

REGIONE BASILICATA

DOWNSCALING DELLE PREVISIONI METEOROLOGICHE PER APPLICAZIONI IN AGRICOLTURA DI PRECISIONE

0.9

0.8

و بر 2007

BSN 0.6

1.8

1.6 1.4

°¥ 1.2

BSW 1.0

0.8

0.6

0.4

15

업

Nodes

•

Final configuration nodes: [6 - 9]

MSE: 0.52 K²

Nell'ambito del progetto *On Demand Services For Smart Agriculture* (OD4SA), finanziato della Regione Basilicata con il programma PO FESR 2014-2020, è stato sviluppato un servizio di previsioni meteorologiche ad alta risoluzione spaziale, per applicazioni in agricoltura di precisione, basato sul modello numerico di previsione *Weather Research and Forecasting* (WRF) Advanced Research core (ARW). Nonostante negli ultimi anni siano stati raggiunti buoni progressi nella modellazione a microscala, gli elevati costi computazionali impediscono previsioni operative al di sotto della scala chilometrica, che risulta un po' troppo grossolana per supportare i processi decisionali per l'agricoltura di precisione. Questo limite risulta particolarmente marcato nel sud Italia, dove l'orografia complessa rende le principali grandezze meteorologiche altamente variabili con lo spazio. Per superare questo limite è stato sviluppato un algoritmo a Reti Neurali Artificiali (ANN) per l'aumento di risoluzione spaziale delle previsioni WRF, portandole dalla risoluzione nativa di 1,2 km alla più alta risoluzione di 240 m.

Unione Europea

ndo Europeo di Sviluppo Regiona

Database di addestramento

OD4SA

Il database di addestramento delle ANN è stato costruito utilizzando il modello WRF in modalità *Large Eddy Simulation* (WRF-LES), che consente di raggiungere la scala subchilometrica risolvendo esplicitamente i processi turbolenti all'interno dello strato limite planetario. Sono state effettuate 12 simulazioni eseguite in modalità rianalisi per 12 gioratte nel 2017, con 2 domini annidati a 1,2 km e 240 m. Iconfronto delle simulazioni nei due domini mostra valori leggermente diversi nell'area di sovrapposizione, impendendo quindi l'uso dei campi a 1,2 km e a 240 m come coppie di input-output per la fase di addestramento delle ANN. Per superare questa limite sono stati creati campi meteorologici sintetici a 1,2 km u tilizzando quelli a 240 m degradandone la risoluzione spaziale attraverso un'operazione di convoluzione con un kernel gaussiano bidimensionale, e successivamente riducendone il campionamento spaziale attraverso un processo di decimazione. L'intero processo è stato guidato dall'analisi dei campi meteorologici nel dominio delle frequenze spaziali, per dimensionare opportunamente la dimensione del kernel di *smoothing*. A valle di quest'operazione sono stati sviluppati due database per ogni ANN, il primo composto da 10⁶ pattern di addestramento per l'analisi degli input e la definizione dell'architettura, il secondo con 10⁶



Analisi degli input per T2 – ANN con 6 e 9 nodi sul primo

imo... Jella tempera... ` •viperflui e qui `•to cr

80 77 181 18 removed input [#]

0.05

-0.05

-0.1

-0.15

-0.2

0.8 2

MSE

ra iterative per la configurazione dei nodi per T2

ing criteria: positive gradient for 10 til

training set validation se

moving average moving average

Nodes [#]

8 1 Nodes in H1

XXV Convegno Nazionale di Agrometeorologia

Matera 14 - 16 Giugno 2023

Definizione delle ANN

Sono state sviluppate 7 diverse ANN, una per ciascuna delle seguenti variabili meteorologiche: (*i*) temperatura a 2m (T2), (*ii*) rapporto di mescolanza del vapore acqueo a 2 m (C2), (*iii*) pressione superficiale (PSFC), (*iv*) irradianza solare globale (SWD), (*v*) precipitazione cumulata oraria (RAIN), (*vi*, *vii*) componenti orizzontali del vento a 10 m (U10 e V10). Per ogni ANN l'output è definito a valle dell'Analisi delle Componenti Principali (PCA) effettuata ordinando da sinistra a destra e dal basso verso l'alto i punti della griglia (7x7) a 240 m della variabile meteorologica considerata (simboli circolari verdi in figura), selezionando infine le prime componenti che spiegano complessivamente il 99,9% della varianza del segnale.

Scelta ed analisi degli input

La scelta degli input delle ANN, è stata effettuata in due fasi. Nella prima fase sono state prese in considerazione alcune variabili meteorologiche potenzialmente utili all'aumento di risoluzione e quindi utili all'addestramento delle ANN, o perché direttamente correlate con l'output o perché in grado di caratterizzare le condizioni atmosferiche e le caratteristiche della superficie. L'input preliminare considerato per ciascuna ANN è costituito da: (i) le PC della stessa variabile di cui si vuole aumentare la risoluzione spaziale calcolate sul box (5x5) a 1,2 km (simboli a croce gialla e rossa nella figura in alto a destra); (ii) le PC degli altri campi meteorologici sopra riportati calcolate sul box (3x3) a 1,2 km, utili per caratterizzare l'atmosfera (simbolo a croce rossa); (iii) le PC dell'altitudine della superficie (HGT), della rugosità superficiale (SRL) e dalle tre componenti del versore normale alla superficie terrestre (VN e VE paralleli alla griglia WRF, rispettivamente diretti verso nord e verso est, VZ ortogonale sia a VN che a VE e diretto verso quote crescenti), calcolate nello stesso box (7x7) a 240 m considerato per l'output. Mediante quest'approccio tuttavia sono stati selezionati degli input alle ANN il cui contributo può essere trascurabile o addirittura controproducente perché aumentano il rischio di overfitting. Per questo motivo, nella seconda fase è stata adottata una procedura iterativa per la valutazione del contributo di ciascun input e la sua eventuale successiva rimozione.

Architettura delle ANN

Un'ANN con uno strato nascosto, con un numero sufficiente di nodi, con una funzione di attivazione per lo strato nascosto arbitraria, limitata e non costante, e con una funzione di attivazione lineare per lo strato di output, è un approssimatore universale di ogni funzione continua sotto assunzioni abbastanza generali (Funahashi 1989; Hornik et al. 1989). Tuttavia, è preferibile scegliere due strati nascosti perché forniscono generalmente alcuni vantaggi, come un'architettura con meno nodi totali, una maggiore efficienza degli algoritmi di addestramento e una migliore capacità di generalizzazione rispetto a un singolo strato nascosto (Thomas et al., 2017). Non vi è inoltre alcuna ragione teorica, per lo studio proposto, di ricorrere a tre o più strati nascosti (Sharma et al., 2013). Scegliendo quindi la tangente iperbolica come funzione di attivazione per gli strati nascosti, l'unico parametro che deve essere definito è il loro numero di nodi. A tal fine è stato utilizzato un approccio iterativo che, partendo da un solo nodo su entrambi i livelli nascosti, confronta per ogni iterazione due diverse configurazioni di nodi, ciascuna ottenuta aumentandone il numero di una unità separatamente sul primo e sul secondo strato nascosto, scegliendo per ogni iterazione la configurazione tra le due che produce il minor MSE minimo calcolato su un dataset di validazione. La procedura iterativa si è conclusa quando l'aggiunta progressiva di nodi sui due strati nascosti ha prodotto per 10 volte un gradiente positivo della media mobile a 3 punti del MSE, indice che un'ulteriore aumento dei nodi non comporta più un miglioramento significativo nelle prestazioni.



Risultati

Utilizzando un set di dati WRF indipendente, i risultati dell'aumento di risoluzione spaziale sono stati confrontati con gli output a 240 m in termini di RMSE, che risulta sempre inferiore alla deviazione standard del target, assunta come valore di riferimento della variabilità intrinseca del campo.

