

Introduzione e motivazione del lavoro

Nel medio termine la previsione di eventi intensi, e potenzialmente pericolosi, può essere mancata dalla corsa deterministica dei modelli meteorologici numerici. Per ovviare a questo inconveniente è ormai consolidato l'utilizzo di previsioni di ensemble che, oltre a fornire un ventaglio di possibili scenari diversi, provvedono anche a dare una misura della probabilità di accadimento degli eventi meteorologici. Le probabilità sono solitamente calcolate seguendo una metodica di tipo combinatorio, attraverso cioè il conteggio degli elementi dell'ensemble considerati tutti come equiprobabili. Nella pratica si adottano anche altre metodologie più complesse, come l'analisi a cluster, che presuppongono la possibilità di distribuzioni multimodali delle previsioni, ossia della tendenza degli elementi dell'ensemble a concentrarsi attorno a un numero ridotto di scenari alternativi. Il procedimento di inferenza statistica, con il quale è possibile dedurre da un campione le caratteristiche dell'intera popolazione, si basa sull'utilizzo di un corretto metodo di campionamento, mediante il quale si cerca di ridurre al minimo la possibilità di errore nella selezione degli elementi. Per le previsioni di ensemble ciò fa emergere almeno due tipi di questioni metodologiche:

- gli elementi di un ensemble rappresentano un campione non distorto di tutti i possibili scenari in grado di evolvere da una determinata situazione meteorologica?
- gli elementi di un ensemble sono statisticamente indipendenti tra loro?

Per costruzione, le previsioni di ensemble Ecmwf danno una risposta positiva alle due questioni poste ma gli stimatori della probabilità, basate sulle distribuzioni campionarie di un numero finito di elementi, possono però essere statisticamente distorti, soprattutto quando si guarda agli eventi marginali della distribuzione. In questo lavoro si è cercato, mediante un metodo non parametrico di geometria stocastica, di verificare se le singole distribuzioni empiriche degli elementi dell'ensemble possano invece essere statisticamente coerenti a una distribuzione limite, da cui poi calcolare le probabilità di accadimento degli eventi meteorologici.

I dati

Sono state prese in esame le previsioni di ensemble dell'Ecmwf dell'altezza di geopotenziale a 500 hPa, definite sull'area europea e a +120 ore di previsione.

Analisi delle componenti principali

Questa analisi statistica (Preisendorfer, 1988) è stata applicata alle previsioni di ensemble ECMWF relative ai 90 giorni della stagione invernale 2016-17. È stato calcolato che per questi 90 giorni la varianza spiegata dalle prime due componenti principali variava dal 46% al 69% della varianza totale.

I point-process

Gli elementi dell'ensemble sono stati trasformati in point processes riducendo i campi dell'altezza di geopotenziale a punti nel piano aventi per coordinate le prime due componenti principali standardizzate (**ensemble point processes**). I 90 ensemble point processes sono stati quindi confrontati con 99 realizzazioni casuali di tre modelli di point processes ottenute mediante un metodo Montecarlo: gaussiano, poissoniano e poissoniano a cluster. Per questo ultimo modello il numero dei clusters veniva fissato prima della generazione dei punti e fatto variare casualmente nelle diverse realizzazioni da un minimo di due a un massimo di sei, in analogia all'analisi a cluster compiuta sugli elementi dell'ensemble Ecmwf.

Il confronto tra i diversi point processes è avvenuto comparando le stime corrispondenti della nearest-neighbour distribution function D(r) e della reduced second order momentum K(r). La D(r) descrive la probabilità che la distanza di un punto del piano scelto a caso dal punto più vicino sia minore o uguale a un certo valore. La K(r) function stima il numero dei punti contenuti in circonferenze di raggio crescente e usa, perciò, informazioni di più larga scala rispetto a quelle considerate nella nearest-neighbour distribution function. Se i punti sono distribuiti in modo omogeneo, la derivata della K function è approssimativamente lineare al variare del raggio della circonferenza: $L(r) = [K(r)/\pi]^{1/2}$. Una sua deviazione positiva dalla linearità indica la tendenza dei punti a una sorta di clusterizzazione nel piano, mentre una tendenza negativa è rappresentativa di una distribuzione più regolare (Cressie N., 1991; Stoyan et al., 1987).

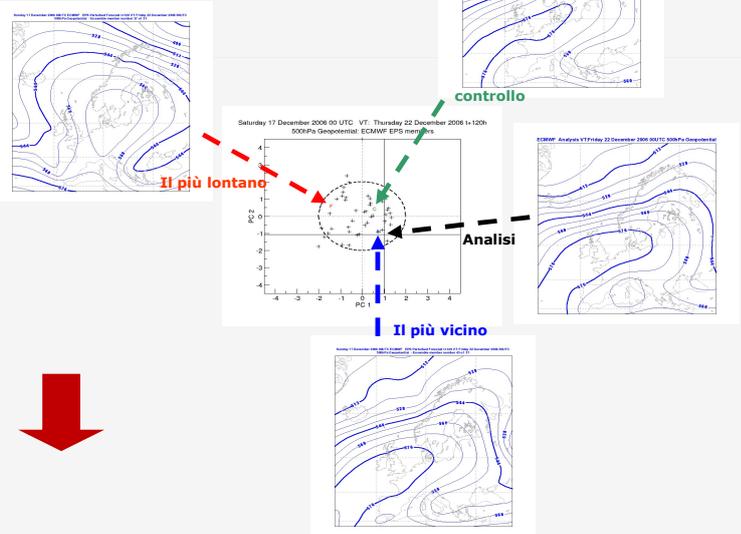
Risultati e Conclusioni

Le stime delle due funzioni riguardanti i 90 ensemble point processes sono statisticamente coerenti con i valori delle funzioni corrispondenti ai 99 point processes gaussiani; ciò comporta che le singole distribuzioni dei 90 ensemble point processes tendono in media a una distribuzione limite normale standardizzata e che le deviazioni da questa distribuzione limite sono contenute nell'intervallo di confidenza dell'errore statistico di campionamento (vedi figura, quadro a sinistra). Allo stesso tempo le stime non confermano, invece, una tendenza degli ensemble point processes né verso una distribuzione a cluster né verso una distribuzione omogenea (vedi figura, quadro a destra). Se i risultati di questo lavoro fossero successivamente confermati anche da altri test statistici, si potrebbe allora calcolare la probabilità di accadimento di un evento proporzionalmente al volume sotteso dalla funzione di distribuzione normale sulla superficie areale occupata dalla rappresentazione di questo evento nel piano. Sebbene le rappresentazioni delle prime due componenti principali producano ricostruzioni soltanto approssimate dei campi meteorologici di partenza, un possibile utilizzo di questo algoritmo di calcolo sarebbe quello di tracciare sul piano altri punti, eventualmente corrispondenti a scenari di eventi intensi non previsti dall'ensemble, e stimarne la loro probabilità di accadimento.

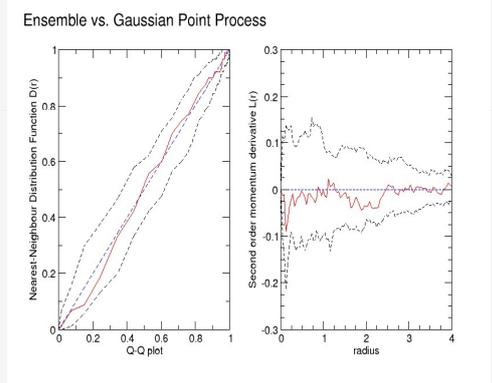
Analisi delle componenti principali



Riduzione degli ECMWF EPS a Point Processes



Confronto tra Ensemble e Gaussian Point Processes



Riferimenti bibliografici

Cressie N., 1991. Statistics for spatial data. Eds. J. Wiley & Sons.

Marsigli C., Montani A., Nerozzi F., Paccagnella T., Tibaldi S., Molteni F., Buizza R., 2001. A strategy for high-resolution ensemble prediction. Part II: Limited-area experiments in four Alpine flood events. Q. J. R. Meteorol. Soc., 127, 2069-2094.

Molteni F., Buizza R., Palmer T.N., Petroliagis T., 1996. The ECMWF Ensemble Prediction System: Methodology and validation. Q.J.R. Meteorol. Soc. 122, 73-119.

Molteni F., Buizza R., 1999. Validation of ECMWF Ensemble Prediction System using empirical orthogonal function. Mon. Wea. Rev., 127, 2346-2358.

Molteni F., Buizza R., Marsigli C., Montani A., Nerozzi F., Paccagnella T., 2001. A strategy for high-resolution ensemble prediction. Part I: Definition of Representative Members and Global Model Experiments. Q. J. R. Meteorol. Soc., 127, 2069-2094.

Preisendorfer R.W., 1988. Principal component analysis in meteorology and oceanography. Eds. Curtis D. Mobley.

Stoyan D., Kendall W.S., Mecke J., 1987. Stochastic geometry and its applications. Eds. J. Wiley & Sons.