

Analisi del modello Random Walk per le variazioni di temperatura e studio della correlazione con i gas serra

Tommaso Di Luciano

Introduzione

Il report ha come finalità studiare i comportamenti delle anomalie di temperatura nell'ipotesi che si comportino come un Random Walk e la loro correlazione con concentrazioni di gas serra su diverse scale temporali. I dataset utilizzati nello studio¹ sono relativi a diversi span temporali

Premesse e note metodologiche

- Misure dirette della temperatura sono state effettuate solo negli ultimi 150 anni, dati relativi ad anni precedenti al 1880 risalgono a ricostruzioni effettuate attraverso misure indirette. In particolare dati utilizzati in questa analisi sono stati ricavati da carotaggi del ghiaccio in antartide e misure delle concentrazioni di isotopi dell'ossigeno negli oceani. I dati globali provengono da medie effettuate in diverse parti del mondo da parte dei ricercatori che hanno composto il dataset.
- I dati utilizzati per analizzare i gas serra a grande distanza temporale provengono da carotaggi. sono stati paragonati i dati di gas serra e variazioni di temperatura solo con dataset provenienti dalla stessa fonte. Per quanto riguarda i dati locali sono stati paragonati dati relativi allo stesso sito di carotaggio ed allo stesso periodo temporale.
- In ognuna delle analisi effettuate sono state confrontate diverse serie temporali solo se la densità delle misure nell'arco di tempo della serie permetteva di confrontarle tra di loro.

Risultati

- Lo studio del MSD per le anomalie mostra comportamenti differenti sulle diverse scale temporali.

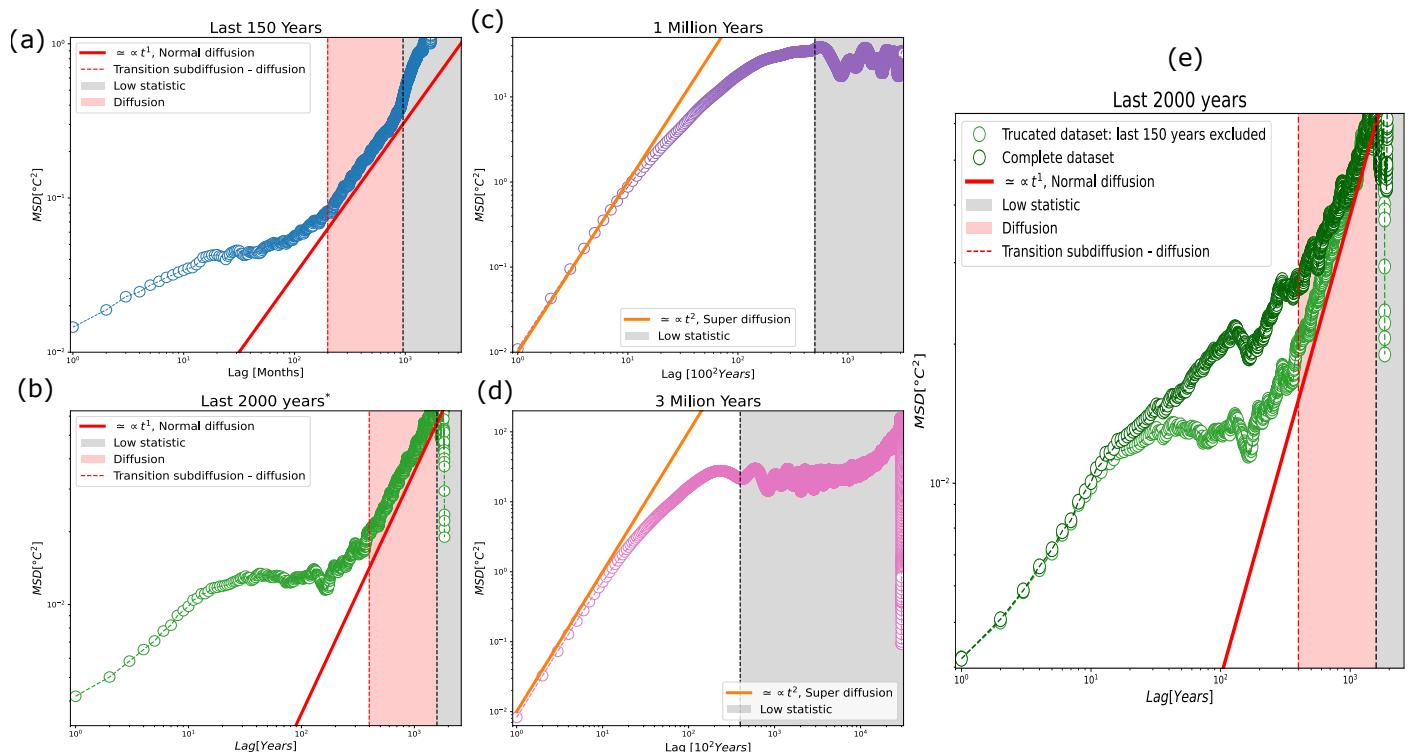


Figura 1: I primi 4 pannelli sono relativi agli MSD delle anomalie di temperatura in funzione di un lag temporale (a) per gli ultimi 150 anni con lag mensile, (b) per gli anni dallo 0 a.C. 1850 con lag annuale, (c) per l'ultimo milione anni con lag di 100 anni, (d) per gli ultimi 3 milioni di anni con lag di 100 anni. (e) Confronto dell'MSD visto al pannello (b): in scuro MSD per gli ultimi 2000 anni, in chiaro vengono scartati gli ultimi 150 anni. Tutti vengono mostrati in scala log-log.

In figura 1 vengono mostrati i risultati del calcolo del *mean square displacement* (MSD) per i diversi dataset in funzione di un lag temporale che varia per ogni dataset:

$$MSD(\tau) = \langle (x(t + \tau) - x(t))^2 \rangle$$

Dove $x(t)$ è la anomalia di temperatura al tempo t e τ è il lag². Questa osservabile ci permette di caratterizzare la diffusione all'interno di un processo stocastico³ e recentemente ha mostrato interessanti risultati in climatologia⁴. Come è noto, una diffusione *normale* è caratterizzata da uno scaling lineare con il tempo, tale che $MSD \propto t^\alpha, \alpha = 1$. Questo è un comportamento che emerge in molti sistemi come conseguenza del *Teorema del Limite Centrale* (CLT) ed è un indicatore del fatto che i singoli step hanno una distribuzione compatibile con le ipotesi di questo teorema (indipendenza, $\sigma < +\infty$).

Come risulta chiaro dai pannelli (a),(b), nel momento in cui si considerano **scale temporali ‘brevi’ per il lag τ si tende ad osservare un comportamento subdiffusivo nel breve periodo**, questo può accadere quando valori successivi del processo stocastico $x_{t+dt} - x_t$ sono correlati negativamente: ciò rende possibile osservare grandi valori di $x_t - x_{t=0}$. Successivamente **con l'aumentare di τ si osserva il recupero di un comportamento diffusivo** per il CLT, ovvero il sistema non sembra mostrare correlazioni su larga scala. In aggiunta il pannello (e) mostra come la convergenza a questo tipo di andamento venga rallentata nel momento in cui vengono presi in considerazione anche i dati relativi agli ultimi 150 anni.

Per quanto riguarda le **scale temporali più ampie, sui brevi lag si osserva un comportamento balistico** ($\propto t^\alpha, \alpha \simeq 2$): le anomalie tendono a ‘repellersi’: questo può accadere nel caso in cui la distribuzione mostra delle *heavy tails* oppure nel caso di correlazioni positive su brevi lag τ . Su lag maggiori invece il sistema transisce verso uno stato subdiffusivo. In ogni caso questa semplice analisi mostra che su queste scale il sistema mostra una certa *memoria*, anche sul lungo periodo, che sembrerebbe non permettere l'innestarsi del CLT e per la quale si tende ad ‘esplorare’ lo spazio delle x più di quanto accadrebbe nel caso di un normale moto browniano (*memoryless*) non mostrando comunque invarianza in scala temporale.

2. Uno studio sugli scarti temporali fa emergere un comportamento tipo *white noise* per le anomalie.

Nel verificare l'ipotesi che le anomalie di temperatura seguissero un andamento di tipo **random walk** sono stati analizzati gli scarti temporali relativi a diverse scale corrispondenti sia a valori locali che globali. Per scarto temporale si intende il valore dell'anomalia al netto della quota raggiunta al tempo precedente nella serie.

Se le variazioni di temperatura nella serie temporale di un elemento rispetto al precedente fossero distribuite in maniera casuale, come previsto in un random walk, presenterebbero un comportamento di tipo **white noise**: cioè una distribuzione simmetrica di oscillazioni casuali attorno alla media e varianza che tende ad un valore finito costante.

Per un Random Walk con drift ciò che mi aspetto di trovare è una distribuzione di questo tipo:

$$x(t) - x(t - dt) = \delta + wn$$

dove δ è il coefficiente di drift del Random Walk e wn sta per white noise. Avendo il white noise media nulla la media delle variazioni dovrebbe restituire il coefficiente di drift.

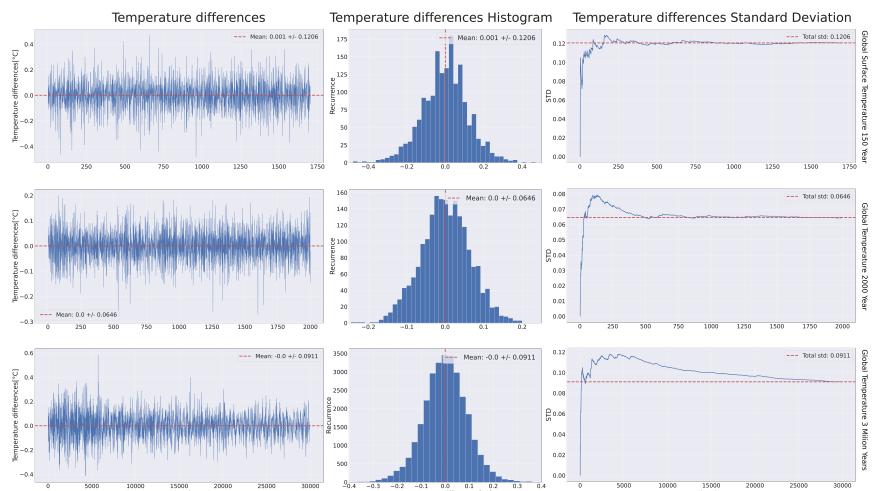


Figura 2: Scarti della temperatura su 3 diverse scale temporali: gli ultimi 150 anni, gli ultimi 2 mila anni e 3 milioni di anni fa

Dallo studio effettuato emerge una **distribuzione simmetrica negli scarti** per tutte le serie temporali analizzate, con **media nulla e deviazione standard che tende ad un valore finito** come visibile in figura 2.

Questo non comporta necessariamente che la variazione di temperatura degli ultimi anni sia assimilabile ad un Random Walk privo di drift ma potrebbe significare che questo comportamento non sia ancora evidentemente apprezzabile e potrebbe emergere in seguito.

3. A diverse scale temporali la CO_2 risulta essere il gas serra maggiormente correlato con la temperatura.

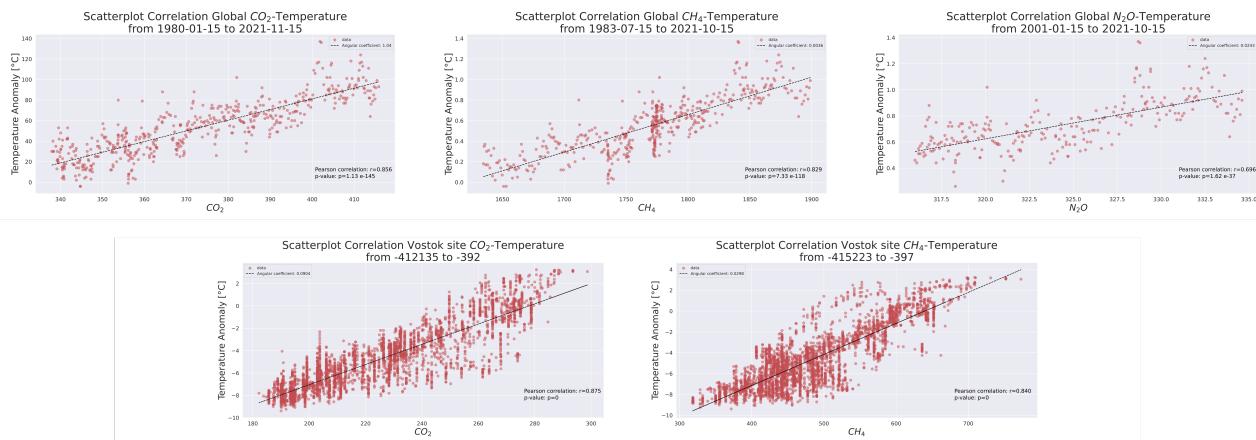


Figura 3: Scatterplot Gas Serra-Temperatura, in alto sono presentati i dataset globali, in basso quelli locali relativi al sito Vostok in Antartide

Nell'analizzare le concentrazioni di gas serra ho utilizzato serie temporali relative agli ultimi 40 anni per concentrazioni globali di CO_2 e CH_4 , una serie temporale relativa agli ultimi 20 anni di concentrazione di N_2O e serie lunghe 40 mila anni di concentrazioni di CO_2 e CH_4 , ottenute attraverso carotaggi effettuati in Antartide nella località chiamata Vostok.

Partendo dall' ipotesi che la correlazione tra la temperatura e i gas serra fosse in prima approssimazione lineare ho calcolato la correlazione di Pearson per tutti i dataset ottenendo i seguenti valori:

$$\begin{aligned} CO_2: & r=0.856 \text{ anni [1980, 2021] D.C.} \\ & r=0.875 \text{ anni [-41221335, -392] A.C.} \\ CH_4: & r=0.829 \text{ anni [1983, 2021] D.C} \\ & r=0.840 \text{ anni [-415223, -397] A.C} \\ N_2O: & r= 0.696 \text{ anni [2001, 2021] D.C} \end{aligned}$$

Dall' analisi emerge che a tutte le scale temporali **il gas che correla meglio con la temperatura risulta essere la CO_2** .

Per quanto riguarda la concentrazione di Metano N_2O non è stato possibile ottenere dati nel Dataset relativo a 40 mila anni, mentre sulla concentrazione negli ultimi anni era disponibile una serie temporale dal 2001 al 2021 D.C. .

Tutti i valori di correlazione calcolati possiedono un corrispondente p-value inferiore a 0.05, questo permette di affermare che i valori ottenuti non sono riconducibili a fluttuazioni statistiche dovute agli errori intrinseci al processo di misura.

Notes

¹I datasets relativi a temperatura e gas serra negli ultimi 150 sono reperibili sul sito Our World in Data e fanno riferimento a misure dirette disponibili sul sito della nasa e del noaa.

I dati relativi a 1 e 3 milioni sono reperibili nella sezione paleoclimatology del sito del noaa cercando rispettivamente "Global 1Ma Temperature, Sea Level, and Ice Volume Reconstructions" e "Global 3Ma Temperature, Sea Level, and Ice Volume Reconstructions". Lo stesso vale per i dati relativi alla località Vostok di cui sono disponibili vari files ognuno relativo ad una osservabile specifica. I dati relativi ai duemila anni sono infine erogati dal consorzio Page2sk

²In particolare, qui è stato computato il l'MSD ad un dato lag mediando su tutte le coppie di dati che presentano questo dato lag. In questo modo a lag più grandi l'MSD risulta avere una statica inferiore e quindi essere meno robusto.

³Gordon, A. H. "Global warming as a manifestation of a random walk." Journal of Climate 4.6 (1991): 589-597.

⁴Elnar, Allan RB, et al. "Great Barrier Reef degradation, sea surface temperatures, and atmospheric CO_2 levels collectively exhibit a stochastic process with memory." Climate Dynamics 57.9 (2021): 2701-2711.