



ISPRA

Istituto Superiore per la Protezione
e la Ricerca Ambientale

Controlli di qualità delle serie di temperatura e precipitazione





ISPRA

Istituto Superiore per la Protezione
e la Ricerca Ambientale

Controlli di qualità delle serie di temperatura e precipitazione

Informazioni legali

L'istituto Superiore per la Protezione e la Ricerca Ambientale (ISPRA) e le persone che agiscono per conto dell'Istituto non sono responsabili per l'uso che può essere fatto delle informazioni contenute in questo rapporto.

ISPRA – Istituto Superiore per la Protezione e la Ricerca Ambientale
Via Vitaliano Brancati, 48 – 00144 Roma
www.isprambiente.gov.it

ISPRA, Stato dell'Ambiente 66/2016

ISBN 978-88-448-0780-1

Riproduzione autorizzata citando la fonte

Elaborazione grafica

ISPRA

Grafica di copertina: Franco Iozzoli

Foto di copertina: Franco Desiato, Franco Iozzoli

Coordinamento editoriale:

Daria Mazzella

ISPRA - Settore Editoria

Luglio 2016

Autori

Guido Fioravanti, Piero Fraschetti, Walter Perconti, Emanuela Piervitali, Franco Desiato.

INDICE

1. INTRODUZIONE	6
2. PROCEDURA DI CONTROLLO DELLA QUALITÀ DEI DATI	7
2.1 Test di integrità di base	7
2.2 Test per l'identificazione dei valori anomali	10
2.3 Test di consistenza interna e temporale	14
2.4 Test di consistenza spaziale	16
3. APPLICAZIONE DELLA PROCEDURA: STATISTICHE RIASSUNTIVE	23
3.1 Dati.....	23
3.2 Risultati dei controlli di qualità.....	29
4. CONCLUSIONI	33
BIBLIOGRAFIA	34

1. INTRODUZIONE

La disponibilità di lunghe serie di dati strumentali di buona qualità è di fondamentale importanza per lo studio della variabilità del clima e in particolare degli eventi estremi, che sono spesso causa di impatti negativi sull'ambiente e la società. Negli ultimi anni è stata rivolta un'attenzione crescente proprio allo studio delle variazioni della frequenza e dell'intensità degli estremi climatici a scala globale e a scala regionale e locale (Alexander et al. 2006, Bartholy e Pongracz, 2007; Toreti e Desiato, 2008; Efthymiadis et al., 2011; Donat et al., 2013; Fioravanti et al., 2016; Hartmann et al., 2013). Una valutazione periodica e il più possibile completa del clima e delle variazioni climatiche a scala nazionale rende necessaria la realizzazione di un dataset climatico nazionale di dati giornalieri di qualità controllata, che integri le serie disponibili da diverse fonti e reti osservative sul territorio. La quantità elevata e la grande varietà delle fonti dei dati richiedono un approccio sistematico, sia nella fase di raccolta delle serie sia nell'applicazione delle procedure di controllo della qualità. Come raccomandato dall'Organizzazione Meteorologica Mondiale (WMO, 2011) i dati dovrebbero essere considerati validi, e quindi inseriti nell'archivio permanente, solo dopo aver superato adeguati controlli di qualità. Per garantire la buona qualità dei dati è necessario sottoporre regolarmente le serie temporali a rigorosi controlli che assicurino che ogni singolo valore di una serie temporale non sia affetto da errori introdotti nella fase in osservazione, trascrizione e archiviazione del valore stesso.

Il Sistema nazionale per la raccolta, l'elaborazione e la diffusione di dati Climatologici di Interesse Ambientale (SCIA, www.scia.isprambiente.it) assicura da tempo la disponibilità di dati, statistiche e indici climatici di qualità controllata, che originano dalle principali reti osservative nazionali e regionali, con passo temporale decadale, mensile e annuale (Desiato 2007; Desiato et al., 2006, 2007, 2011). La crescente disponibilità di serie di dati giornalieri rende possibile incrementare la risoluzione temporale di dati e indici, attraverso la realizzazione di un dataset nazionale di serie temporali di temperatura e precipitazione con passo giornaliero.

In questo rapporto viene descritta la metodologia adottata dall'ISPRA per il controllo delle serie giornaliere. La procedura, che prevede una sequenza di controlli automatici, è stata elaborata sulla base di un'analisi dei controlli di qualità applicati da varie Istituzioni europee ed internazionali che sono responsabili della raccolta e della diffusione di dati climatici, in particolare il National Climatic Data Center (NCDC) della National Oceanic and Atmospheric Administration (NOAA). I test sviluppati e descritti nel Capitolo 2 hanno l'obiettivo di individuare diverse tipologie di errore, quali i valori impossibili, i dati duplicati, i valori anomali (*outlier*), le inconsistenze interne o temporali.

Nel Capitolo 3 viene illustrata l'applicazione della procedura di controllo al dataset nazionale con passo giornaliero in corso di realizzazione in ISPRA; il dataset include le serie temporali provenienti da diverse fonti: il Servizio Meteorologico dell'Aeronautica Militare (dal database globale della NOAA), alcuni Servizi Meteorologici o Centri Funzionali di Protezione Civile regionali, alcuni Servizi Idrografici o Idrologici regionali (dal Sistema Informativo Idrologico HIScentral, www.hiscentral.isprambiente.gov.it); quest'ultimo permette di recuperare i dati direttamente dalle fonti tramite opportune interfacce, fornendo ulteriori capacità operative, quali un registro nazionale di servizio dati, catalogati utilizzando le norme e le procedure del Geospatial Consortium (www.opengeospatial.org/standards/) e dell'Organizzazione Meteorologica Mondiale. I risultati dell'applicazione delle procedure di controllo sono illustrati tramite statistiche riassuntive (grafici e tabelle).

2. PROCEDURA DI CONTROLLO DELLA QUALITÀ DEI DATI

Per garantire l'affidabilità di un dataset climatico, le procedure di controllo di qualità dei dati devono essere applicate in maniera sistematica e regolare. Quando la mole di dati è elevata, è necessario sviluppare test che operino in modalità automatica e che permettano di individuare in breve tempo i dati errati o sospetti (dati che potrebbero essere errati). L'Organizzazione Meteorologica Mondiale (WMO, 2011) suggerisce di affiancare a queste procedure una valutazione manuale degli elenchi di dati sospetti generati automaticamente, per ulteriori azioni da intraprendere, tra cui l'eventuale correzione del dato.

Tenendo presente questa indicazione, è stata sviluppata una procedura automatica finalizzata ad individuare valori palesemente errati o molto improbabili da un punto di vista climatico; sono stati invece rinviati a una seconda fase i test per individuare gli *outlier* meno estremi, che, seppure con il supporto di software applicativi, richiedono comunque una valutazione soggettiva e manuale da parte dell'operatore.

La procedura automatica prevede una sequenza di controlli, mutuati da quelli applicati al dataset globale di osservazioni meteorologiche giornaliere (Global Historical Climatological Network – GHCN) del National Climatic Data Center (NCDC) della NOAA (Durre et al., 2010; Arguez et al., 2012), opportunamente modificati tenendo conto di alcune caratteristiche peculiari del clima italiano. Ciascun controllo contrassegna con un *flag* di dato errato i dati che non superano il test prestabilito e li esclude da tutti i controlli successivi. Fanno eccezione i controlli di consistenza spaziale che, invece, contrassegnano con un *flag* di dato dubbio i valori anomali individuati, prevedendo un'ulteriore verifica manuale. La sequenza dei test è stata stabilita in modo da individuare il maggior numero di errori e ridurre al minimo il rischio di eliminare dati validi (falsi positivi).

I controlli della NOAA sono concepiti per garantire la buona qualità delle serie giornaliere di temperatura massima e minima, di precipitazione (pioggia e neve) e di altezza del manto nevoso. Alcuni test vengono applicati a valori medi (nel caso della temperatura) o cumulati giornalieri (precipitazione e altezza manto nevoso), mentre altri prevedono come input le serie di osservazioni orarie. Nel nostro caso, ai fini della realizzazione del dataset italiano di serie climatiche giornaliere e sulla base dei parametri climatici disponibili, sono stati considerati e applicati solo i controlli sui dati giornalieri di temperatura massima e minima e di precipitazione cumulata.

L'insieme dei controlli si può dividere in quattro categorie generali e viene eseguito secondo l'ordine seguente: test di integrità di base, test per l'identificazione dei valori anomali, test di consistenza interna e temporale, test di consistenza spaziale.

2.1 Test di integrità di base

I test di integrità di base verificano la presenza di valori ripetuti, mesi o anni duplicati, valori sospetti uguali a zero e valori impossibili (tabella 1).

Il primo controllo identifica sequenze di 20 o più valori consecutivi identici, che vengono contrassegnati con un *flag* di dato errato. Nelle sequenze di temperatura vengono trascurati i valori mancanti, in quelle di precipitazione vengono trascurati i dati mancanti e i valori nulli. Nella figura 1 è riportato un esempio di serie con 27 valori ripetuti uguali a zero.

Nel caso delle serie di precipitazione, un ulteriore controllo esclude le sequenze di 180 o più valori nulli consecutivi. Questo controllo, non compreso fra i test della NOAA, è stato aggiunto per eliminare lunghe sequenze di valori nulli associate erroneamente a dati mancanti. Il test successivo identifica le duplicazioni di sequenze di dati giornalieri in diversi periodi di tempo, in particolare: due anni diversi che hanno esattamente gli stessi dati oppure due mesi diversi (nello stesso anno o in anni diversi) che hanno esattamente gli stessi dati (figura 2). Il test si applica sia alla temperatura (minima e massima), che alla precipitazione. In quest'ultimo caso, per evitare di contrassegnare erroneamente come duplicati due periodi asciutti, si richiede la presenza di almeno cinque valori non nulli.

Per la temperatura un ulteriore controllo identifica le sequenze di 10 o più giorni consecutivi in cui la

temperatura massima è uguale alla temperatura minima, contrassegnando con un *flag* di dato errato sia la temperatura massima che la temperatura minima.

Poiché spesso lo zero viene erroneamente utilizzato per indicare un dato mancante, per le serie di temperatura un ulteriore controllo esclude i valori di temperatura massima e temperatura minima giornaliera che sono entrambi nulli.

L'ultimo controllo individua i valori di temperatura e precipitazione impossibili: temperatura massima superiore a 50°C o inferiore a -30°C, temperatura minima superiore a 40°C o inferiore a -40°C, precipitazione cumulata superiore a 800 mm o con valore negativo. Queste soglie sono state fissate in base ai valori massimi e minimi assoluti di temperatura e precipitazione massima giornaliera presenti nel dataset delle osservazioni giornaliere italiane, validati manualmente anche tramite il confronto con i valori registrati nelle stazioni limitrofe. Gli estremi validati sono: +48.1°C per la temperatura massima, -29.0°C per la temperatura minima, 538.2 mm per la precipitazione cumulata.

Tabella 1: Test di integrità di base.

Controllo	Condizione per invalidare	Variabili contrassegnate	Commenti
Valori ripetuti	10 o più valori consecutivi identici	Tmax, Tmin, Prec	Tmin, Tmax: non vengono considerati i dati mancanti Prec: non vengono considerati i dati mancanti e gli zeri
Persistenza zeri	180 o più valori nulli consecutivi	Prec	
Anni duplicati	Tutti i valori giornalieri di un anno uguali a tutti i valori di un altro anno.	Tmax, Tmin, Prec	Prec: si richiede la presenza di almeno 5 valori non nulli
Mesi duplicati	<u>All'interno dello stesso anno:</u> tutti i valori giornalieri di un mese uguali a tutti i valori di un altro mese. <u>In anni diversi:</u> tutti i valori giornalieri di un mese uguali a tutti i valori del mese corrispondente	Tmax, Tmin, Prec	Prec: si richiede la presenza di almeno 5 valori non nulli
Tmax uguale a Tmin	Tmax uguale a Tmin per 10 o più giorni consecutivi	Tmax e Tmin	
Valori nulli di Tmax e Tmin	Tmax e Tmin entrambe uguali a 0°C	Tmax e Tmin	Identifica gli zeri erroneamente utilizzati al posto del codice per i dati mancanti
Valori impossibili	Tmax > 50°C; Tmax < -30°C Tmin > 40°C; Tmin < -40 °C Prec > 800 mm; Prec < 0 mm	Tmax, Tmin, Prec	

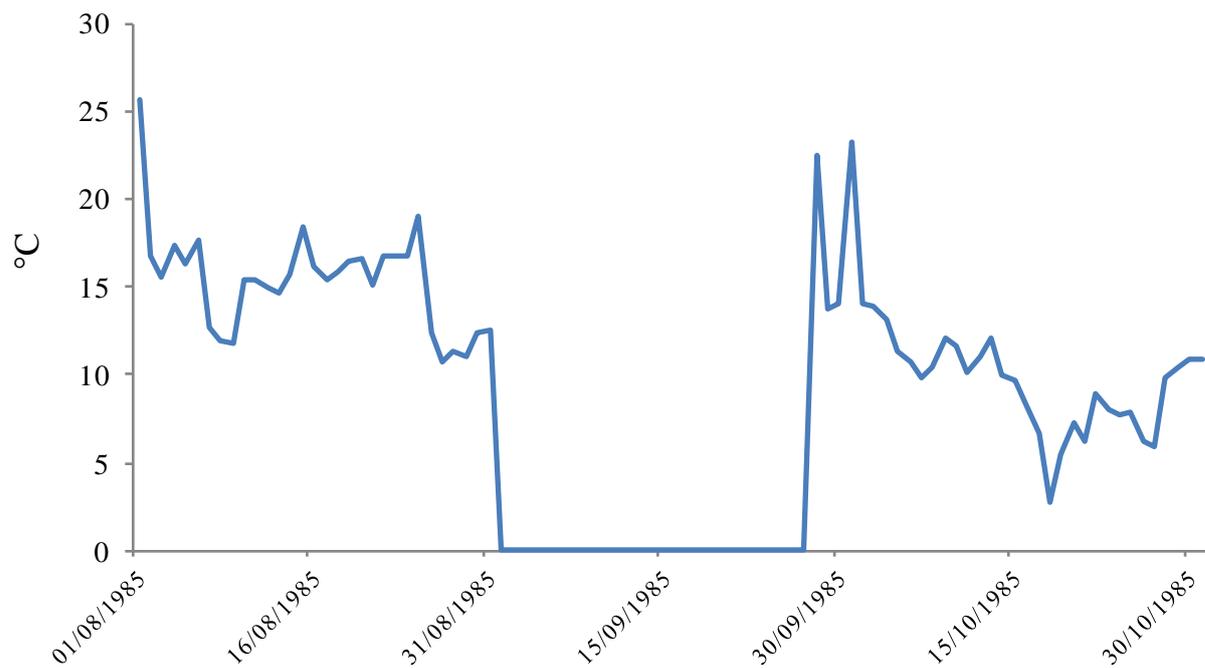


Figura 1: Serie di temperatura minima giornaliera da agosto a ottobre 1985, nella stazione di San Rossore in Toscana. Il grafico mostra un esempio di identificazione di una sequenza di valori uguali ripetuti.

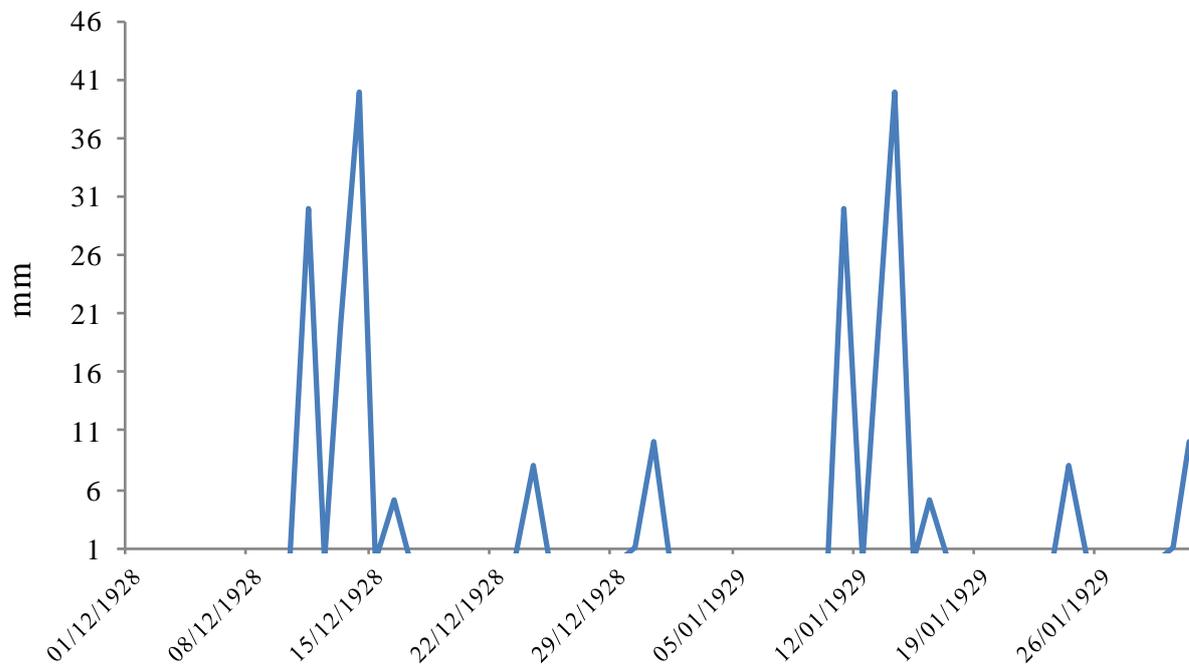


Figura 2: Serie di precipitazione della stazione di Fiumefreddo Bruzio (Calabria) nei mesi di dicembre 1928 e gennaio 1929. Esempio di duplicazione di dati in due mesi consecutivi.

2.2 Test per l'identificazione dei valori anomali

Per identificare i valori anomali vengono utilizzati due approcci diversi (tabella 2): un test che individua un eventuale salto nella serie di dati (*gap check*) e un controllo climatologico.

Il *gap check*, che si applica sia alle serie di temperatura sia a quelle di precipitazione, identifica i valori che si trovano nelle code della corrispondente distribuzione, quando sono eccessivamente distanti dal resto dei valori e quindi probabilmente errati.

Il controllo climatologico si basa invece sul confronto di ogni valore della serie da testare con alcuni parametri ricavati dall'intera serie, per lo stesso periodo dell'anno: test *z-score* per la temperatura e test basato sui percentili per la precipitazione.

Tabella 2: Test per l'identificazione dei valori anomali.

Controllo	Condizione per invalidare	Variabili contrassegnate	Commenti
<i>Gap check</i> temperatura	Differenza $\geq 10^{\circ}\text{C}$ fra due valori nella distribuzione mensile della temperatura	Tmax, Tmin	Nelle code della distribuzione (oltre il gap)
<i>Gap check</i> precipitazione	Differenza ≥ 300 mm fra due valori non nulli nella distribuzione mensile di Prec	Prec	Valori oltre il gap
<i>z-score</i>	$ z \geq 6$	Tmax, Tmin	Si richiedono almeno 100 valori su tutto il periodo nella finestra di 15 giorni
95° percentile (T media ≥ 0 o non disponibile)	Prec ≥ 9 volte il 95° percentile	Prec	Si richiedono almeno 20 valori non nulli su tutto il periodo, nella finestra di 29 giorni
95° percentile (Tmedia $< 0^{\circ}\text{C}$)	Prec ≥ 5 volte il 95° percentile e	Prec	Si richiedono almeno 20 valori non nulli su tutto il periodo, nella finestra di 29 giorni

Gap check

Fissato un mese dell'anno, il *gap check* esamina la distribuzione di frequenza delle osservazioni in quel mese, su tutti gli anni disponibili. Il test individua i valori alle code della distribuzione che risultano eccessivamente distanti e quindi separati in maniera poco realistica dal resto della distribuzione. L'algoritmo prima di tutto ordina tutti i valori dal più piccolo al più grande e calcola le differenze tra ogni coppia di valori consecutivi. Se un valore è separato dal valore precedente di una quantità superiore alla soglia fissata (*gap*) allora tutti gli elementi successivi vengono contrassegnati come errati.

Per la precipitazione il massimo *gap* consentito è stato fissato a 300 mm. La figura 3 mostra l'applicazione del test ad una serie di precipitazione e l'individuazione di un *outlier*.

Per la temperatura il test analizza separatamente le due metà della distribuzione: dalla mediana verso i valori superiori e dalla mediana verso i valori inferiori. In questo caso la differenza massima fra due valori consecutivi è fissata a 10°C .

Vale la pena osservare che per la precipitazione il test viene applicato utilizzando solo i giorni con precipitazione non nulla e che per entrambi i parametri le soglie di scostamento non cambiano nel corso dell'anno.

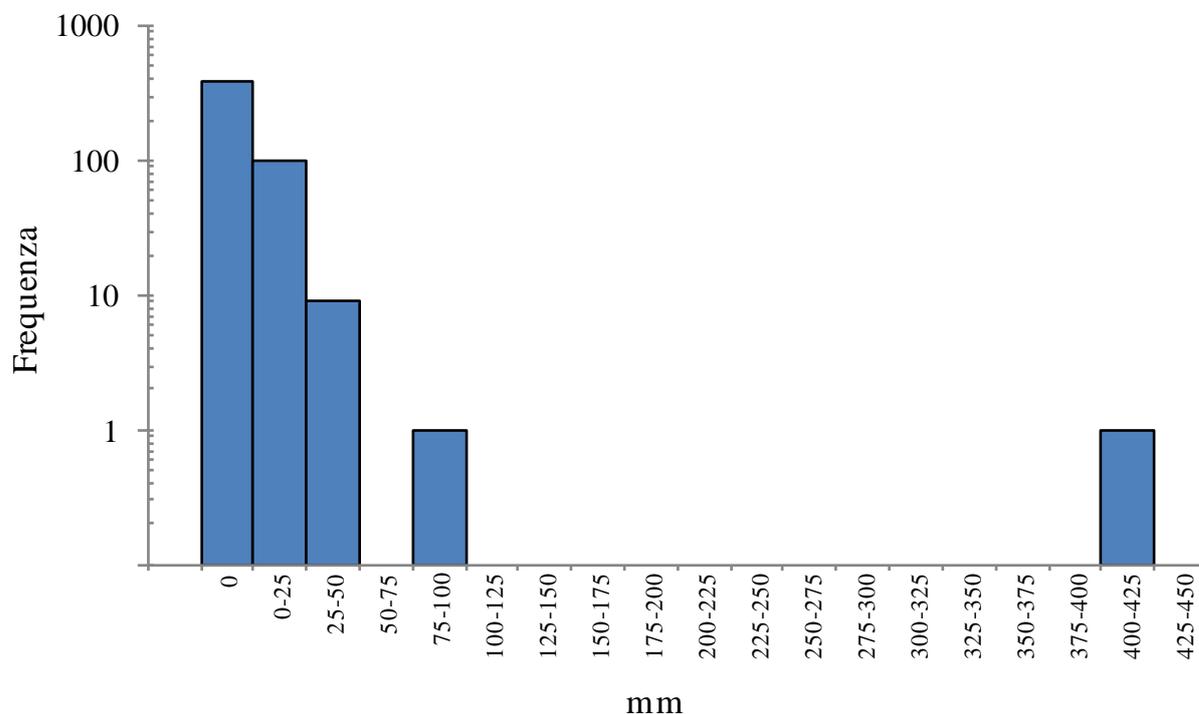


Figura 3: Esempio di un outlier identificato con il gap check, serie di precipitazione di Cerreto in Toscana. L'istogramma rappresenta la distribuzione di frequenza (in scala logaritmica) dei totali giornalieri di precipitazione nel mese di agosto (periodo 1998-2015). Il test individua come dato errato il valore 410 mm, perché la sua differenza con il valore più elevato del resto della distribuzione (93.6 mm) è superiore a 300 mm.

Controllo climatologico

La tecnica più comune per identificare i valori anomali (*outlier*) nei dati meteo-climatici è il test *z-score*, che consiste nel normalizzare i dati (utilizzando la media e la deviazione standard relativa ad uno specifico intervallo di tempo) e individuare i valori il cui *z-score* (valore normalizzato) eccede una soglia prefissata. Va tenuto presente che questa metodologia può essere applicata solo a serie sufficientemente lunghe come di seguito spiegato. Poiché questo tipo di test in distribuzioni anche leggermente asimmetriche potrebbe determinare un numero eccessivo di falsi positivi (*overflagging*) il controllo delle serie è stato applicato solo alla temperatura massima e minima, per le quali è generalmente plausibile un'approssimazione mediante la distribuzione normale.

Per la temperatura massima e minima la media (*biweight mean*) e la deviazione standard (Lanzante, 1996) richiesti per il test *z-score* sono calcolati su tutti i dati che cadono all'interno di una finestra di 15 giorni centrata sul giorno in esame, su tutto il periodo a disposizione. Affinché il controllo sia valido si richiede un totale di almeno 100 valori. Un valore di temperatura viene invalidato se supera di 6 deviazioni standard la media di lungo periodo. Il valore di questa soglia è stato scelto abbastanza alto allo scopo di individuare il maggior numero possibile di errori e nel contempo limitare il rischio di falsi positivi, considerando che uno degli obiettivi dell'analisi dei dati giornalieri è proprio l'analisi degli eventi estremi.

Nella figura 4 è mostrato un esempio di *outlier* individuato applicando il test alla serie di temperatura minima della stazione di Nizza Monferrato in Piemonte.

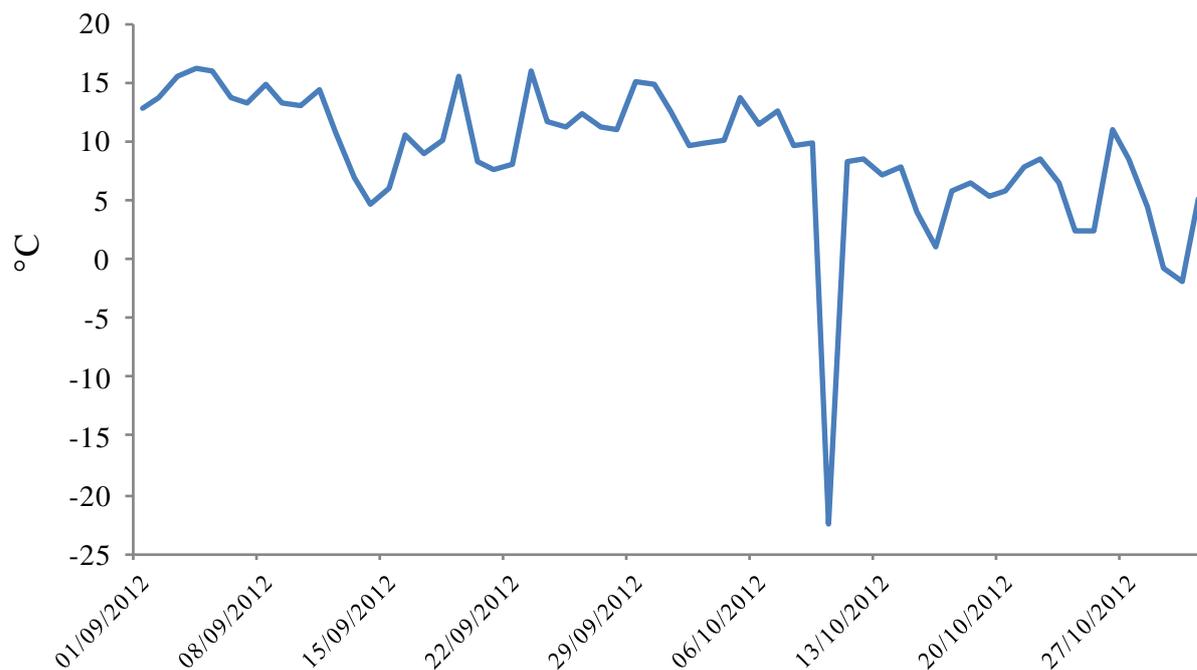


Figura 4: Esempio di un outlier identificato con il test z-score, serie di temperatura minima giornaliera di Nizza Monferrato in Piemonte. Il test è stato applicato utilizzando tutti i dati che cadono nella finestra di 15 giorni centrata sul giorno 10/10/2012, su tutto il periodo disponibile (1993-2015).

Alle serie di precipitazione viene invece applicato un test basato sui percentili. Per ogni giorno dell'anno viene definita una finestra di 29 giorni centrata sul giorno in esame e viene calcolato il 95° percentile della distribuzione di tutti i valori diversi da zero che cadono all'interno della finestra, su tutto il periodo disponibile (si richiede la presenza di almeno 20 valori). Nel caso della precipitazione la finestra è più ampia rispetto a quella definita per la temperatura (15 giorni), perché vengono presi in considerazione solo i valori di precipitazione non nulli. Il test identifica i valori di precipitazione che eccedono un multiplo prefissato del 95° percentile. Il multiplo dipende dalla temperatura: se la temperatura media giornaliera, $(T_{min} + T_{max})/2$, è maggiore o uguale a zero o non è disponibile, allora il multiplo viene fissato uguale a 9 volte il 95° percentile, se invece la temperatura media è inferiore allo zero, il multiplo viene fissato uguale a 5. In figura 5 è mostrato un esempio di applicazione di questo test.

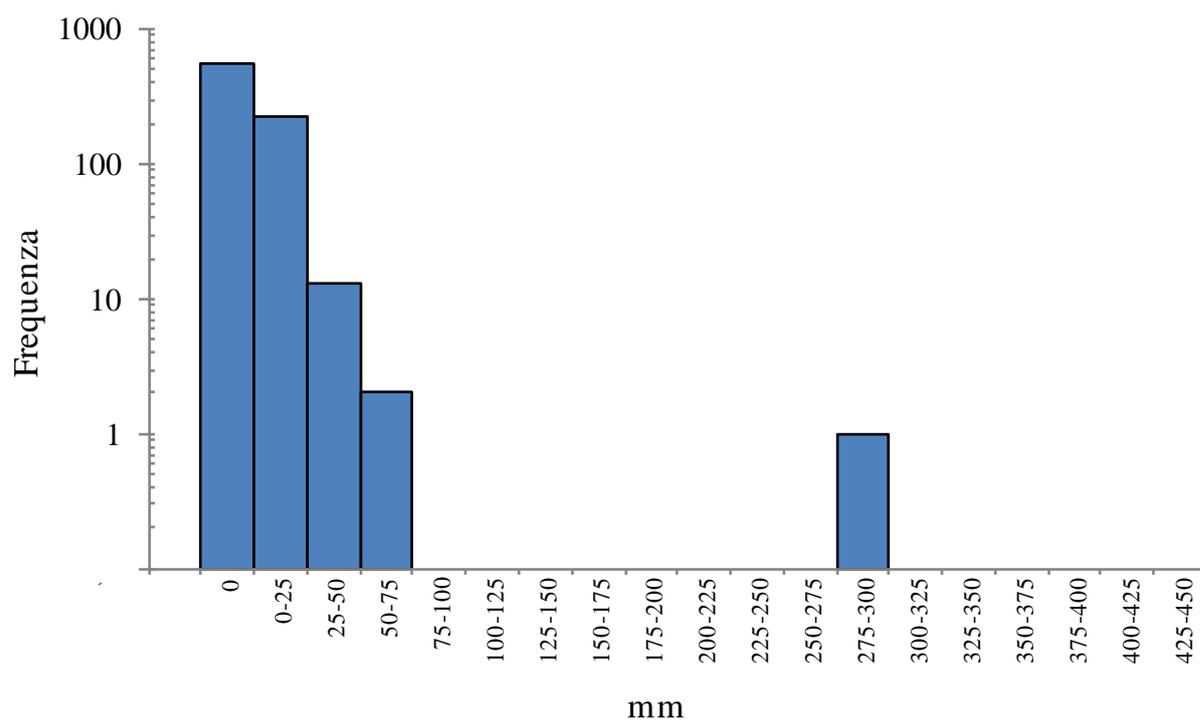


Figura 5: Esempio di un outlier identificato con il test del 95° percentile, serie di precipitazione di Fara Filiorum Petri in Abruzzo. L'istogramma rappresenta la distribuzione di frequenza dei totali giornalieri di precipitazione che cadono nella finestra di 29 giorni, centrata sul giorno 14/02/2012 per tutto il periodo disponibile (1987-2015). Il valore di 297.8 mm del 14/02/2012 viene contrassegnato come dato errato perché supera di 9 volte il 95° percentile della distribuzione (26.3 mm).

2.3 Test di consistenza interna e temporale

I test di consistenza interna e temporale si riferiscono alle serie di temperatura massima e minima e sono riportati nella tabella 3.

I test di consistenza interna controllano le violazioni fisiche fra due o più parametri. I test di consistenza temporale confrontano invece ogni valore di una variabile con quello del giorno precedente e successivo.

Il primo test identifica i valori di temperatura (massima o minima) che sono almeno 20°C più alti o più bassi del valore registrato nel giorno precedente o successivo (figura 6).

Il secondo test individua e contrassegna come dati errati i valori di temperatura massima e minima che verificano la condizione $T_{max} < T_{min}$ (figura 7).

Il terzo e ultimo test (*lagged range test*), individua le variazioni eccessive e repentine nelle serie di temperatura tramite il confronto tra temperatura massima e minima nel giorno in esame, nel giorno precedente e in quello successivo. Se applicato alla serie di temperatura massima, il test calcola la differenza tra la temperatura massima in un giorno $T_{max}(0)$ e la più alta temperatura minima fra quelle riportate nel giorno precedente, $T_{min}(-1)$, lo stesso giorno, $T_{min}(0)$ e il successivo, $T_{min}(1)$. Se la differenza supera i 35°C vengono contrassegnati come dati errati il valore di temperatura massima $T_{max}(0)$ e i tre valori di temperatura minima considerati.

Analogamente, per la serie di temperatura minima viene calcolata la differenza tra la temperatura minima in un giorno $T_{min}(0)$ e la più bassa temperatura massima fra quelle riportate nel giorno precedente, $T_{max}(-1)$, lo stesso giorno, $T_{max}(0)$, e il successivo, $T_{max}(1)$. Se la differenza supera i 35°C vengono contrassegnati come dati errati il valore di temperatura minima $T_{min}(0)$ e i tre valori di temperatura massima considerati.

Tabella 3: Test di consistenza interna e temporale.

Controllo	Condizione per invalidare	Variabili contrassegnate
Salto di temperatura	$T_{max}(0)$ almeno 20°C più alta o almeno 20°C più bassa dei valori $T_{max}(-1)$ e $T_{max}(+1)$ (stesso controllo per $T_{min}(0)$)	$T_{max}(0)$, $T_{min}(0)$
T_{max_Tmin}	$T_{max} < T_{min}$	T_{max} e T_{min}
<i>Lagged range test</i> T_{max}/T_{min}	$T_{max}(0) \geq \max[T_{min}(-1:1)] + 35^\circ\text{C}$	$T_{max}(0)$ e $T_{min}(-1:1)$
<i>Lagged range test</i> T_{min}/t_{max}	$T_{min}(0) \geq \max[T_{max}(-1:1)] - 35^\circ\text{C}$	$T_{min}(0)$ e $T_{max}(-1:1)$

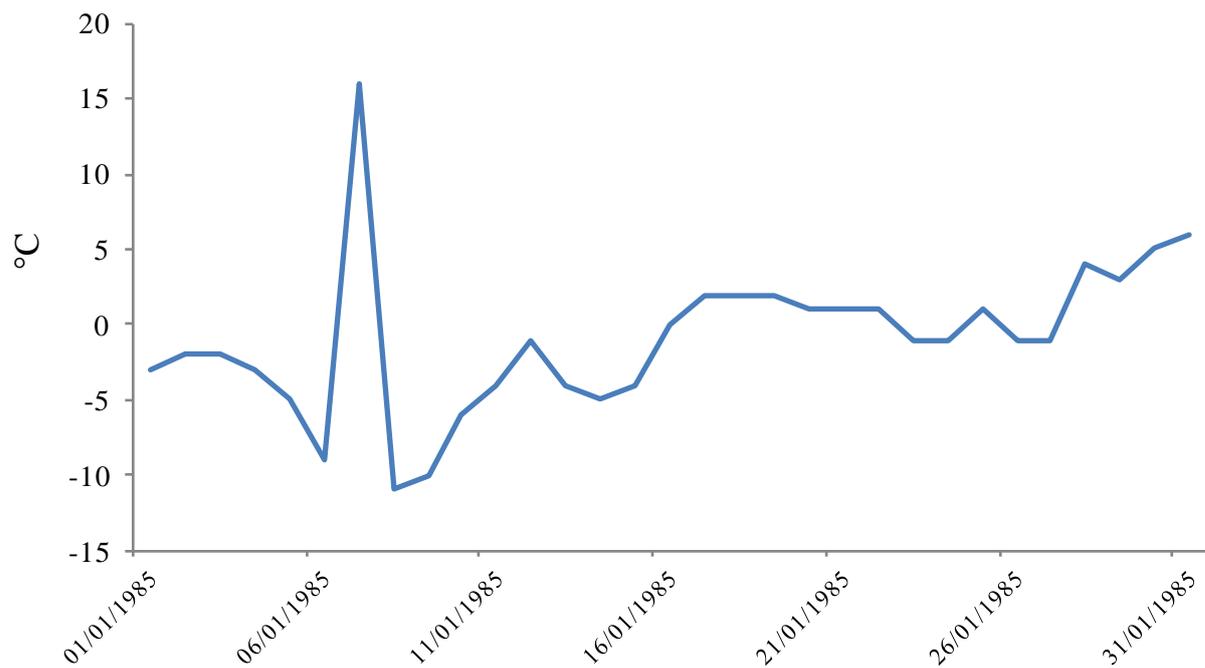


Figura 6: Serie di precipitazione della stazione di Favogna di Sotto (Bolzano) nel mese di gennaio 1985. Esempio di dato errato individuato verificando le differenze di temperatura superiori a 20°C rispetto al giorno precedente e al successivo.

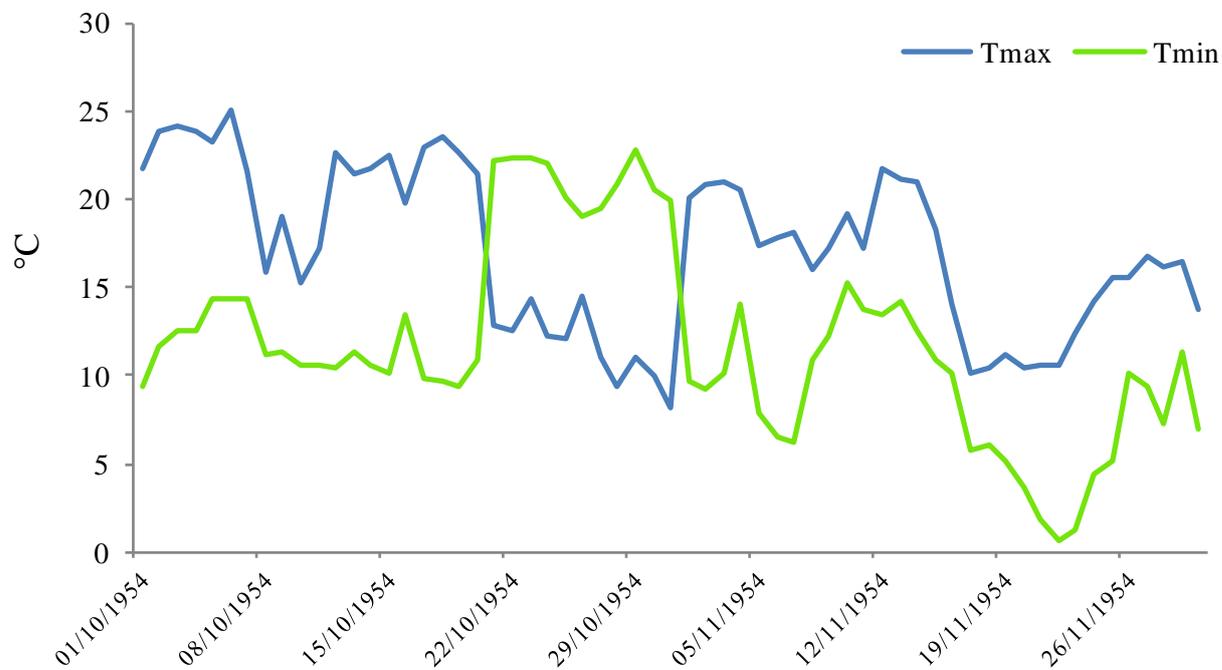


Figura 7: Serie di temperatura minima e massima della stazione di Roma – Collegio Romano, ottobre e novembre 1954. Il test individua come dati errati i valori dal 21 al 31 ottobre, per i quali $T_{min} > T_{max}$.

2.4 Test di consistenza spaziale

Un'altra tipologia di controlli è rappresentata dai test di consistenza spaziale, che hanno l'obiettivo di identificare i valori anomali non individuati dai precedenti test, tramite il confronto tra il valore registrato nella stazione target e le osservazioni rilevate nelle stazioni limitrofe. Questo approccio utilizza generalmente tecniche statistiche (quali la regressione) per stimare la variabile climatica nella stazione target, utilizzando i valori osservati nelle stazioni limitrofe e individuare poi le osservazioni che si discostano eccessivamente da questa stima.

In particolare, per la temperatura, le stime basate sulla regressione sono solitamente più accurate di quelle prodotte usando altri metodi (Hubbard et al., 2007).

Tuttavia, questa metodologia potrebbe non funzionare correttamente in situazioni di elevata variabilità orografica o durante il passaggio di un fronte meteorologico; inoltre essa riesce difficilmente a fornire una stima accurata di grandezze caratterizzate da un'elevata variabilità spaziale, come la precipitazione giornaliera.

Un'alternativa è rappresentata dal *corroboration* test, che si basa sul confronto a due a due tra il valore osservato nella stazione target e quello delle stazioni limitrofe, contrassegnando come dato anomalo il valore target se non viene confermato (corroborato) da tutte le osservazioni delle stazioni vicine.

Le serie di temperatura massima e minima sono state sottoposte sia al test di regressione spaziale sia al *corroboration* test, mentre alle serie di precipitazione è stato applicato solo il *corroboration* test (tabella 4).

I controlli spaziali sono stati inseriti nella procedura automatica di controllo, tuttavia i valori anomali individuati da questi test non vengono contrassegnati con un *flag* di dato errato, come avviene per i test precedenti, ma con un *flag* di dato dubbio. Un'ulteriore verifica manuale di questi dati anomali permette poi di valutare se escluderli o meno dal dataset. Da una prima analisi dell'esito di questi test sembra tuttavia che i controlli spaziali forniscano buoni risultati per la temperatura, individuando valori anomali effettivamente errati, e risultati incerti per la precipitazione, che devono essere valutati caso per caso.

Tabella 4: *Test di consistenza spaziale.*

Controllo	Condizione per invalidare	Variabili contrassegnate
Test di regressione	I residui e i residui standardizzati nella stazione target superano una determinata soglia	Tmax, Tmin
<i>Corroboration</i> test temperatura	L'anomalia nella stazione target differisce più di 10°C dalle anomalie delle stazioni limitrofe (nei giorni -1, 0,1)	Tmax, Tmin
<i>Corroboration</i> test precipitazione	La differenza tra il valore nella stazione target e i valori delle stazioni limitrofe (nei giorni -1, 0,1) supera una determinata soglia	Prec

Test di regressione

Questo test identifica per ogni mese di ciascun anno un modello di regressione tra la serie di temperatura (minima o massima) nella stazione target e una corrispondente serie di temperatura, stimata a partire dalle serie delle stazioni limitrofe opportunamente selezionate.

Le stazioni limitrofe (da un minimo di tre a un massimo di sette) vengono individuate in base alla loro completezza, alla distanza dalla stazione target (massimo 35 Km) e ad un indice di *agreement*, che valuta l'accordo tra le osservazioni di ogni stazione con quelle della stazione target.

Per ogni giorno dell'anno, il valore di temperatura nella stazione target viene stimato come media dei valori osservati nelle stazioni limitrofe, nella finestra di tre giorni centrata sul giorno in esame, pesati

con i coefficienti di regressione lineare semplice (rispetto alla serie target) e l'indice di *agreement*. I dati così stimati (variabile indipendente) vengono utilizzati per costruire un modello di regressione rispetto alla serie target (variabile dipendente). I residui della serie target rispetto al modello di regressione vengono quindi utilizzati per identificare eventuali dati anomali. Il valore osservato nella stazione target viene contrassegnato come dato anomalo se sia il residuo, sia il residuo standardizzato eccedono determinate soglie (rispettivamente 8°C per il residuo e 4 per il residuo standardizzato). Il test può essere applicato se il coefficiente di correlazione tra la serie osservata e la serie stimata è maggiore o uguale a 0.8. Per un approfondimento di questo test si rimanda all'articolo di Durre et al. (2010).

Un esempio di dato anomalo individuato dal test di regressione spaziale viene mostrato nella figura 8. La serie temporale della temperatura massima di una stazione con un valore contrassegnato dal test con un *flag* di dato dubbio viene mostrata insieme a quella delle stazioni limitrofe. Nella mappa di figura 9 è riportata la localizzazione della stazione target e di quelle limitrofe. La figura 10 illustra invece la serie osservata nella stazione target, insieme a quella stimata dal *regression* test, come media pesata dei valori di temperatura delle stazioni limitrofe (in tabella 5 sono riportati i coefficienti che concorrono al calcolo dei pesi). Poiché il coefficiente di correlazione tra serie osservata e serie stimata è uguale a 0.9, si procede con il calcolo dei residui e dei residui standardizzati, che risultano in questo caso entrambi superiori alla soglia (rispettivamente 9.1°C e 4.3).

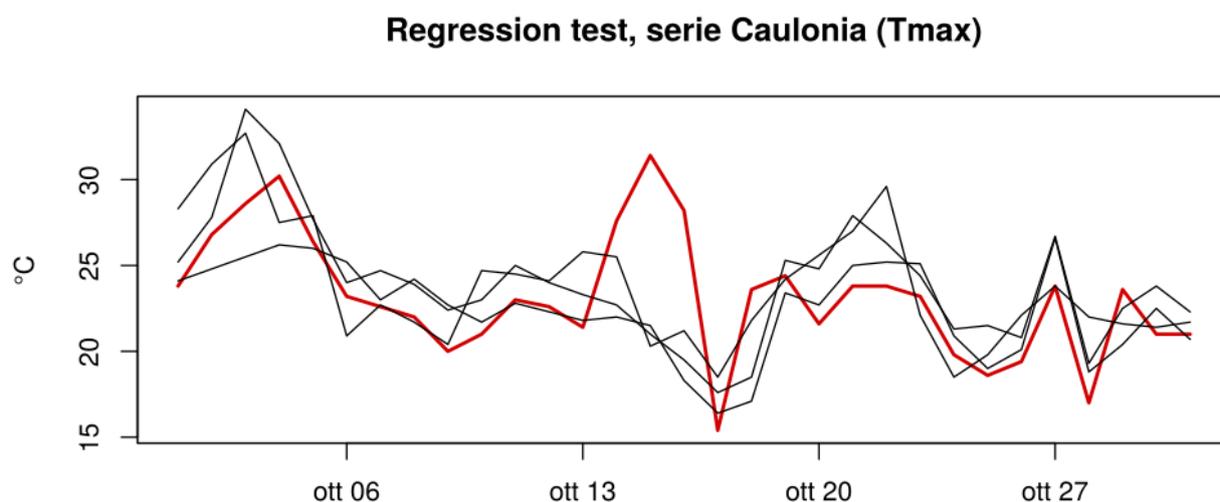


Figura 8: Serie di temperatura massima giornaliera ottobre 2003, nella stazione di Caulonia in Calabria (stazione target) e di tre stazioni limitrofe. Identificazione di un outlier il giorno 15 con il test di regressione spaziale. Il valore di temperatura di 31.4°C del giorno 15 è stato contrassegnato con flag di dato dubbio.



Figura 9: Localizzazione della stazione di Caulonia (in rosso) e delle stazioni limitrofe utilizzate dal test di regressione spaziale.

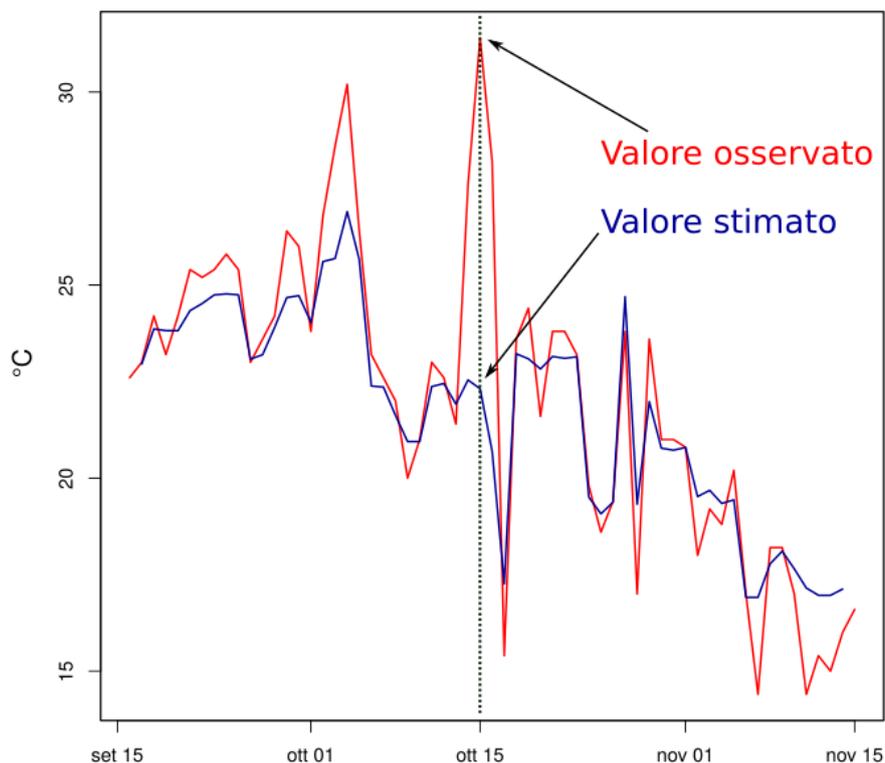


Figura 10: Temperatura massima nella stazione di Caulonia: serie osservata (in rosso) e serie stimata (media pesata dei valori osservati nelle stazioni limitrofe).

Tabella 5: Temperatura massima nelle stazioni limitrofe alla stazione target, nel giorno 15 ottobre 2003 e relativi coefficienti di agreement e di regressione (slope e intercetta).

Stazione	Tmax 15/10/2003 (°C)	Indice di agreement	Slope	Intercetta (°C)
Monasterace - Punta Stilo	21.5	0.73	0.99	0.23
Antonimina	21.0	0.59	0.84	2.31
Feroletto della Chiesa	20.3	0.60	0.74	4.59

Corroboration test

Il *corroboration* test verifica se il valore da testare cade significativamente fuori dall'intervallo dei valori registrati nelle stazioni vicine.

Le stazioni limitrofe (da un minimo di tre a un massimo di sette) in questo caso sono selezionate solo in base alla loro maggiore vicinanza alla stazione target e alla disponibilità di dati; si ammette una distanza massima dalla stazione target di 20 Km e una differenza massima di quota di 100 m. Ogni osservazione della stazione target viene confrontata con le osservazioni delle stazioni limitrofe all'interno di una finestra di tre giorni centrata sul giorno in esame.

Per la temperatura minima e massima il test confronta le anomalie rispetto alla media a lungo termine. Il calcolo di quest'ultima è analogo a quello relativo al test per individuare gli *outlier* basato sullo *z-score*. Un valore di temperatura viene contrassegnato con un *flag* di dato dubbio se la relativa anomalia si discosta di almeno 10°C dalle anomalie di tutte le stazioni limitrofe, cioè se nessuna anomalia di temperatura delle stazioni limitrofe, selezionate nella finestra di tre giorni, corrobora

l'anomalia di temperatura osservata nella stazione target.

Il *corroboration* test e il controllo di regressione spaziale per la temperatura si completano a vicenda in un modo importante. Il *corroboration* test può essere applicato in giorni e luoghi in cui il controllo di regressione spaziale non può essere utilizzato, per esempio per mancanza di serie sufficientemente complete per il calcolo della regressione, oppure nei casi in cui si ottiene una correlazione troppo bassa fra la serie osservata della stazione target e quella stimata. D'altra parte il test di regressione spaziale, quando applicabile, riesce a rilevare inconsistenze spaziali di ampiezza minore rispetto a quelle che riesce ad individuare il *corroboration* test.

Un esempio di applicazione del *corroboration* test ad una serie di temperatura è riportato nella figura 11, dove viene mostrato l'andamento della temperatura minima giornaliera di dicembre 2003 delle stesse stazioni (target e limitrofe) di figura 9. Il test identifica come dato anomalo il dato registrato il giorno 11 nella stazione target. Come mostrato nella tabella 6, infatti, l'anomalia rispetto alla media di lungo periodo nella stazione target (15.4°C) si discosta più di 10°C da quelle delle stazioni limitrofe.

Nel caso della precipitazione, il *corroboration* test viene applicato ai totali giornalieri. Il valore osservato nella stazione target viene contrassegnato con un *flag* di dato dubbio se si discosta dal più basso o dal più alto valore registrato nelle stazioni limitrofe di una quantità superiore ad una soglia prefissata. Questa soglia dipende in maniera inversa dalla differenza tra i ranghi climatologici percentuali del totale di precipitazione della stazione target e delle stazioni limitrofe (per un approfondimento si veda Durre et al., 2010). Questa dipendenza implica che il test accetti differenze più elevate fra il valore della precipitazione della stazione target e i valori nelle stazioni limitrofe quando la differenza fra i ranghi percentuali è piccola (cioè quando si verifica un episodio di pioggia forte su tutte le stazioni), rispetto al caso in cui la differenza fra i ranghi percentuali è grande (cioè quando si registra un totale elevato di precipitazione solo nella stazione target).

In figura 12 è mostrato un esempio di individuazione di un *outlier* nella serie di precipitazione della stazione di Viganego, in Liguria. Le stazioni limitrofe utilizzate per il *corroboration* test sono riportate nella mappa di figura 13. Il valore di precipitazione giornaliera di 403.6 mm del giorno 12 dicembre 2003 viene contrassegnato da *flag* di dato dubbio perché si discosta di 291.6 mm dal più alto valore registrato nelle stazioni limitrofe nei giorni 11, 12 e 13, come riportato nella tabella 7 (in questo caso la soglia è uguale a 269.24 mm).

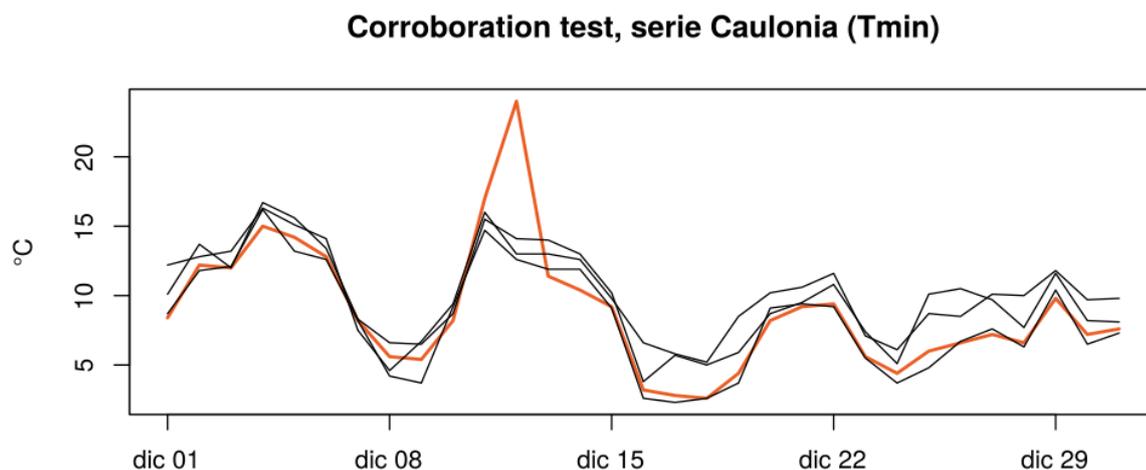


Figura 11: Serie di temperatura minima giornaliera dicembre 2003, nella stazione di Caulonia in Calabria (stazione target) e di tre stazioni limitrofe. Identificazione di un outlier il giorno 11 con il *corroboration* test. Il valore di temperatura di 24°C è stato contrassegnato con un *flag* di dato dubbio.

Tabella 6: Valore di temperatura minima del giorno 12 dicembre 2003, media climatologica e relativa anomalia per la stazione target e le limitrofe.

Stazione	Tmin 12/12/2003 (°C)	Media climatologica (°C)	Anomalia (°C)
Caulonia (target)	24.0	8.7	15.4
Monasterace - Punta Stilo	13.0	10.4	2.6
Antonimina	14.1	10.6	3.5
Feroletto della Chiesa	12.6	8.5	4.1

Corroboration test, serie Viganego (Prec)

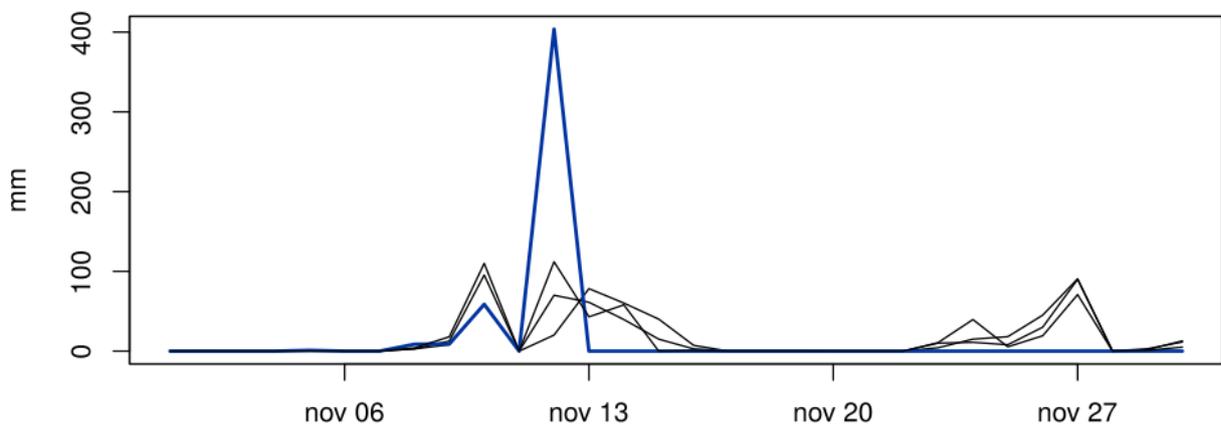


Figura 12: Serie di precipitazione giornaliera di novembre 1961, nella stazione di Viganego in Liguria (stazione target) e di tre stazioni limitrofe. Identificazione di outlier il giorno 12 con il corroboration test. Il valore di precipitazione di 403.6 mm è stato contrassegnato con un flag di dato dubbio.

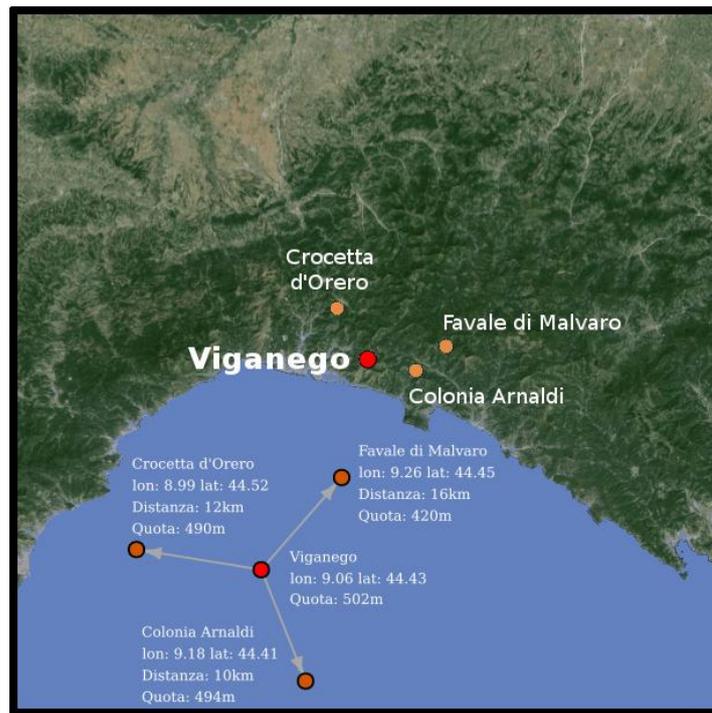


Figura 13: Localizzazione della stazione di Viganego (in rosso) e delle stazioni limitrofe utilizzate dal corroboration test.

Tabella 7: Precipitazione cumulata giornaliera del giorno 12 novembre 1961 nella stazione target e dei giorni 11, 12 e 13 nelle stazioni limitrofe.

Stazione	Prec 11/11/1961 (mm)	Prec 12/11 (mm)	Prec 13/11 (mm)
Viganego (target)		403.6	
Colonia Arnaldi	0	20.1	78.3
Crocetta d'Orero	0	70.0	61.2
Favale di Malvaro	0	112.0	43.0

3. APPLICAZIONE DELLA PROCEDURA: STATISTICHE RIASSUNTIVE

3.1 Dati

Il dataset climatico nazionale con passo giornaliero in corso di realizzazione, al quale è stata applicata la procedura di controllo di qualità dei dati, è costituito, alla data di pubblicazione del presente rapporto, dalle serie di 2650 stazioni di temperatura (massima e minima) e 4740 stazioni di precipitazione. Le fonti dei dati sono riportate in tabella 8. Le stazioni sono distribuite sul territorio nazionale in maniera abbastanza uniforme, ad eccezione di due regioni. Nelle mappe delle figure 14 e 15, sono riportate le stazioni con disponibilità di serie di temperatura (minima e massima) e di precipitazione. Il numero totale di stazioni con dati disponibili per ogni anno viene riportato nelle figure 16 (temperatura) e 17 (precipitazione): per ogni anno sono state incluse nel conteggio solo le stazioni che hanno almeno il 90% di dati validi. Come mostrato nelle figure le osservazioni più datate risalgono al decennio 1910-1920. Per la temperatura si osserva un aumento nel tempo del numero di stazioni disponibili, con un totale di più di 1000 stazioni dopo il 2000.

Il numero di stazioni con dati di precipitazione è molto più elevato rispetto a quelle di temperatura, con circa 1000 stazioni disponibili già nel 1930, che aumentano negli anni (a parte l'intervallo 1942-1950) per arrivare a più di 1500 stazioni a partire dal 1951. Una parte consistente delle stazioni origina dall'ex Servizio Idrografico e Mareografico Nazionale (SIMN), istituito nel 1917. Queste stazioni (con strumentazione meccanica) sono state integrate o sostituite dalle nuove reti regionali, costituite da strumenti automatici, a partire dagli anni '90 e successivamente, con il trasferimento delle competenze del SIMN alle Regioni (nel 2002), sono state progressivamente sostituite nel tempo da stazioni automatiche.

La copertura temporale dei dati per regione e per la rete sinottica nazionale viene riportata nelle figure 18 (temperatura minima e massima) e 19 (precipitazione).

Tabella 8: Fonti di dati che contribuiscono al dataset di temperatura e precipitazione con passo giornaliero.

Fonte	Sito Web
Servizio Idrografico e Mareografico – Regione Abruzzo	http://www.regione.abruzzo.it/xIdrografico/
Centro Funzionale – Regione Autonoma Valle d’Aosta	http://cf.regione.vda.it/
Protezione Civile – Regione Basilicata	http://www.protezionecivilebasilicata.it/protcivbas/home.jsp
Provincia Autonoma di Bolzano – Alto Adige	http://www.provincia.bz.it/meteo/
Arpacal, Centro Funzionale multirischi – Regione Calabria	http://www.cfd.calabria.it/
Protezione Civile – Regione Campania	http://centrofunzionale.regione.campania.it/
Arpae Emilia Romagna	http://www.arpae.it/
Protezione Civile – Regione Autonoma Friuli Venezia Giulia	http://www.protezionecivile.fvg.it/
Centro Funzionale Regionale – Regione Lazio	http://www.idrografico.roma.it/
Repertorio Cartografico – Regione Liguria	http://www.cartografiarl.regione.liguria.it/CartoWebNet/CwMappa.aspx
ARPA Lombardia	http://www2.arpalombardia.it/siti/arpalombardia/meteo
Protezione Civile – Regione Marche	http://www.protezionecivile.marche.it/
Protezione Civile – Regione Molise	http://www.protezionecivile.molise.it/index.php/centro-funzionale.html
Centro Funzionale – ARPA Piemonte	https://www.arpa.piemonte.gov.it/rischinaturali/centro-funzionale
Protezione Civile – Regione Puglia	http://www.protezionecivile.puglia.it/centro-funzionale
Osservatorio Acque – Regione Sicilia	http://www.sias.regione.sicilia.it/
Settore Idrologico Regionale – Regione Toscana	http://www.sir.toscana.it/
Servizio Idrografico regionale – Regione Umbria	http://www.regione.umbria.it/ambiente/servizio-idrografico
ARPAV – Dipartimento Regionale per la Sicurezza del Territorio	http://www.arpa.veneto.it/temi-ambientali/idrologia/
Centro Funzionale Protezione Civile – Provincia Autonoma di Trento	http://www.meteotrentino.it/
Servizio Meteorologico dell’Aeronautica Militare (Rete sinottica)	https://www.ncdc.noaa.gov/data-access/
Sistema Informativo Idrologico HIScentral	http://www.hiscentral.isprambiente.gov.it/hiscentral/default.aspx

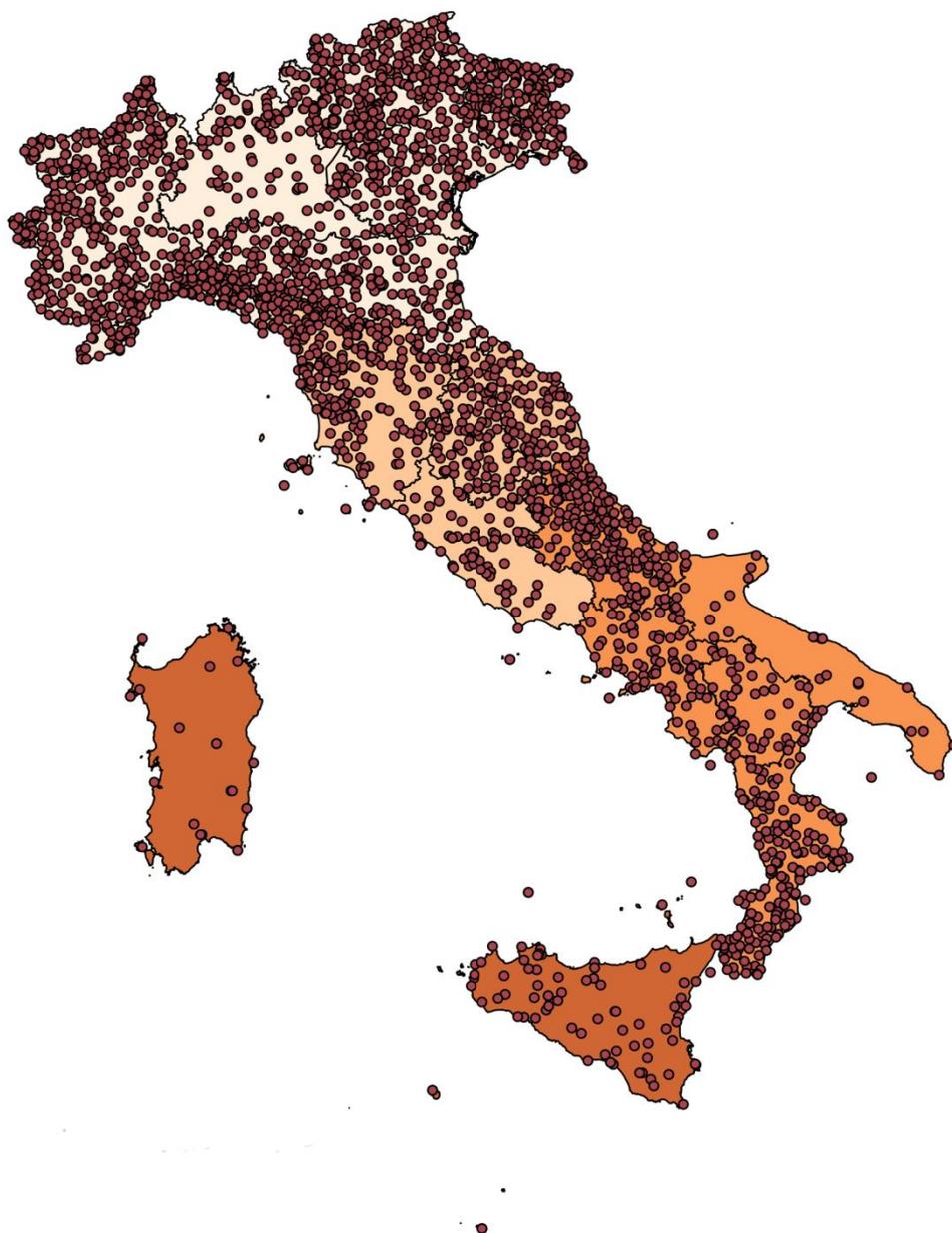


Figura 14: *Stazioni di temperatura minima e massima.*

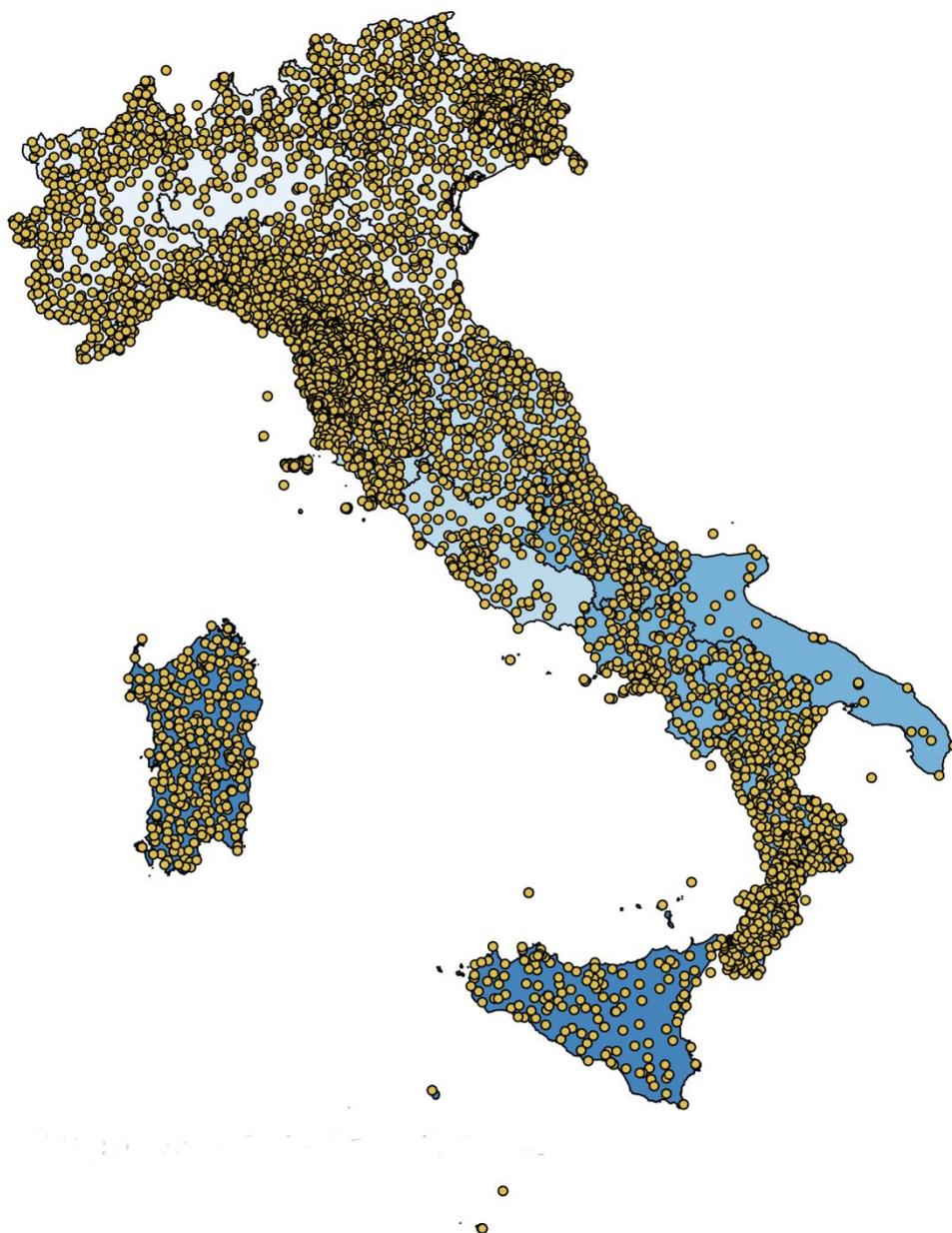


Figura 15: *Stazioni di precipitazione cumulata.*

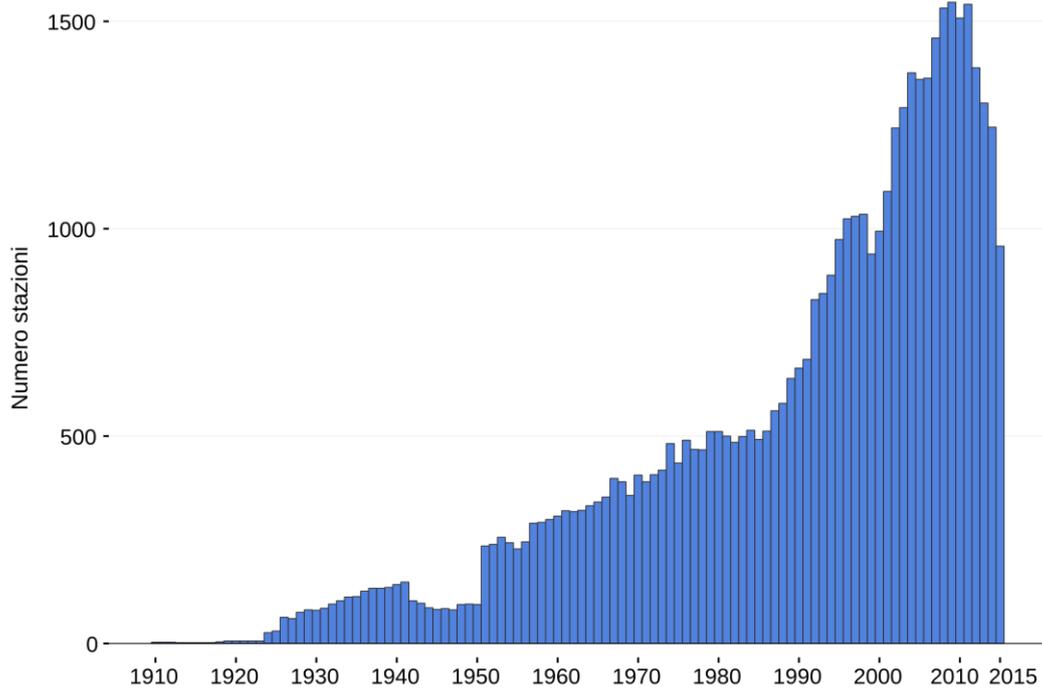


Figura 16: *Numero di stazioni di temperatura massima e minima disponibili per ogni anno.*

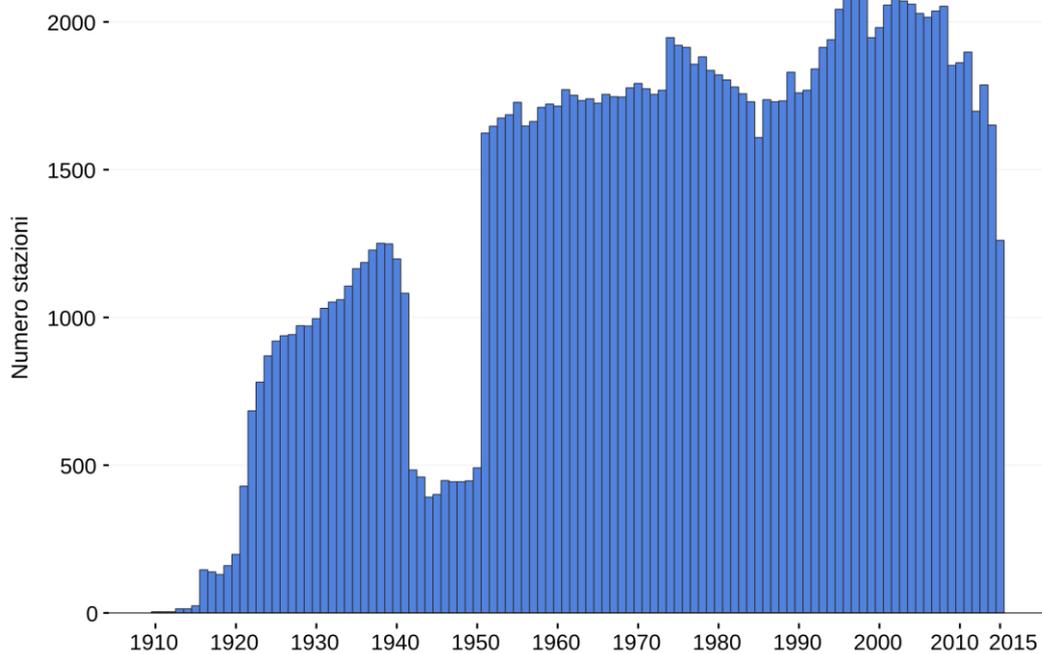


Figura 17: *Numero di stazioni di precipitazione disponibili per ogni anno.*

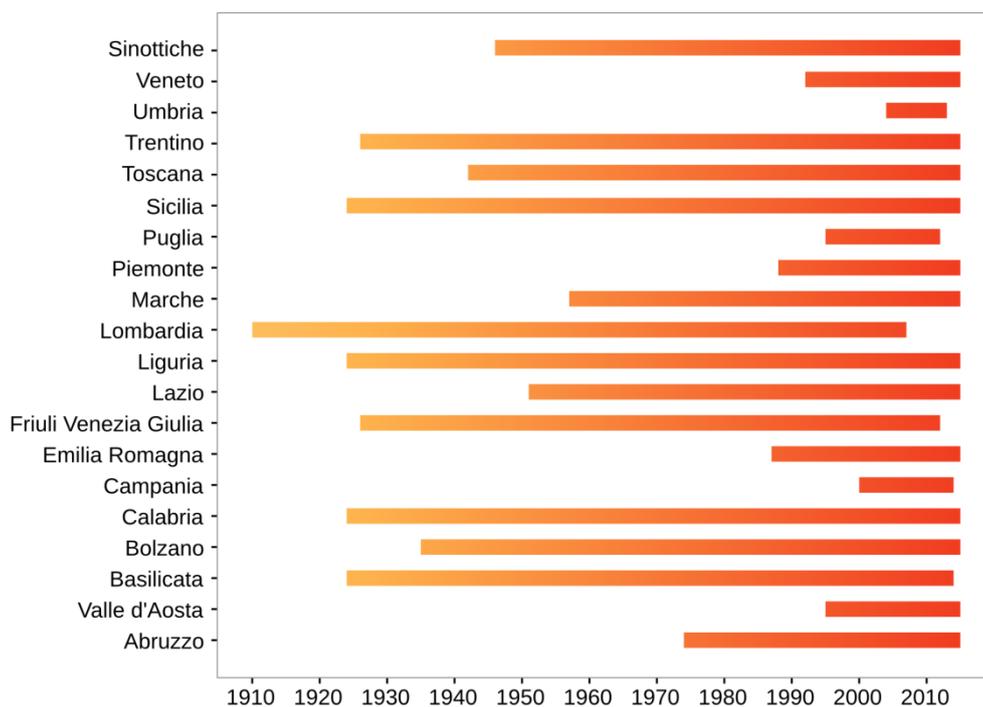


Figura 18: Copertura temporale dei dati di temperatura per ogni regione e per la rete sinottica.

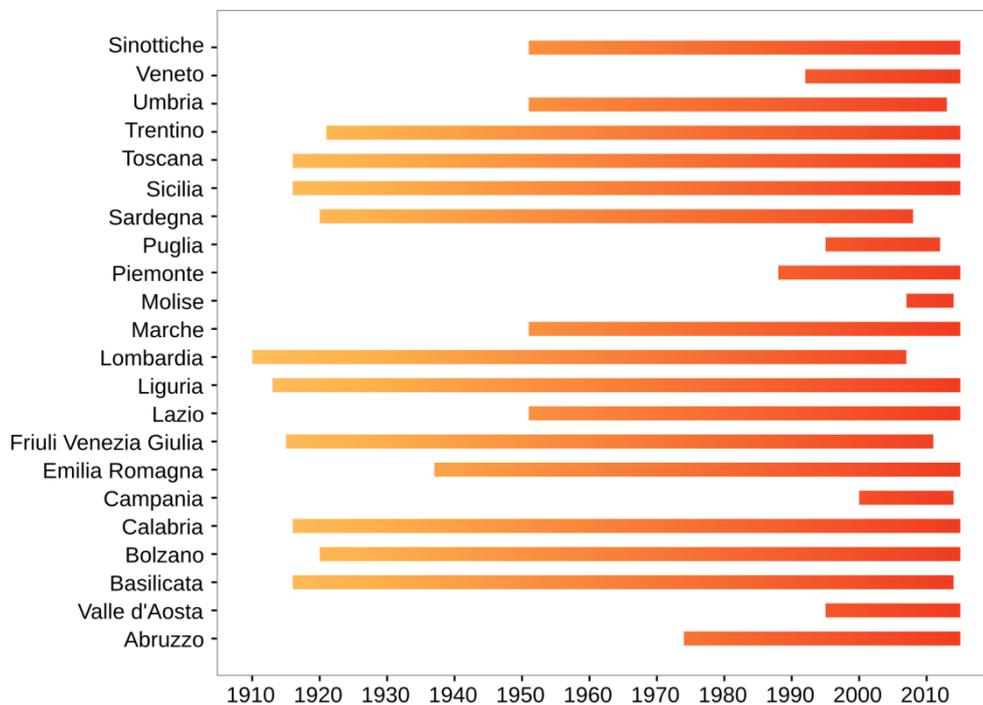


Figura 19: Copertura temporale dei dati di precipitazione per ogni regione e per la rete sinottica.

3.2 Risultati dei controlli di qualità

La procedura automatica di controllo di qualità è stata applicata all'intero dataset di serie giornaliere, escludendo in questa prima fase i controlli di consistenza spaziale, che prevedono anche una valutazione soggettiva dei dati anomali individuati da parte di un operatore. I risultati sono stati esaminati e sintetizzati per ogni tipologia di controllo.

Complessivamente sono stati contrassegnati da *flag* di dato errato, e quindi esclusi dal dataset finale, un totale di 124165 valori, pari allo 0.13% delle osservazioni giornaliere (tabella 9). La percentuale maggiore di errori è stata individuata dai test di integrità di base, sia per la temperatura minima e massima (0.06%) che per la precipitazione (0.18%), in sintonia con quanto evidenziato da studi analoghi (Brunet et al., 2008). Per le serie di precipitazione il numero maggiore di *flag* di dato errato (93357 valori, pari allo 0.16% del totale) proviene dal controllo di persistenza degli zeri (180 o più zeri ripetuti), che individua lunghe sequenze di dati mancanti, erroneamente indicati con zero.

Tabella 9: Risultati dei controlli di qualità per ogni gruppo di test e per ogni parametro (percentuale di dati errati rispetto al totale di dati esaminati).

Gruppo di procedure	Flag (%)	Tmin	Flag (%)	Tmax	Flag (%)	Prec	Totale flag %
Test di integrità di base	0.060		0.058		0.1824		0.132
Test per identificare i valori anomali	0.003		0.003		0.0005		0.002
Test di consistenza interna e temporale	0.006		0.005				0.002
Totale dati errati	0.069		0.068		0.1829		0.132

I controlli per identificare gli *outlier* restituiscono le percentuali di dati errati più basse fra tutti i test. Questo risultato è dovuto al fatto che questi controlli sono stati predisposti con limiti di tolleranza alti, con l'obiettivo di eliminare dal dataset, tramite procedura automatica, solo i dati palesemente errati o molto improbabili da un punto di vista climatico, limitando in questo modo il rischio di falsi positivi. Nella figure 20, 21 e 22 viene mostrata la distribuzione statistica delle precipitazioni giornaliere per ogni regione. In ogni grafico, il boxplot di sinistra comprende anche i dati anomali identificati dal test del 95° percentile (in rosso). La maggior parte degli *outlier* corrisponde a valori estremi o elevati per il mese dell'anno in cui sono stati registrati. Il boxplot a destra evidenzia le stesse distribuzioni, dopo l'eliminazione degli outlier dal dataset.

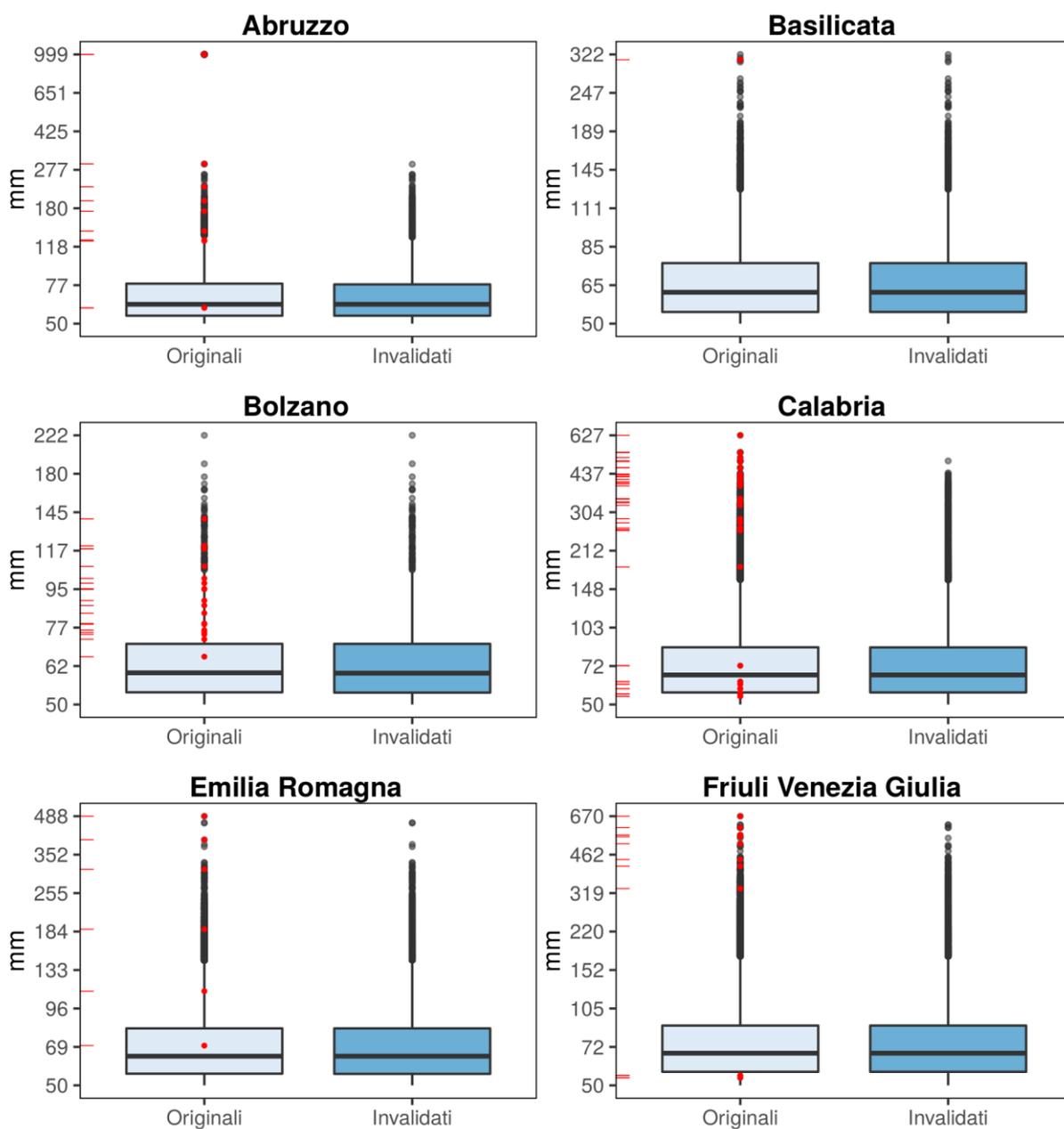


Figura 20: Box-plot delle precipitazioni giornaliere per regione (Abruzzo, Basilicata, Bolzano, Calabria, Emilia Romagna, Friuli Venezia Giulia): dati originali (a sinistra), dove in rosso sono riportati i dati anomali identificati dal test del 95° percentile, e dopo l'eliminazione degli outlier (a destra).

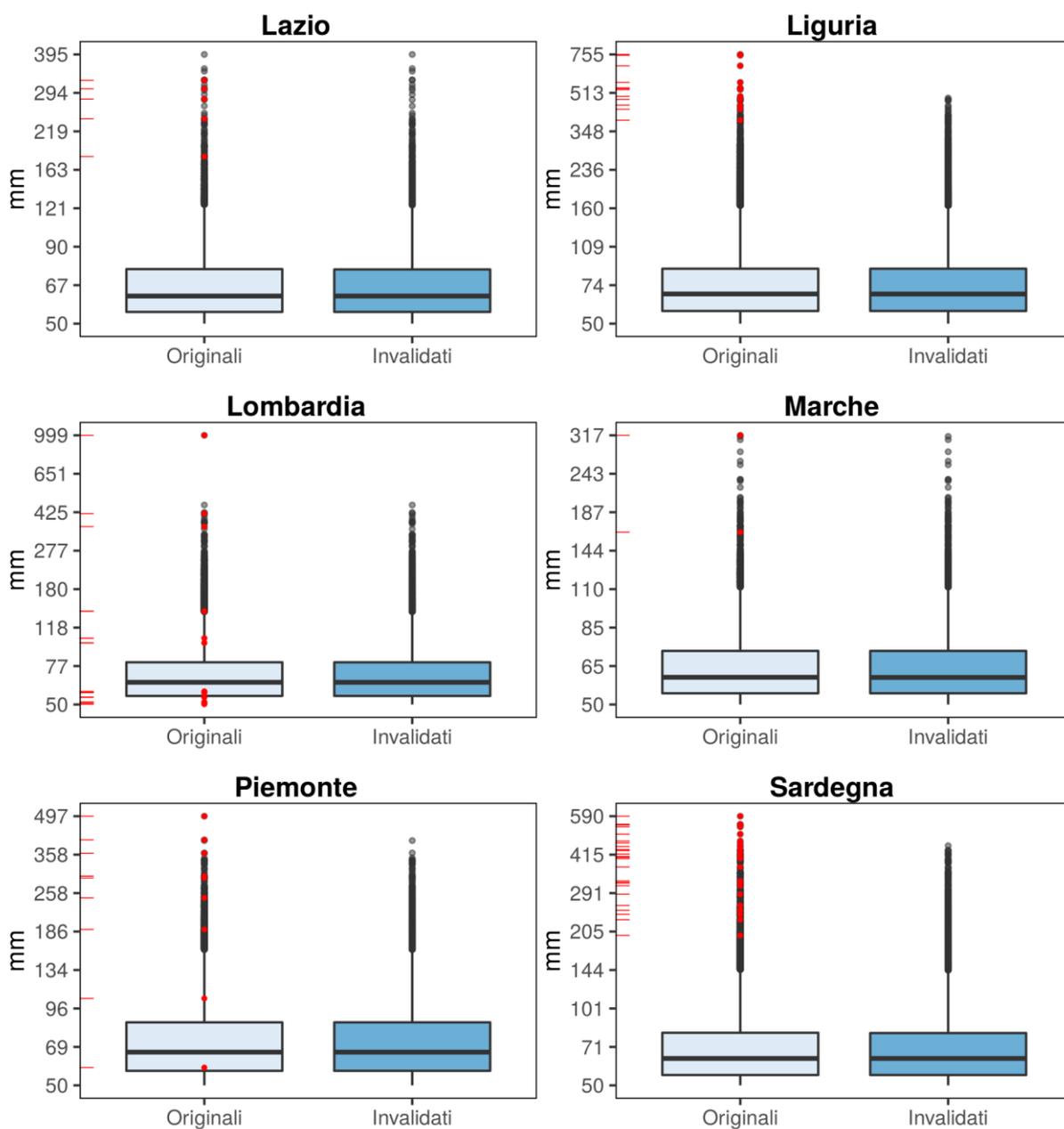


Figura 21: Come in figura 20, per Lazio, Liguria, Lombardia, Marche, Piemonte e Sardegna.

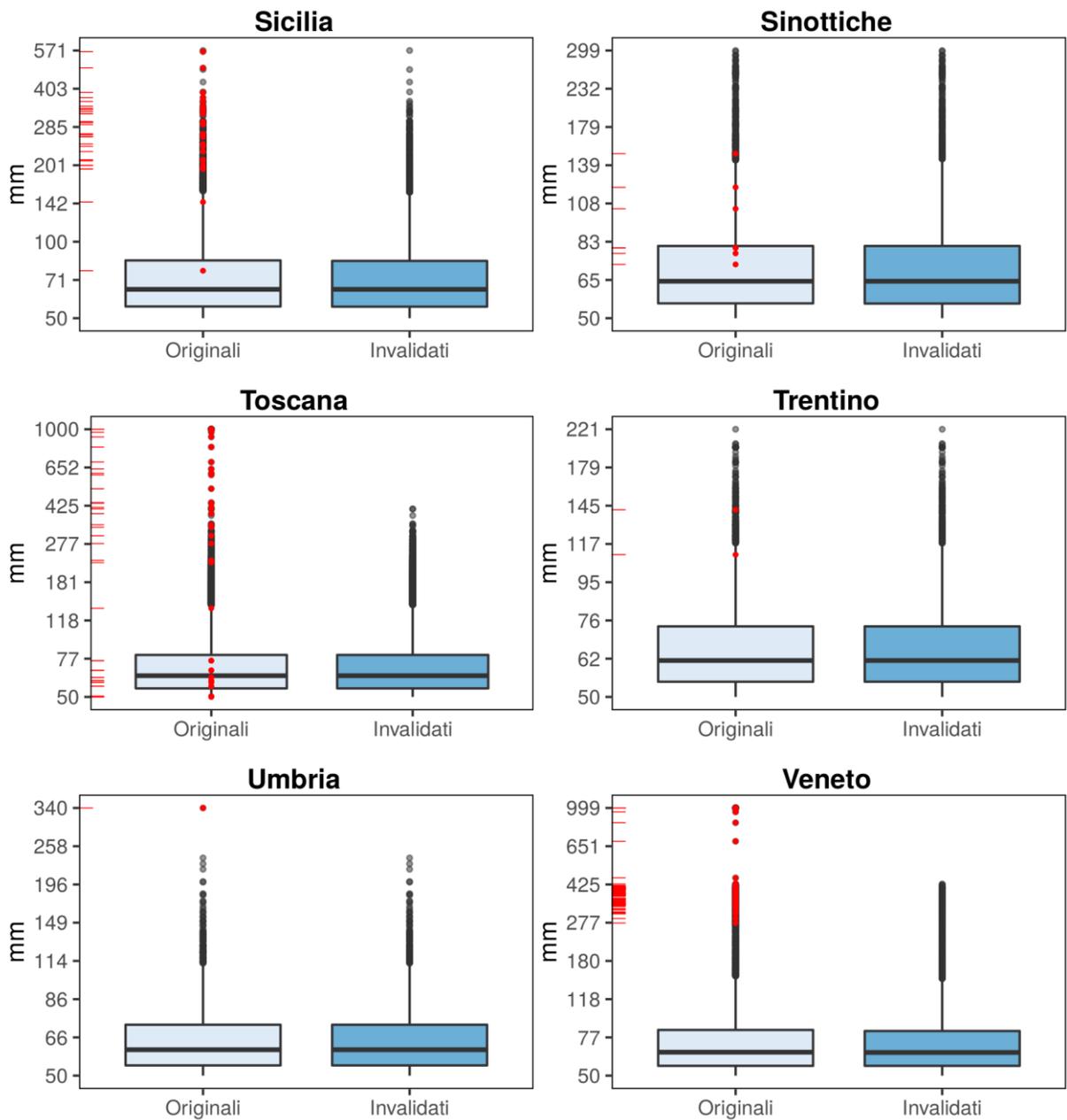


Figura 22: Come in figura 20, per Sicilia, stazioni sinottiche, Toscana, Trentino, Umbria, Veneto.

4. CONCLUSIONI

I controlli di qualità delle serie temporali di dati climatici sono propedeutici a qualsiasi utilizzo delle serie stesse. In particolare, nel nostro caso è stata sviluppata una procedura standardizzata di controllo delle serie temporali con passo giornaliero di temperatura e precipitazione. La sequenza di controlli è stata applicata alle serie pluridecennali ricostruite a partire da dati provenienti da fonti e reti di osservazione diverse e che hanno operato in periodi diversi.

I risultati dell'applicazione dei test di controllo mostrano che complessivamente sono stati contrassegnati da *flag* di dato errato un totale di 124165 valori, pari allo 0.13% delle osservazioni giornaliere di cui si dispone attualmente. La percentuale maggiore di errori è stata individuata dai test di integrità di base, sia per la temperatura minima e massima (0.06%) che per la precipitazione (0.18%). Per le serie di precipitazione il numero maggiore di *flag* di dato errato (93357 valori, pari allo 0.16% del totale) proviene dal controllo di persistenza degli zeri (180 o più zeri ripetuti), che individua lunghe sequenze di dati mancanti, erroneamente indicati con zero.

Si ritiene che l'eliminazione dalle serie dei dati errati, nonché il riconoscimento e la verifica degli *outlier* (dati sospetti) con procedure semiautomatiche o manuali, sia di grande valore ed anzi indispensabile per il successivo calcolo delle statistiche, degli indici e degli indicatori di variazione del clima in Italia. Pertanto, l'applicazione della procedura di controllo di qualità al dataset con passo giornaliero in corso di realizzazione, sarà replicata sistematicamente in fase di aggiornamento e integrazione dei dati con cadenza annuale.

BIBLIOGRAFIA

- Alexander L.V., Zhang X., Peterson T.C., Caesar J., Gleason B., Tank A., Haylock M., Collins D., Trewin B., Rahimzadeh F., Tagipour A., Kumar K.R., Revadekar J., Griffiths G., Vincent L., Stephenson D.B., Burn J., Aguilar E., Brunet M., Taylor M., New M., Zhai P., Rusticucci M., Vazquez-Aguirre J.L., 2006. Global observed changes in daily climate extremes of temperature and precipitation. *J. Geophys. Res.*, 111: D05109. doi:10.1029/2005JD006290.
- Arguez A., Durre I., Applequist S., Vose R.S., Squires M.F., Yin X., Heim R.R. Jr. and Owen T.W., 2012. NOAA's 1981-2010 U.S. Climate Normals: An Overview. *Bull. Amer. Meteor. Soc.*, 93, 1687-1697, doi:10.1175/BAMS-D-11-00197.1.
- Bartholy J., Pongracz R., 2007. Regional analysis of extreme temperature and precipitation indices for the Carpathian Basin from 1946 to 2001. *Global Planet Change* 57: 83-95. doi:10.1016/j.gloplacha.2006.11.002.
- Brunet M., Saladie O., Jones P.D., Sigro J., Aguilar E., Moberg A., Lister D., Walther A. and Almaraz C., 2008. A case study/guidance on the development of long-term daily adjusted temperature datasets. WCDMP-66/WMO-TD-1425, Geneva.
- Desiato F., Lena F. e Toreti A., 2006. Un sistema per tutti – climatologia: i dati italiani. *Sapere*, Anno 72, n. 2, 62-69.
- Desiato F., Lena F. e Toreti A., 2007. SCIA: a system for a better knowledge of the Italian climate, *Bollettino di Geofisica Teorica ed Applicata*, 48, n. 3, 351-358.
- Desiato F., 2007. SCIA. un sistema nazionale unico per i dati climatologici, *ARPA Rivista*, N. 1, gennaio-febbraio 2007.
- Desiato F., Fioravanti G., Frascchetti P., Perconti W. e Toreti A., 2011. Climate indicators for Italy: calculation and dissemination, *Adv. Sci. Res.*, 6, 147-150.
- Donat M.G., Alexander L.V., Yang H., Durre I., Vose R., Dunn R., Willett K., Aguilar E., Brunet M., Caesar J., Hewitson B., Jack C., Klein Tank A.M.G., Kruger A.C., Marengo J., Peterson T.C., Renom M., Oria Rojas C., Rusticucci M., Salinger J., Sekele S.S., Srivastava A.K., Trewin B., Villarroya C., Vincent L.A., Zhai P., Zhang X., Kitching S., 2013. Updated analyses of temperature and precipitation extreme indices since the beginning of the twentieth century: The HadEX2 dataset, *J. Geophys. Res.* 118: 2098-2118. doi: 10.1002/jgrd.50150.
- Durre I., Menne M.J., Gleason B.E., Houston T.G. and Vose R.S., 2010. Comprehensive automated quality assurance of daily surface observations, *J. Appl. Meteor. Climatol.*, 8, 1615–1633.
- Efthymiadis D., Goodess C.M., Jones P.D., 2011. Trends in Mediterranean gridded temperature extremes and large-scale circulation influences. *Nat. Hazards Earth Syst. Sci.* 11: 2199-2214. doi:10.5194/nhess-11-2199-2011.
- Fioravanti G., Piervitali E., Desiato F., 2016. Recent changes of temperature extremes over Italy: an index-based analysis. *Theor. Appl. Climatol.*, 123, 473-486. doi:10.1007/s00704-014-1362-1.
- Hartmann D.L., Klein Tank A.M.G., Rusticucci M., Alexander L.V., Brönnimann S., Charabi Y., Dentener F.J., Dlugokencky E.J., Easterling D.R., Kaplan A., Soden B.J., Thorne P.W., Wild M., Zhai P.M., 2013. Observations: atmosphere and surface. In: Stocker T.F., Qin D., Plattner G.-K., Tignor M., Allen S.K., Boschung J., Nauels A., Xia Y, Bex V., Midgley P.M. (eds.) *Climate change 2013: the physical science basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press, Cambridge, pp 159–254.
- Hubbard K.G., Guttman N.B., You J., and Chen Z., 2007. An improved QC process for temperature in the daily cooperative weather observations. *J. Atmos. Oceanic Technol.*, 24, 201–213.
- Lanzante J. R., 1996. Resistant, robust and non-parametric techniques for the analysis of climate data: theory and examples, including applications to historical radiosonde station data. *Int. J. Climatol.*, 16: 1197–1226. doi: 10.1002/(sici)1097-0088(199611)16:11<1197::aid-joc89>3.0.co;2-l.
- Toreti A., Desiato F., 2008. Changes in temperature extremes over Italy in the last 44 years. *Int. J. Climatol.* 28: 733-745. doi: 10.1002/joc.1576.
- WMO, 2011. *Guide to Climatological Practices*. WMO-No.100. Geneva.