



Downscaling Models for Early Warning Forecast of Extreme Hydrological Events

Francesco Cioffi, Tommaso Lapini and
Federico Rosario Conticello

EasyChair preprints are intended for rapid dissemination of research results and are integrated with the rest of EasyChair.

December 9, 2020

MODELLI DI DOWNSCALING PER LA PREVISIONE A MEDIO TERMINE DI EVENTI CLIMATICI ESTREMI LEGATI ALLE INONDAZIONI

Francesco Cioffi^{1*}, *Tommaso Lapini*¹ & *Federico Rosario Conticello*

(1) Dipartimento di Ingegneria Civile, Edile e Ambientale, Università “La Sapienza” di Roma

*email: francesco.cioffi@uniroma1.it

ASPETTI CHIAVE

- Previsioni a medio termine (10 giorni) dell'occorrenza di eventi alluvionali estremi.
- NHMM ed Event Synchronization; modelli stocastici condizionati dalle circolazioni atmosferiche.
- Condivisione di parametri tra i due modelli per migliorare l'efficacia della previsione.
- Applicazione ad un caso di studio: isola caraibica di Saint Lucia.

1 PREMESSA

L'indagine, nonché lo sviluppo di metodi, in grado di formalizzare la relazione di causa effetto fra le dinamiche atmosferiche alla grande scala e le piogge intense a scala di bacino o regionale è fondamentale nella previsione a breve e medio termine di eventi alluvionali, così come nelle proiezioni sul lungo periodo per gli ipotizzati scenari di riscaldamento globale.

Diversi approcci sono stati sviluppati in passato per formalizzare tale relazione. Fra questi i metodi di downscaling statistico, quali il Non-homogeneous Hidden Markov Model (NHMM). Numerosi autori, fra i quali (Kirshner, 2005; Khalil et al., 2010; Robertson et al., 2004; Cioffi & Conticello, 2016,2020) hanno dimostrato l'efficacia del NHMM nel quantificare le principali caratteristiche delle precipitazioni a scala locale in funzione di opportuni predittori atmosferici, quali campi di geopotenziale, temperatura, venti meridionali e zonal. Cioffi et al. (2017) hanno comunque evidenziato le difficoltà del metodo nel quantificare con sufficiente accuratezza le piogge estreme, quelle cioè potenzialmente in grado di generare alluvioni. Per superare tale insufficienza Conticello et al. (2018) hanno proposto un nuovo metodo, che combina, nel quadro di un approccio di tipo Complex network, Self Organized Map (SOM) ed Event Synchronization. Sinteticamente, tale metodo consente l'individuazione di curve di regressioni fra la probabilità di superamento di una prefissata soglia di altezza di pioggia di data durata, in un determinato punto di un bacino, ed il valore locale della variabile atmosferica “Integrated Vapor Transport” (IVT), essendo tale probabilità condizionata al verificarsi delle configurazioni atmosferiche classificate mediante le SOM. Tali configurazioni fanno riferimento a mappe di geopotenziale ad 850 hPa (GPH850). L'applicazione di tale approccio al caso delle precipitazioni intense nella regione Lazio (Conticello et al., 2018) e delle piene in Germania (Cioffi et al., 2020) ha dimostrato la robustezza e l'efficacia del metodo nell'individuare relazioni statistiche significative fra estremi idrologici e le variabili atmosferiche IVT e GPH850.

Tali risultati suggeriscono di verificare se tale approccio metodologico possa essere utilizzato in un contesto di Early Warning per prevedere sul breve e medio periodo (da 1 a 10 giorni) la probabilità del verificarsi di piogge intense in grado di produrre eventi alluvionali e se fosse possibile attraverso l'impiego di un diverso sistema di classificazione, alternativo alle SOM, migliorarne l'affidabilità della previsione.

In questa nota, con riferimento al caso studio relativo all'isola di St. Lucia nei Caraibi, si tenta di dare risposta alle due questioni suddette, combinando gli approcci NHMM ed Event Synchronization, utilizzando il primo per classificare gli stati del tempo (in funzione dei due predittori IVT e GPH850), il secondo per individuare le curve di regressione fra probabilità di superamento di una soglia di altezza di pioggia giornaliera ed il valore dell'IVT locale. Prove di validazione del sistema di modelli suddetto, effettuate nell'ambito del caso di studio, sembrano confermare l'efficacia e l'affidabilità del metodo proposto.

2 MATERIALI E METODI

2.1 Dati

Le serie temporali dei campi delle variabili atmosferiche, con una risoluzione spaziale orizzontale di $0.25^\circ \times 0.25^\circ$ e compresi fra la latitudine di 45°N e 10°S e fra la longitudine di 110°W e 10°W , sono state estratte dall'ECMWF Reanalysis Dataset "ERA5". In particolare, le variabili atmosferiche selezionate sono: l'altezza del Geopotenziale a 850hPa, il vento meridionale, il vento zonale e l'umidità specifica da 1000hPa a 300hPa (passo 100hPa). Dalle ultime tre è possibile calcolare l'integrale del vapore trasportato (IVT) nella colonna di atmosfera da 1000hPa a 300hPa. Per quanto riguarda gli output sono stati reperiti 40 anni di registrazioni (1979-2018), effettuate da 18 misuratori pluviometrici giornalieri, dall'Agenzia per la gestione delle risorse idriche di St. Lucia.

2.2 Non-homogeneous Hidden Markov Model

Le linee principali del modello non omogeneo di Hidden Markov (NHMM) sono riassunte qui di seguito, per completezza e maggiori dettagli si rimanda il lettore a questi articoli (*Kirshner, 2005; Khalil et al., 2010; Robertson et al., 2004; Cioffi et al., 2017*). L'NHMM è un modello doppiamente stocastico in cui sono generate serie temporali multivariate, utilizzando una certa distribuzione, condizionatamente ad uno stato nascosto individuato; inoltre gli stati nascosti subiscono transizioni markoviane. In questa applicazione idroclimatica, uno stato nascosto è pensato come una particolare configurazione meteorologica che influenza l'accadimento e la quantità della precipitazione in un certo numero di località contemporaneamente. Per determinare il miglior set di variabili in input al modello nonché i parametri dello stesso (es. numero di stati nascosti) sono stati addestrati molti modelli differenti. Sulla base del confronto fra i valori delle Log-Likelihood per i diversi modelli si sono selezionati come ottimali 4 modelli distinti associati uno ad ogni stagione, i quali identificano ognuno 6 stati nascosti ed utilizzano in input: 12 componenti principali di GPH-850hPa ed 8 componenti principali di IVT-1000/300hPa. Tutti i dettagli sulla procedura utilizzata per la stima dei parametri si trovano in (*Kirshner et al., 2004*) e (*Robertson et al., 2004*). Infine, l'algoritmo di Viterbi è utilizzato per identificare la sequenza più probabile di stati nascosti associati alla sequenza delle osservazioni (*Viterbi, 1967*). Si sottolinea che a ciascun stato nascosto sono associate le probabilità di accadimento e funzioni di densità di probabilità di altezza giornaliera di pioggia ed i corrispondenti campi compositi delle variabili atmosferiche.

2.3 Modello di precipitazioni intense tramite Event Synchronization

Si definiscono "eventi di precipitazioni intense" (HPE) i giorni in cui le precipitazioni superano una soglia di allerta prefissata. Tale soglia per la nostra area di studio è stata identificata da simulazioni condotte con un modello idraulico bidimensionale ed è risultata essere pari a 60mm giornalieri. Imponendo la condizione di superamento di questa soglia alle serie storiche di ogni pluviometro si ricava una serie temporale binaria dove 1 identifica il verificarsi di HPE e 0 altrimenti (dal 1979 al 2018). Successivamente, allo scopo di ridurre le configurazioni di Geopotenziale ad un set di campi rappresentativi di tutte le condizioni verificatesi negli ultimi 40 anni, sono stati calcolati i campi medi di GPH associati ad ogni stato nascosto identificato dai modelli NHMM stagionali. In questo modo abbiamo ottenuto un set di 24 mappe di Geopotenziale (GSs) rappresentative della variabilità del clima nell'area di studio. Applicando quindi il k-nearest neighbor, (*Cover & Hart, 1967*), la sequenza temporale dei campi di GPH viene trasformata in una matrice binaria assegnando, ad ogni giorno, il GS più simile all'effettiva configurazione. Viene quindi calcolata la sincronizzazione tra le serie binarie ottenute (superamento della soglia da parte di una certa stazione pluviometrica ed il verificarsi di una determinata mappa di GS) secondo la procedura descritta in *Conticello et al. (2018)*. Gli stati di geopotenziale caratteristici (GSs) più sincronizzati sono quelli che hanno maggiore probabilità di produrre piogge intense e di conseguenza eventi alluvionali. Allo scopo di quantificare numericamente la probabilità di insorgenza di HPE associati ad un determinato GS ed in funzione dell'IVT previsto, utilizziamo la regressione locale logistica basata sul modello lineare generalizzato con risposta di Bernoulli. Per gli approfondimenti sullo sviluppo di questo metodo si rimanda a *Conticello & Cioffi, 2018* e *Loader, 1999*.

2.4 Struttura del sistema di previsione

La combinazione dei metodi sopra descritti costituisce la struttura del sistema di previsione a medio termine. I dati di input sono le previsioni fino a 10 giorni dei campi di GPH e IVT forniti dall'ECMWF. Gli output generati dai modelli in relazione ad ogni giorno di previsione sono i seguenti: l'NHMM identifica lo stato nascosto più probabile associato a quel giorno e il valore delle precipitazioni giornaliere previste per ogni stazione pluviometrica, l'Event Synchronization identifica il GS più simile al GPH previsto per quel giorno e

fornisce la relativa curva di regressione che descrive la probabilità di superare la soglia critica di precipitazioni in funzione dell'IVT previsto sull'isola.

3 RISULTATI

L'analisi dei risultati ottenuti mostra che il modello NHMM identifica per ogni stagione lo stato nascosto associato alle precipitazioni più intense, altri stati intermedi a media intensità di pioggia ed almeno uno stato secco. Ciascuna configurazione di geopotenziale associata agli stati nascosti di tutti e quattro i modelli NHMM stagionali viene utilizzata per calcolare la sincronizzazione con gli HPE. Nella figura 1a è mostrata la mappa di geopotenziale associata alla configurazione GS-19, risultata essere la più sincronizzata (figura 1c). Tale configurazione tipica dei mesi di Settembre, Ottobre e Novembre (SON), presenta un centro di bassa pressione localizzato sull'arcipelago delle piccole Antille di cui fa parte l'isola di Saint Lucia; al contrario è presente un'estesa zona di alta pressione sul nord Atlantico che comporta il trasporto di umidità verso sud alimentando proprio il vortice di bassa pressione con grandi quantitativi di vapore d'acqua. La figura 1b mostra l'IVT associato alla stessa configurazione. Dalla figura è evidente come la direzione principale dei venti orizzontali formi un vero e proprio ciclone quasi perfettamente centrato sull'isola di Saint Lucia (cerchio nero). Quando si verifica questa configurazione la probabilità di superare la soglia di pioggia critica (60 mm) incrementa significativamente all'aumentare dell'IVT, cosicché per un valore non troppo raro di IVT superiore ai 1200 ($\text{kg m}^{-1}\text{s}^{-1}$) la probabilità sarà pari al 60% (figura 1d).

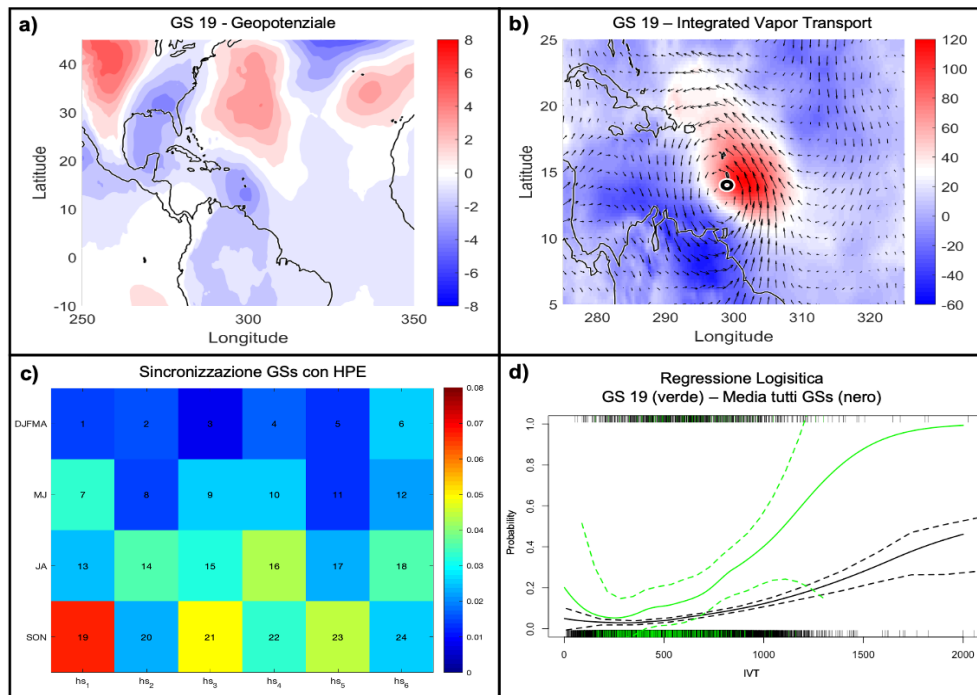


Figura 1 Nel pannello a) è riportata la differenza tra il campo di geopotenziale della configurazione GS-19 e quello medio stagionale. Il pannello b) illustra analogamente la differenza del campo di IVT della configurazione GS-19 rispetto alla media stagionale. Il pannello c) mostra la sincronizzazione delle 24 GSs con gli eventi estremi. Il pannello d) infine mostra la curva di regressione logistica associata alla configurazione GS-19 (in verde) rispetto a quella media relativa a tutte le configurazioni (in nero).

La validazione dei modelli, effettuata su un quinquennio di dati indipendenti, ha dimostrato un'ottima capacità dei modelli di identificare l'accadimento degli eventi estremi. In particolare, l'NHMM relativo alla stagione Settembre-Ottobre-Novembre è stato in grado di prevedere l'occorrenza dello stato nascosto critico l'86% delle volte che questo si era verificato, con picchi del 97% se guardiamo i soli giorni in cui si sono verificati passaggi di Ciclone Tropicali. Per quanto riguarda l'Event Synchronization abbiamo registrato probabilità di superare la soglia critica predette sempre superiori all'80% nei giorni in cui si sono verificati eventi alluvionali.

4 CONCLUSIONI

L'utilizzo combinato dell'approccio NHMM, e dell'event synchronization si è dimostrato un metodo in grado di fornire con accettabile accuratezza la probabilità di accadimento di eventi estremi a partire dalle previsioni a medio termine (10 giorni) fornite dall'ECMWF. La regressione logistica associata agli specifici GS permette di valutare accuratamente la probabilità di superare una soglia identificata come critica in relazione ad un potenziale pericolo alluvionale. Affinamenti della metodologia e la verifica della applicabilità di questa in altre area del pianeta sono in corso di sviluppo, così come la verifica della possibilità della stessa di essere utilizzata nel contesto delle problematiche associate ai cambiamenti climatici.

RIFERIMENTI BIBLIOGRAFICI

- Cioffi, F., Conticello, F., & Lall, U. (2016). Projecting changes in Tanzania rainfall for the 21st century. *International Journal of Climatology*, 36(13), 4297-4314
- Cioffi, F., Conticello, F., Lall, U., Marotta, L., & Telesca, V. (2017). Large scale climate and rainfall seasonality in a Mediterranean Area: Insights from a non-homogeneous Markov model applied to the Agro-Pontino plain. *Hydrological Processes*, 31(3), 668-686.
- Cioffi, F., Conticello, F. R., & Lall, U. (2020). Stochastic Scenarios for 21st Century Rainfall Seasonality, Daily Frequency, and Intensity in South Florida. *Journal of Water Resources Planning and Management*, 146(8), 04020058
- Conticello, F., Cioffi, F., Merz, B., & Lall, U. (2018). An event synchronization method to link heavy rainfall events and large-scale atmospheric circulation features. *International Journal of Climatology*, 38(3), 1421-1437.
- Conticello, F. R., Cioffi, F., Lall, U., & Merz, B. (2020). Synchronization and Delay Between Circulation Patterns and High Streamflow Events in Germany. *Water Resources Research*, 56(4), e2019WR025598.
- Cover T, Hart P. 1967. Nearest neighbor pattern classification. *IEEE Transactions on Information Theory* 13(1): 21–27.
- Khalil AF, Kwon H-H, Lall U, Kaheil YH. 2010. Predictive downscaling based on non-homogeneous hidden Markov models. *Hydrological Sciences Journal* 55(3): 333–350.
- Kirshner S. 2005. Modeling of Multivariate Time Series Using Hidden Markov Models.
- Robertson A, Kirshner S, Smyth P. 2004. Downscaling of daily rainfall occurrence over northeast Brazil using a hidden Markov model. *Journal of climate* 4407–4424.
- Viterbi A. 1967. Error bounds for convolutional codes and an asymptotically optimum decoding algorithm. *IEEE Transactions on Information Theory* 13(2): 260–269.