temperatura istantanea dell'aria a 180 cm di altezza (deve essere convertita perché misurata in C ed estratta in K dai GRIB) [C o K]
- `PRES`: pressione istantanea dell'aria al livello di stazione (deve essere convertita perché misurata in hPa ed estratta in Pa) [hPa]
- `WIND`:
 - `WIND_U` e `WIND_V`: componenti u e v della velocità istantanea del vento a 10 m di altezza (estratte dai GRIB ma da calcolare a partire da intensità e direzione per i dati misurati) [m/s]
 - `WIND_I`: velocità istantanea del vento a 10 m di altezza (misurata dalle stazioni ma da calcolare a partire dalle sue componenti nei dati estratti dai GRIB) [m/s]
- `RADT` \neq 0: radiazione solare totale (misurata dalle stazioni ma da convertire da $kJ/h * mq$ a W/mq , da calcolare nei GRIB sommando i contributi della radiazione longwave e shortwave) [W/mq]

ATTENZIONE: l'ordine con cui vengono indicate le variabili scelte per il confronto in appositi array (uno per le variabili ed uno per le sotto-variabili della variabile `WIND`) nel file di inizializzazione di famiglia è fissato: NON può essere invertito.

L'ordine è:

1. `PREC`
2. `TEMP`
3. `HUMI`
4. `WIND`
5. `PRES`
6. `RADT`

Estraibili da netCDF

1. `EVTR`: evapotraspirazione giornaliera [mm/mq] (deve essere calcolata a partire dal flusso orario estratto dai netCDF, variabile `QFX`)

OSSERVAZIONE: per la variabile `WIND`, essendo questa descritta tramite basi diverse nei due set di dati, l'elaborazione è diversa per misure e GRIB mentre è analoga per tutte le altre variabili (ad eccezione di cambi di unità di misura).

OSSERVAZIONE: anche per la variabile `RADT` l'elaborazione è stata scritta in modo differente in quanto si pensava di dover calcolare la radiazione totale (corta+lunga), invece, dato che quella contenuta nelle misure è solo radiazione corta e va quindi confrontata con questa componente, è stato modificato il codice del filtro FORTRAN che invece di sommare le due componenti attribuisce alla variabile `rad tot` la solo componente corta. Il resto del codice NON è stato modificato quindi viene estratta anche la componente lunga e creata una tabella con le due componenti: per ottimizzare il codice andrebbe modificata tutta la trattazione della variabile `RADT`.

OSSERVAZIONE: per una corretta analisi della `RADT` bisognerebbe calcolare dai GRIB, tramite UPP, il valore integrato di questa, questo deve infatti essere

confrontato con la radiazione oraria misurata. Per poter ottenere questa variabile tramite UPP (post-processing della suite NAUSICA_ver) è necessario aggiungere il flag relativo nella produzione dei netCDF (namelist del wrf). Nei casi di netCDF costruiti senza questo flag il confronto può essere fatto calcolando dal valore delle misure il valore al secondo (questa operazione avviene e si può controllare all'interno dell'eseguibile f_serie_temp

UNITA' DI MISURA: Queste vengono indicate nel file di inizializzazione di famiglia in vettori paralleli a quelli degli elenchi. Ad eccezione dell'unità di misura della temperatura, la quale viene utilizzata per la scelta della conversione (sono disponibili gradi Celsius e gradi Kelvin), le altre vengono semplicemente riportate nei plot.

E' consigliato scegliere la temperatura in °C in quanto questa scelta riduce il tempo di calcolo del task t_Aext_osmer drasticamente.

OSSERVAZIONE: nel caso di operazioni di taglio (PREC, HUMI e RADT) è stato necessario approssimare i valori estratti dagli output del modello allo stesso numero di cifre disponibile nei valori misurati.

G.2.3 Famiglie di estrazione NAUSICA_ver_"codice"_"periodo"

Ogni famiglia di estrazione definita nel file elenco confronti è caratterizzata da

- periodo di estrazione
- area di estrazione
- variabili da confrontare
- tipo di analisi da eseguire

OSSERVAZIONI GENERALI:

- Dato che in ogni area possono essere contenute più stazioni le cui serie temporali devono essere estratte separatamente ma in modo analogo, si utilizza la tecnica del "job-array" lanciando contemporaneamente un job di estrazione per ogni stazione sia per i dati che per i GRIB. Anche l'operazione di combinazione e di confronto utilizzano questo metodo. Dove necessario vengono create sottocartelle di lavoro per i singoli job dell'array identificate tramite il nome della stazione.
- Ogni task si crea una sotto directory di run identificata dal nome del task all'interno di una a livello superiore caratterizzata dal nome della famiglia. I file che devono essere utilizzati da più di un task vengono creati o spostati a livello della directory di famiglia.
- In caso di mancanza di elementi necessari al confronto: file GRIB del periodo selezionato o stazioni OSMER nell'area selezionata, si accende un "event" visibile sull'ecflowviewer e l'esecuzione passa alla famiglia di confronto successiva (tramite comando ecflow "complete").

"Famiglia ESTRAZIONE"

"task t_ext_osmer":

1. Crea la sotto directory di lavoro
2. Controlla che nell'area scelta (tramite elenco dei vertici del bordo nel file accessorio AREA) sia contenuta almeno una stazione (utilizzando l'eseguibile point_in_polygon). Le coordinate delle stazioni per il controllo sono lette nel file accessorio elenco stazioni OSMER. Crea un file con l'elenco di quelle da estrarre per la famiglia. Se nell'area scelta non sono contenute stazioni si accende l'ecflow event "Fuori AREA" e l'intera famiglia di confronto viene settata come completata (NON vengono quindi attivati altri task di questa) altrimenti il numero di stazioni da estrarre è riportato in una label.

3. Viene preparato il job-array tramite template ("`../wrf_model/src/ NAUSICA_ver_ext_osmer.tpl.job`") e lanciato. Questo estrae per ogni stazione la serie temporale delle variabili richieste nel file di inizializzazione di famiglia utilizzando l'eseguibile `estrai` e le scrive in una forma analoga a quella che risulterà dall'estrazione dei GRIB con cui devono essere combinate e confrontate (in particolare si fa la sostituzione delle ore 24 con le ore 00).
OSSERVAZIONE: dato che il file di input dell'eseguibile `estrai` deve avere lo stesso nome per ogni esecuzione, viene creata una directory temporanea per ogni stazione e poi rimossa.
4. Nel caso di richiesta di analisi dei venti il job crea anche una tabella in cui sono estratte le sole variabili relative a questi (direzione "dir" ed intensità "I").
5. Nel caso di richiesta di confronto relativamente alla radiazione totale crea una tabella specifica in cui sono riportati i valori di questa.
6. Vengono spostate a livello superiore della directory di lavoro le tabelle di estrazione dei dati (una, o due nel caso dei venti, per ogni stazione)

"task `t_ext_simu_punto`":

Per questo task esistono tre versioni che fanno uso per l'estrazione dai GRIB rispettivamente di CDO, GRIB TOOLS e GRIB API. A causa della struttura dei GRIB prodotti da UPP è risultato più funzionale in questo caso l'utilizzo della versione con `cdo`.

"Versione con utilizzo di CDO":

1. Crea la sotto directory di lavoro e carica il modulo CDO
2. Controlla la presenza dei GRIB necessari Se non sono presenti tutti i GRIB si accende l'ecflow event "ASSENZA GRIB" e l'intera famiglia di confronto viene settata come completata (NON vengono quindi attivati altri task di questa).
3. Se ci sono tutti i GRIB prepara il job-array tramite template ("`../wrf_model/src/ NAUSICA_ver_ext_simu_p_cdo.tpl.job`") e lo lancia. Questo estrae la serie temporale per ogni stazione (utilizzando il comando "remapdis" delle CDO) creando una tabella per ogni variabile letta nel file di inizializzazione di famiglia (utilizzando il comando "outputtab" delle CDO).
4. Se si vogliono analizzare anche i venti crea una tabella unica contenente le due componenti di questi.
5. Se si vuole confrontare anche la radiazione vengono estratte sia la long-wave che la shortwave e combinate in un'unica tabella.
6. Vengono spostate a livello superiore della directory di lavoro le tabelle di estrazione dei GRIB (una, o due nel caso dei venti, per ogni stazione e per ogni variabile)

"Versione con utilizzo di GRIB TOOLS (obsoleta)":

OSSERVAZIONE: Per utilizzare questa versione si deve:

- Cambiare il `job.tpl` nel file ini della suite
 - Cambiare il codice delle var da estrarre nel file.ini di famiglia
1. Crea la sotto directory di lavoro e carica il modulo GRIB_API
 2. Controlla la presenza dei GRIB necessari
Se non sono presenti tutti i GRIB si accende l'ecflow event "ASSENZA GRIB" e l'intera famiglia di confronto viene settata come completata (NON vengono quindi attivati altri task di questa).
 3. Se ci sono tutti i GRIB prepara il job-array tramite template ("`../wrf_model/src/ NAUSICA_ver_ext_simu_p_api.tpl.job`") e lo lancia.

Questo estrae la serie temporale per ogni stazione (utilizzando il comando "grib_copy" degli API tools) creando una tabella per ogni variabile letta nel file di inizializzazione di famiglia (utilizzando il comando "grib_get" degli API tools).

ATTENZIONE: i GRIB prodotti da UPP hanno il valore della key "jScansNegatively" (verso di lettura della griglia) settato ad 1 invece che a 0, viene quindi cambiato tramite il comando "grib_set" degli API tools prima della creazione della tabella.

4. Se si vogliono analizzare anche i venti crea una tabella unica contenente le due componenti di questi.
5. Se si vuole confrontare anche la radiazione vengono estratte sia la longwave che la shortwave e combinate in un'unica tabella.
6. Vengono spostate a livello superiore della directory di lavoro le tabelle di estrazione dei GRIB (una, o due nel caso dei venti, per ogni stazione e per ogni variabile)

"Versione con utilizzo di GRIB_API (obsoleta)"

OSSERVAZIONE: Per utilizzare questa versione si deve:

- Cambiare il job.tpl nel file ini della suite
- Cambiare il codice delle var da estrarre nel file.ini di famiglia
- Creare un GRIB maschera con 1 nei punti di griglia interni all'area scelta e 0 per quelli esterni

1. Crea la sotto directory di lavoro, carica il modulo GRIB_API e copia nella work dir l'eseguibile extraction.exe
2. Controlla la presenza dei GRIB necessari Se non sono presenti tutti i GRIB si accende l'ecflow event "ASSENZA GRIB" e l'intera famiglia di confronto viene settata come completata (NON vengono quindi attivati altri task di questa).
3. Se ci sono tutti i GRIB prepara il job-array tramite template ("../wrf_model/src/ NAUSICA_ver_ext_simu_p_api_F90.tpl.job") e lo lancia. Questo estrae la serie temporale per ogni stazione (utilizzando le opzioni -f -var -tyl -lev dell'eseguibile extraction.exe) creando una tabella per ogni variabile letta nel file di inizializzazione di famiglia (utilizzando l'opzione -mask dell'eseguibile).

ATTENZIONE: i GRIB prodotti da UPP hanno il valore della key "jScansNegatively" (verso di lettura della griglia) settato ad 1 invece che a 0, viene quindi cambiato tramite il comando "grib_set" degli API tools prima della creazione della tabella.

4. Se si vogliono analizzare anche i venti crea una tabella unica contenente le due componenti di questi.
5. Se si vuole confrontare anche la radiazione vengono estratte sia la longwave che la shortwave e combinate in un'unica tabella.
6. Vengono spostate a livello superiore della directory di lavoro le tabelle di estrazione dei GRIB (una, o due nel caso dei venti, per ogni stazione e per ogni variabile)

"task t_comb_serie":

1. Crea la sotto directory di lavoro
2. Prepara un job-array tramite template ("../wrf_model/src/NAUSICA_ver_comb.tpl.job") che combina le serie temporali estratte dai dati OSMER con quelle estratte dai GRIB in un'unica tabella (eventualmente cambiando le unità di misura e ricavando le grandezze non direttamente disponibili tramite l'eseguibile f_serie_temp) e lo lancia (elimina le linee contenenti "Missing Value").

3. Sposta a livello superiore della directory di lavoro le tabelle di combinazione tra le estrazioni dei dati e dei GRIB (una per ogni stazione e per ogni variabile)

ATTENZIONE: Questo task deve essere ottimizzato sostituendo i due cicli for annidati con un "grep"

"task t_Aext_osmer":

1. Crea la sotto directory di lavoro
2. Prepara e lancia un job ("../wrf_model/src/ NAUSICA_ver_Aext_osmer.tpl.job") che concatena in un unico file le serie temporali estratte dalle misure ARPA OSMER dal task t_ext_osmer.ecf di tutte le stazioni contenute nell'area richiesta

"task t_ext_simu_area":

1. Crea la sotto directory di lavoro
2. Prepara e lancia un job ("../wrf_model/src/ NAUSICA_ver_ext_simu_a.tpl.job") che crea tabelle delle estrazioni dei valori su tutti i punti di griglia dell'area scelta e, nel periodo indicato dalle variabili di famiglia, per le var indicate nel file .ini di famiglia.
3. Vengono eliminate le linee contenenti i "Missing Value" GRIB da tutte le tabelle create

Famiglie di estrazione NAUSICA_nc_"codice"_"periodo"

"task t_ET_comb_tab" (opzionale)

Task per confrontare la variabile EVTR tra netCDF e dati Cicogna che crea le tabelle di combinazione delle serie temporali (questo task è opzionale in modo che l'operazione possa anche essere svolta a mano per altre variabili)

1. Crea la sottodirectory di lavoro
2. Cerca tra le stazioni per le quali sono disponibili i dati dell'EVTR quali sono contenute nell'area considerata (utilizza l'eseguibile point_in_polygon.exe)
3. Lancia il job costruito dal template "/wrf_model/src/ NAUSICA_ver_extraction_NC.tpl.job" che estrae dai wrfout i valori orari dell'evapotraspirazione giorno per giorno
4. Calcola i valori giornalieri sommando i valori giornalieri
5. Estrae dai dati a disposizione i valori giornalieri utilizzando le CDO [24]
6. Scrive le tabelle di combinazione tra valori estratti e misurati per ogni stazione (tabella_comb_stz_EVTR.tab) nella directory ini_ver

"task t_A_ET_comb_tab" (opzionale)

Task per confrontare la variabile EVTR tra netCDF e dati Cicogna che crea le tabelle di estrazione sulle aree (questo task è opzionale in modo che l'operazione possa anche essere svolta a mano per altre variabili)

1. Crea la sottodirectory di lavoro
2. Cerca tra le stazioni per le quali sono disponibili i dati dell'EVTR quali sono contenute nell'area considerata (utilizza l'eseguibile point_in_polygon.exe)
3. Lancia il job costruito dal template "/wrf_model/src/ NAUSICA_ver_A_extraction_NC.tpl.job" che estrae dai wrfout i valori orari dell'evapotraspirazione per mezzo dell'eseguibile GRIB_API [26] extraction_V3.exe e li somma utilizzando il comando timesum delle CDO [24]
4. Crea la tabella di estrazione da simulazione con un' unica colonna di valori (tabella2_A_codice-area.tab) nella directory ini_ver

5. Concatena tutte le serie temporali estratte dal task precedente relative a stazioni contenute nell'area in una tabella con tre colonne: data (formato anno-mese-giorno), ora (formato hh:mm:ss) e valore (tabella1_A_codice-area.tab) nella directory ini_ver

"task t_comb_ini"

Inizializza la suite generando le tabelle di input per il resto della suite a partire da quelle create dai task opzionali o dall'utente nella directory ini_ver. Tabelle necessarie alla suite nella directory ini_ver:

- tabella_comb_stz_EVTR.tab per ogni stazione da considerare a 4 colonne: data (formato anno-mese-giorno), ora (formato hh:mm:ss), valore simulazione e valore misure
- tabella1_A_codice-area.tab per ogni area da considerare a 3 colonne: data (formato anno-mese-giorno), ora (formato hh:mm:ss) e valore (é la tabella delle misure)
- tabella2_A_codice-area.tab per ogni area da considerare a 1 colonna di valori (é la tabella dei valori estratti dalla simulazione)

"Famiglia CONFRONTO":

Questa famiglia contiene un task per ogni analisi da eseguire (per ora sono implementati quelli relativi alla "regressione lineare", ai "diagrammi di Taylor" e al "test di Kolmogorov_Smirnov") e un task che crea un file ASCII per ogni confronto contenente le informazioni rilevanti ricavate dai test.

Nel caso sia richiesta anche l'analisi delle distribuzioni, vengono create una "famiglia confronto_serie" e una "famiglia confronto_distrib", se invece è richiesta solo l'analisi puntuale nelle stazioni vengono creati solo i task che sarebbero altrimenti contenuti nella famiglia confronto_serie:

"task t_lin_reg"

1. Crea la sotto directory di lavoro
2. Prepara e lancia un job-array tramite template ("../wrf_model/src/ NAUSICA_ver_lin_reg.tpl.job") il quale a sua volta crea per sostituzione da template.R un file.R che crea i plot di regressione lineare e gli altri associati a questo definiti nello script di R (in tutto sono 8 plot per ogni estensione) per ogni variabile di ogni stazione e un file txt contenente le informazioni statistiche ("../wrf_model/src/ NAUSICA_ver_Reg_lin.tpl.R").
3. Sposta a livello superiore della directory di lavoro i plot creati e li archivia anche nella directory di archivio plot all'interno di cartelle con nome del codice della famiglia di confronto insieme ai file contenenti la statistica ("/data/ver_plot/ "codice-famiglia/""")

"task t_taylor"

1. Crea la sotto directory di lavoro
2. Prepara e lancia un job-array tramite template ("../wrf_model/src/ NAUSICA_ver_taylor.tpl.job") il quale a sua volta crea un file.R che genera il diagramma di Taylor per ogni variabile di ogni stazione in ogni formato richiesto utilizzando l'eseguibile Taylor ("../wrf_model/src/ NAUSICA_ver_Taylor.R").
3. Sposta a livello superiore della directory di lavoro i diagrammi di Taylor creati e li archivia anche nella directory di archivio plot all'interno di cartelle con nome del codice della famiglia di confronto ("/data/ver_plot/ "codice-famiglia/""")

"task t_kolmogorov_smirnov_punto"

1. Crea la sotto directory di lavoro

2. Prepara e lancia un job-array tramite template ("`../wrf_model/src/ NAUSICA_ver_kol_smf.tpl.job`") che esegue il test di Kolmogorov-Smirnov per ogni serie temporale estratta nelle coordinate di stazione utilizzando l'eseguibile `kst2sample`, salvando i dati rilevanti in file "`Test_Kol_smf_codice-famiglia_data-inizio_data-fine_nome-stazione_nome-var.txt`". Inoltre Prepara e lancia un secondo job-array tramite template ("`../wrf_model/src/ NAUSICA_ver_kol_smf_plot.tpl.job`") il quale esegue i plot utilizzando "gnuplot" sia delle distribuzioni che delle differenze calcolate nel test.
3. Sposta a livello superiore della directory di lavoro i file txt e i plot creati e li archivia anche nella directory di archivio plot all'interno di cartelle con nome del codice della famiglia di confronto ("`/data/ver_plot/ "codice-famiglia/"`")

"task t_spettro"

1. Crea la sotto directory di lavoro
2. Prepara e lancia un job-array tramite template ("`../wrf_model/src/ NAUSICA_ver_spettro.tpl.job`") il quale a sua volta crea un file.R che genera gli spettri relativi ai dati e ai GRIB per ogni variabile di ogni stazione in ogni formato richiesto
3. Sposta a livello superiore della directory di lavoro gli spettri creati e li archivia anche nella directory di archivio plot all'interno di cartelle con nome del codice della famiglia di confronto ("`/data/ver_plot/ "codice-famiglia/"`")

OSSERVAZIONE: Per rendere lo spettro leggibile viene utilizzato il metodo di smoothing non parametrico "modified Daniell Kernel" con parametri $m=(2,2)$ (<https://onlinecourses.science.psu.edu/stat510/node/80>), attribuendo cioè ad ogni frequenza il valore calcolato facendo una media pesata delle densità spettrali nell'intervallo di frequenze $[f-4;f+4]$ che corrisponde ad un periodo di 4 ore. I coefficienti con cui vengono pesati i valori nella media sono i seguenti:

```
mDaniell(1,1)
-coef[-2] = 0.0625
-coef[-1] = 0.2500
-coef[ 0] = 0.3750
-coef[ 1] = 0.2500
-coef[ 2] = 0.0625
```

"task t_stat_info"

1. Per ogni stazione e variabile di ogni famiglia controlla che nei file .txt (file "`Stat_reg_lin.txt`" creati dal task `t_lin_reg` e file "`Test_Kol_smf.txt`" creati dal task `t_kolmogorov_smirnov_punto`) contenenti le informazioni rilevanti ricavate dai test ci siano le informazioni relative a stazione, variabile e periodo attese
2. Crea un file ASCII unico per ogni stazione, variabile e periodo con le informazioni da salvare e li archivia: "`info_stat.txt`"
3. Archivia i file prodotti nella directory di archivio plot all'interno di cartelle con nome del codice della famiglia di confronto ("`/lustre/arpa/goglioa/data/ver_plot/ "codice-famiglia/"`")

"task t_kolmogorov_smirnov_area"

1. Crea la sotto directory di lavoro

2. Prepara e lancia un job tramite template ("../wrf_model/src/ NAUSICA_ver_kol_smf_a.tpl.job") che esegue il test di Kolmogorov-Smirnov per ogni variabile utilizzando l'eseguibile kst2sample, salvando i dati rilevanti in file "Test_Kol_smf_A_codice-famiglia_data-inizio_data-fine_nome-stazione_nome-var.txt". Inoltre Prepara e lancia un secondo job tramite template ("../wrf_model/src/ NAUSICA_ver_kol_smf_a_plot.tpl.job") il quale esegue i plot utilizzando "gnuplot" sia delle distribuzioni che delle differenze calcolate nel test.

3. Sposta a livello superiore della directory di lavoro i file txt e i plot creati e li archivia anche nella directory di archivio plot all'interno di cartelle con nome del codice della famiglia di confronto ("/lustre/arpa/gogloia/data/ver_plot/ "codice-famiglia/"")

"task t_Astat_info"

1. Per ogni variabile di ogni famiglia crea un file.txt (file "sum_A.txt") contenente le informazioni statistiche sulla distribuzione tramite un job ("NAUSICA_ver_Astat_info.tpl.job") che lancia uno script di R ("NAUSICA_ver_A_stat_info.tpl.R") che legge i file creati nei due task di estrazione sull'area

2. Crea un file "info_stat_A.txt" per ogni variabile che contiene le informazioni rilevanti ricavate da "Test_Kol_smf.txt" creato dal task t_kolmogorov_smirnov_area e da "sum_A.txt" creato precedentemente all'interno del task stesso.

3. Archivia i file prodotti nella directory di archivio plot all'interno di cartelle con nome del codice della famiglia di confronto ("/data/ver_plot/ "codice-famiglia/"")

Famiglie di controllo

"Famiglia SENTINEL"

Questa è definita all'esterno delle singole famiglie di confronto.

"task tsen_run_cleaner"

Pulisce le cartelle di LOG (importante perché la ricerca di eventuali plot non eseguiti correttamente avviene all'interno di questi), e la cartella di lavoro generale delle famiglie indicate da eseguire nel file elenco confronti

"Famiglia SPY"

È definita una di queste all'interno di ogni singola famiglia di confronto.

"task tspy_STZ_TAB"

Crea delle tabelle CSV al fine di controllare le estrazioni fatte. In particolare calcola il numero di valori estratti attesi in base alle date e all'intervallo di estrazione e conta il numero di righe dei file confronto, che dovrebbero corrispondere al numero di valori estratti ottenuto dai task di estrazione.

"task tspy_STZ_TAB_HTML":

Visualizza le tabelle CSV create dal task precedente in html alla pagina http://grido.mercuriofv.it/*/tab_comb.html. La tabella visualizzata alla pagina indicata è relativa soltanto all'ultima famiglia di estrazione (vengono sovrascritte). Ogni valore relativo ad una stazione e ad una variabile compare in verde se il numero di righe atteso e calcolato coincidono, altrimenti in rosso.

"task tspy_controllo_plot":

Conteggia il numero di file prodotti senza errori nell'esecuzione di R (sommando tutti i tipi di plot, tutte le stazioni e tutte le variabili) e lo confronta con quello atteso (**ATTENZIONE** se uno script di R deve produrre più plot ma c'è un errore nella produzione anche di uno solo tra questi non ne viene prodotto nessuno). Il risultato è riportato in una label in modo da essere leggibile direttamente dall'ecflowview. Esegue inoltre l'identificazione degli eventuali plot non prodotti (tramite "grep" all'interno dei log).

"Famiglia LOG"

È definita una di queste all'interno di ogni singola famiglia di confronto.

"task tlog_TAR"

Copia l'output flusso, jobout e joberror di tutti i task in un file compresso.

Cancella inoltre i file di inizializzazione di famiglia creati dai template.

ATTENZIONE: la cartella relativa a tlog output va rimossa a mano alla fine delle simulazioni

G.2.4 Eseguibili e file accessori

Elenco di tutti i file accessori necessari alla suite (compresi quelli già citati e descritti):

- File elenco confronti
("../wrf_model/etc/ NAUSICA_ver_elenco_confronti.csv"): contiene tutte le informazioni necessarie all'esecuzione delle famiglie di confronto, come descritto in modo dettagliato nel paragrafo relativo.
- File di inizializzazione della suite ("NAUSICA_ver.ini"): definizione dei parametri generali della simulazione.
- Template dei file di inizializzazione delle singole famiglie da cui lo shell script di creazione genera un file per ogni famiglia: "../wrf_model/etc/ NAUSICA_ver_TPL_A.ini" per le analisi su punti ed aree (codice WW=01) e "../wrf_model/etc/ NAUSICA_ver_TPL_P.ini" per le analisi solo su punti di stazione (codice WW=02). I file generati dallo shell script sono "/scratch/ini_ver/NAUSICA_ver_"codice-famiglia"_"data-inizio"_"data-fine".ini." e contengono la definizione dei parametri relativi alle singole famiglie, in particolare l'elenco delle variabili da estrarre (vedere il paragrafo variabili confrontabili e i formati dei plot da produrre. Se si vogliono settare questi parametri in modo diverso per ogni famiglia basta editare questi file.
- File elenco stazioni OSMER ("NAUSICA_ver_elenco_stz_osmer.txt"): contiene l'elenco di tutte le stazioni OSMER con coordinate e H sul livello del mare
- Template del file di input per l'eseguibile "estrai"
("../wrf_model/etc/ NAUSICA_ver_STZ_input.dat"): contiene il template in vengono indicate dai task di estrazione le stazioni e variabili estrarre.
- File elenco vertici poligonale
("../wrf_model/etc/ NAUSICA_ver_AREA_"codice-area-famiglia".TPL.txt"): contiene le coordinate dei vertici della poligonale delimitante l'area che si desidera analizzare. Se è necessario definire una nuova area di estrazione basta crearne uno nuovo con un codice (codice XX del codice di famiglia) diverso da quelli esistenti.

Elenco degli eseguibili necessari alla suite:

- Shell script di creazione della suite
("ecf_script/ NAUSICA_ver_Suite_Ini_def.sh"): crea il file di definizione della suite leggendo dal file elenco confronti e crea i file di inizializzazione di famiglia dai template relativi.

OSSERVAZIONE: all'interno dello shell script di creazione della suite è anche possibile inserire "date" e "cron" per programmare il run delle singole famiglie. Dato che la regione è stata suddivisa in 7 zone meteorologicamente uniformi da analizzare, a partire dal giorno di inizio vengono eseguite 7 famiglie di run al giorno.

- "estrai" ("../wrf_model/src/ NAUSICA_ver_estrai"): esegue l'estrazione delle misure OSMER in base alle stazioni e alle variabili richieste nel file di input

- "poynt_in_polygon" ("`../wrf_model/src/ NAUSICA_ver_point_in_polygon`"): seleziona le stazioni OSMER contenute all'interno dell'area delimitata dalla poligonale descritta nel file elenco vertici poligonale
- "f_serie_temp" ("`../wrf_model/src/ NAUSICA_ver_f_serie_temp`"): cambia le unità di misura quando necessario per il confronto (per le var TEMP e PRES) e ricava le grandezze non direttamente disponibili per il confronto (le componenti dei venti U e V nel caso delle misure e l'intensità del vento nel caso dei GRIB)
- "kstzsample" ("`../wrf_model/src/ NAUSICA_ver_kstzsample`"): esegue il test di Kolmogorov_Smirnov su una coppia file di campioni
- "extraction" ("`../wrf_model/src/ NAUSICA_ver_extraction`"): estrattore che utilizza le GRIB API per estrarre, interpolare e fare selezioni spaziali in base ad un GRIB maschera.

H

Struttura ed utilizzo delle suite

NAUSICA_rds e NAUSICA_srds

La suite esegue un confronto statistico tra misure dei radiosondaggi e campi estratti dal modello lungo la verticale in corrispondenza di aree centrate sui livelli verticali del modello, archiviando informazioni statistiche (medie, mediane, percentili, valori massimi e minimi, bias, test di ipotesi Wilcoxon) e test di Kolmogorov-Smirnov. Il metodo utilizzato per l'estrazione dei campioni da confrontare, le variabili disponibili e i tipi di analisi eseguite sono descritti nel paragrafo relativo (Par. 4.4)

La struttura della suite e la classificazione dei confronti seguono la metodologia utilizzata per la suite NAUSICA_ver (e NAUSICA_nc).

Il confronto viene eseguito

- per ogni stazione di lancio (elencate in un file di elenco stazioni)
- per ogni livello del modello (scelti da file di inizializzazione)
- per ogni variabile (scelte da file di inizializzazione)
- per ogni intervallo temporale, anche formato da intervalli disgiunti (scelto da file csv)
- per ogni scadenza di lancio (indicata da file csv)

Codici file csv elenco confronti

Questo file è utilizzato per creare il file di definizione della suite da uno shell-script e deve essere della forma:

```
f;XXYYZZWW;DIi,...DIj;DFi,...,DFf;descrizione estesa
```

dove:

- f = flag 0,1 di attivazione del confronto
- XXYYZZWW = codice confronto:
 - XX = stazione di lancio del radiosondaggio
 - YY = run del WRF da cui si estrae
 - ZZ = numero del run
 - WW = scadenza del lancio (hh del lancio)
- DI_j = date di inizio dei singoli intervalli temporali da considerare
- DF_j = date di fine dei singoli intervalli temporali da considerare

Ogni singola famiglia di confronto avrà quindi nome:

NAUSICA_rds_XXYYZZWW_DI_DFf. Le date indicano gli estremi dell'intervallo temporale, l'informazione è già contenuta nel codice ma esplicitata in questo modo aumenta la leggibilità dei nomi delle famiglie e dei prodotti archiviati.

File inizializzazione di famiglia

Nei file di inizializzazione, che saranno creati dallo shell script di creazione della suite a partire da un template, andranno indicati:

- I livelli del modello sui quali vogliamo eseguire il confronto (quanti e quali)
- La larghezza Delta_z del range verticale di altezze dei radiosondaggi da confrontare per singolo livello del modello con i dati estratti dalle simulazioni
- Il numero di punti di griglia (sulle superfici a z costante) da considerare nella statistica relativa al campione estratto dal modello per singolo livello, che come calcolato deve essere $\geq 3 * 3$

H.o.1 Descrizione dei compiti dei singoli task

Famiglia SENTINEL

- Task tsen_run_cleaner: Pulisce le cartelle dei log e di lavoro

Famiglia ESTRAZIONE

- Task t_mod_lev: Legge l'altezza dei livelli del modello Z_lev in corrispondenza delle coordinate della stazione di lancio del radiosondaggio
- Task t_ext_rds: Estrae per tutte le scadenze contenute nell'intervallo temporale indicato (parallelamente per ciascun livello del modello) i valori delle variabili misurate dalla sonda nel range di altezze indicato
- Task t_ext_mod: Estrae per tutte le scadenze contenute nell'intervallo temporale indicato (parallelamente per ogni livello del modello) i valori dei campi per ogni punto di griglia appartenente alla superficie a z costante di superficie indicata

Famiglia CONFRONTO

- Task t_kol_smv: Esegue i test di Kolmogorov-Smirnov e salva i plot in file PNG e PS
- Task t_info_stat: Calcola medie, mediane, percentili, valori massimi e minimi, bias, test di ipotesi Wilcoxon e li salva in un file txt

Famiglia SPY

- Task tspy_controllo: Controlla che i plot e i file prodotti siano quelli attesi

Famiglia LOG

- Task tlog_TAR: Archivia log e prodotti della suite

Script accessori:

- "NAUSICA_rds_bias_plot_Perc.job": Grafica il bias, differenza tra venticinquesimo (in alto) e settantacinquesimo (in basso) percentile tra distribuzione relativa al modello e distribuzione relativa alle misure per la temperatura in corrispondenza di ogni livello del modello
- "NAUSICA_rds_bias_plot_vert.job": Grafica il profilo verticale medio della temperatura alle 12:00 UTC e alle 24:00 UTC per modello e misure dei radiosondaggi
- "NAUSICA_rds_bias_grad.job": Confronta i gradienti medi di temperatura misurati e calcolati per la alle 12 e 24 UTC per ciascun livello del modello sulla verticale nel punto di lancio del radiosondaggio

NAUSICA_srds

Questa suite è analoga alla precedente ma, invece di fare statistica sui periodi interi (stagioni), la fa giorno per giorno. Si ottengono quindi dei plot analoghi a quelli di NAUSICA_rds con un punto per ogni giornata invece del valor medio. Un esempio di output è riportato in Figura H.0.1, lo script accessorio che esegue il plot in questo caso è "NAUSICA_srds_bias_boxplot.job".

Questa analisi NON è valida sia perché affetta dagli stessi problemi della suite NAUSICA_rds che perché non tiene conto del fatto che nell'output delle simulazioni viene scritto il valore istantaneo per i campi: confrontare questo con quello letto dalle misure e associato ad un istante e attendersi un bias nullo risulta troppo stringente tenendo conto che l'intervallo di scrittura temporale del modello sul dominio regionale è di 1 ora.

Si sta infatti pretendendo che la simulazione abbia risoluzione temporale tendente a 0!

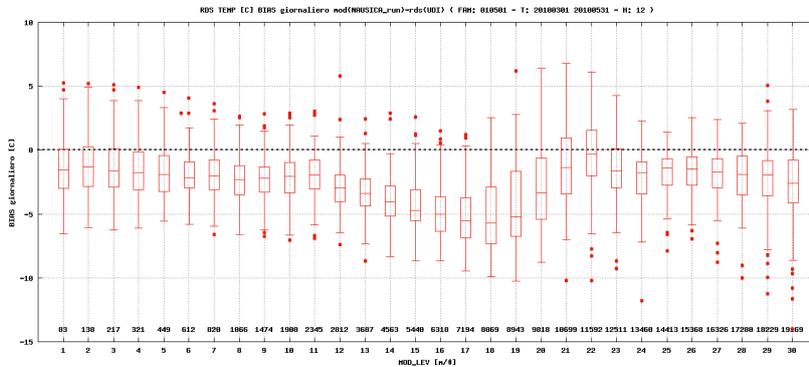


Figura H.1: Esempio di analisi della suite NAUSICA_srds (ERRATA nel metodo come spiegato nel Paragrafo relativo). Rappresenta il risultato del calcolo giornaliero dei bias per ogni livello del modello per la primavera del 2010 a Rivolto alle 12:00 UTC.

I

Grandezze cumulate e differenziali

I.0.1 Problema relativo alle grandezze cumulate

Durante la simulazione della famiglia di run 01-15/01/2015 è stato riscontrato un problema nella creazione dei GRIB di entrambi i tipi (PRS_LEV e MOD_LEV) successivi alla scadenza 11/01/2015, ore 18 comprese, su tutti i domini. Questi vengono creati da entrambi i task senza nessuna segnalazione particolare nè su log.out nè su log.err ma non vengono letti dalle CDO durante l'esecuzione del task SCAN.

Analizzando questo run (ripetendo anche l'esecuzione dei task UPP) si è osservato che

- I wrfout vengono tutti letti correttamente anche dalle CDO
- I GRIB segnalati vengono letti correttamente dal wgrib ma non dalle cdo
- Splittando i GRIB segnalati nelle parti relative alle singole variabili vengono letti anche dalle CDO
- Si osserva però che gli intervalli temporali relativi alle grandezze cumulate (piogge e neve) nei GRIB successivi al problema di post-processing sono diversi da quelli delle altre variabili (che riportano correttamente l'intervallo dall'inizio della simulazione). In particolare l'intervallo delle grandezze cumulate oltre un certo limite si azzerava.
- Ripetendo tutta la simulazione dall'ultima data con produzione corretta di GRIB, tramite un run di un unico giorno i GRIB ottenuti non presentano più il problema segnalato.
- Ripetendo una simulazione di 15 giorni analoga alla precedente su un periodo successivo alla scadenza problematica (dal 12 al 26) si è ripresentato lo stesso problema a partire dalle ore 18 del giorno 21/01/2015.
- Stessa cosa del caso precedente per una simulazione del solo giorno 21/01/2016: è presente il problema sempre a partire dalle ore 18 del giorno 21/01/2015.
- Tagliando record per record i GRIB problematici si osserva che il problema è legato alle variabili relative alle grandezze cumulate (neve e pioggia) che risultano accumulate su sole 2 ore invece che sull'intervallo totale simulato
- Analizzando i wrfout con ncdump si vede che su ogni dominio è settato di default un limite massimo per grandezze cumulate (nel nostro caso si è osservato che questo era di 100 mm sul do1 e 300 mm sul do3) oltre al quale il totale viene automaticamente azzerato
- Dalla documentazione risulta che per eliminare la soglia sul valore è necessario aggiungere alla namelist del WRF `bucket_mm=-1` e `bucket_J = -1`
- Eseguendo una simulazione di 15 giorni con partenza il 01/01/2015 aggiungendo `bucket_mm=-1` e `bucket_J=-1` alla namelist del WRF risulta risolto il problema sui wrfout: le grandezze cumulate non presentano più

un limite massimo oltre il quale vengono resettate a zero ma persiste il problema sul post-processing che si è interrotto sempre in corrispondenza delle ore 18 del 11/01/2015.

Soluzione

I problemi riscontrati sono stati risolti come segue:

1. Problema su limite massimo delle grandezze cumulate: per evitare l'imposizione di un limite massimo oltre il quale queste grandezze vengono resettate a zero è necessario aggiungere alla namelist del WRF `bucket_mmm=-1` e `bucket_J=-1`
2. Problema limite UPP 2.2: il fatto che UPP V2.2 scriva che l'intervallo di integrazione delle variabili cumulate sia sempre compreso tra 0 e 255 h è implicito nella destinazione d'uso data ai byte 19 e 20 della codifica GRIB1. Il byte 21 definisce cosa stanno ad indicare i due precedenti e nel caso il codice sia 4, ovvero integrazione, il valore del byte 19 (sezione 1) sta per inizio e il 20 (sezione 1) per fine periodo. Per questo motivo, i campi non integrati, che hanno byte 21 (sezione 1) settato a 10 (che sta per: il byte 19 e 20 compongono il forecast time dall'inizio della simulazione), riescono ad andare con i tempi fino a 256×256 unità di tempo, (ore per le simulazioni). Per esempio nel caso del run a 384hr gli accumulati indicano 0-128 hr. Questo è voluto esplicitamente nella routine di UPPV2.2 che scrive i GRIB1, la: `UPPV2.2/src/unipost/GRIBIT.f`
OSSERVAZIONE: NON utilizzare UPP V3.0 perché, a differenza della versione 2.2, in questa non è possibile eliminare il problema
3. Il problema creato dagli applicativi CDO nell'estrazione di informazioni da record GRIB aventi byte 21 della sezione PDS diverso, quando sono concatenati in un unico file, è stato risolto costruendo un wrapper da applicare alle chiamate CDO presenti negli script e job templates utilizzati per l'estrazione e l'analisi dei dati.

I.0.2 Calcolo delle grandezze differenziali a partire dalle cumulate

Essendo interessati a confrontare direttamente le grandezze differenziali e non le cumulate è stato inizialmente creato un programma FORTRAN [16] che per mezzo degli applicativi GRIB API [26] crea dei record GRIB differenziali a partire da campi cumulati: `acc2dif.exe`. Questo veniva utilizzato tramite un job per creare file GRIB con prefisso `PREC` nelle cartelle relative alle scadenze considerate che venivano utilizzati dalla suite di verifica per leggere i campi da confrontare con le misure. Per settare come missing i record GRIB relativi alla scadenza iniziale (non avendo a disposizione il valore precedente per calcolare la differenza) è stato creato anche l'eseguibile `miss.exe`. Si riportano le caratteristiche e l'help dell'eseguibile `acc2dif.exe`:

- Eseguibile `f90` (formato standard) che permette, a partire da record GRIB di tipo NAUSICA contenenti campi integrati, di creare nuovi record GRIB contenenti il rispettivo campo orario.
- I GRIB in input devono essere di tipo Nausica e contenere un unico record per variabile (altrimenti viene utilizzato solo il primo).
- Devono essere forniti in input anche i 4 codici necessari all'identificazione univoca del campo scelto (per la precipitazione: `var=61, tyl=1, lev=0, tytr=4`)
- E' possibile (opzionale) fornire anche nuovi valori per settare le chiavi del nuovo GRIB (data e ora di riferimento, i 4 parametri della variabile, i due byte contenenti il tempo di forecast). Se si decide di modificarne almeno uno vanno però indicati anche i valori di tutti gli altri.

- Walltime: circa 10 secondi per mese

-----PROGRAM ACC2DIF V 0.1-----

This program builds a differential field
between two NAUSICA GRIB records

Available options:

```

-newf      new record GRIB
-oldf      old record GRIB
-diff      dif file.GRIB

-var       variable name
-tyl       variable type of level
-lev       variable level
-tytr      variable time range type

-refdate   new ref date value      (optional)
-reftime   new ref time value      (optional)
-ovar      new var name value      (optional)
-otyl      new var type of level value (optional)
-otytr     new var level value      (optional)
-otytr     new time range type value (optional)
-time1     new first time byte value (optional)
-time2     new second time byte value (optional)

-v         verbose option

-help      program help

```

EXIT STATUS:

```

EXIT STATUS = 0   OK
EXIT STATUS = 1   Input error
EXIT STATUS = 2   GRIB error
EXIT STATUS = 3   F90 error
EXIT STATUS = 4   Mask file error
EXIT STATUS = 5   GRIB API error

```

Help dell'eseguibile miss.exe:

-----PROGRAM MISS V 0.1-----

This program sets a complete field as missing
in NAUSICA GRIB files with one record per var

Available options:

```

-f         input file.GRIB
-out       output file.GRIB

-var       variable name
-tyl       variable type of level
-lev       variable level
-tytr      variable time range type
-refdate   record date
-reftime   record time
-otrange   record time range type

```

```

-time1      record first time byte
-time2      record second time byte

-v          verbose option

-help       program help

```

EXIT STATUS:

```

EXIT STATUS = 0    OK
EXIT STATUS = 1    Input error
EXIT STATUS = 2    GRIB error
EXIT STATUS = 3    F90 error
EXIT STATUS = 4    Mask file error
EXIT STATUS = 5    GRIB API error

```

I.o.3 Introduzione della produzione diretta di grandezze differenziali

Per semplificare le operazioni sono state successivamente introdotte le modifiche nella namelist del WRF e nel file cntr.parm di UPP per produrre direttamente i campi differenziali al posto di quelli cumulati.

- Modifiche nella namelist del WRF: aggiunta in &physics di `prec_acc_dt = 360,180,60` Questa aggiunta produce i seguenti nuovi campi nei wrfout : `prec_acc_c`, `prec_acc_nc` e `snow_acc_nc`.
- Sostituzioni nel file cntrl.parm di UPP:
 - acm total precip sostituito con BUCKET TOTAL PRECIP
 - acm convctiv precip sostituito con BUCKET CONV PRECIP

Come osservabile in Figura 3.8 i due metodi, diretto ed indiretto, per il calcolo delle grandezze differenziali producono risultati coerenti tra loro e con le misure.

J

Conversione di coordinate

J.o.1 Problema relativo alla conversione in coordinate lat/lon

Durante il post-processing dei netCDF relativi al 2015 sono stati riscontrati problemi nella creazione dei GRIB sia sui livelli del modello (per ognuna delle famiglie mensili del 2015, su scadenze e domini diversi) che sui livelli isobarici (da Novembre 2015).

Errore nell'esecuzione del *copygb* (UPP), esempio di errore restituito

```
Floating point exception(core dumped)
/lustre/arpa/./wrf3.7.1/UPPV2.2/bin/copygb.exe -xg"\$PRO\_DETAIL" \$GRB\_F \$TMP\_GRB\_F'
```

Analizzando i GRIB si è ricavato che

- L'errore nell'esecuzione del *copygb* è sempre relativo a record dei GRIB contenenti il campo della DEW POINT TEMPERATURE (calcolata da UPP), ma in corrispondenza di diversi livelli.
- Analizzando i valori max/min del campo con le CDO questo risulta essere costante su gran parte del dominio (in particolare di valore 192, 125)
- Visualizzando con GRADS il campo si osserva che diverge in un punto della griglia (per es. nel caso specifico osservato: *mod_lev* 20150905_06 su *do1 lev 26 var17* diverge nell'intervallo di coordinate lon/lat -1W - 0.5W/37.4N37.7N)

Soluzione Il problema è stato risolto eliminando dal file dei parametri di UPP la DEW POINT TEMPERATURE e rieseguendo il post-processing a partire dalla prima famiglia con scadenza che presentava il problema.

J.o.2 Problema generale relativo alla conversione in coordinate lat/lon (*copygb.exe*)

Il problema riscontrato per la Dew Point Temperature si è ripresentato anche per altri campi. Questo è risultato essere causato dall'eseguibile *copygb.exe* nella conversione in coordinate Lat/Lon.

Soluzione

Si è quindi deciso di reintrodurre la produzione della Dew Point Temperature ma di post processare gli output del modello in Lambert Conformal (operazione che riduce anche i tempi di produzione dei GRIB).

Elenco delle figure

- 1.1 Nausicaa incontra Odisseo. 9
- 1.2 Dominio delle simulazioni ad alta risoluzione. 10

- 2.1 Schema a blocchi di WPS e WRF, sono riportati solo gli elementi coinvolti nell'ambito del progetto NAUSICA. 14
- 2.2 Esempio dello stato di esecuzione, visualizzato tramite *ecflowview*, di alcune *suite ecFlow* relative al progetto. Anche senza entrare nei dettagli del funzionamento del software si nota che in questo run c'è stato un problema nel flusso all'interno del nodo di colore rosso, i nodi gialli sono quelli la cui esecuzione è terminata, quelli grigi sono stati caricati ma non mandati in esecuzione, quelli azzurri aspettano il segnale del trigger per attivarsi e quelli arancioni sono stati sospesi. Si possono inoltre osservare indicatori di stato in percentuale di un processo (GRIB in figura) o indicatori che si accendono al verificarsi di una determinata condizione (ASSENZA netCDF e ASSENZA GRIB in figura rispettivamente acceso e spento), definiti *meter* ed *event*. 17
- 2.3 Diagramma di GANTT relativo allo stato di avanzamento di NAUSICA a Dicembre 2016. Si possono notare ad esempio l'interruzione delle attività avvenuta ad Ottobre per problemi contrattuali e la successiva ripresa a fine mese solo di alcune sezioni del progetto a causa della mancanza di risorse disponibili per il completamento di tutte le attività, prima tra queste la produzione del database. 18
- 2.4 Esempio di utilizzo del repository GIT per confrontare una parte di codice tra due versioni di questo. Come si può osservare questo strumento è caratterizzato da un'ottima leggibilità delle modifiche. 18

- 3.1 Scalabilità del modello WRF relativa alla coda *hp* eseguita con la configurazione operativa attuale, ovvero prima dell'estensione del dominio regionale descritta nel Paragrafo 3.2 e delle modifiche ai parametri descritte nel Paragrafo 3.3. Si può osservare che il modello scala bene all'aumentare delle risorse di calcolo utilizzate, ovvero che il *wall-time* è inversamente proporzionale al numero di nodi di calcolo utilizzati. 22
- 3.2 Ore di simulazione all'aumentare del numero di giorni simulato con il WRF utilizzando 30 processori per 6 nodi. Si nota stabilità nei tempi di produzione. 23

- 3.3 Confini spaziali dei tre domini utilizzati dai flussi operativi (prima dell'estensione del dominio regionale descritta nel Par. 3.2) 24
- 3.4 Fasi di estensione del dominio regionale. La linea blu corrisponde all'area attualmente utilizzata nelle simulazioni operative, la linea verde al primo tentativo di estensione e la linea rossa al dominio scelto. Le caratteristiche dei tre domini sono riassunte in Tabella 3.3. 25
- 3.5 Campi di precipitazione cumulata (in alto) e vento a 10 metri (in basso) relativi al dominio nazionale do2. Si nota come, nel caso della precipitazione, i campi del database relativo al dominio do2 contengano un'evidente discontinuità in corrispondenza dell'area del dominio interno do3. Questo è dovuto all'opzione del WRF di *feedback* e rende necessaria estrema cautela nell'utilizzo dei campi generati dal modello per quanto riguarda i domini di raccordo. 27
- 3.6 Descrizione del contenuto di alcuni record di un GRIB relativo ad un'unica scadenza (ottenuto tramite *wgrib*) affetto dal problema causato dall'insufficiente numero di ottetti destinati da UPP alla scrittura dell'intervallo temporale di integrazione. Si nota infatti che l'intervallo attribuito alle grandezze cumulate (2014 h) è diverso dall'indicazione del tempo di forecast (4812 h) di tutte le altre. 28
- 3.7 Serie temporale della pioggia cumulata [mm] in unità di *time-step* corrispondenti a 6 ore per unità letta dai file *wrfout.nc* nel caso in cui il limite massimo del valore di questa sia quello di default del WRF (100 mm per il do1). Si osserva infatti che in corrispondenza del superamento di questa soglia, indicato dalla freccia nera, il valore viene resettato a 0. 29
- 3.8 Confronto qualitativo tra le serie della precipitazione differenziale in Gennaio 2016 delle misure (linea verde), dei campi differenziali ottenuti tramite l'eseguibile *acc2diff.exe* a partire dai GRIB contenenti i valori cumulati (linea blu) e del campo differenziale prodotto direttamente dal WRF (linea rossa). Si osserva un accordo ritenuto soddisfacente. 30
- 3.9 Serie temporale della pioggia cumulata prodotta dal WRF per il do1 e descrizione del contenuto di alcuni record GRIB derivanti dal post-processamento dei *wrfout* per il do3. Si osserva che come atteso le grandezze relative alla precipitazione sono differenziali. 30

- 3.10 Confronto tra gli istogrammi relativi alla temperatura a due metri dal suolo calcolata dal modello e misurata in corrispondenza della stazione ARPA-OSMER di Trieste rispettivamente nella prima settimana di Agosto con la configurazione attualmente utilizzata in operatività: modello di suolo TDS in assenza di aggiornamento SST (sinistra) e nei primi 15 giorni di Agosto con la configurazione scelta per i run del modello NAUSICA: modello di suolo Noah LSM e aggiornamento SST ogni 6 ore (destra). Nel primo caso il modello sottostima la temperatura media di circa 17 °C, inoltre la temperatura massima raggiunta dal modello è inferiore di circa 5°C rispetto alla minima misurata mentre con la seconda parametrizzazione utilizzata si è ottenuta una simulazione con valori delle temperature verosimili. La frequenza che compare in ordinata corrisponde al numero di occorrenze di valori appartenenti al range della classe dell'istogramma tra le scadenze orarie disponibili sul periodo selezionato. 32
- 3.11 Schema del *Noah Land Surface Model*: il modello di suolo utilizzato per le simulazioni del progetto. La linea rossa corrisponde alla profondità massima raggiunta dallo schema precedentemente utilizzato: il *Thermal Diffusion Scheme* che, a differenza del Noah LSM, considerava unicamente la conduzione di calore tra i diversi livelli. 33
- 3.12 Serie temporali delle temperature in °C del suolo nei GRIB di analisi ECMWF (fonte delle condizioni iniziali ed al contorno della simulazione) alle quattro profondità descritte (fasce di profondità: 0-7 cm, 7-28 cm, 28-100cm, 100-255cm) relative all'intero 2015. Come atteso l'ampiezza di oscillazione e la frequenza di variazione della temperatura aumentano avvicinandosi alla superficie. 33
- 3.13 Illustrazione del processo di aggiornamento delle SST nello schema del WRF, gli elementi evidenziati in verde sono quelli coinvolti nel processo di aggiornamento. 34
- 3.14 Effetto di un suolo caratterizzato da diversi gradi di umidità. Data la differenza tra i flussi di calore nelle diverse condizioni risulta comprensibile la necessità di utilizzare un modello di suolo, come il *Noah LSM model*, che introduca il campo relativo all'umidità di questo. 35

- 3.15 Categorizzazione in classi di run in base alla durata del periodo simulato. A seconda della classe di appartenenza è indicata una specifica configurazione parametrica che include l'aggiornamento dei campi necessari alla corretta descrizione dell'evoluzione temporale sul periodo scelto. Il periodo scelto per la produzione di NAUSICA attualmente disponibile è di un anno, linea blu, per il quale è consigliato includere umidità del suolo, copertura nevosa, ghiaccio marino e frazione stagionale di vegetazione oltre all'aggiornamento delle SST. La variazione della temperatura dello strato più profondo di suolo non risulta necessaria per questo tipo di simulazioni ed effettivamente, come verrà descritto nel paragrafo relativo, non si sono osservati effetti sui valori dei campi prodotti attivando questa opzione. 36
- 3.16 In alto: confronto tra le distribuzioni empiriche relative alla temperatura a due metri della zona di costa del Friuli Venezia Giulia nella prima settimana di Agosto 2015 costruiti considerando la misure e nel primo caso, grafico a sinistra, un run di restart inizializzato il primo Gennaio 2015 mentre nel secondo caso, grafico a destra, un run inizializzato il primo Agosto 2015 (nei run si è utilizzato il TDS). Si nota che in entrambi la forma della distribuzione generata dal modello non è corretta ma solo nel secondo il range di temperature simulate è verosimile. È quindi confermato che processo di restart (lettura dei *wrfrst* affetti da errore dovuto al modello di suolo) causa la traslazione della distribuzione delle temperature verso il basso. In basso: effetto sulle temperature legato al TDS e dovuto al processo di restart negli output operativi. Si nota che in corrispondenza del restart quotidiano la temperatura minima viene inizializzata a valori molto inferiori rispetto a quelli calcolati dall'evoluzione ma che l'effetto viene corretto in tempi brevi e non costituisce quindi un problema per simulazioni relative ad intervalli della durata di qualche giorno. 38
- 3.17 Serie temporale della temperatura a due metri nel mese di Agosto in corrispondenza delle coordinate di Udine con schemi di suolo TDS (verde) e Noah LSM (viola). Si nota che nel secondo caso le temperature riprodotte dalla simulazioni sono verosimili. 39
- 3.18 Confronto tra gli istogrammi relativi alla temperatura a due metri dal suolo calcolata dal modello e misurata in corrispondenza della stazione ARPA-OSMER di Udine nel mese di Luglio in assenza di aggiornamento SST rispettivamente con modello di suolo TDS (sinistra) e con modello di suolo Noah LSM (destra). Si nota che la parametrizzazione di suolo Noah LSM risolve l'elevata differenza nelle temperature osservata nel caso di schema TDS. La frequenza che compare in ordinata corrisponde al numero di occorrenze di valori appartenenti al range della classe dell'istogramma tra le scadenze orarie disponibili sul periodo selezionato. 39

3.19 Confronto tra gli istogrammi relativi alla temperatura a due metri dal suolo calcolata dal modello con schema di suolo Noah LSM e misurata in corrispondenza della stazione ARPA-OSMER BOA nel mese di Maggio rispettivamente in assenza di aggiornamento SST (destra) e con aggiornamento delle SST ogni 6 ore (sinistra). Si nota che attivando l'aggiornamento delle SST come descritto nei precedenti paragrafi si è ottenuta una configurazione del modello che descrive in modo soddisfacente i dati. La frequenza che compare in ordinata corrisponde al numero di occorrenze di valori appartenenti al range della classe dell'istogramma tra le scadenze orarie disponibili sul periodo selezionato. 40

3.20 Campi di SKINTEMP (sinistra) ed SST (destra) nelle simulazioni NAUSICA. Il primo è ben definito anche se si nota che i valori sono interpolati a partire da condizioni al contorno a bassa risoluzione ("cuneo" in zona golfo di Trieste) e si intuisce la presenza di una discontinuità in corrispondenza della superficie su cui sono lette le SST (rettangolo a sud del dominio). Per il campo SST, entrambi i difetti sono molto marcati ed inoltre si osserva che non è stata settata la prevaricazione della terra sul mare dove presenti entrambi (parte dell'Istria scompare). In entrambe le figure la temperatura è espressa in Kelvin. 41

3.21 Campi di SKINTEMP (sinistra) ed SST (destra) nelle simulazioni eseguite per la tesi di Pierluigi Masai (WRF 3.9, WPS 3.9.1). In queste, oltre ad essere stata usata la risoluzione più alta (0.1° invece di 0.5°) per le condizioni al contorno della SST, è stato mascherato il mare per la definizione della SKINTEMP e la terra per la definizione del campo SST. Inoltre è stata attivata l'opzione per far prevaricare land rispetto a sea dove entrambi sono definiti. 41

3.22 Confronto tra le serie temporali relative alla temperatura superficiale in presenza (linea viola) e assenza (linea gialla) dell'aggiornamento della temperatura profonda del suolo (TLAG) in configurazione TDS senza aggiornamento SST. Si nota che l'utilizzo di questo parametro non influenza i campi al di sopra del livello del suolo. 42

- 3.23 Evoluzione della temperatura dello strato di suolo inferiore al variare di più parametri: le linee viola e azzurra sovrapposte descrivono il valore della variabile in configurazione di suolo TDS rispettivamente prima (Old WRF) e dopo (New WRF) la correzione del bug al WRF, la quale doveva riguardare infatti soltanto l'utilizzo del `tmn_update` che in questi casi era disattivato. La linea gialla descrive nella prima parte il valore in configurazione Noah LSM con `tmn_update` disattivata: si osserva che, rispetto a tutti gli altri casi in TDS, la temperatura dello strato profondo è di circa dieci gradi più elevata. La seconda e la terza fase descritte dalla linea gialla rappresentano l'evoluzione temporale in caso di modello TDS con attivazione di `tmn_update` ma tempi di modulazione calcolati sul valore della skin temperature su periodi differenti (`lagday` [giorni] rispettivamente di 5 e 150 giorni) ottenute dopo la correzione del bug. 43
- 3.24 Schematizzazione del fenomeno di *upwelling* descrivibile come il rimescolamento delle acque causato dall'interazione della superficie marina con venti da terra molto intensi. Nel periodo estivo, essendo maggiore il gradiente di temperatura dell'acqua, il fenomeno ha un effetto più importante sulla diminuzione delle temperature superficiali e di conseguenza dell'intera zona di costa. 44
- 3.25 Localizzazione stazioni ARPA OSMER di Trieste (lat/lon 45.651;13.753) e Lignano (lat/lon 45.703;13.148) 45
- 3.26 Confronto tra il valore della temperatura superficiale del mare ricavato dall'interpolazione del campo SST dei GRIB ECMWF nelle coordinate relative alla stazione di Trieste nei GRIB ECMWF e i valori misurati presso la centralina ARPA di Trieste nell'anno 2015. Le frecce indicano alcuni episodi di discostamento GRIB/dati riconducibili al fenomeno di *upwelling* (Fig. 3.24) causato da raffiche di Bora. I due episodi evidenziati in rosso sono analizzati in dettaglio in Fig. 3.27 (sinistra) 46
- 3.27 Confronto tra il valore della temperatura superficiale del mare ricavato dall'interpolazione del campo SST nelle coordinate relative al punto di griglia più vicino alla stazione di Trieste nei GRIB ECMWF e i valori misurati presso la centralina affiancati da velocità e direzione del vento misurati in Luglio 2015 (sinistra) e Novembre 2015 (destra). Le zone evidenziate corrispondono ad episodi di Bora, in particolare quelle evidenziate in rosso corrispondono agli stessi episodi evidenziati in Fig. 3.4 47
- 3.28 Confronto tra il valore della temperatura superficiale del mare ricavato dall'interpolazione del campo SST nelle coordinate relative alla stazione di Lignano nei GRIB ECMWF e i valori misurati presso la centralina nell'anno 2015. 48

- 3.29 Confronto tra il valore della temperatura superficiale del mare ricavato dall'interpolazione del campo SST nelle coordinate relative alla stazione di Lignano nei GRIB ECMWF e i valori misurati presso la centralina affiancati da velocità e direzione del vento misurati relativi a Luglio 2015. Le zone evidenziate sono episodi di Bora corrispondenti agli stessi episodi evidenziati in Fig. 3.4 e Fig. 3.27 A relativi alla stazione di Trieste. 49
- 3.30 Plot di regressione lineare relativo al campo SST in GRIB ECMWF e misure ARPA per le stazioni di Trieste (in alto) e Lignano (in basso) 50
- 3.31 Diagramma di Taylor per la valutazione della stima delle misure ARPA da parte dei GRIB ECMWF relativamente alla SST presso le stazioni di Trieste (in alto) e Lignano (in basso) 51
- 4.1 Esempio di settaggio dei parametri di inizializzazione delle suite NAUSICA_run e NAUSICA_pp (per i dettagli sui singoli parametri fare riferimento alla descrizione in Appendice. 54
- 4.2 Schema della struttura logico/computazionale della suite NAUSICA_run (e NAUSICA_pp) i cui dettagli sulla struttura e l'utilizzo sono esaustivamente riportati in Appendice. 55
- 4.3 Esempio di pagina html per il controllo della produzione dei CRMA GRIB tramite suite NAUSICA_run o NAUSICA_pp sui livelli isobarici. Nella tabella superiore si legge lo stato delle scadenze attese (presenza indicata dal colore verde, assenza dal colore rosso). Nella tabella inferiore vengono riportati data, acronimo, nome, livello, tipo di livello, massimo, minimo e media di ciascun campo contenuto nei file analizzati. 57
- 4.4 Schema della struttura logico-computazionale della suite WRF_ana_post i cui dettagli sulla struttura e l'utilizzo sono esaustivamente riportati in Appendice. 58
- 4.5 Esempio di plot di controllo della produzione di GRIB della suite WRF_ana_post. Per alcune grandezze scelte da file di inizializzazione sono graficati i valori massimo (linea rossa) e minimo (linea blu) letti nelle varie scadenze dei file GRIB. 58
- 4.6 Esempio di attivazione della suite NAUSICA_ver tramite file csv. Ogni riga del file contiene le informazioni necessarie ad un run della suite codificate secondo la convenzione descritta in Appendice. 60
- 4.7 Schema della struttura logico/computazionale della suite NAUSICA_ver i cui dettagli sulla struttura e l'utilizzo sono esaustivamente riportati in Appendice. 61

- 4.8 Esempi di alcuni plot prodotti dei test statistici di confronto tra misure e simulazioni eseguiti. In particolare in (A) è riportato il plot di regressione lineare tra le serie temporali relative alla pressione a livello di stazione nelle coordinate corrispondenti alla stazione ARPA di Grado nel mese di Gennaio 2015. In (B) sono mostrati gli istogrammi relativi alle serie temporali della componente $u(EW)$ del vento presso la stazione ARPA di Crodriipo nel mese di Gennaio 2015. La frequenza che compare in ordinata corrisponde al numero di occorrenze di valori appartenenti al range della classe dell'istogramma tra le scadenze orarie disponibili sul periodo selezionato. La (C) è il risultato del test di Kolmogorov-Smirnov applicato alle distribuzioni empiriche relative all'area di media pianura, definita nel Paragrafo 4.3.3, per la variabile wind $v(NS)$ in gennaio 2015. Il plot in (D) contiene il confronto tra gli spettri della radiazione visibile relativi alle serie temporali nel punto di stazione di Udine in Luglio 2015. In (E) c'è invece un esempio di diagramma di Taylor (stazione di Trieste, pressione a livello di stazione, Luglio 2015). 62
- 4.9 Effetto del taglio dei valori nulli sulle distribuzioni empiriche per precipitazione oraria e radiazione e del taglio al di sopra della soglia del 95% per l'umidità relativa. 65
- 4.10 Rappresentazione delle aree meteorologicamente omogenee definite per l'analisi delle distribuzioni. 66
- 4.11 Esempio di tabella html per il controllo della presenza del numero di scadenze attese per l'estrazione di misure e campi dai GRIB nei punti di stazione, il verde corrisponde alla presenza di tutte le scadenze per variabile per stazione mentre il rosso indica che il numero di scadenze è inferiore a quello atteso. La tabella viene aggiornata a conclusione dell'analisi relativa a ciascuna area. 68
- 4.12 Pagina HTML realizzata per la visualizzazione dei prodotti della valutazione della qualità delle simulazioni WRF realizzate nell'ambito del progetto NAUSICA. 68
- 4.13 Flusso di lavoro della suite NAUSICA_rds. 69
- 4.14 Schematizzazione dei due campioni confrontati in NAUSICA_rds per ogni livello estratti rispettivamente da modello (tutti i punti di griglia interni all'area selezionata, in blu) e radiosondaggio (tutte le misure effettuate nella parte di traiettoria contenuta all'interno della banda rossa). Come descritto nel Paragrafo 4.4.2 questo tipo di analisi è valido solo per campi che variano linearmente spostandosi lungo la verticale. 70
- 4.15 Profilo verticale medio della temperatura alle 12:00 UTC e 24:00 UTC nella primavera del 2010 a Rivolto per modello e misure dei radiosondaggi. 71
- 4.16 Bias (al centro), differenza tra venticinquesimo (in alto) e settantacinquesimo (in basso) percentile tra distribuzione relativa al modello e distribuzione relativa alle misure per la temperatura in corrispondenza di ogni livello del modello, per la primavera del 2010 a Rivolto. 72

- 4.17 Andamento della differenza tra valore del campo atteso a $\Delta z/2$ e valore di aspettazione del campo sull'intero Δz in funzione del Δz , nell'ipotesi che la pressione abbia una dipendenza dalla quota del tipo $P(z) = P_0 + e^{-az}$, dove a è dell'ordine di $10^{-4}[\text{m}^{-1}]$. Le discontinuità che si osservano in Figura 4.18 sono dovute al fatto che il valore di Δz scelto per l'analisi (in particolare per i livelli dall'ottavo in poi) corrisponde ad un valore non nullo di questa funzione e di conseguenza aggiunge un contributo dominante al bias rendendo trascurabile il contributo dovuto alle differenze tra campi simulati e misurati a cui siamo interessati. 73
- 4.18 Alcuni risultati delle analisi svolte con la suite NAUSICA_rds da cui emerge la problematica descritta, nel grafico superiore è riportata l'analisi valida del bias e differenza tra 25° e 75° percentile delle distribuzioni di modello e dati raccolti da radiosondaggi del campo di temperatura potenziale nella primavera del 2010 a Rivolto utilizzando bande verticali di 50 m per i primi 8 livelli e di 200 m per i livelli successivi, mentre in quello inferiore è riportata la stessa analisi relativa alla pressione. In questo secondo caso l'analisi è da ritenersi NON attendibile per i livelli successivi all'ottavo come spiegato nel Paragrafo 74
- 4.19 Analisi analoga a quella riportata in Figura 4.18) (bias e differenza tra 25° e 75° percentile delle distribuzioni di modello e dati raccolti da radiosondaggi del campo di temperatura potenziale nella primavera del 2004 a Rivolto) ottenuta utilizzando bande verticali di 50 m per tutti i livelli. Si può notare la scomparsa delle discontinuità presenti nel caso precedente. 75
- 5.1 Diagramma di GANTT relativo allo stato di avanzamento del progetto NAUSICA a Giugno 2017. Si possono notare le interruzioni delle attività avvenute ad Ottobre 2016 e Febbraio 2017 per problemi contrattuali (periodi evidenziati in rosso) e l'attesa delle risorse di calcolo necessarie per la produzione, disponibili solo a partire dal 15 marzo 2017 (periodo evidenziato in verde), che ha causato ritardi nella produzione del database. 78
- 5.2 Confronto tra distribuzioni empiriche di modello e misure della temperatura sull'area di alta pianura in estate (Giugno, Luglio e Agosto 2010) e in inverno (Dicembre 2010, Gennaio e Febbraio 2011). Questi due confronti rappresentano uno dei casi di miglior riproduzione della distribuzione dei valori misurati e uno dei casi con bias misure/modello maggiore. Si può inoltre notare che le distribuzioni prodotte dal modello sono leggermente più strette di quelle relative alle misure (si veda anche Figura 5.3). 82
- 5.3 Bias (differenza tra i valori medi delle simulazioni e quelli delle misure) per le temperature stagionali degli anni 2010 e 2011. Il bias è sempre inferiore ai 2°C. Si può notare che tendenzialmente i valori medi risultano sottostimati nelle stagioni calde e sovrastimati in quelle fredde. 83

- 5.4 Istogrammi della temperatura di simulazioni e misure relativi all'estate 2014 a Lignano e all'autunno 2014 a Fossalon. E' evidenziata la tendenza delle simulazioni a sottostimare le temperature massime, specialmente in estate, e a sovrastimare le minime, specialmente invernali. Questa tendenza è analizzata in generale tramite lo studio dei 1° e 99° percentile riportati in Tabella 5.5. 83
- 5.5 Differenza tra 99° e 1° percentili di simulazioni e dati per le temperature stagionali per gli anni 2010 e 2011. Il 99° percentile risulta sottostimato in tutte le aree e in tutte le stagioni mentre il 1° percentile tendenzialmente sovrastimato. 84
- 5.6 Confronto tra distribuzioni empiriche di simulazioni e modello della temperatura sull'area di media pianura per il mese di Luglio 2015. Si nota una sottostima delle alte temperature e l'andamento di crescita discontinuo nella che indica che si sta considerando un'area contenente zone con comportamenti meteorologici differenti. 84
- 5.7 Spettro in potenza (sopra) e forma d'onda su 10 giorni (sotto) della temperatura presso la stazione di Palazzolo dello Stella in Primavera (Marzo, Aprile e Maggio 2010). Si nota un ottimo accordo tra simulazione e dati. Nello spettro compare, oltre ai picchi dovuti all'oscillazione diurna ed armoniche successive, anche il picco relativo all'oscillazione stagionale. 85
- 5.8 Confronto delle distribuzioni di misure e modello dell'intensità del vento sull'area di bassa pianura in autunno (Settembre, Ottobre e Novembre 2010). Rappresenta un buon esempio della generale tendenza del modello a sovrastimare i venti. 86
- 5.9 Confronto tra distribuzioni simulata e misurata dell'intensità del vento sull'area di costa in autunno (Settembre, Ottobre e Novembre 2010). Come si può anche osservare confrontandola con il caso di bassa pianura di Figura 5.8, in quest'area l'accuratezza nella riproduzione dei dati da parte del modello risulta maggiore. 86
- 5.10 Confronto delle distribuzioni misurata e simulata dell'intensità del vento presso le stazioni di Cividale (sopra) e Faedis (sotto) in primavera (Marzo, Aprile e Maggio 2010). L'accuratezza nella riproduzione dell'intensità del vento risulta molto variabile localmente. 87
- 5.11 Intensità del vento presso la stazione di Cividale, indicata dal Δ , in una scadenza per la quale la simulazione prevede un massimo molto superiore a quello misurato (i massimi di misure e simulazioni sono riportati nella tabella di destra). Si può osservare che la discrepanza è dovuta al fatto che la stazione si trova nella zona periferica del temporale chiaramente visibile in figura, il valore ricavato dalle simulazioni è quindi influenzato anche dai valori in punti di griglia interessati dal temporale mentre la stazione non è stata interessata dal fenomeno. 88

- 5.12 Bias (differenza tra la media delle simulazioni e quella relativa ai dati) dell'intensità del vento per la primavera 2010, 2011 e 2014. Si osserva una dipendenza dall'anno considerato. 89
- 5.13 Spettro di potenza, ricavato da misure e modello, dell'intensità del vento a Cividale nella stagione estiva del 2010. I picchi corrispondono alle frequenze delle tipiche brezze di valle, fenomeni locali che risultano ben riprodotti dal WRF. Di contro si osserva una sottostima del modello per quanto riguarda le basse frequenze. 90
- 5.14 Bias (differenza tra la media delle simulazioni e quella relativa ai dati) dell'umidità relativa stagionale del 2010. Come si può notare il bias è inferiore al 5% in tutte le aree e stagioni considerate ad eccezione della costa che risulta affetta da un bias sistematico di circa il 10% dovuto alla presenza di numerosi punti di griglia di tipo mare. 90
- 5.15 Confronto tra distribuzione empirica delle misure e simulata relativo all'umidità relativa sull'area alpina in estate (Giugno, Luglio e Agosto 2010). 91
- 5.16 Confronto delle distribuzioni empiriche ricavate rispettivamente da modello e misure dell'umidità relativa sulle aree di alta pianura e montagna rispettivamente in inverno (Dicembre 2010, Gennaio e Febbraio 2011) ed estate (Giugno, Luglio e Agosto 2010). La distribuzione simulata risulta più stretta di quella relativa alle misure, in particolare si notano in pianura sovrastima valori bassi e in montagna sottostima dei valori alti. 91
- 5.17 Confronto delle distribuzioni di misure e modello della precipitazione oraria sull'area di bassa pianura in primavera (Marzo, Aprile, Maggio 2010). In generale il bias nelle distribuzioni relative alla precipitazione è inferiore ad 1 mm orario in tutte le stagioni e su tutte le aree. 92
- 5.18 Confronto tra distribuzione simulata e misurata della precipitazione oraria sull'area di bassa pianura in primavera (Marzo, Aprile, Maggio 2011) e, in alto, istogramma di una delle stazioni contenute in questa area (Gorgo) nel medesimo periodo. Si nota una sottostima del range 2mm/h -10 mm/h comune alla maggior parte delle aree e stazioni analizzate. 93
- 5.19 Confronto tra distribuzioni estratte dal modello e misurate della precipitazione oraria di intensità superiore ai 10 mm orari sulle aree di alta e bassa pianura per l'anno 2010. Il bias come si può notare sempre inferiore ad 1 mm/h. 94
- 5.20 Confronto tra distribuzioni simulata e misurata della precipitazione oraria di intensità superiore ai 25 mm orari sulle aree di montagna, media e alta pianura per l'anno 2010. Si osservano nelle simulazioni una sottostima su media pianura e una sovrastima su area di montagna di entità contenuta. 94

- 5.21 Confronto delle distribuzioni modello/misure della pressione sull'area Alpi in primavera (Marzo, Aprile, Maggio 2010) e relativo al punto di stazione del monte Lussari nello stesso periodo. Si nota la sovrapposizione di tre zone caratterizzate da andamenti differenti nell'area considerata che rendono poco significativa tutta l'analisi. 95
- 5.22 Confronto delle distribuzioni modello/misure della pressione nella stazione Boa in inverno (Dicembre 2010, Gennaio e Febbraio 2011) ed estate (Giugno, Luglio ed Agosto 2010). La distribuzione calcolata dal modello è più larga e ha bias variabile tra 0.5 e 7.0 hPa. Si notano inoltre una sottostima nella simulazione dei valori bassi nelle stagioni fredde e una sovrastima dei valori alti nelle stagioni calde. 95
- 5.23 Confronto tra distribuzioni simulata e misurata della pressione sull'area di costa per il mese di Luglio 2015. Si nota una sovrastima da parte del modello dell'ordine del 5%. 96
- 5.24 Forma d'onda su 10 e 90 giorni (sinistra) e spettro in potenza (destra) della pressione presso la stazione di Palazzolo dello Stella in Primavera (Marzo, Aprile e Maggio 2010). Si nota un ottimo accordo tra simulazione e dati. Lo spettro è caratterizzato dal picco relativo alle modulazioni su scala sinottica. 96
- 5.25 Confronto tra distribuzioni ricavata dal modello e misurata della radiazione nell'area di media pianura in inverno (Dicembre 2010, Gennaio e Febbraio 2011) ed estate (Giugno, Luglio ed Agosto 2010). Si vede il bias dovuto alla sovrastima del modello rispetto alle misure. Nella forma della distribuzione si osserva una sovrastima dei valori bassi nelle stagioni fredde mentre le stagioni calde sono caratterizzate da una distribuzione di forma analoga a quella delle misure ma più larga, come evidenziato anche in Figura 5.26. 97
- 5.26 Istogrammi delle distribuzioni della radiazione in una delle stazioni appartenenti all'area di media pianura (Talmasson) in inverno (Dicembre 2010, Gennaio e Febbraio 2011) ed estate (Giugno, Luglio ed Agosto 2010). Nella forma della distribuzione è evidenziata la sovrastima dei valori bassi nelle stagioni fredde e la tendenza nelle stagioni calde ad avere una distribuzione di forma analoga a quella delle misure ma più larga. 98

5.27 Analisi dei risultati ricavati riguardo lo spettro in potenza della radiazione visibile. In Fig. in alto a destra è riportato un singolo ciclo diurno relativo ad un giorno caratterizzato da bassa copertura nuvolosa estratto dalla serie temporale dei netCDF (1 Luglio 2015, Udine): si può osservare che è caratterizzato da un valore nullo durante le ore notturne. In alto a sinistra è rappresentata la serie temporale della radiazione ad onda corta estratta dai netCDF relativi alle prime 200 ore del mese di Luglio 2015, la parte finale dell'intervallo temporale riportato è interessata dagli effetti di copertura nuvolosa che modificano la forma del segnale. Nella Fig. centrale sono sovrapposti gli spettri delle serie temporali della radiazione visibile ricavati da misure e simulazioni nella stazione di Gradisca nel trimestre Settembre, Ottobre e Novembre 2015. L'accordo tra dati e simulazioni è ritenuto soddisfacente soprattutto per quanto riguarda la riproduzione dei picchi relativi alla frequenza fondamentale, dovuta al ciclo diurno e di conseguenza di periodo 24 ore, e alle prime armoniche indicate dalle frecce in figura. La discordanza che si osserva per le alte frequenze è dovuta al rumore introdotto dalla copertura nuvolosa e al fatto che nemmeno in assenza di nuvolosità l'onda è quadra, vedere anche La Figura in alto a destra di destra. Nella Fig. in basso sono invece riportati un esempio di onda quadra e il relativo spettro la cui costruzione è illustrata nello schema a destra. 99

5.28 Spettro in potenza della radiazione visibile relativo a 15 mesi (Gennaio 2015-Marzo 2016) ricavato dalla serie temporale dei netCDF nelle coordinate della stazione BOA: si nota la frequenza caratteristica dell'oscillazione stagionale della radiazione indicata dalla freccia azzurra. 100

- L'estate è la stagione con accordo migliore mentre l'inverno è caratterizzato da una sottostima da parte del modello. 101
- 5.29 Confronto delle distribuzioni simulata/misurata dell'evapotraspirazione nell'area di costa in estate (Giugno, Luglio ed Agosto 2010). Si nota il bias dovuto alla presenza all'interno dell'area di molti punti di griglia di tipo mare. Il problema è stato ridotto considerando solo i punti di tipo *land*. 101
- 5.31 Istogrammi delle distribuzioni dell'evapotraspirazione in una delle stazioni appartenenti all'area di media pianura (Gradisca) in inverno (Dicembre 2010, Gennaio e Febbraio 2011) e in zona montagna (Gemona) in estate (Giugno, Luglio ed Agosto 2010). Si notano una leggera sovrastima dei valori alti da parte del modello in estate e la sottostima dei valori centrali della distribuzione in inverno. 102
- 5.32 Confronto tra distribuzioni empiriche misurata e simulata dell'evapotraspirazione nelle stazioni di Faedis e Fagagna, entrambe appartenenti all'area di alta pianura, in inverno (Dicembre 2010, Gennaio e Febbraio 2011). Si osserva che la qualità della riproduzione locale è molto variabile. 102
- 5.33 Bias e differenza tra 25° e 75° percentile delle distribuzioni di modello e dati misurati da radiosondaggi del campo di temperatura potenziale nell'ottobre del 2015 a Rivolto utilizzando bande verticali di 50 m per i primi 8 livelli e di 200 m per i livelli successivi. Si può notare che il modello riproduce molto bene le misure almeno fino ad una quota di circa 10 km oltre la quale la distribuzione dei dati simulati risulta un po' più stretta di quella reale anche se il bias risulta sempre inferiore ai 2°C. La divergenza per i livelli più alti può essere attribuita ad effetti di bordo. 103
- 5.34 Profilo verticale medio stagionale della temperatura alle 12:00 UTC e alle 24:00 UTC nell'ottobre del 2015 a Rivolto per modello e misure dei radiosondaggi. Nonostante l'ampio range della scala (necessario alla visualizzazione dell'intero profilo) che non rende quantificabile la differenza, l'accordo appare qualitativamente buono. 104
- 5.35 Confronto tra i gradienti medi di pressione e temperatura misurati e calcolati per l'autunno del 2015 alle 12 e 24 UTC per ciascun livello del modello sulla verticale nel punto di lancio del radiosondaggio di Rivolto. Qualitativamente l'accuratezza delle simulazioni NAUSICA è molto buona. I livelli con maggiore bias sul gradiente risultano essere i due livelli inferiori. Il segno del gradiente resta comunque sempre ben riprodotto. 105

5.36 Confronto tra i gradienti medi di temperatura potenziale ed intensità del vento misurati e calcolati per l'autunno del 2015 alle 12 e 24 UTC per ciascun livello del modello sulla verticale nel punto di lancio del radiosondaggio di Rivolto. Qualitativamente l'accuratezza delle simulazioni NAUSICA è molto buona. I livelli con maggiore bias sul gradiente risultano essere i due livelli inferiori. Il segno del gradiente resta comunque sempre ben riprodotto. 106

H.1 Esempio di analisi della suite NAUSICA_srds (ERRATA nel metodo come spiegato nel Paragrafo relativo). Rappresenta il risultato del calcolo giornaliero dei bias per ogni livello del modello per la primavera del 2010 a Rivolto alle 12:00 UTC. 149

Elenco delle tabelle

- 2.1 Elenco degli strumenti di pre e post elaborazione utilizzati nel progetto NAUSICA oltre a WPS e WRF. Gli eseguibili e gli script nominati sono stati prodotti specificamente per il progetto e saranno descritti nei capitoli successivi (Appendici E, G). 16
- 3.1 Prove dei tempi di calcolo della coda *Julia* per simulazioni WRF di 10 giorni, da questi risultati si è deciso di riservare 180 processori (30 per sei nodi) per le simulazioni del WRF. 22
- 3.2 Risoluzione spaziale e temporale e intervallo temporale di salvataggio dei tre domini innestati utilizzati. 23
- 3.3 Caratteristiche dei tre domini utilizzati nei test per la definizione del dominio regionale. I tempi riportati corrispondono a run del WRF di 10 giorni in configurazione operativa su 4 nodi da 32 core ciascuno. 25
- 4.1 Elenco delle caratteristiche dei prodotti delle suite utilizzate nel progetto NAUSICA. A causa delle problematiche riscontrate nella suite NAUSICA_rds descritte non è ancora stata completata la struttura della suite. Per questo motivo non è attualmente possibile fornire informazioni complete su dimensioni e tempi di produzione. 56
- 4.2 Variabili utilizzabili per il confronto eseguito dalla suite NAUSICA_ver, operazioni necessarie per poterle confrontare e unità di misura disponibili per l'analisi. 64
- 4.3 Dimensioni dei prodotti di NAUSICA_ver per area per trimestre considerando tutte le stazioni, tutte le aree, tutte le variabili, tutte la analisi disponibili e tutti i formati di output disponibili per i plot (ps e png). 67
- 5.1 Elenco dei prodotti del progetto disponibili 78
- B.1 Codici ZZ dei prodotti della suite NAUSICA_ver 114

Bibliografia

- [1] Wikipedia. Voce "Nausicaa" di Wikipedia.
<https://it.wikipedia.org/wiki/Nausicaa>.
- [2] CRMA. Pag. intranet ARPA FVG Progetto NAUSICA.
http://ms05lxarpa.arpa.fvg.it/wiki/index.php/Project_NAUSICA.
- [3] WRF. Weather Research and Forecasting Model (WRF).
<http://www.wrf-model.org/index.php>.
- [4] The DPSIR framework used by the EEA.
<https://www.eea.europa.eu/publications/92-9167-059-6-sum/page002.html>.
- [5] ECMWF. European centre for medium-range weather forecasts.
<http://www.ecmwf.int/>.
- [6] Formato netCDF. https://www.unidata.ucar.edu/software/netcdf/docs_rc/.
- [7] World Meteorological Organization WMO. A GUIDE TO THE CODE FORM FM 92-IX Ext. GRIB. .
- [8] National Centers for Environmental Prediction (NCEP). Unified post processor (UPP). <http://www.dtcenter.org/upp/users/> .
- [9] ARPA FVG. Relazioni qualità aria.
http://www.arpa.fvg.it/cms/tema/aria/utilita/Documenti_e_presentazioni/tecnico_scientifici.html#Relazioni_qualita_aria.
- [10] Pagina intranet wiki-crma relativa al cluster regionale Fenice.
<http://ms05lxarpa.arpa.fvg.it/wiki/index.php/FENICE>.
- [11] CRMA. Pagina intranet wiki-crma relativa all'operatività.
<http://ms05lxarpa.arpa.fvg.it/wiki/index.php/WRF:Operativita%27>.
- [12] WRF Parametrizations - Sound Climate Change Scenario Modeling. .
- [13] WRF Preprocessing System (WPS).
http://www2.mmm.ucar.edu/wrf/users/docs/user_guide/users_guide_chap3.html.
- [14] WMO. Sito della World Meteorological Organization (WMO).
http://www.wmo.int/pages/index_en.html.

- [15] Brian Fox (Free Software Foundation) Chet Ramey (Case Western Reserve University). Bash reference manual. .
- [16] Raucci Biagio. Il linguaggio fortran 90/95. .
- [17] David Dent (ECMWF) Michael Gertz (DWD) Jean Louis Ricard (Meteo France) Phillip Andrews (UKMO), Gerard Cats (KNMI/HIRLAM). European Standards for Writing and Documenting Exchangeable FORTRAN 90 Code. .
- [18] COSMO. Modello COSMO. <http://www.cosmo-model.org/>.
- [19] David De Roure Carole Goble. The Impact of Workflow Tools on Data-centric research. .
- [20] ECMWF. Workflow manager ecFlow. <https://software.ecmwf.int/wiki/display/ECFLOW/Home>.
- [21] Trac - tracking system for software development projects. <https://trac.edgewall.org/>.
- [22] Jenkins. <https://jenkins.io/index.html>.
- [23] Git Repository. <https://git-scm.com/>.
- [24] Climate Data Operators (CDO). <https://code.zmaw.de/projects/cdo>.
- [25] CRMA. Pagina intranet wiki-crma relativa al progetto NAUSICA contenente la descrizione di vari eseguibili utilizzati. .
- [26] ECMWF. Applicativi GRIB API. <https://software.ecmwf.int/wiki/display/GRIB/What+is+GRIB-API>.
- [27] Fei Chen. The Noah Land Surface Model in WRF - A short tutorial. .
- [28] Run climatici WRF. http://www2.mmm.ucar.edu/wrf/OnLineTutorial/Class/cases/climate_run.php.
- [29] Tutorial aggiornamento SST nel WRF. <http://www2.mmm.ucar.edu/wrf/OnLineTutorial/DATA/-SST/index.html> <http://www2.mmm.ucar.edu/wrf/OnLineTutorial/CASES/SST/index.html>.
- [30] Bug relativo all'aggiornamento della TLAG (temperatura dello strato più profondo di suolo) nelle versioni del WRF successive alla 3.7 e soluzione. .
- [31] Pierluigi Masai. Quality evaluation of different atmospheric boundary layer parameterizations implemented in the meteorological wrf model. an annual focus study over friuli venezia giulia. Tesi di Laurea Magistrale svolta presso il CRMA.
- [32] Gridpoint Statistical Interpolation (GSI) System. <https://dtcenter.org/com-GSI/users/>.

- [33] CRMA. Classificazione delle simulazioni crma, pag. intranet wiki-crma. http://ms05lxarpa.arpa.fvg.it/wiki/index.php/Elenco_simulazioni.
- [34] R. <https://www.r-project.org/>.
- [35] Gnuplot. <http://www.gnuplot.info/>.