



ISPRA

Istituto Superiore per la Protezione
e la Ricerca Ambientale

Elaborazione delle serie temporali per la stima delle tendenze climatiche



STATO DELL'AMBIENTE

Elaborazione delle serie temporali per la stima delle tendenze climatiche

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.

La Legge 133/2008 di conversione, con modificazioni, del Decreto Legge 25 giugno 2008, n. 112, pubblicata sulla Gazzetta Ufficiale n. 195 del 21 agosto 2008, ha istituito l'ISPRA - Istituto Superiore per la Protezione e la Ricerca Ambientale.

L'ISPRA svolge le funzioni che erano proprie dell'Agenzia per la Protezione dell'Ambiente e per i servizi Tecnici (ex APAT), dell'Istituto Nazionale per la Fauna Selvatica (ex INFS) e dell'Istituto Centrale per la Ricerca scientifica e tecnologica Applicata al Mare (ex ICRAM).

La presente pubblicazione fa riferimento ad attività svolte in un periodo antecedente l'accorpamento delle tre Istituzioni e quindi riporta ancora, al suo interno, richiami e denominazioni relativi ai tre Enti soppressi.

ISPRA – Istituto Superiore per la protezione e la ricerca ambientale
Via Vitaliano Brancati, 48 – 00144 Roma
www.isprambiente.it

ISPRA, Stato dell'Ambiente 32/2012

ISBN 978-88-448-0559-3

Riproduzione autorizzata citando la fonte

Elaborazione grafica

ISPRA

Grafica di copertina: Franco Iozzoli

Foto di copertina: Franco Desiato

Coordinamento tipografico:

Daria Mazzella

ISPRA - Settore Editoria

Luglio 2012

Autori

Franco Desiato, Guido Fioravanti, Piero Frascchetti, Walter Perconti, Emanuela Piervitali (ISPRA)

Indice

| | |
|---|----|
| 1. Introduzione | 6 |
| 2. Criteri e controlli di validita' dei dati e degli indicatori climatici | 7 |
| 3. Selezione delle serie temporali | 10 |
| 4. Omogeneizzazione delle serie di temperatura | 16 |
| 4.1 Introduzione | 16 |
| 4.2 Metodi di omogeneizzazione..... | 17 |
| 4.3 Omogeneizzazione delle serie del sistema SCIA | 18 |
| 5. Calcolo dei valori normali e delle anomalie..... | 25 |
| 5.1 Temperatura | 25 |
| 5.2 Precipitazione | 27 |
| 6. Stima delle tendenze | 30 |
| 6.1 Temperatura | 31 |
| 6.2 Precipitazione | 34 |
| Bibliografia..... | 36 |

1. INTRODUZIONE

Per favorire l'elaborazione, la rappresentazione e la disponibilità di indicatori sul clima in Italia, l'ISPRA ha realizzato da tempo il Sistema nazionale per la raccolta, l'elaborazione e la diffusione di dati Climatologici di Interesse Ambientale (SCIA), in collaborazione con il Servizio Meteorologico dell'Aeronautica Militare, l'Unità di Ricerca per la Climatologia e la Meteorologia applicate a l'Agricoltura (CMA-CRA), numerose Agenzie Regionali per la Protezione dell'Ambiente e i Servizi Agrometeorologici Regionali della Sicilia e delle Marche.

Attraverso l'elaborazione delle serie temporali di dati misurati da diverse reti osservative, il sistema SCIA (Desiato 2007; Desiato et.al., 2006, 2007, 2011) rende disponibili i valori statistici decadali, mensili e annuali (denominati convenzionalmente indicatori) di numerose variabili meteo-climatiche; esso consente inoltre di aggiornare periodicamente le informazioni con una procedura standardizzata. Le informazioni prodotte da SCIA sono accessibili attraverso un sito web dedicato nel portale del Sistema Nazionale Conoscitivo e di Informazione Ambientale (SINANET) dell'ISPRA, all'indirizzo: www.scia.sinanet.apat.it.

Sulla base delle serie temporali di indicatori elaborati dal sistema SCIA è stato possibile, a partire dal 2006, calcolare e presentare numerosi indicatori climatici aggregati, rappresentativi delle tendenze del clima in Italia negli ultimi 50 anni. Tali indicatori sono alla base delle informazioni diffuse annualmente attraverso i rapporti ISPRA – Stato dell'Ambiente della serie “Gli indicatori del clima in Italia” (ISPRA 2011a il più recente) e, per quanto attiene al capitolo “Cambiamenti climatici ed energia”, attraverso l'Annuario di dati ambientali dell'ISPRA (ISPRA 2011b il più recente).

Il passaggio dalle serie di dati rilevate dalle reti osservative alla stima delle tendenze delle variabili climatiche, comporta l'adozione e l'applicazione di criteri rigorosi di controllo, selezione ed elaborazione delle serie temporali stesse. Questi criteri, e il risultato della loro applicazione agli indicatori climatici italiani, vengono descritti nel presente rapporto, suddivisi in capitoli che rappresentano in sequenza le diverse fasi di elaborazione.

2. CRITERI E CONTROLLI DI VALIDITA' DEI DATI E DEGLI INDICATORI CLIMATICI

In assenza di una terminologia condivisa nel contesto della elaborazione dei dati climatici, in questo rapporto vengono adottate convenzionalmente le seguenti definizioni.

Con il termine dati si intendono i valori delle osservazioni climatiche rilevate dalle reti di monitoraggio, sottoposti ai controlli di validità e archiviati autonomamente dagli organismi nazionali o regionali titolari delle reti e delle proprie basi dati. Con le serie temporali di dati messe a disposizione dalle fonti vengono alimentati i programmi di calcolo degli indicatori elaborati dal sistema SCIA.

Con il termine indicatori si intendono quindi i valori statistici (medie, estremi, frequenze, ecc.) delle variabili climatiche riferite a una stazione di misura in un determinato intervallo di tempo, tipicamente un mese o un anno. Per esempio, sono indicatori la temperatura media del mese di gennaio di Roma Ciampino e la precipitazione massima giornaliera nel 2010 di Milano Linate.

Con il termine indici si intendono i valori delle grandezze aggregate sia in senso temporale che spaziale, calcolati applicando agli indicatori opportuni criteri, metodi o modelli statistici. Per esempio, sono indici i valori climatologici normali di temperatura o precipitazione in un trentennio di riferimento e le serie di anomalie annuali, sia riferite a una singola stazione che rappresentative della media sull'intero territorio nazionale.

Adottando queste definizioni, nel presente capitolo sono descritti i controlli e i criteri di validità dei dati e degli indicatori climatici che vengono applicati nel sistema SCIA.

Controlli di validità dei dati in ingresso

In generale, i controlli di validità dei dati rientrano nei compiti e nelle attività delle fonti dei dati stessi. In altre parole, il sistema SCIA si basa sul presupposto che i dati degli organismi titolari delle reti di osservazione meteorologica con cui viene alimentato il sistema siano validi, o comunque siano riconoscibili i dati non validi come esito di controlli di validità effettuati dalla fonte. In quest'ultimo caso l'esito dei controlli è rappresentato da un *flag* di validità (uno per dato) presente sui file di ingresso. Il programma di calcolo degli indicatori di SCIA prende in considerazione, sia per la determinazione del valore dell'indicatore decadale, mensile o annuale, sia per il conteggio dei dati utili alla determinazione del valore stesso, solo i dati con *flag* di dato valido.

Per quanto riguarda i dati sinottici (derivati dai messaggi triorari SYNOP e riepilogativi giornalieri SYREP), il programma di calcolo degli indicatori SCIA procede a un'ulteriore serie di controlli di validità sulle variabili contenute nei messaggi. In particolare, vengono utilizzate e applicate in sequenza due diverse classi di controllo:

- a) controllo climatologico debole;
- b) controllo di consistenza interna.

Dopo ogni controllo il dato originale viene restituito senza essere modificato, insieme ad un *flag* che ne attribuisce la qualità di dato valido o non valido.

Il controllo climatologico debole si basa sul requisito di non superamento di soglie minime e massime (abbastanza blande) di accettazione del dato. Le soglie sono state definite a priori utilizzando dei criteri di ragionevolezza (tab. 2.1). Il non superamento di questo controllo vieta qualsiasi utilizzo successivo del dato.

Il controllo di consistenza interna prevede il controllo reciproco di più variabili in uno stesso istante temporale. Ad esempio, la temperatura dell'aria viene messa a confronto con la temperatura di rugiada: ovviamente, non potrà mai verificarsi che quest'ultima sia più grande della temperatura dell'aria. Anche in questo caso, il non superamento di questo controllo pregiudica per sempre l'utilizzo successivo del dato.

Tabella 2.1 - Soglie di accettazione per i valori delle variabili dei messaggi sinottici

| Variabile | Unità di misura | Descrizione | Soglia minima | Soglia massima |
|-----------|-------------------------|--------------------------------------|---------------|-------------------------------------|
| HH | M | Altezza nubi più basse | 0 | 2500 |
| VIS | Hm | Visibilità orizzontale | 0 | 9999 |
| NTOT | 1/8 | Copertura nuvolosa totale | 0 | 9 |
| DD | Gradi | Direzione (di provenienza) del vento | 0 | 360 |
| FF | Nodi | Forza del vento | 0 | 199 |
| TATA | C/10 | Temperatura dell'aria | -290 | 490 |
| TDTD | C/10 | Temperatura di rugiada | -290 | 490 |
| UMID | % | Umidità relativa | 20 | 100 |
| PRES | hPa/10 | Pressione al livello del mare | 9600 | 10600 |
| PRECI | mm/10 | Precipitazione | 0 | 9890 |
| TEMPPR | | Tempo presente | 0 | 99 |
| TEMPAS | | Tempo passato | 0 | 99 |
| TETE | °C/10 | Temperatura massima o minima | -290 | 490 |
| RADSOL | cal/cm ² min | Radiazione solare | 0 | Funzione di giorno, ora, latitudine |

Controlli di validità degli indicatori

Ogni indicatore calcolato attraverso i programmi del sistema SCIA e poi caricato sul database nazionale degli indicatori climatici, è contrassegnato, oltre che dal suo valore, dal numero di giorni con dati da cui ha avuto origine e da un *flag* di validità dell'indicatore stesso, che può assumere diversi valori; un valore maggiore o uguale a 1 ha il significato di indicatore valido o consolidato; un valore uguale a 0 ha il significato di indicatore generato da un numero insufficiente di dati validi.

In generale, il criterio per l'attribuzione del *flag* di indicatore valido consiste nella disponibilità di almeno il 75% di dati elementari validi che concorrono al calcolo dell'indicatore stesso. Questo criterio viene adottato per tutti gli indicatori (tipicamente, valori medi ed estremi) per i quali è ragionevole ipotizzare che la mancanza del 25% o meno dei dati di origine non alteri significativamente il valore dell'indicatore, ed è motivato anche dall'obiettivo di non scartare un numero troppo elevato di indicatori. Un'eccezione è costituita dai valori di precipitazione cumulata che, per la natura stessa dell'indicatore, devono considerarsi validi solo se è disponibile una percentuale dei dati di origine più elevata. Per le precipitazioni cumulate (decadali, mensili, annuali) la soglia di dati disponibili e validi oltre la quale viene assegnato il *flag* di indicatore valido è fissata al 90%.

Il programma di calcolo degli indicatori controlla in prima battuta la presenza dei dati necessari al calcolo di un certo indicatore a livello giornaliero; se la percentuale di dati presenti e validi nella giornata è pari almeno al 75% (per esempio, per i dati contenuti nei messaggi sinottici triorari, almeno 6 osservazioni su 8) l'indicatore giornaliero viene considerato valido. La temperatura media giornaliera viene calcolata, secondo gli standard dell'OMM, come valore medio delle temperature minima e massima giornaliera. Tuttavia, viene calcolata anche la temperatura media come media di tutti i valori rilevati nell'arco della giornata; in questo caso, si aggiunge un controllo sulla distribuzione dei dati tra giorno e notte (presenza di almeno 75% dei dati nell'intervallo diurno e di almeno 75% dei dati nell'intervallo notturno). L'attribuzione del *flag* di validità all'indicatore decadale, mensile o annuale viene poi effettuata controllando ancora che siano presenti almeno il 75% di dati giornalieri validi. Per alcuni indicatori come la temperatura media annuale il *flag* di validità viene assegnato solo dopo aver effettuato un ulteriore controllo sul popolamento dei valori giornalieri rispetto alle stagioni. Infatti, nel caso in cui i valori giornalieri mancanti fossero sbilanciati tra estate e inverno, la temperatura media annuale calcolata risulterebbe significativamente alterata e quindi da considerare non valida. Il controllo che viene effettuato considera il valore assoluto della differenza tra

il numero di dati giornalieri nei tre mesi invernali (gennaio, febbraio e dicembre) ed estivi (giugno, luglio e agosto). Un valore maggiore di 20 pregiudica la validità dell'indicatore temperatura media annuale. Lo stesso tipo di controllo e di criterio di assegnazione del *flag* di validità viene applicato agli indicatori annuali di altre variabili caratterizzate da forte variabilità stagionale: media delle temperature minima e massima giornaliera, eliofania, radiazione globale incidente, evapotraspirazione potenziale e gradi giorno.

Sul database degli indicatori climatici, accanto al valore dell'indicatore viene conservato, oltre al *flag* di validità, il numero dei giorni con dati validi che hanno generato l'indicatore stesso; esso risulta pertanto compreso tra 0 e 30 (o 31, o 28/29 per febbraio) per un mese.

A seconda della variabile climatica e del significato dell'indicatore stesso, l'assegnazione del *flag* di validità può differire dai criteri generali finora descritti. Ad esempio, per il "numero dati giorni nevosi" e il "numero dati giorni temporaleschi" il *flag* giornaliero viene impostato a 1 non appena appare su uno dei report di un certo giorno il valore del codice "tempo presente" corrispondente a "neve" o "temporale"; nel caso in cui questo non si verifichi (cioè nessun report indica neve o temporale) il *flag* viene impostato comunque a 1 quando si raggiunge il 75% dei dati "tempo presente" disponibili, seguendo il criterio generale. Per il "numero dati eventi temporaleschi" invece, il *flag* viene impostato a 1 comunque quando si raggiunge il 75% dei dati "tempo presente" disponibili, prima su base giornaliera e poi su base decadale, mensile o annuale.

Nonostante i controlli dei dati effettuati dalle fonti all'origine e, per quanto riguarda i dati sinottici, i controlli climatologico debole e di consistenza interna, può accadere che un certo numero di dati errati sfugga ai controlli e concorra a determinare il valore di indicatori che sono evidentemente, in prima battuta, errati. L'entità e l'evidenza dell'errore dipende dall'entità dell'errore del dato originale e dal tipo di indicatore. In generale, un dato errato risulta poco visibile dall'analisi di indicatori medi annuali, mentre può essere meglio evidenziato dall'analisi dei valori estremi. L'individuazione degli indicatori non validi passa necessariamente attraverso l'analisi delle serie di indicatori mediante l'utilizzo di "filtri" mirati sulla base della stagione e delle caratteristiche climatiche della stazione (altezza sul livello del mare, latitudine). Si applicano cioè dei criteri oggettivi di ricerca di indicatori con valori sospetti (*outlier*) o palesemente errati (Baffo et al., 2005).

A seguito del rilevamento degli *outlier*, si risale al dato o ai dati originali da cui essi sono stati originati; in qualche caso viene richiesto un riscontro o una conferma alla fonte; più spesso, attraverso l'analisi della sequenza temporale dei dati o di dati concomitanti di stazioni limitrofe, si verifica la validità o meno del dato originale. Nei casi in cui si conferma che l'indicatore non è valido, vengono assegnati valori negativi al *flag* di validità (indicatore non valido) : -9 se viene invalidato un indicatore con *flag* di partenza 1; -8 se viene invalidato un indicatore con *flag* di partenza 0 (cioè generato da un numero insufficiente di dati validi). Con l'assegnazione dei nuovi *flag*, gli indicatori non validi non partecipano alla determinazione degli indici climatologici (valori normali, anomalie, ecc.). Solo in casi limitati, qualora si renda disponibile dalla fonte originale del dato errato il suo valore corretto, si ripete la procedura di calcolo degli indicatori e il caricamento degli stessi sul database.

3. SELEZIONE DELLE SERIE TEMPORALI

Dopo aver popolato il database di indicatori statistici delle diverse variabili climatiche, il passo successivo necessario a pervenire alla stima delle variazioni e delle tendenze del clima in Italia è costituito dalla selezione delle serie temporali di indicatori utili e idonee a questo scopo. In questa fase, l'obiettivo è disporre di un insieme abbastanza ricco e rappresentativo del clima italiano, di stazioni con serie temporali che rispondono a requisiti di qualità, completezza e continuità tali da garantire una stima affidabile delle variazioni nel tempo e quindi delle tendenze. Nel seguito viene trattata in particolare la selezione delle serie annuali di temperatura e precipitazione.

Per quanto riguarda la qualità, la selezione riguarda solo gli indicatori annuali cui è stato assegnato un *flag* di dato valido al termine dei controlli descritti nel capitolo precedente. Per quanto riguarda la completezza e la continuità delle serie, deve essere anzitutto fissato l'intervallo di tempo sul quale effettuare la stima delle tendenze. Ai fini di questo rapporto, anche in considerazione della copertura temporale dei dati che alimentano il sistema SCIA, l'intervallo è stato fissato ai 50 anni che vanno dal 1961 al 2010. Poiché il numero di stazioni con serie complete è limitato e consentirebbe di stimare le tendenze climatiche solo su pochi punti del territorio nazionale, il requisito di completezza viene rilasciato e reso meno stringente applicando tre criteri: il primo assicura che la serie disponga di una percentuale elevata (anche se non necessariamente il 100%) di indicatori annuali validi; il secondo, che non ci sia un numero consistente di anni consecutivi senza indicatori validi; il terzo, che gli anni "mancanti" non siano concentrati nel periodo più recente. Quest'ultimo criterio deriva dall'obiettivo di aggiornare agli anni più recenti la stime delle tendenze.

E' stata quindi effettuata una interrogazione alla base dati del sistema SCIA con i seguenti criteri di selezione delle serie temporali: periodo: 1961-2010; numero massimo di anni consecutivi mancanti: 4; ultimo anno presente nella serie: ≥ 2007 . Il risultato, per quanto riguarda le serie di temperatura media e di precipitazione cumulata, è rappresentato dalle tabelle 3.1 e 3.2. L'elenco delle stazioni è stato ordinato per numero di anni con indicatore valido, ed è stato troncato a 43 anni presenti su 50 (corrispondenti all' 86%). In questo modo si ottengono un insieme di 97 stazioni per la temperatura media e un insieme di 116 stazioni per la precipitazione cumulata, rappresentate nelle mappe delle figure 3.1 e 3.2.

Tabella 3.1 – Stazioni con serie temporali di temperatura media selezionate secondo i seguenti criteri: periodo 1961-2010; numero massimo di anni consecutivi mancanti = 4; Ultimo anno presente nella serie: almeno 2007; almeno 43 anni di dati disponibili

| Nome | Regione | Codice Rete * | Numero anni | Primo anno | Ultimo anno |
|---------------------|---------------------|---------------|-------------|------------|-------------|
| Verona/Villafranca | Veneto | 11 | 50 | 1961 | 2010 |
| Monte Cimone | Emilia-Romagna | 11 | 50 | 1961 | 2010 |
| Latina | Lazio | 11 | 50 | 1961 | 2010 |
| Campobasso | Molise | 11 | 50 | 1961 | 2010 |
| Ponza | Lazio | 11 | 50 | 1961 | 2010 |
| Brindisi | Puglia | 11 | 50 | 1961 | 2010 |
| S. Maria di Leuca | Puglia | 11 | 50 | 1961 | 2010 |
| Messina | Sicilia | 11 | 50 | 1961 | 2010 |
| Cozzo Spadaro | Sicilia | 11 | 50 | 1961 | 2010 |
| Cagliari/Elmas | Sardegna | 11 | 50 | 1961 | 2010 |
| Vipiteno | Trentino-Alto Adige | 35 | 50 | 1961 | 2010 |
| Anterselva di Mezzo | Trentino-Alto Adige | 35 | 50 | 1961 | 2010 |
| Bolzano | Trentino-Alto Adige | 35 | 50 | 1961 | 2010 |
| Sesto | Trentino-Alto Adige | 35 | 50 | 1961 | 2010 |
| Todi | Umbria | 5 | 49 | 1961 | 2010 |
| Dejola | Valle D'Aosta | 6 | 49 | 1961 | 2010 |
| Dobbiaco | Trentino-Alto Adige | 11 | 49 | 1961 | 2010 |

| Nome | Regione | Codice Rete * | Numero anni | Primo anno | Ultimo anno |
|-------------------------|-----------------------|---------------|-------------|------------|-------------|
| Torino/Bric della Croce | Piemonte | 11 | 49 | 1961 | 2010 |
| Piacenza | Emilia-Romagna | 11 | 49 | 1961 | 2010 |
| Brescia/Ghedi | Lombardia | 11 | 49 | 1961 | 2010 |
| Treviso/Istrana | Veneto | 11 | 49 | 1962 | 2010 |
| Treviso/S. Angelo | Veneto | 11 | 49 | 1962 | 2010 |
| Albenga | Liguria | 11 | 49 | 1961 | 2010 |
| Rimini | Emilia-Romagna | 11 | 49 | 1961 | 2010 |
| Pisa/S. Giusto | Toscana | 11 | 49 | 1961 | 2010 |
| Arezzo | Toscana | 11 | 49 | 1961 | 2010 |
| Grosseto | Toscana | 11 | 49 | 1961 | 2010 |
| Termoli | Molise | 11 | 49 | 1961 | 2010 |
| Roma/Ciampino | Lazio | 11 | 49 | 1961 | 2010 |
| Frosinone | Lazio | 11 | 49 | 1962 | 2010 |
| Trevico | Campania | 11 | 49 | 1961 | 2010 |
| Capo Palinuro | Campania | 11 | 49 | 1961 | 2010 |
| Gioia del Colle | Puglia | 11 | 49 | 1961 | 2010 |
| Latronico | Basilicata | 11 | 49 | 1961 | 2010 |
| Lecce | Puglia | 11 | 49 | 1961 | 2010 |
| Monte Scuro | Calabria | 11 | 49 | 1962 | 2010 |
| Ustica | Sicilia | 11 | 49 | 1961 | 2010 |
| Trapani/Birgi | Sicilia | 11 | 49 | 1962 | 2010 |
| Fiè Allo Sciliar | Trentino-Alto Adige | 35 | 49 | 1961 | 2010 |
| Piacenza | Emilia-Romagna | 6 | 48 | 1961 | 2008 |
| Salo | Lombardia | 6 | 48 | 1961 | 2008 |
| Paganella | Trentino-Alto Adige | 11 | 48 | 1961 | 2010 |
| Firenze/Peretola | Toscana | 11 | 48 | 1961 | 2010 |
| Frontone | Marche | 11 | 48 | 1961 | 2010 |
| Roma Fiumicino | Lazio | 11 | 48 | 1961 | 2010 |
| Pratica di Mare | Lazio | 11 | 48 | 1961 | 2010 |
| Monte S. Angelo | Puglia | 11 | 48 | 1962 | 2010 |
| Bressanone | Trentino-Alto Adige | 35 | 48 | 1961 | 2010 |
| Terni | Umbria | 6 | 47 | 1961 | 2008 |
| Imperia | Liguria | 6 | 47 | 1961 | 2008 |
| Tarvisio | Friuli-Venezia Giulia | 11 | 47 | 1961 | 2010 |
| Vicenza | Veneto | 11 | 47 | 1961 | 2007 |
| Punta Marina | Emilia-Romagna | 11 | 47 | 1961 | 2008 |
| Roma/Urbe | Lazio | 11 | 47 | 1961 | 2010 |
| Amendola | Puglia | 11 | 47 | 1961 | 2009 |
| Bonifati | Calabria | 11 | 47 | 1962 | 2008 |
| Crotone | Calabria | 11 | 47 | 1961 | 2010 |
| Catania/Sigonella | Sicilia | 11 | 47 | 1964 | 2010 |
| Pantelleria | Sicilia | 11 | 47 | 1961 | 2010 |
| Lampedusa | Sicilia | 11 | 47 | 1961 | 2010 |
| Decimomannu | Sardegna | 11 | 47 | 1962 | 2010 |
| Capo Bellavista | Sardegna | 11 | 47 | 1961 | 2010 |
| Plata | Trentino-Alto-Adige | 35 | 47 | 1961 | 2010 |
| Pian Fedaia (Diga) | Trentino-Alto Adige | 32 | 46 | 1961 | 2010 |
| Aviano | Friuli-Venezia Giulia | 11 | 46 | 1961 | 2010 |
| Trieste | Friuli-Venezia Giulia | 11 | 46 | 1961 | 2010 |
| Capo Mele | Liguria | 11 | 46 | 1964 | 2010 |
| Monte Argentario | Toscana | 11 | 46 | 1962 | 2010 |
| Pescara | Abruzzo | 11 | 46 | 1961 | 2010 |
| Guidonia | Lazio | 11 | 46 | 1962 | 2010 |
| Grazzanise | Campania | 11 | 46 | 1964 | 2010 |

| Nome | Regione | Codice Rete * | Numero anni | Primo anno | Ultimo anno |
|------------------------|---------------------|---------------|-------------|------------|-------------|
| Mongiana | Calabria | 5 | 45 | 1961 | 2010 |
| Taranto | Puglia | 6 | 45 | 1961 | 2008 |
| Parma | Emilia-Romagna | 6 | 45 | 1961 | 2008 |
| Catanzaro | Calabria | 31 | 45 | 1961 | 2010 |
| Passo Rolle | Trentino-Alto Adige | 11 | 45 | 1961 | 2010 |
| Milano/Malpensa | Lombardia | 11 | 45 | 1961 | 2010 |
| Bergamo/Orio al Serio | Lombardia | 11 | 45 | 1961 | 2010 |
| Monte Terminillo | Lazio | 11 | 45 | 1961 | 2010 |
| Napoli/Capodichino | Campania | 11 | 45 | 1961 | 2010 |
| Prizzi | Sicilia | 11 | 45 | 1961 | 2010 |
| Dobbiaco | Trentino-Alto Adige | 35 | 45 | 1961 | 2010 |
| Ridanna | Trentino-Alto Adige | 35 | 45 | 1965 | 2010 |
| Forlì | Emilia-Romagna | 6 | 44 | 1961 | 2010 |
| Alassio | Liguria | 6 | 44 | 1961 | 2008 |
| Milano/Linate | Lombardia | 11 | 44 | 1961 | 2010 |
| Gela | Sicilia | 11 | 44 | 1966 | 2010 |
| Alghero | Sardegna | 11 | 44 | 1961 | 2010 |
| Monte Maria | Trentino-Alto Adige | 35 | 44 | 1967 | 2010 |
| S. Maddalena in Casies | Trentino-Alto Adige | 35 | 44 | 1967 | 2010 |
| Cittanova | Calabria | 31 | 43 | 1961 | 2010 |
| Careser (Diga) | Trentino-Alto Adige | 32 | 43 | 1961 | 2010 |
| S. Valentino alla Muta | Trentino-Alto Adige | 11 | 43 | 1961 | 2010 |
| Passo della Cisa | Toscana | 11 | 43 | 1961 | 2009 |
| Grottaglie | Puglia | 11 | 43 | 1962 | 2010 |
| Marina Di Ginosa | Puglia | 11 | 43 | 1968 | 2010 |
| Vernago | Trentino-Alto Adige | 35 | 43 | 1967 | 2010 |

* 11 – Rete sinottica del Servizio Meteorologico dell’Aeronautica Militare e dell’Ente Nazionale di Assistenza al Volo;
4, 5, 6 – Rete del CRA-CMA (ex UCEA);
da 20 in poi – Reti regionali

Tabella 3.2 – Stazioni con serie temporali di precipitazione cumulata selezionate secondo i seguenti criteri: periodo 1961-2010; numero massimo di anni consecutivi mancanti = 4; Ultimo anno presente nella serie: almeno 2007; almeno 43 anni di dati disponibili

| Nome | Regione | Cod. rete | Numero anni | Primo anno | Ultimo anno |
|-----------------------------|---------------------|-----------|-------------|------------|-------------|
| Latina | Lazio | 11 | 50 | 1961 | 2010 |
| Cagliari/Elmas | Sardegna | 11 | 50 | 1961 | 2010 |
| Monte Maria | Trentino-Alto Adige | 35 | 50 | 1961 | 2010 |
| Silandro | Trentino-Alto Adige | 35 | 50 | 1961 | 2010 |
| Plata | Trentino-Alto Adige | 35 | 50 | 1961 | 2010 |
| Vipiteno | Trentino-Alto Adige | 35 | 50 | 1961 | 2010 |
| Dobbiaco | Trentino-Alto Adige | 35 | 50 | 1961 | 2010 |
| Bolzano | Trentino-Alto Adige | 35 | 50 | 1961 | 2010 |
| Sesto | Trentino-Alto Adige | 35 | 50 | 1961 | 2010 |
| Diga di Zoccolo S. Valburga | Trentino-Alto Adige | 35 | 50 | 1961 | 2010 |
| Diga di Vizzate | Trentino-Alto Adige | 35 | 50 | 1961 | 2010 |
| Ridanna | Trentino-Alto Adige | 35 | 50 | 1961 | 2010 |
| S. Maddalena in Casies | Trentino-Alto Adige | 35 | 50 | 1961 | 2010 |
| Fundres | Trentino-Alto Adige | 35 | 50 | 1961 | 2010 |
| Mongiana | Calabria | 5 | 49 | 1961 | 2010 |
| Montebello Ionico | Calabria | 31 | 49 | 1961 | 2010 |
| Scilla | Calabria | 31 | 49 | 1961 | 2010 |
| Torino/Bric della Croce | Piemonte | 11 | 49 | 1961 | 2010 |

| Nome | Regione | Cod. rete | Numero anni | Primo anno | Ultimo anno |
|------------------------|-----------------------|-----------|-------------|------------|-------------|
| Rimini | Emilia-Romagna | 11 | 49 | 1961 | 2010 |
| Grosseto | Toscana | 11 | 49 | 1961 | 2010 |
| Ponza | Lazio | 11 | 49 | 1961 | 2010 |
| Gioia del Colle | Puglia | 11 | 49 | 1961 | 2010 |
| Messina | Sicilia | 11 | 49 | 1961 | 2010 |
| Anterselva di Mezzo | Trentino-Alto Adige | 35 | 49 | 1961 | 2010 |
| Fiè Allo Sciliar | Trentino-Alto Adige | 35 | 49 | 1961 | 2010 |
| Sarentino | Trentino-Alto Adige | 35 | 49 | 1961 | 2010 |
| Palermiti | Calabria | 31 | 48 | 1961 | 2009 |
| Rocella Ionica | Calabria | 31 | 48 | 1961 | 2010 |
| Cittanova | Calabria | 31 | 48 | 1961 | 2010 |
| Arena | Calabria | 31 | 48 | 1961 | 2010 |
| Piacenza | Emilia-Romagna | 11 | 48 | 1961 | 2010 |
| Verona/Villafranca | Veneto | 11 | 48 | 1961 | 2010 |
| Treviso/Istrana | Veneto | 11 | 48 | 1961 | 2010 |
| Arezzo | Toscana | 11 | 48 | 1961 | 2010 |
| Frontone | Marche | 11 | 48 | 1961 | 2010 |
| Vigna di Valle | Lazio | 11 | 48 | 1961 | 2010 |
| Termoli | Molise | 11 | 48 | 1961 | 2010 |
| Campobasso | Molise | 11 | 48 | 1961 | 2010 |
| Trevico | Campania | 11 | 48 | 1961 | 2010 |
| Monte Scuro | Calabria | 11 | 48 | 1962 | 2010 |
| S. Maria di Leuca | Puglia | 11 | 48 | 1961 | 2010 |
| Cozzo Spadaro | Sicilia | 11 | 48 | 1961 | 2010 |
| Bressanone | Trentino-Alto Adige | 35 | 48 | 1961 | 2010 |
| Vernago | Trentino-Alto Adige | 35 | 48 | 1961 | 2009 |
| Ponte Gardena | Trentino-Alto Adige | 35 | 48 | 1961 | 2010 |
| Todi | Umbria | 5 | 47 | 1961 | 2010 |
| Casale Monferrato | Piemonte | 28 | 47 | 1961 | 2010 |
| Cropalati | Calabria | 31 | 47 | 1961 | 2010 |
| Crotone | Calabria | 31 | 47 | 1961 | 2010 |
| Catanzaro | Calabria | 31 | 47 | 1961 | 2010 |
| Gioiosa Ionica | Calabria | 31 | 47 | 1961 | 2010 |
| Staiti | Calabria | 31 | 47 | 1961 | 2010 |
| Filadelfia | Calabria | 31 | 47 | 1961 | 2010 |
| Tiriolo | Calabria | 31 | 47 | 1961 | 2010 |
| Santagata del Bianco | Calabria | 31 | 47 | 1961 | 2010 |
| Cogolo Pont (Centrale) | Trentino-Alto Adige | 32 | 47 | 1961 | 2009 |
| Trieste | Friuli-Venezia Giulia | 11 | 47 | 1961 | 2010 |
| Frosinone | Lazio | 11 | 47 | 1964 | 2010 |
| Pratica di Mare | Lazio | 11 | 47 | 1963 | 2010 |
| Grazzanise | Campania | 11 | 47 | 1964 | 2010 |
| Monte S. Angelo | Puglia | 11 | 47 | 1962 | 2010 |
| Brindisi | Puglia | 11 | 47 | 1961 | 2010 |
| Lecce | Puglia | 11 | 47 | 1961 | 2010 |
| Naturno | Trentino-Alto Adige | 35 | 47 | 1961 | 2010 |
| Potenza | Basilicata | 24 | 46 | 1961 | 2010 |
| Luserna S. Giovanni | Piemonte | 28 | 46 | 1961 | 2010 |
| Cassano Allo Ionio | Calabria | 31 | 46 | 1961 | 2010 |
| San Sosti | Calabria | 31 | 46 | 1961 | 2010 |
| Cariati Marina | Calabria | 31 | 46 | 1961 | 2010 |
| Botricello | Calabria | 31 | 46 | 1961 | 2010 |
| Cropani | Calabria | 31 | 46 | 1961 | 2010 |
| Sinopoli | Calabria | 31 | 46 | 1961 | 2010 |

| Nome | Regione | Cod. rete | Numero anni | Primo anno | Ultimo anno |
|--------------------------|-----------------------|-----------|-------------|------------|-------------|
| Ioppolo | Calabria | 31 | 46 | 1961 | 2010 |
| Campotenese | Calabria | 31 | 46 | 1961 | 2010 |
| Bieno | Trentino-Alto Adige | 32 | 46 | 1961 | 2009 |
| Dobbiaco | Trentino-Alto Adige | 11 | 46 | 1961 | 2010 |
| Brescia/Ghedi | Lombardia | 11 | 46 | 1961 | 2010 |
| Vicenza | Veneto | 11 | 46 | 1961 | 2007 |
| Capo Mele | Liguria | 11 | 46 | 1964 | 2010 |
| Pisa/S. Giusto | Toscana | 11 | 46 | 1961 | 2010 |
| Ustica | Sicilia | 11 | 46 | 1961 | 2010 |
| Trapani/Birgi | Sicilia | 11 | 46 | 1962 | 2009 |
| Pantelleria | Sicilia | 11 | 46 | 1961 | 2010 |
| Borgomanero | Piemonte | 28 | 45 | 1961 | 2010 |
| Varallo | Piemonte | 28 | 45 | 1961 | 2010 |
| Oropa | Piemonte | 28 | 45 | 1961 | 2010 |
| Domodossola | Piemonte | 28 | 45 | 1961 | 2010 |
| Torano Scalo | Calabria | 31 | 45 | 1961 | 2010 |
| Nocelle - Arvo | Calabria | 31 | 45 | 1961 | 2010 |
| Fabrizia - Cassari | Calabria | 31 | 45 | 1961 | 2010 |
| Mileto | Calabria | 31 | 45 | 1961 | 2010 |
| Treviso/S. Angelo | Veneto | 11 | 45 | 1964 | 2010 |
| Roma/Ciampino | Lazio | 11 | 45 | 1961 | 2008 |
| Capo Palinuro | Campania | 11 | 45 | 1961 | 2010 |
| Enna | Sicilia | 11 | 45 | 1961 | 2010 |
| Chiaravalle Centrale | Calabria | 31 | 44 | 1961 | 2010 |
| Capo Spartivento | Calabria | 31 | 44 | 1961 | 2010 |
| Rogliano | Calabria | 31 | 44 | 1961 | 2010 |
| Amantea | Calabria | 31 | 44 | 1961 | 2010 |
| Plati | Calabria | 31 | 44 | 1961 | 2010 |
| Mezzana | Trentino-Alto Adige | 32 | 44 | 1961 | 2010 |
| Passo Rolle | Trentino-Alto Adige | 11 | 44 | 1961 | 2010 |
| Tarvisio | Friuli-Venezia Giulia | 11 | 44 | 1961 | 2010 |
| Viterbo | Lazio | 11 | 44 | 1961 | 2010 |
| Prizzi | Sicilia | 11 | 44 | 1961 | 2010 |
| Decimomannu | Sardegna | 11 | 44 | 1962 | 2010 |
| Cosenza | Calabria | 31 | 43 | 1961 | 2010 |
| Castrovillari | Calabria | 31 | 43 | 1961 | 2010 |
| Serra San Bruno | Calabria | 31 | 43 | 1961 | 2010 |
| Molotio | Calabria | 31 | 43 | 1961 | 2010 |
| Rosarno | Calabria | 31 | 43 | 1961 | 2010 |
| Pieve Tesino (O.P. Enel) | Trentino-Alto Adige | 32 | 43 | 1961 | 2010 |
| Punta Marina | Emilia-Romagna | 11 | 43 | 1961 | 2008 |
| Guidonia | Lazio | 11 | 43 | 1964 | 2008 |
| Latronico | Basilicata | 11 | 43 | 1961 | 2010 |
| Catania/Sigonella | Sicilia | 11 | 43 | 1964 | 2010 |

* 11 – Rete sinottica del Servizio Meteorologico dell’Aeronautica Militare e dell’Ente Nazionale di Assistenza al Volo;
4, 5, 6 – Rete del CRA-CMA (ex UCEA);
da 20 in poi – Reti regionali

La distribuzione spaziale delle stazioni riflette la disomogeneità e l’irregolarità delle reti di monitoraggio degli organismi nazionali e regionali che, ad oggi, alimentano con i propri dati il sistema SCIA. In particolare, dalla mappa relativa alle serie di precipitazione è evidente l’elevata densità di stazioni utili della Calabria e della provincia di Trento, i cui dati di precipitazione giornaliera sono aggiornati con cadenza annuale e scaricabili direttamente dal web.

Le serie annuali delle stazioni elencate nella tabella 3.1 rappresentano la base di partenza per il calcolo degli indici climatici, quali i valori normali e le anomalie annuali, nonché per la stima delle tendenze della temperatura media in Italia. Per quanto riguarda le precipitazioni invece, sia a causa della maggiore variabilità spaziale e temporale di questa variabile, che per la minore uniformità della base dati disponibile (sia in senso spaziale che temporale), ogni serie annuale tra quelle elencate nella tabella 3.2 è utile alla stima delle tendenze di precipitazione nella singola località; per una stima della

tendenza media in Italia o su un'area geoclimatica estesa è necessario un trattamento delle serie con metodi geostatistici di aggregazione spaziale (v. cap. 5.2).



Figura 3.1 – *Distribuzione geografica delle stazioni riportate in tabella 3.1 (temperatura media).*



Figura 3.2 – *Distribuzione geografica delle stazioni riportate in tabella 3.2 (precipitazione cumulata).*

4. OMOGENEIZZAZIONE DELLE SERIE DI TEMPERATURA

4.1 Introduzione

Lo studio della variabilità del clima e la valutazione delle tendenze climatiche hanno come prerequisito la disponibilità di lunghe serie omogenee di dati climatici. Una serie omogenea è una serie temporale nella quale le variazioni sono da attribuire esclusivamente a fattori climatici (Conrad e Pollack, 1950). Nella realtà, le serie di dati spesso presentano delle perturbazioni (o disomogeneità) che sono la conseguenza di fattori esterni non climatici. Nelle serie di temperatura bruschi cambiamenti del valore medio si possono ad esempio attribuire allo spostamento della stazione di misura, alla sostituzione della strumentazione, a cambiamenti nell'esposizione dello strumento o all'adozione di nuove procedure di elaborazione dei dati. Cambiamenti nell'ambiente circostante la stazione (urbanizzazione o crescita di vegetazione) e cambiamenti graduali nella calibrazione o nelle caratteristiche fisiche dello strumento di misura possono invece essere fonte di disomogeneità che si manifestano come *trend* artificiale nella serie (Peterson et al., 1998; Aguilar et al., 2003).

L'istante temporale in cui la serie comincia a manifestare una perturbazione viene solitamente indicato con il termine "*breakpoint*" (talvolta "*changepoint*"). Non è raro che serie di dati secolari siano perturbate da tre o più *breakpoint*. Poiché eventuali disomogeneità possono alterare in maniera rilevante i risultati ottenuti dall'analisi delle serie climatiche, l'individuazione e la rimozione dei *breakpoint* costituisce un aspetto non secondario dell'analisi climatica.

La disponibilità di informazioni (metadati) che documentano la storia delle stazioni di rilevamento facilita lo studio delle disomogeneità di una serie. Un documento del WMO (Aguilar et al., 2003) fornisce una descrizione del supporto che possono fornire i metadati, essendo i dati meteorologici "...influenzati da una vasta gamma di pratiche di rilevazione. I dati dipendono dallo strumento, dalla sua esposizione, dal processo di registrazione e da molti altri fattori. E' necessario tenere traccia di tutti questi metadati al fine di sfruttare al meglio i dati." Spesso tuttavia gli studiosi del clima si trovano a dover utilizzare serie di dati la cui documentazione risulta incompleta se non del tutto assente.

Per identificare le disomogeneità non documentate e correggere il loro effetto sulle serie, sono state sviluppate diverse metodologie, per lo più di carattere statistico. Il documento del WMO descrive sinteticamente quattordici metodi utili per l'omogeneizzazione delle serie mensili, annuali e stagionali. Tra di essi, nel presente rapporto vengono sinteticamente descritti lo Standard Normal Homogeneity Test (SNHT) (Alexandersson, 1986; Alexandersson e Moberg, 1997) e il Two Phase Regression (Lund and Reeves, 2002; Wang et al., 2007), utilizzati per omogeneizzare le serie annuali e mensili di temperatura del sistema SCIA.

Ben più complessa rimane l'omogeneizzazione dei dati giornalieri e orari. Le difficoltà in questo caso consistono nell'elevata variabilità dei dati e nella presenza di valori estremi che pongono non pochi problemi, trattandosi per definizione di eventi rari. Vincent e Zhang (2001) descrivono un processo di omogeneizzazione delle serie di temperatura giornaliera minima e massima per il Canada, sulla base delle correzioni effettuate sulle serie mensili e annuali. Un approccio più sofisticato è descritto da Kuglitsch et al. (2009) per l'omogeneizzazione di 174 serie di temperatura massima disseminate su una vasta area centrata sul Mediterraneo. Il metodo di Kuglitsch, proprio nell'ottica dello studio degli eventi estremi e dei cambiamenti climatici, non si limita a correggere il valor medio della distribuzione, ma corregge la distribuzione della serie anche nei momenti di ordine superiore al primo (varianza e *skewness*).

La varietà dei metodi proposti riflette il fatto che ogni variabile climatica richiede una trattazione specifica, così come la risoluzione e la scala temporale su cui viene svolta l'analisi dei dati. La maggior parte dei lavori pubblicati riguarda l'omogeneizzazione delle serie di temperatura, essendo caratterizzata da una variabilità minore rispetto a quella di altri parametri ed essendo maggiormente sensibile ad eventuali disomogeneità delle serie, nell'ambito della valutazione degli effetti dei cambiamenti climatici.

Un'ulteriore complicazione deriva dal fatto che l'utilizzo di metodi differenti può produrre risultati talvolta discordanti per una stessa serie di dati; per questo motivo, è opinione diffusa che uno studio attento delle disomogeneità di una serie debba prendere in esame i risultati di più metodi statistici, con il supporto dei metadati qualora essi siano disponibili (Wijngaard et al., 2003).

La maggior parte delle metodologie statistiche oggi utilizzate si basano sul confronto fra la serie oggetto di studio (serie candidata) e serie di riferimento. Per serie di riferimento si intende una serie rappresentativa del clima della regione in cui si trova la stazione candidata e nello stesso tempo priva di disomogeneità. Poiché nella realtà si ha difficilmente a che fare con serie siffatte, e poiché l'uso di una serie non omogenea come serie di riferimento può inficiare i risultati del processo di omogeneizzazione (Reeves et al., 2007), è prassi comune ricorrere a serie di riferimento artificiali costruite combinando in maniera opportuna il segnale climatico di serie di stazioni limitrofe sufficientemente correlate con la candidata. In questo modo, eventuali disomogeneità presenti in una delle serie dovrebbero divenire trascurabili nella serie aggregata.

La costruzione di una serie di riferimento pone nuovi problemi: quale deve essere il livello minimo di correlazione tra le serie limitrofe e la serie candidata? Quante stazioni sono necessarie per la costruzione di una serie di riferimento affidabile? Come costruire la serie di riferimento in modo di poter evidenziare eventuali segnali non climatici nella serie candidata? Ovviamente non è possibile dare risposte univoche a queste domande; come già precedentemente osservato, sarà cura del ricercatore prendere delle decisioni valutando di volta in volta studio nella situazione specifica.

Sebbene siano stati sviluppati metodi per l'individuazione dei *breakpoint* che prescindono dall'uso di serie di riferimento (un esempio è il filtro adattivo di Kolmogorov-Zurbenko, [Zurbenko et al., 1996] che fornisce un supporto visivo all'individuazione di possibili disomogeneità di una serie), il confronto tra serie candidata e serie di riferimento facilita il riconoscimento delle disomogeneità artificiali che possono nascondere un reale segnale climatico e aumenta pertanto la potenza del test statistico utilizzato.

4.2 Metodi di omogeneizzazione

L'omogeneizzazione delle serie di temperatura consiste sostanzialmente nell'individuazione dei *breakpoint* e nella correzione della serie. Poiché gli effetti tipici dei *breakpoint* riguardano principalmente il valore medio della serie (Wang, 2003), gli approcci presi in esame assumono che in corrispondenza di un *breakpoint* non vi siano cambiamenti nel *trend* della serie. In questo rapporto vengono presi in esame i metodi che riguardano le serie di temperatura mensili e annuali, utilizzati per omogeneizzare le serie di temperatura del sistema SCIA.

Standard Normal Homogeneity Test (SNHT)

Si tratta di un metodo basato sul test della massima verosimiglianza. L'idea di base è che la serie standardizzata (Z_t) costruita sottraendo o dividendo la serie candidata dalla corrispondente serie di riferimento, sia per lo più stazionaria in caso di omogeneità. Il primo passo consiste quindi nel riscaldare la serie candidata in una serie (Q_t) con *trend* nullo e con distribuzione gaussiana. Standardizzando la serie Q_t , si ottiene una serie Z_t che nell'ipotesi di omogeneità (H_0) si distribuisce secondo una normale con media $\mu = 0$ e varianza unitaria. Indicando la distribuzione normale con $N(\mu, \sigma^2)$, dove μ è la media e σ^2 è la varianza, si ha:

$$H_0: Z_t \sim N(0,1), \quad 1 \leq t \leq n$$

dove t è il tempo e n è il numero di elementi della serie.

Nell'ipotesi (H1) di discontinuità nel punto $t = c$, Z_t segue una normale con valor medio μ_1 e μ_2 rispettivamente prima e dopo il *breakpoint*.

$$H1: \begin{cases} Z_t \sim N(\mu_1, 1), & 1 \leq t \leq c \\ Z_t \sim N(\mu_2, 1), & c+1 \leq t \leq n \end{cases}$$

L'SNHT è facilmente traducibile in programmi software; ai fini dei risultati presentati in questo rapporto ne sono state utilizzate due versioni che differiscono nella costruzione delle serie di riferimento e nel numero di test condotti su ciascuna serie candidata.

Penalized t test e penalized F test

Questi due test differiscono tra di loro essenzialmente nell'uso della statistica test (la T di Student per il t-test e la F di Fisher per l'F test) utilizzata per l'identificazione di un probabile *breakpoint*. Il punto di partenza per il *penalized F test* (Wang 2008a, 2008b) è un modello regressivo del tipo:

$$X_t = \begin{cases} \mu_1 + \beta t + \varepsilon_t, & 1 \leq t \leq c \\ \mu_2 + \beta t + \varepsilon_t, & c+1 \leq t \leq n \end{cases}$$

dove X_t è la serie candidata; β rappresenta il *trend* della serie, che rimane costante prima e dopo il *breakpoint* c ; μ_1 e μ_2 rappresentano il valore medio della serie detrendizzata ($X_t - \beta t$), rispettivamente prima e dopo il *breakpoint*; ε_t è la componente di errore.

L'ipotesi che si vuole testare è $H_0: \mu_1 = \mu_2$, $H_1: \mu_1 \neq \mu_2$.

Il *penalized t test* (Wang et al. 2007, Wang, 2008b) prende in esame un modello di regressione privo di *trend*. L'idea è che, avendo a disposizione una serie di riferimento, sia possibile riscaldare la serie candidata in modo da eliminare il *trend*.

Il termine *penalized* che accompagna la dicitura di entrambi i test si riferisce al fatto che *breakpoint* troppo vicini agli estremi di un segmento temporale in esame vengono "penalizzati" dal test in modo da evitare un numero eccessivo di falsi *breakpoint* (problema, questo, comune anche allo SNHT).

4.3 Omogeneizzazione delle serie del sistema SCIA

Per omogeneizzare le serie annuali di temperatura del sistema SCIA è stato utilizzato il pacchetto software RHtestV3, basato sul *penalized maximal t test* (Wang et al. 2007) e il *penalized maximal F test* (Wang, 2008a). Il pacchetto RHtestV3 è scaricabile dal sito del gruppo di lavoro CCI/CLIVAR/JCOMM costituito da esperti nel settore dei cambiamenti climatici – ETCCDI (Expert Team on Climate Change Detection and Indices) : <http://cccma.seos.uvic.ca/ETCCDMI/software.shtml>.

Lo studio delle disomogeneità per il set di dati SCIA è stato condotto utilizzando il *penalized t test*, basato sul confronto tra serie candidata e serie di riferimento.

Creazione della serie di riferimento

La serie di riferimento per ogni serie di temperatura candidata è stata costruita secondo il procedimento descritto da Peterson e Easterling (1994). Per ciascuna serie candidata è stato individuato un set di stazioni relativamente vicine, rappresentative della regione climatica a cui appartiene la serie candidata. Le stazioni sono state selezionate tenendo conto della distanza (non superiore ai 200 km), della differenza di quota (non oltre i 400 m) e del grado di correlazione con la serie candidata (coefficiente di correlazione non inferiore a +0.6). Tali requisiti rappresentano un compromesso tra la necessità di disporre di un numero sufficiente (minimo tre, massimo cinque) di serie utili alla costruzione della serie di riferimento e la necessità che tali serie siano rappresentative della climatologia della serie candidata, considerando la variabilità spaziale media della temperatura.

Le serie individuate sono state successivamente combinate in una serie di riferimento tramite un'operazione di media pesata, in cui i pesi associati a ciascuna serie limitrofa corrispondono al quadrato del coefficiente di correlazione precedentemente calcolato. Questa procedura dovrebbe garantire che eventuali disomogeneità presenti in ciascuna delle serie limitrofe risultino trascurabili nella serie di riferimento.

I coefficienti di correlazione tra la serie candidata e le serie di stazioni utilizzate per la serie di riferimento, sono state calcolati nel periodo di riferimento 1961-1990. Inoltre, poiché il valore del coefficiente di correlazione può essere influenzato dalla presenza di discontinuità artificiali negli stessi anni in due serie diverse, il calcolo di tale coefficiente è stato effettuato considerando non i valori assoluti delle serie di temperatura, bensì le serie delle differenze prime, in cui ciascun valore viene sostituito con la sua differenza rispetto al valore dell'anno precedente (Peterson e Easterling, 1994). Per l'anno i -esimo si ha:

$$(dT/dt)_i = T_i - T_{i-1}$$

Applicando questa metodologia alle 97 serie di temperatura media annuale della tabella 3.1, è stato possibile costruire una serie di riferimento per 77 stazioni. Le restanti 20 stazioni non sono state sottoposte al test di omogeneità.

Omogeneizzazione

Il pacchetto software RhtestV3 utilizzato per l'omogeneizzazione è stato sviluppato per il linguaggio statistico R ed è provvisto di un'interfaccia grafica che ne semplifica l'utilizzo. Dovendo testare un numero rilevante di serie, si è però preferito modificare il codice originale in modo di rendere parzialmente automatico il test, senza dover passare per l'interfaccia grafica per ogni singola stazione candidata.

L'aspetto interessante del pacchetto RhtestV3 è costituito dall'ampia gamma di informazioni statistiche che restituisce per i vari *breakpoint* individuati. A seconda dei risultati del test statistico (t test o F test), i *breakpoint* vengono distinti in: 1) *breakpoint* statisticamente significativi; 2) *breakpoint* non statisticamente significativi e quindi da scartare; 3) *breakpoint* dubbi sulla cui validità si rimanda al giudizio dell'esperto.

Nel caso in cui si presentino *breakpoint* di tipo 2) e 3) all'utente è richiesto di ripetere il test più volte mediante un processo iterativo, eliminando di volta in volta i *breakpoint* non significativi o quelli dubbi che l'utente ritiene opportuno eliminare; il processo termina quando tutti i *breakpoint* sono stati eliminati o quando si ha a che fare solo con *breakpoint* statisticamente significativi. Nell'analisi delle serie di indicatori SCIA, i casi dubbi sono stati trattati affiancando al risultato del t-test i risultati dello SNHT e un'analisi visiva della serie tramite il filtro di Kolmogorov-Zubenko.

Di seguito vengono riportati due esempi di omogeneizzazione per le stazioni di Arezzo e Ciampino. L'uso del t-test per queste due stazioni ha prodotto in prima istanza dei *breakpoint* dubbi. Per la stazione di Arezzo, i *breakpoint* dubbi vengono inizialmente individuati nel 1987 e nel 2003. La serie in figura 4.1 rappresenta l'andamento della differenza fra la serie candidata e la serie di riferimento, mentre la linea rossa mette in evidenza l'entità dello *shift* prima e dopo ciascun *breakpoint*. Poiché queste disomogenietà non hanno trovato conferma con lo SNHT, è stato effettuato nuovamente il t-test imponendo al software di trascurare il *breakpoint* di entità minore (1987). Il risultato di questo secondo *run* è che il 2003 non viene più segnalato come *breakpoint* e la serie risulta omogenea.

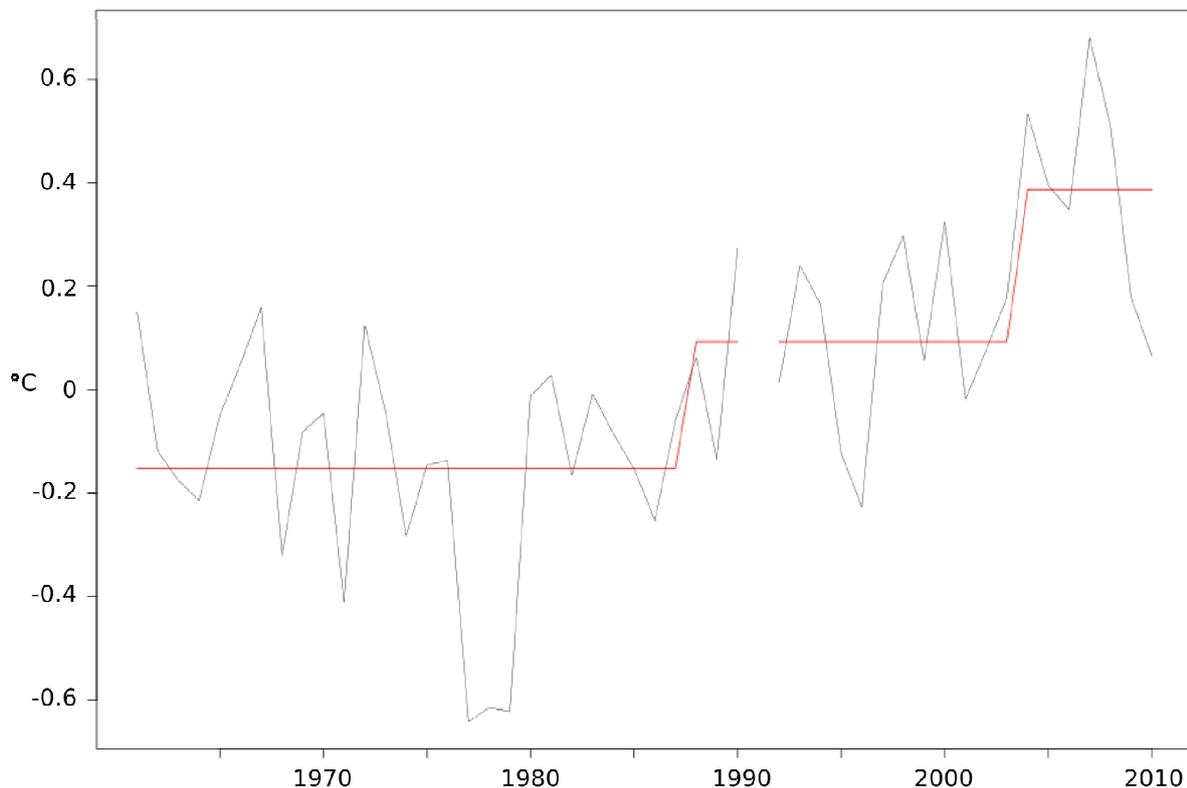


Figura 4.1 – Andamento della differenza fra la serie candidata (Arezzo) e la serie di riferimento dopo la prima applicazione del test. La linea rossa rappresenta il valore medio della serie prima e dopo li *breakpoint*.

Per la serie di Ciampino, la prima applicazione del t-test identifica tre *breakpoint*: il primo (dubbio) nel 1965 e gli altri due nel 1972 e nel 1977, statisticamente significativi (fig.4.2). Poiché il *breakpoint* del 1965 viene supportato dal metadato disponibile per questa stazione, esso è stato confermato e la serie è stata corretta come mostrato in fig.4.3.

I risultati del test di omogeneità applicato alle 77 stazioni per le quali è stato possibile creare la serie di riferimento sono riportati nella tabella 4.1, dove viene indicato per ogni stazione l'esito del test, il numero di *breakpoint* (nel caso di serie disomogenee) e i relativi anni. In sintesi, 45 stazioni sono risultate omogenee, 33 stazioni presentano uno o più *breakpoint* e quindi sono state corrette. La stazione di Campobasso, pur essendo priva di serie di riferimento, risulta omogenea secondo il *penalized maximal F test* ed è stata inserita comunque nel set di stazioni utili al calcolo dei valori normali e delle anomalie annuali, essendo rappresentativa di un'area relativamente vasta dell' Italia centrale non coperta da altre serie di dati utili

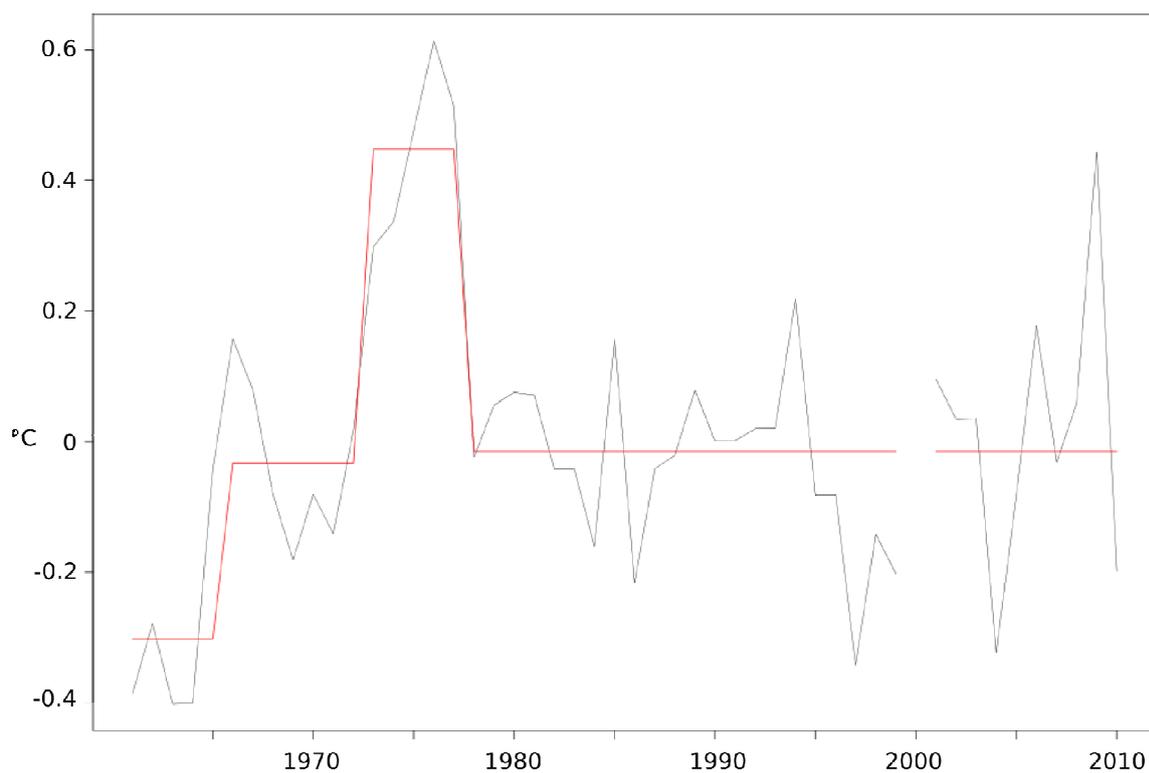


Figura 4.2 – Andamento della differenza fra la serie candidata (Ciampino) e la serie di riferimento, , dopo la prima applicazione del test. La linea rossa rappresenta il valore medio della serie prima e dopo il breakpoint.

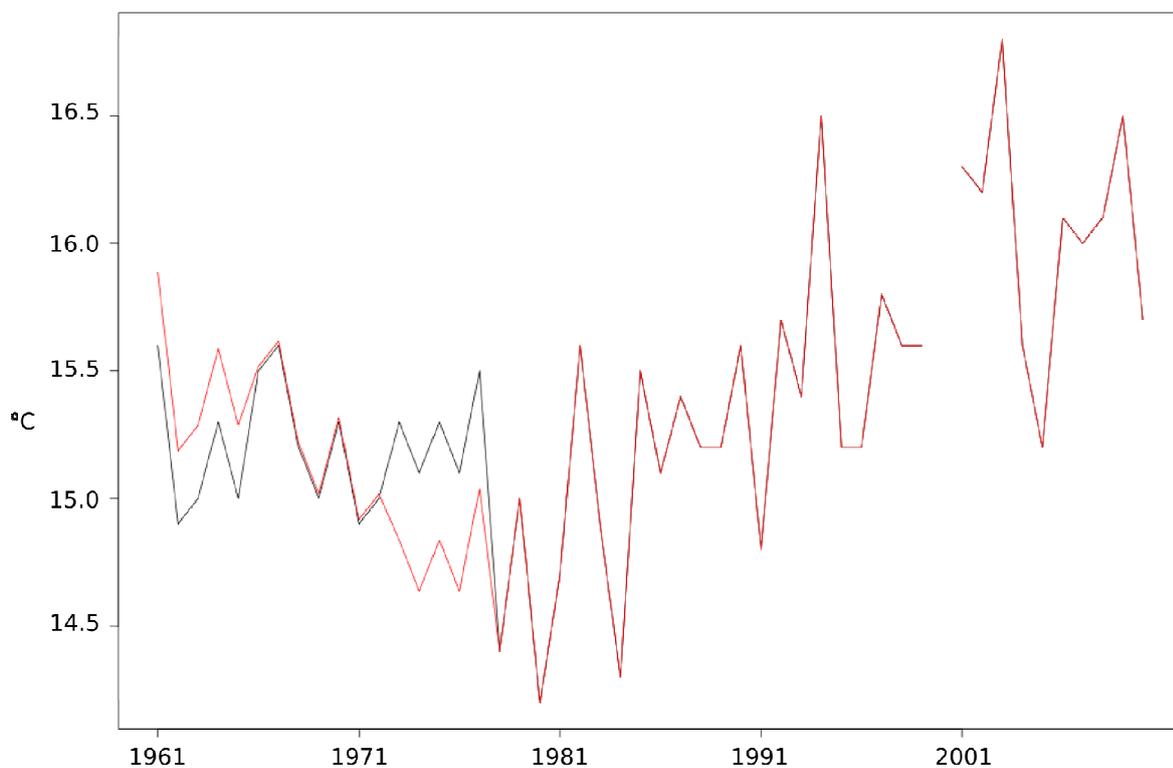


Figura 4.3 – Andamento della serie di temperatura originale e della serie omogeneizzata, stazione di Ciampino.

Tabella 4.1 – Risultati del test di omogeneizzazione (*t test*) sulle serie di temperatura delle 77 stazioni per le quali è stato possibile creare la serie di riferimento.

| Nome | Regione | Codice Rete* | RisultatoTest | Numero <i>breakpoint</i> | Anni |
|-------------------------|-----------------------|--------------|--------------------------|--------------------------|----------------|
| Todi | Umbria | 5 | disomogenea | 1 | 2002 |
| Terni | Umbria | 6 | omogenea | 0 | NA |
| Forlì | Emilia-Romagna | 6 | disomogenea | 2 | 1982,1990 |
| Parma | Emilia-Romagna | 6 | omogenea | 0 | NA |
| Piacenza | Emilia-Romagna | 6 | omogenea | 0 | NA |
| Alassio | Liguria | 6 | omogenea | 0 | NA |
| Imperia | Liguria | 6 | disomogenea | 1 | 1987 |
| Salò | Lombardia | 6 | omogenea | 0 | NA |
| Catanzaro | Calabria | 31 | omogenea | 0 | NA |
| S. Valentino alla Muta | Trentino-Alto Adige | 11 | disomogenea | 3 | 1979,1986,1998 |
| Passo Rolle | Trentino-Alto Adige | 11 | disomogenea | 1 | 1986 |
| Dobbiaco | Trentino-Alto Adige | 11 | disomogenea | 2 | 1974,1995 |
| Aviano | Friuli-Venezia Giulia | 11 | omogenea | 0 | NA |
| Tarvisio | Friuli-Venezia Giulia | 11 | disomogenea | 1 | 1991 |
| Torino/Bric Della Croce | Piemonte | 11 | omogenea | 0 | NA |
| Milano/Malpensa | Lombardia | 11 | omogenea | 0 | NA |
| Bergamo/Orio al Serio | Lombardia | 11 | disomogenea | 1 | 1987 |
| Milano/Linate | Lombardia | 11 | disomogenea | 1 | 1991 |
| Piacenza | Emilia-Romagna | 11 | omogenea | 0 | NA |
| Brescia/Ghedi | Lombardia | 11 | omogenea | 0 | NA |
| Verona/Villafranca | Veneto | 11 | omogenea | 0 | NA |
| Vicenza | Veneto | 11 | omogenea | 0 | NA |
| Treviso/Istrana | Veneto | 11 | omogenea | 0 | NA |
| Treviso/S. Angelo | Veneto | 11 | omogenea | 0 | NA |
| Trieste | Friuli-Venezia Giulia | 11 | omogenea | 0 | NA |
| Punta Marina | Emilia-Romagna | 11 | disomogenea | 1 | 1983 |
| Rimini | Emilia-Romagna | 11 | disomogenea | 2 | 1976,1984 |
| Capo Mele | Liguria | 11 | disomogenea | 1 | 1973 |
| Pisa/S. Giusto | Toscana | 11 | disomogenea | 2 | 1965,1996 |
| Monte Argentario | Toscana | 11 | disomogenea | 2 | 1975,1987 |
| Firenze/Peretola | Toscana | 11 | omogenea | 0 | NA |
| Arezzo | Toscana | 11 | omogenea | 0 | NA |
| Frontone | Marche | 11 | disomogenea | 1 | 1988 |
| Grosseto | Toscana | 11 | omogenea | 0 | NA |
| Pescara | Abruzzo | 11 | omogenea | 0 | NA |
| Termoli | Molise | 11 | disomogenea | 1 | 2000 |
| Guidonia | Lazio | 11 | omogenea | 0 | NA |
| Roma/Urbe | Lazio | 11 | omogenea | 0 | NA |
| Roma/Ciampino | Lazio | 11 | disomogenea | 3 | 1965,1972,1977 |
| Roma Fiumicino | Lazio | 11 | omogenea | 0 | NA |
| Latina | Lazio | 11 | disomogenea | 1 | 1984 |
| Frosinone | Lazio | 11 | omogenea | 0 | NA |
| Pratica Di Mare | Lazio | 11 | disomogenea | 1 | 1985 |
| Campobasso | Molise | 11 | omogenea (F test) | 0 | NA |
| Grazzanise | Campania | 11 | disomogenea | 2 | 1974,1993 |
| Monte S. Angelo | Puglia | 11 | omogenea | 0 | NA |
| Amendola | Puglia | 11 | omogenea | 0 | NA |
| Treviso | Campania | 11 | omogenea | 0 | NA |

| Nome | Regione | Codice Rete* | RisultatoTest | Numero <i>breakpoint</i> | Anni |
|------------------------|---------------------|--------------|-----------------|--------------------------|----------------|
| Ponza | Lazio | 11 | omogenea | 0 | NA |
| Napoli/Capodichino | Campania | 11 | disomogenea | 1 | 1975 |
| Capo Palinuro | Campania | 11 | omogenea | 0 | NA |
| Gioia Del Colle | Puglia | 11 | disomogenea | 1 | 1985 |
| Latronico | Basilicata | 11 | omogenea | 0 | NA |
| Brindisi | Puglia | 11 | omogenea | 0 | NA |
| Marina Di Ginosa | Puglia | 11 | omogenea | 0 | NA |
| Lecce | Puglia | 11 | disomogenea | 1 | 1977 |
| Bonifati | Calabria | 11 | omogenea | 0 | NA |
| Crotone | Calabria | 11 | omogenea | 0 | NA |
| S. Maria Di Leuca | Puglia | 11 | omogenea | 0 | NA |
| Messina | Sicilia | 11 | disomogenea | 1 | 1974 |
| Trapani/Birgi | Sicilia | 11 | disomogenea | 1 | 1985 |
| Gela | Sicilia | 11 | disomogenea | 1 | 1996 |
| Catania/Sigonella | Sicilia | 11 | disomogenea | 2 | 1980,2003 |
| Cozzo Spadaro | Sicilia | 11 | omogenea | 0 | NA |
| Alghero | Sardegna | 11 | disomogenea | 1 | 1985 |
| Decimomannu | Sardegna | 11 | omogenea | 0 | NA |
| Capo Bellavista | Sardegna | 11 | disomogenea | 2 | 1976,1987 |
| Cagliari/Elmas | Sardegna | 11 | omogenea | 0 | NA |
| Monte Maria | Trentino-Alto Adige | 35 | omogenea | 0 | NA |
| Plata | Trentino-Alto Adige | 35 | omogenea | 0 | NA |
| Vipiteno | Trentino-Alto Adige | 35 | omogenea | 0 | NA |
| Dobbiaco | Trentino-Alto Adige | 35 | omogenea | 0 | NA |
| Anterselva di Mezzo | Trentino-Alto Adige | 35 | disomogenea | 2 | 1981,1996 |
| Bressanone | Trentino-Alto Adige | 35 | disomogenea | 3 | 1968,1977,1990 |
| Bolzano | Trentino-Alto Adige | 35 | omogenea | 0 | NA |
| Sesto | Trentino-Alto Adige | 35 | disomogenea | 3 | 1966,1985,2000 |
| Ridanna | Trentino-Alto Adige | 35 | omogenea | 0 | NA |
| S. Maddalena in Casies | Trentino-Alto Adige | 35 | disomogenea | 3 | 1971,1996,2003 |

* 11 – Rete sinottica del Servizio Meteorologico dell’Aeronautica Militare e dell’Ente Nazionale di Assistenza al Volo;
4, 5, 6 – Rete del CRA-CMA (ex UCEA);
da 20 in poi – Reti regionali

Ai fini del calcolo degli indicatori rappresentativi dei valori medi sull’intero territorio nazionale, sono state privilegiate le stazioni che garantiscono una distribuzione spaziale abbastanza uniforme sul territorio e che sono risultate omogenee ai test di omogeneità. In tal modo il set di 77 stazioni è stato ridotto alle 65 mostrate nella mappa di fig. 4.4 , dove sono anche indicati con colori diversi i risultati del test di omogeneità.



Figura 4.4 – *Distribuzione geografica delle 65 stazioni selezionate per il calcolo degli indici di temperatura medi sul territorio nazionale.*

5. CALCOLO DEI VALORI NORMALI E DELLE ANOMALIE

5.1 Temperatura

Con le serie omogenee di temperatura media annuale vengono calcolati i valori climatologici normali nel trentennio di riferimento 1961-1990 e le serie annuali di anomalia rispetto ai valori normali nell'intero periodo 1961-2010.

Il valore climatologico normale per una stazione è, per definizione, la media dei trenta valori medi annuali tra il 1961 e il 1990. Poiché non tutte le 65 serie omogenee selezionate sono complete dal 1961 al 1990, è necessario anzitutto determinare i valori climatologici normali delle serie incomplete in maniera affidabile. Infatti, la mancanza di uno o più anni con temperatura media che scarta significativamente dalla media climatologica, può determinare una sovrastima o una sottostima del valore normale. Un errore di questo tipo può causare, successivamente, un errore sistematico di segno positivo o negativo sulla serie delle anomalie annuali.

Per limitare l'effetto degli anni mancanti si può adottare un metodo di "riempimento dei buchi", basato sulla osservazione che la distribuzione spaziale delle anomalie annuali di temperatura è piuttosto uniforme. Infatti, la variabilità interannuale della temperatura è determinata principalmente da sistemi e fenomeni a grande scala. In mancanza di un dato di temperatura media della stazione i per l'anno a , esso può essere quindi utilmente sostituito dal suo valore normale più l'anomalia media di quell'anno calcolata sulle altre stazioni:

$$t_{i,a} = (T_{clima,i} + \bar{A}_a)$$

Partendo dalla tabella di n stazioni, di cui k con serie complete:

| | 1961 | 1962 | 1963 | 1964 | 1965 | 1966 | | 1990 | |
|---------------------|------------|----------------|----------------|--------------|----------------|----------------|--------------|------|----------------|
| Stazioni complete | staz 1 | $t_{1,1961}$ | $t_{1,1962}$ | $t_{1,1963}$ | $t_{1,1964}$ | $t_{1,1965}$ | $t_{1,1966}$ | | $t_{1,1990}$ |
| | staz 2 | $t_{2,1961}$ | $t_{2,1962}$ | $t_{2,1963}$ | $t_{2,1964}$ | $t_{2,1965}$ | $t_{2,1966}$ | | $t_{2,1990}$ |
| | staz 3 | $t_{3,1961}$ | $t_{3,1962}$ | $t_{3,1963}$ | $t_{3,1964}$ | $t_{3,1965}$ | $t_{3,1966}$ | | $t_{3,1990}$ |
| | | | | | | | | | |
| | staz k | $t_{k,1961}$ | $t_{k,1962}$ | $t_{k,1963}$ | $t_{k,1964}$ | $t_{k,1965}$ | $t_{k,1966}$ | | $t_{k,1990}$ |
| Stazioni incomplete | staz (k+1) | $t_{k+1,1961}$ | $t_{k+1,1962}$ | | $t_{k+1,1964}$ | $t_{k+1,1965}$ | | | $t_{k+1,1990}$ |
| | | | | | | | | | |
| | staz n | $t_{n,1961}$ | $t_{n,1962}$ | $t_{n,1963}$ | | | $t_{n,1966}$ | | |

il valore climatologico normale di una stazione i con serie completa è:

$$\{T_{clima}\}_i = \frac{1}{30} \sum_{a=1961}^{1990} t_{i,a} \quad i=1,k$$

L'anomalia dell'anno a di una stazione i con serie completa è:

$$A_{i,a} = t_{i,a} - \{T_{clima}\}_i \quad i=1,k$$

L'anomalia media dell'anno a calcolata su tutte le stazioni con serie complete è:

$$\bar{A}_a = \frac{1}{k} \sum_{i=1}^k A_{i,a}$$

Il valore normale climatologico incognito della generica stazione i è:

$$T_{clima,i} = \frac{1}{30} \sum_{a=1961}^{1990} t_{i,a}$$

$$= \frac{1}{30} \left[\sum_{a=\text{anni con dato presente}} t_{i,a} + \sum_{a=\text{anni con dato assente}} (T_{clima,i} + \bar{A}_a) \right]$$

Risolvendo per $T_{clima,i}$:

$$T_{clima,i} \left(1 - \frac{N}{30}\right) = \frac{1}{30} \sum_{a=\text{anni con dato presente}} t_{i,a} + \frac{1}{30} \sum_{a=\text{anni con dato assente}} \bar{A}_a$$

dove N è il numero di anni con dato assente per la stazione i .

Infine:

$$T_{clima,i} = \frac{1}{30 - N} \left[\sum_{a=\text{anni con dato presente}} t_{i,a} + \sum_{a=\text{anni con dato assente}} \bar{A}_a \right]$$

Il procedimento descritto è stato applicato al calcolo dei valori climatologici normali di temperatura media delle 65 stazioni selezionate con serie omogenee. Di queste, 30 serie sono complete nel periodo 1961-1990.

Dopo aver calcolato i valori climatologici normali, si passa al calcolo dei valori annuali di anomalia della temperatura media di ciascuna stazione:

$$A_{i,a} = t_{i,a} - T_{clima,i} \quad i=1, n \quad n=65 \quad a=1961, 2010$$

Vogliamo ora calcolare la serie di anomalia media in Italia, come media sul set di 65 stazioni selezionate e distribuite in maniera abbastanza uniforme sul territorio nazionale. La tabella $A_{i,a}$ dei valori di anomalia presenta dei “buchi” in corrispondenza dei valori di $t_{i,a}$ mancanti (ricordiamo che le serie selezionate hanno da un minimo di 43 a un massimo di 50 anni con indicatore valido nel periodo 1961-2010, v. tab. 3.1). Tuttavia, il calcolo dell’anomalia media in Italia di ogni anno a , si può effettuare utilizzando tutte e sole le stazioni con $A_{i,a}$ presente, anche se, anno per anno, in numero inferiore a 65. Di nuovo, ciò è supportato dal fatto che le anomalie di un certo anno sono in generale abbastanza uniformi spazialmente, e che l’assenza di un numero limitato di stazioni non comporta un errore rilevante nella stima dell’anomalia media di quell’anno.

Il risultato del procedimento descritto è mostrato in figura 5.1, dove la serie annuale delle anomalie della temperatura media in Italia viene rappresentata e messa a confronto con la serie analoga della temperatura media globale sulla terraferma.

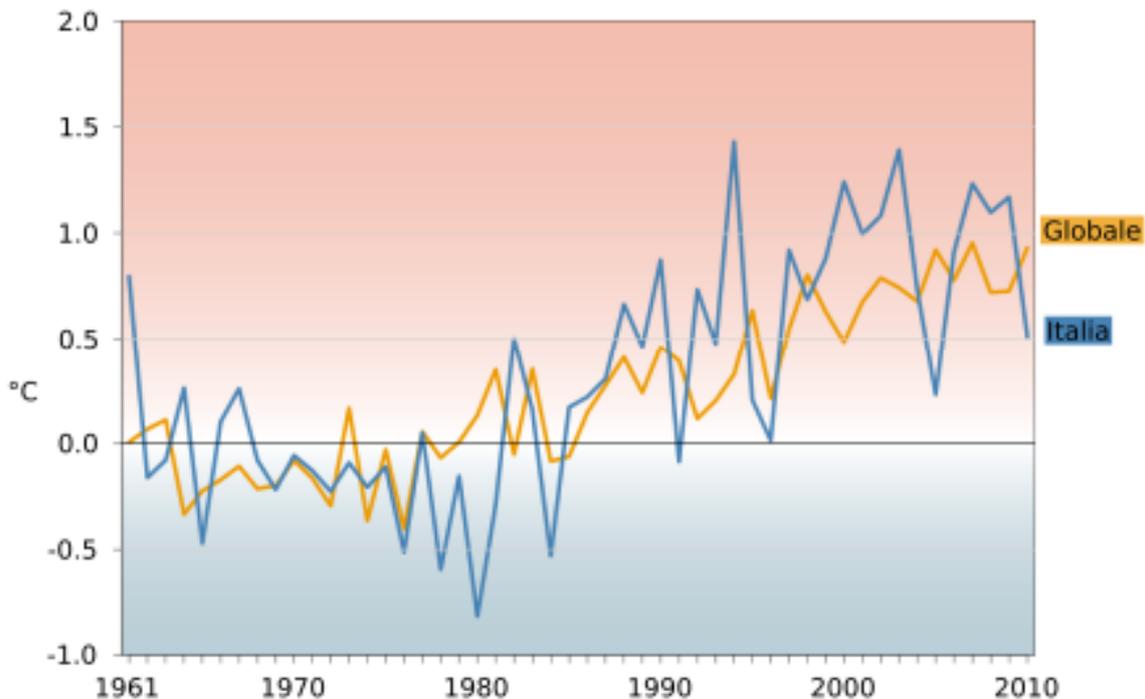


Figura 5.1 – Serie temporale 1961-2010 delle anomalie di temperatura media globale sulla terraferma (fonte: NCDC/NOAA) e in Italia, rispetto ai valori climatologici normali 1961-1990.

5.2 Precipitazione

Come già accennato alla fine del cap. 3, la precipitazione ha una variabilità spaziale e temporale molto più marcata rispetto alla temperatura. La serie dei valori annuali di precipitazione di una stazione viene utilizzata per la stima del valore normale, delle anomalie e delle tendenze di precipitazione sulla località o sull'area geografica di dimensioni limitate, di cui la stazione stessa è rappresentativa. Le 117 stazioni selezionate con serie complete e continue secondo i criteri illustrati nel capitolo 3 non sono sufficienti o idonee alla stima dei valori medi su aree geografiche di grande estensione o addirittura sull'intero territorio nazionale. Al fine di calcolare le medie spaziali degli indicatori, la natura stessa della variabile, la sua elevata variabilità spaziale e temporale e anche le discontinuità dei dati disponibili, suggeriscono invece di interpolare spazialmente i dati prima di effettuare il calcolo dei valori normali e delle anomalie. In tal modo può essere ottimizzato l'uso della base dati disponibile, ai fini della stima delle tendenze di precipitazione in Italia.

Per il calcolo dei valori normali, i metodi di riempimento dei “buchi” delle serie descritti per la temperatura non possono essere adottati a causa della elevata variabilità spaziale delle anomalie annuali. Nel nostro caso, tuttavia, il numero di stazioni con serie utili al calcolo del valore climatologico normale è molto elevato, grazie alle serie di precipitazione della rete dell'ex Servizio Idrografico e Mareografico Nazionale. In particolare, prendendo in considerazione il periodo 1951-1980, si dispone delle serie annuali di circa 2400 stazioni ben distribuite su tutto il territorio nazionale. Grazie alla elevata densità e alla distribuzione spaziale uniforme delle stazioni, è stato possibile calcolare i valori normali di precipitazione 1951-1980 su grigliato regolare alla risoluzione di 1x1 km, mediante *ordinary kriging* (Li and Heap, 2008).

Volendo calcolare la serie annuale dell'anomalia di precipitazione nell'intervallo 1951-2010, per ciascun anno è disponibile un set di stazioni, variabile di anno di anno, con valore di precipitazione annuale P_{ia} . L'anomalia annuale della stazione i (espressa come differenza percentuale rispetto al valore climatologico) viene calcolata come:

$$A_{ia} = 100 (P_{ia} - P_{clima\ i,a}) / P_{clima\ i,a}$$

dove $P_{clima\ i,a}$ è il valore climatologico normale del punto di griglia più vicino alla stazione i , estratto dalla mappa dei valori normali ad alta risoluzione. In tal modo si dispone di un certo numero (il maggior numero possibile per ogni anno) di punti da utilizzare per il calcolo dell'anomalie annuali su aree geografiche di grandi dimensioni (l'intero territorio italiano, oppure suddiviso tra Nord, Centro e Sud) mediante l'impiego di algoritmi geostatistici.

Nel caso delle precipitazioni, tra gli algoritmi più diffusi il più idoneo allo scopo è quello dei "poligoni di Thiessen" o della "tassellatura di Voronoi" (Li and Heap, 2008). Il metodo consiste nel ripartire una regione piana R , a partire da un insieme di N punti appartenenti ad R , in N regioni R_p , con $p=1 \dots N$, ciascuna associata ad un punto p , in modo che tutti i punti appartenenti a R_p siano più vicini a p che ad ogni altro punto in R . I valori annuali di anomalia vengono quindi sottoposti a tassellatura di Voronoi, e l'anomalia annuale precipitazione di una macro-regione (per esempio, l'Italia settentrionale) risulta essere la media pesata delle anomalie sui punti stazione p , dove il peso è dato dal rapporto tra l'area intersezione della regione R_p con l'area totale della macro-regione, e l'area della macro-regione.

E' preferibile effettuare la media pesata dei valori di anomalia percentuale, piuttosto che sui valori di precipitazione, in quanto il campo di anomalia è relativamente più liscio e si riduce quindi l'errore nelle aree a bassa densità di stazioni con dati utili. A titolo di esempio, nella figura 5.2 viene mostrata la ripartizione del territorio italiano nei poligoni corrispondenti alle stazioni con il dato di anomalia di precipitazione dell'anno 2010.

Con il metodo descritto, sono state calcolate le serie annuali di anomalia di precipitazione rispetto al valore climatologico 1951-1980 nell'intervallo 1951-2010 per le tre macro-aree: Nord, Centro, Sud e Isole (figg. 5.3).

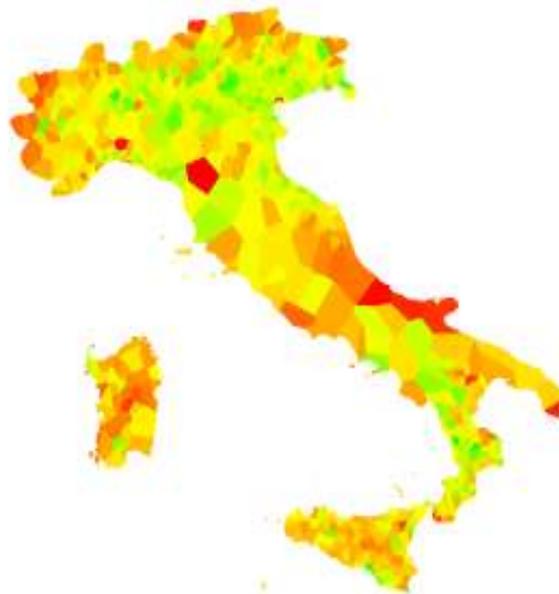


Figura 5.2 – Ripartizione del territorio italiano nei poligoni di Voronoi corrispondenti alle stazioni con il dato di anomalia di precipitazione dell'anno 2010.

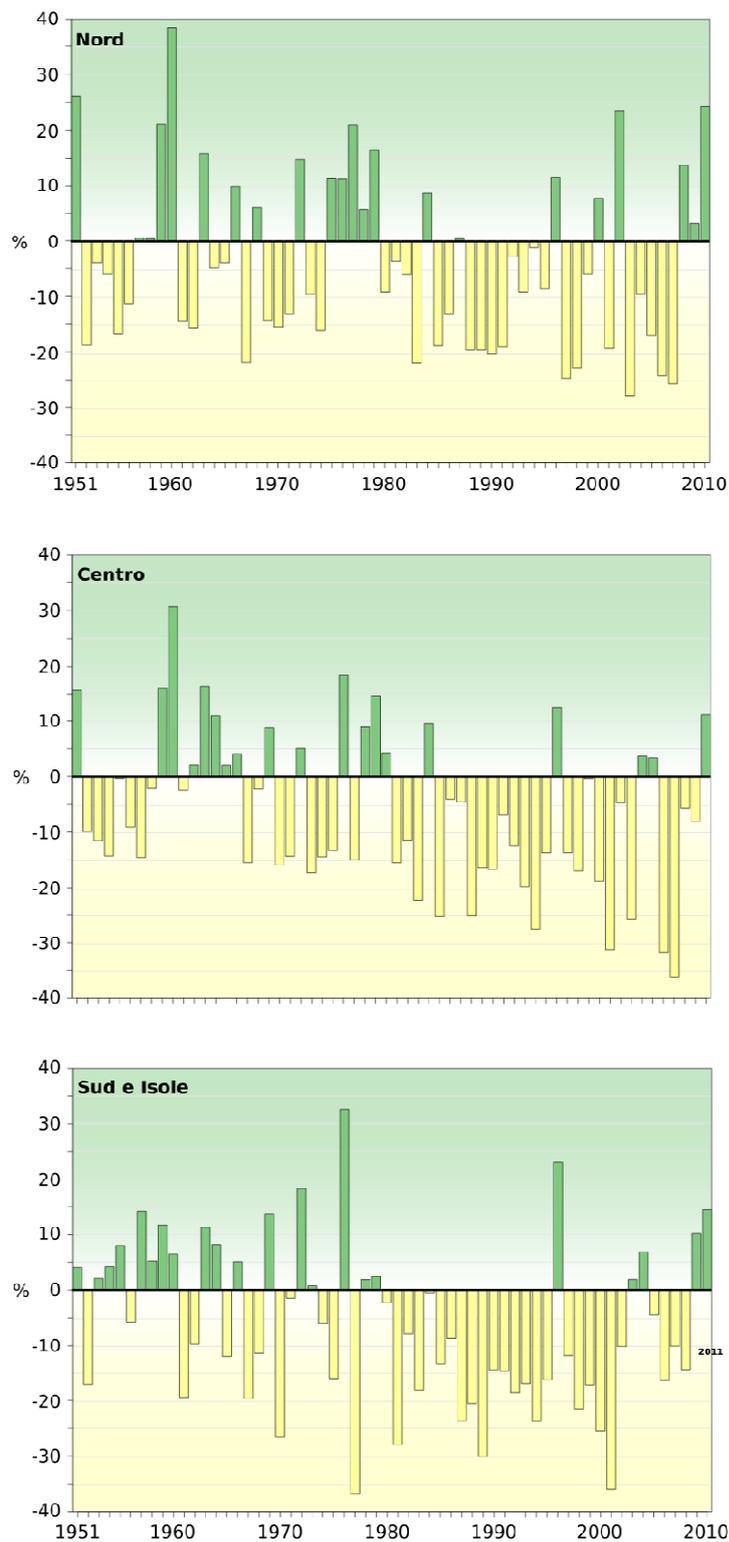


Figura 5.3 – Serie annuali delle anomalie percentuali di precipitazione rispetto al valore climatologico 1951-1980, per le tre macro-aree: Nord, Centro, Sud e Isole.

6. STIMA DELLE TENDENZE

Le serie annuali di anomalia calcolate con i metodi descritti nel capitolo 5, si prestano a valutare la presenza di variazioni significative della temperatura media e delle precipitazioni cumulate in Italia negli ultimi decenni e a stimarne l'entità. In questo capitolo vengono descritti alcuni strumenti analitici utilizzati per la stima delle tendenze di queste variabili; tuttavia, strumenti analoghi possono essere applicati anche allo studio di altre variabili climatiche.

In termini generali, la serie temporale di una variabile climatica $x(t)$ può essere espressa come la somma di due componenti:

$$x(t) = f(t) + e(t)$$

dove $f(t)$ è il valore atteso di $x(t)$, mentre $e(t)$ rappresenta la componente *random* o casuale che è legata principalmente alla variabilità climatica naturale.

Un'ipotesi molto comune è quella di identificare la sequenza $e(t)$ con un processo "white noise", cioè una sequenza di variabili casuali indipendenti (che non si influenzano vicendevolmente) ed identicamente distribuite (il modello probabilistico che descrive tali variabili non cambia nel tempo) con media nulla e varianza costante. Secondo la precedente formula una serie storica non rappresenta solo una sequenza di valori casuali indipendenti fra loro; piuttosto, essi possono essere descritti attraverso una funzione o un modello matematico con i quali vengono messi in relazione i valori attuali con quelli passati. La modellizzazione di una serie storica consiste allora nella scelta del modello analitico più opportuno per descriverne le componenti $f(t)$ ed $e(t)$. Qualunque sia il modello prescelto per la componente $f(t)$, l'analisi delle tendenze (*trend*) ha come obiettivo quello di valutare se essa rimanga costante nel tempo oppure no. E in quest'ultimo caso, quello di quantificare l'entità di tale variazione. Ricorrendo alla terminologia statistica della verifica delle ipotesi, l'analisi dei *trend* mette a confronto le ipotesi:

H0: $f(t)$ costante (invariante nel tempo)

H1: $f(t)$ non costante

Un modello comunemente adottato per rappresentare l'evoluzione dei dati nel tempo è quello lineare. Tale scelta è motivata in primo luogo dalla semplicità analitica di tale modello. Inoltre, nel caso delle serie di anomalia di temperatura e precipitazione, esso consente in prima battuta di verificare la presenza o meno di una tendenza significativa (con il suo segno), a ritmo costante, nell'intervallo di tempo fissato. La variabile $x(t)$ può allora essere espressa come:

$$x(t) = a + bt + e(t) \quad (\text{modello di regressione lineare semplice})$$

con gli errori $e(t)$ indipendenti e identicamente distribuiti (i.i.d.)

Se si assume che gli errori seguano una distribuzione normale, è possibile ottenere una stima ottimale (in senso statistico) per i parametri "a" e "b", rispettivamente l'intercetta e la pendenza (*slope*) della retta. La stima del coefficiente b fornisce una quantificazione del *trend*, la cui significatività statistica può essere facilmente verificata mediante test. Un approccio come quello descritto dalla precedente formula viene detto "parametrico" in quanto basato sull'assunzione che i dati osservati seguano una determinata funzione di distribuzione, in questo caso quella normale.

Un'alternativa non parametrica, impiegata spesso nello studio delle tendenze climatiche (v. per esempio Toreti e Desiato, 2007), è costituita dall'uso combinato del test di Mann-Kendall (Mann, 1945; Kendall, 1975) e dello stimatore di Theil-Sen (Sen, 1968). Lo stimatore di Theil-Sen prescinde dall'ipotesi di normalità dei dati, da cui sovente i dati climatici si discostano ed è robusto rispetto alla

presenza di dati anomali (*outliers*). Poiché lo stimatore di Thiel-Sen non fornisce alcuna informazione riguardo alla significatività del *trend* stimato, si può utilizzare il test non parametrico di Mann-Kendall a tale scopo. Le ipotesi prese in esame dal test sono:

H0: assenza di *trend*

H1: presenza di un *trend* monotono (non necessariamente lineare)

Il test si basa sul concetto di ranghi dei dati, ovvero sulla posizione che ciascun valore occupa all'interno della serie ordinata. La presenza di autocorrelazione positiva è un evento abbastanza comune per le variabili climatiche e ciò può inficiare i risultati del test di Mann-Kendall, favorendo l'ipotesi H1 (*trend* monotono) a discapito dell'ipotesi H0 (assenza di *trend*) (von Storch, 1995; Zhang et al., 2000). Si può ovviare a tale inconveniente facendo precedere l'analisi dei *trend* dalla procedura di "pre-whitening", che permette di rimuovere l'autocorrelazione dalla serie temporale.

L'uso di una retta per rappresentare il *trend* di una serie storica poggia sull'idea che il fenomeno oggetto di studio evolva gradualmente nel tempo, senza cambiamenti repentini (*change point*, talvolta indicati anche come *breakpoint*, da non confondersi con quelli che riguardano l'omogeneizzazione delle serie, v. cap. 4). Tuttavia, diversi lavori, tra cui quello di Tomé and Miranda (2004), Seidel and Lanzante (2004), hanno messo in evidenza che ciò non è sempre vero. Diviene allora necessario prendere in esame modelli più complessi di un semplice *trend* lineare. A tal proposito, Tomé e Miranda (2004), riprendendo un lavoro di Karl et al. (2000), illustrano l'uso di un algoritmo per determinare in maniera ottimale il numero e localizzare eventuali *change point* in cui può verificarsi una variazione del *trend*. Tale approccio consente, secondo gli autori, un calcolo più preciso del *trend* lineare complessivo come media pesata dei *trend* lineari parziali, calcolati per ciascun periodo parziale avente come estremi i *change point* identificati.

Partendo dai risultati di Tomé e Miranda, Toreti e Desiato (2007) hanno confrontato i risultati ottenuti impiegando un modello lineare semplice, contro quelli forniti dall'uso di un modello "piecewise" e "sloped steps" dopo aver identificato un probabile *change point* mediante l'analisi progressiva di Mann-Kendall (Sneyers, 1990).

6.1 Temperatura

La serie di anomalia della temperatura media in Italia (1961-2010), è stata sottoposta in prima battuta all'analisi del *trend* lineare impiegando sia un approccio di tipo parametrico (modello di regressione lineare semplice) che uno non parametrico (stimatore di Theil-Sen e test di Mann-Kendall).

La stima del *trend* lineare (*slope*), ottenuta mediante modello di regressione lineare semplice, è pari a $+0.026$ °C / anno, corrispondente a un incremento complessivo di 1.27 °C nel periodo 1961-2010.

L'uso di un modello di regressione lineare richiede, tra le varie ipotesi, che i dati seguano una distribuzione normale. In effetti, l'analisi dei dati mostra che le anomalie di temperatura presentano un'evidente asimmetria a destra che viola l'ipotesi di normalità. Per questo motivo, il *trend* è stato stimato anche mediante un approccio non parametrico. Lo *slope* fornito dallo stimatore Theil-Sen risulta essere pari a $+0.027$ °C/anno. Come evidenziato dalla Figura 6.1, i risultati ottenuti mediante un approccio di tipo parametrico e uno di tipo non parametrico sono molto simili. La significatività del *trend* lineare è stata verificata mediante il test di Mann-Kendall. L'uso di tale test non ha richiesto il "pre-whitening" dei dati, ovvero la rimozione di un'eventuale correlazione seriale nei residui. L'assenza di correlazione seriale nei residui è stata verificata mediante il test di Wald-Wolfowitz (Sneyers, 1990).

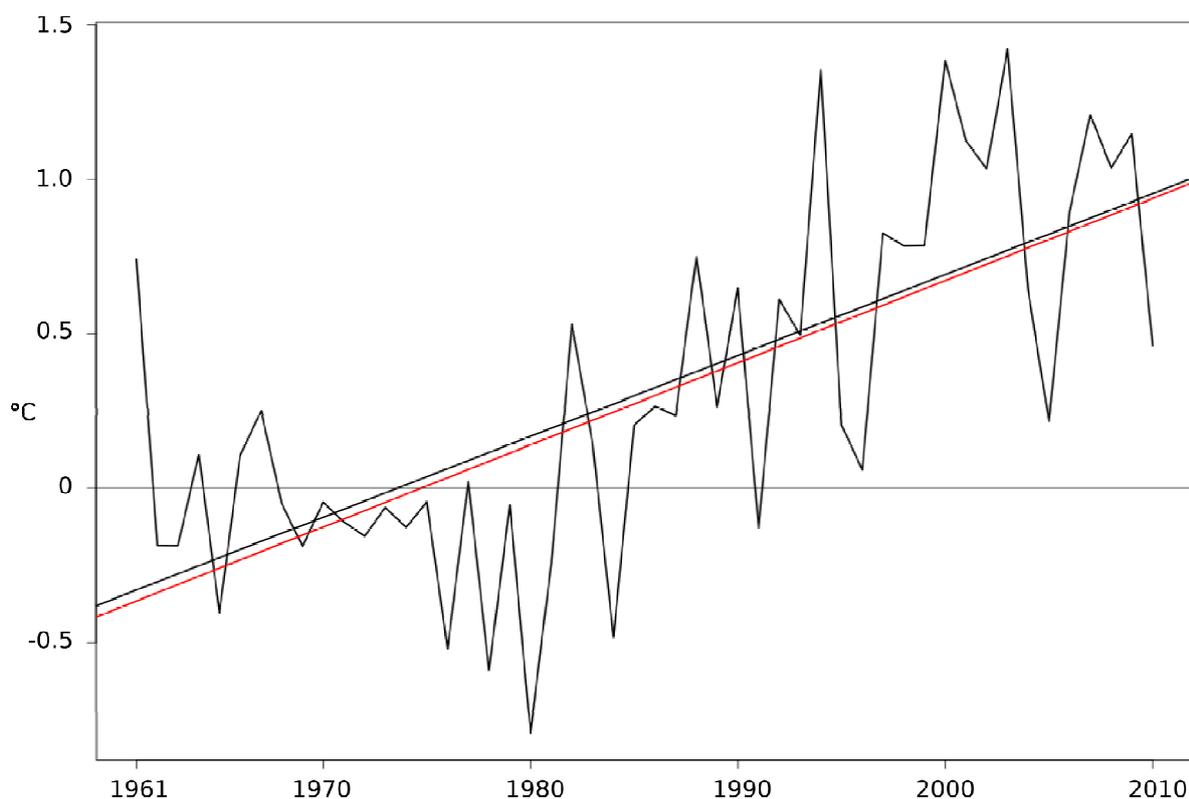


Figura 6.1 – Serie temporale 1961-2010 delle anomalie di temperatura media in Italia e relativo trend lineare stimato con il modello di regressione lineare semplice (linea nera) e con lo stimatore di Theil-Sen (linea rossa).

L'uso di un modello "piecewise" permette di raffinare l'analisi del trend lineare, ammettendo la presenza di uno o più *change point*. Come osservato da Tomé e Miranda (2004), la ricerca di un eventuale *change point* deve prima di tutto essere condotta ispezionando "visivamente" la serie. Il grafico delle anomalie di temperatura media sembra suggerire un'inversione di tendenza verso la fine degli anni '70. Ammettendo la possibilità di un solo *change point* nell'intervallo 1961-2010, il pacchetto R "segmented" ne identifica uno probabile nel 1976 (Figura 6.2). Il modello *piecewise* utilizzato per descrivere il trend nei due periodi separati dal *change point* risulta caratterizzato da un trend negativo ($-0.020^{\circ}\text{C}/\text{anno}$) per la serie delle anomalie dal 1961 al 1976, e da un trend positivo nell'intervallo 1976-2010 ($+0.042^{\circ}\text{C}/\text{anno}$). La variazione della temperatura media in Italia così calcolata per l'intero periodo di 50 anni 1961-2010 risulta essere di 1.13°C . La figura 6.2 evidenzia che l'uso di un modello "piecewise" impone la condizione di continuità delle due rette di regressione in corrispondenza del "change point".

Un'alternativa al modello *piecewise* è il modello "sloped steps" in cui viene meno tale condizione di continuità. La figura 6.3 illustra la serie delle anomalie di temperatura 1961-2010 e il modello *sloped steps* stimato dopo aver identificato il probabile *change point* mediante l'analisi sequenziale di Sneyers. Il punto di intersezione fra le due curve (tratteggiata e continua) del grafico in basso, che rappresenta l'analisi sequenziale di Sneyers, fornisce un'indicazione della posizione del probabile *change point* (1980). Un trend negativo ($-0.029^{\circ}\text{C}/\text{anno}$) caratterizza il periodo 1961-1980, mentre un trend positivo ($+0.035^{\circ}\text{C}/\text{anno}$) caratterizza il periodo 1981-2010. L'incremento complessivo sul cinquantennio 1961-2010 risulta essere di 0.94°C .

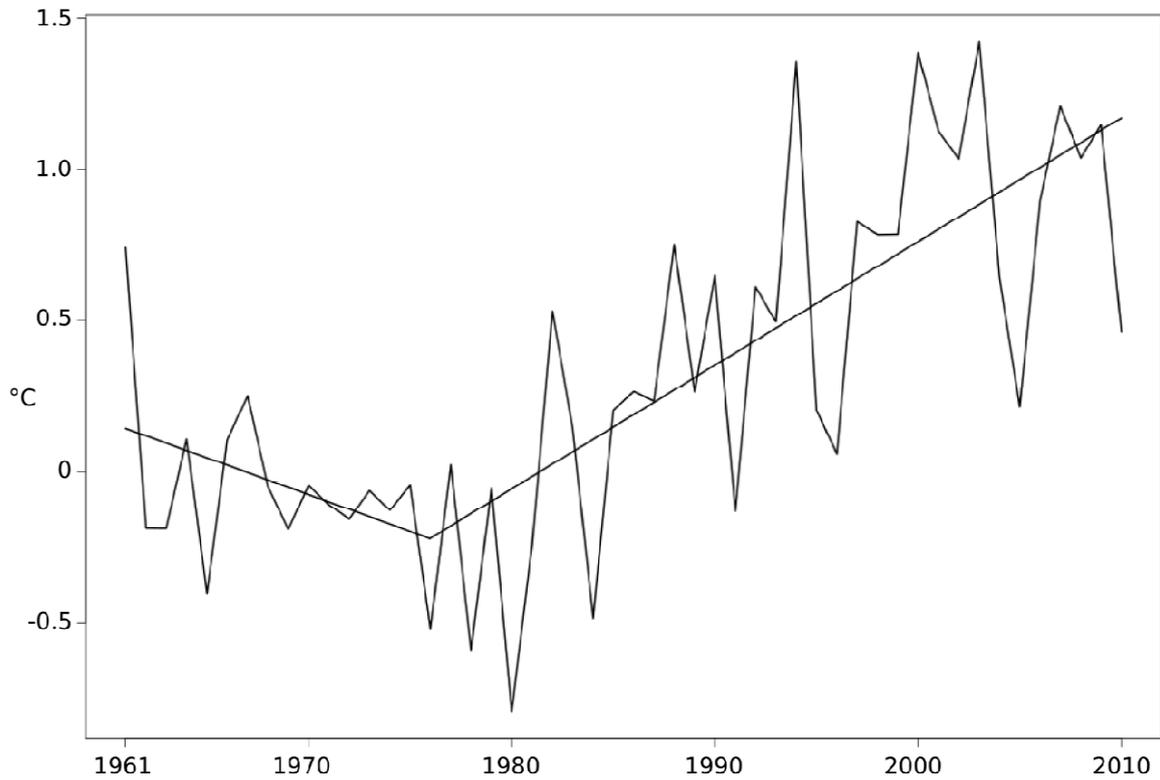


Figura 6.2 – Serie temporale 1961-2010 delle anomalie di temperatura media in Italia. L'utilizzo di un modello "piecewise" mette in evidenza un changepoint nel 1976.

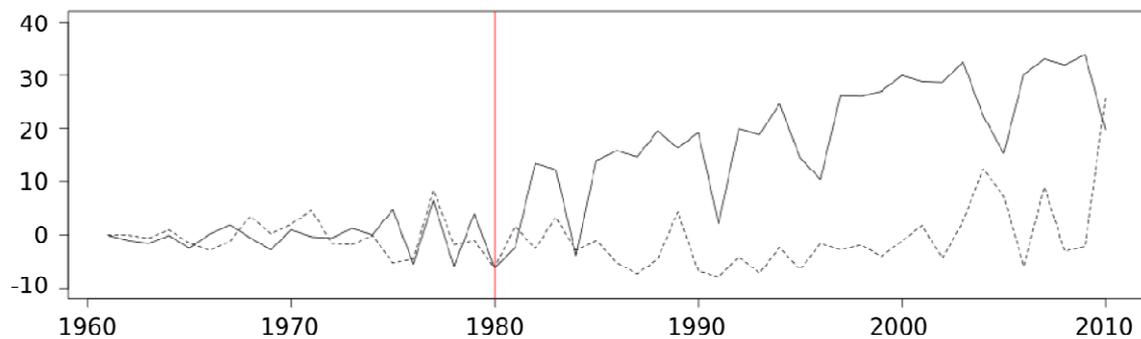
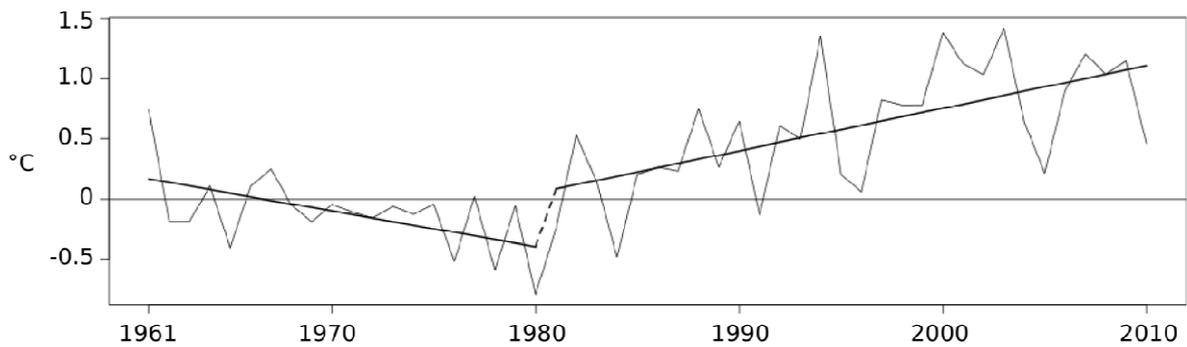


Figura 6.3 – Serie temporale 1961-2010 delle anomalie di temperatura media in Italia. L'utilizzo di un modello "sloped steps" mette in evidenza un changepoint nel 1980. Le due curve del grafico in basso rappresentano l'analisi sequenziale di Sneyers.

I diversi *change point* individuati mediante l'analisi di Sneyers e mediante il pacchetto “*segmented*” di R danno un'indicazione dell'incertezza che caratterizza l'individuazione del *change point*. Dall'analisi dei due risultati si può affermare che una inversione del *trend* della serie di temperatura si è probabilmente verificata tra il 1976 e il 1980. Questa conclusione non si discosta molto dal risultato di uno studio precedente sulla serie di temperatura media in Italia nel periodo 1961-2004 (Toreti e Desiato, 2008), in cui era già stata rilevata la presenza di *trend* analoghi (prima debolmente negativo, poi positivo) con un *change point* collocato nel 1981.

6.2 Precipitazione

L'analisi del *trend* delle anomalie percentuali di precipitazione è stata condotta sulla serie che copre il periodo 1951-2010. L'uso di un modello di regressione lineare semplice fornisce i seguenti risultati: al Nord un *trend* negativo non significativo pari $-0.155\%/anno$; al Centro un *trend* negativo pari a $-0.289\%/anno$; al Sud e sulle Isole, un *trend* negativo pari a $-0.230\%/anno$. Per il Centro, il *trend* stimato si traduce in una diminuzione della precipitazione cumulata annuale del 17.0% dal 1951 al 2010; per il Sud e le Isole si stima una diminuzione del 13.6%. Come per la serie delle anomalie di temperatura, i risultati ottenuti sono stati confrontati con quelli forniti dallo stimatore di Theil-Senn e dal test di Mann-Kendall. Anche in questo caso si ottengono dei risultati in linea con quelli già riportati. Per il Nord un *trend* pari a $-0.179\%/anno$, per il Centro un *trend* pari a $-0.290\%/anno$ e per il Sud un *trend* pari a $-0.270\%/anno$ (fig.6.4). L'uso del test di Mann-Kendall conferma la non significatività del *trend* al Nord, e la significatività dei *trend* al Centro e al Sud. Per quanto riguarda l'eventuale presenza di correlazione seriale nei residui, valgono le stesse considerazioni fatte nel caso delle anomalie di temperatura: assenza di correlazione seriale nei residui e quindi nessuna necessità di adottare tecniche di “*pre-whitening*” prima del test di Mann-Kendall.

Per quanto riguarda la presenza di *change point* e quindi la necessità di adottare un modello “*sloped-step*” invece del più semplice modello lineare, l'analisi dei dati non ha evidenziato alcun punto in cui le tre serie di anomalie manifestano un cambio di *trend*. Tale risultato è in linea con quello ottenuto precedentemente sulle serie di precipitazione nel periodo 1961 – 2006 (Toreti et al., 2009).

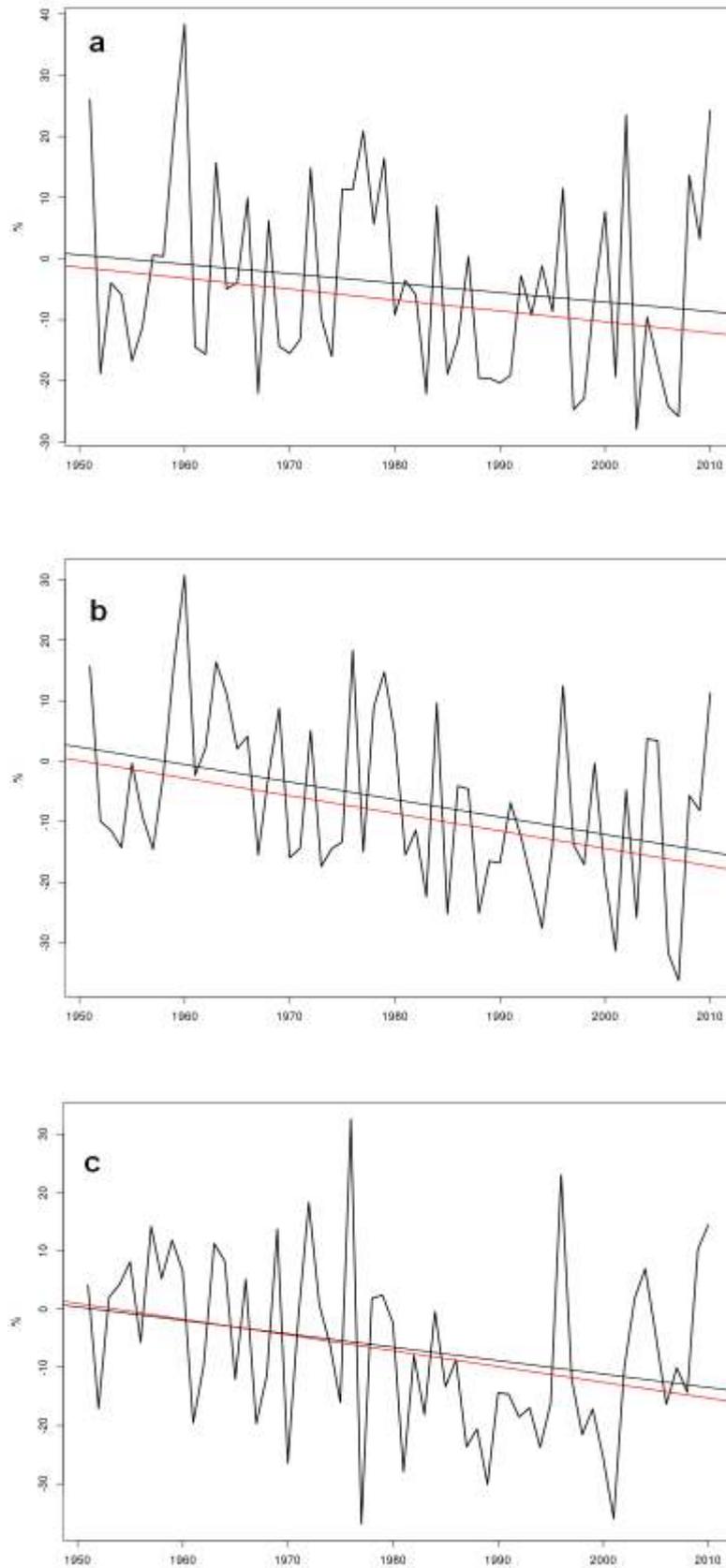


Figura 6.4 - Anomalie percentuali di precipitazione 1951-2010 in Italia e relativo trend lineare stimato con il modello di regressione lineare semplice (linea nera) e con lo stimatore di Theil-Sen (linea rossa): Nord (a), Centro (b), Sud e Isole (c).

BIBLIOGRAFIA

- Aguilar E., Auer I., Brunet M., Peterson T. C., Wieringa J., 2003: "Guidelines on climate metadata and homogenisation". World Meteorological Organization, WMO-TD No. 1186, WCDMP No. 53, Geneva, Switzerland, 55 pp.
- Alexandersson H., 1986: "A homogeneity test applied to precipitation data". *J. Climate*, 6, 661–675.
- Alexandersson H., Moberg A., 1997: "Homogenization of Swedish temperature data. Part I: A homogeneity test for linear trends". *Int. J. Climatol.*, 17, 25–34.
- Baffo F., Suatoni B., Desiato F., 2005: "Indicatori climatici: i controlli di validità e la ricerca dei valori errati". *Bollettino Geofisico*, Anno XXVIII, N. 1-2.
- Conrad V., Pollack C., 1950: "Methods in Climatology". Harvard University Press, 459 pp.
- 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., Toreti A., 2011: "Climate indicators for Italy: calculation and dissemination". *Adv. Sci. Res.*, 6, 147-150.
- Desiato F., Lena F., Toreti A., 2006: "Un sistema per tutti – climatologia: i dati italiani". *Sapere*, Anno 72, n. 2, 62-69.
- Desiato F., Lena F., Toreti A., 2007: "SCIA: a system for a better knowledge of the Italian climate". *Bollettino di Geofisica Teorica ed Applicata*, Vol. 48, n. 3, 351-358.
- ISPRA, 2011a: "Gli indicatori del clima in Italia nel 2010 (anno VI)". *Rapporto ISPRA, Serie Stato dell'Ambiente* 24/2011.
- ISPRA, 2011b: "Tematiche in primo piano, Annuario dei dati ambientali 2010". *Rapporto ISPRA* 20/2011.
- Karl T. R., Knight R. W., Baker B., 2000: "The record breaking global temperatures of 1997 and 1998: Evidence for an increase in the rate of global warming?". *Geophys. Res. Lett.*, 27, 719 – 722.
- Kendall MG., 1975: "Rank Correlation Methods". Griffin: London.
- Kuglitsch F. G., Toreti A., Xoplaki E., Della-Marta P. M., Luterbacher J., Wanner H., 2009: "Homogenization of daily maximum temperature series in the Mediterranean". *J. Geophys. Res.*, 114, D15108.
- Li J., Heap A.D., 2008: "A review of spatial interpolation methods for environmental scientists". *Geoscience Australia, Record* 2008/23, 137 pp.
- Lund R., Reeves J., 2002: "Detection of undocumented changepoints: a revision of two phase regression model". *J. Climate* 15, 2547-2554.
- Mann H.B., 1945: "Nonparametric tests against trend". *Econometrica*, 13, 245–259.
- Peterson T.C., Easterling D.R., 1994: "Creation of homogeneous composite climatological reference series". *Int. J. Climatol.*, 14, 691-679.
- Peterson T.C., Easterling D.R., Karl T.R., Groisman P., Nicholls N., Plummer N., Torok S., Auer I., Böhm R., Gullett D., Vincent L., Heino R., Tuomenvirta H., Mestre O., Szentimrey T., Salinger J., Førland E.J., Hanssen-Bauer I., Alexandersson H., Jones P., Parker D., 1998: "Homogeneity adjustments of in situ atmospheric climate data: a review". *Int. J. Climatol.*, 18, 1493-1517.
- Reeves J., Chen J., Wang X. L., Lund R., Lu Q., 2007: "A review and comparison of changepoint detection techniques for climate data". *J. Appl. Meteorol. Climatol.*, 46, 900–915.
- Seidel D.J., Lanzante J.R., 2004: "An assessment of three alternatives to linear trends for characterizing global atmospheric temperature changes". *J. of Geophys. Res.*, 109, D14108.
- Sen P. K., 1968: "Estimates of the regression coefficient based on Kendall's tau". *Journal of American Statistical Association*, 63, 1379–1389.
- Sneyers R., 1990: "On the statistical analysis of series of observations". *WMO Technical note* no. 143, Geneva, 192 pp.
- Tomé A. R., Miranda P. M. A., 2004: "Piecewise linear fitting and trend changing points of climate parameters". *Geoph. Res. Lett.*, 31, L02207.

-
- Toreti A., Desiato F., 2008: "Temperature trend over Italy from 1961 to 2004". *Theor. Appl. Climatol.*, 91, 51-58.
- Toreti A, Fioravanti G, Perconti W, Desiato F., 2009: "Annual and seasonal precipitation over Italy from 1961 to 2006". *Int. J. Climatol.*, 29, 1976-1987.
- Vincent L., 1998: "A technique for the identification of inhomogeneities in Canadian temperature series". *J. Climate*, 11, 1094–1104.
- Vincent L. A., Zhang X., 2001, "Homogenization of daily temperatures over Canada". *J. Climate*, 15, 1322-1334.
- Von Storch H., 1995: "Misuses of statistical analysis in climate research". In: *Analysis of climate variability: Applications of statistical techniques*. H. Von Storch e A. Navarra (Eds) Springer-Verlag Berlin, pp. 11–26.
- Wang X. L., 2003: "Comments on: Detection of undocumented change points: A revision of the two-phase regression model". *J. Climate*, 16, 3383-3385.
- Wang X. L., 2008a: "Penalized maximal F-test for detecting undocumented mean-shifts without trend-change". *J. Atmos. Oceanic Techn.*, 25, 368-384.
- Wang X. L., 2008b: "Accounting for autocorrelation in detecting mean-shifts in climate data series using the penalized maximal t or F test". *J. Appl. Meteor. Climatol.*, 27, 2423-2444.
- Wang X. L., Q. H. Wen, Wu Y., 2007: "Penalized maximal t test for detecting undocumented mean change in climate data series". *J. Appl. Meteor. Climatol.*, 46, 916-931.
- Wijngaard A. M., Klein Tank G., Können G. P., 2003: "Homogeneity of 20th century European daily temperature and precipitation series". *Int. J. Climatol.* 23, 679–692.
- Zhang X., Vincent L. A., Hogg W. D., Niitsoo A., 2000: "Temperature and precipitation trends in Canada during the 20th century". *Atmos. Ocean*, 38, 395 – 429.
- Zurbenko I., Porter P. S., Rao S. T., Ku J. Y., Gui R., Eskridge, R. E., 1996: "Detecting discontinuities in time series of upper-air data: development and demonstration of an adaptive filter technique". *J. Climate*, 9, 3548-3560.