

5. MODELLI PER LA PREVISIONE DELLA QUALITÀ DELL'ARIA DELLA PROVINCIA DI VICENZA ELABORATI CON RETI NEURONALI

Le reti neurali permettono di utilizzare la multidimensionalità dell'informazione (qualità dell'aria e meteorologia) senza richiedere modelli che interpretino preventivamente le relazioni di causa-effetto tra le variabili e, tramite l'analisi di sensitività delle prestazioni, diventano uno strumento per valutare l'importanza di ogni singola variabile sul risultato finale. Esse, inoltre, permettono di adoperare contemporaneamente l'informazione puntuale relativa a diversi siti.

Per questi motivi e per la complessità del fenomeno in esame, ancora poco conosciuto, le reti neurali sono particolarmente adatte alla costruzione di modelli per lo smog fotochimico, a scala di bacino.

Su questi presupposti sono stati sviluppati dei modelli per la simulazione delle concentrazioni di ozono, misurate nei cinque siti della provincia di Vicenza nei quali si dispongono dati: Bassano del Grappa, Montecchio Maggiore, Schio, Vicenza Ovest e Valdagno.

Le analisi statistiche, effettuate sui dati chimici e meteorologici, hanno messo in evidenza una elevata correlazione tra gli andamenti dell'O₃ dei diversi siti di monitoraggio. Tenendo presenti tali risultati si è studiato un "dispositivo" in grado di segnalare eventi acuti d'inquinamento da ozono, in punti del territorio privi di misuratori di tale inquinante, utilizzando le informazioni raccolte altrove nello stesso territorio. Questo obiettivo è stato raggiunto operando sulle cinque stazioni fisse a disposizione come se ognuna di esse fosse, di volta in volta, un mezzo mobile supportato dalle altre quattro.

Si è costruito, quindi, un congegno in grado di fornire previsioni in tempo reale, utili soprattutto per interpretare l'inquinamento diffuso e per stabilire la struttura a sottobacini dell'area padana, dato che sembra poco ragionevole ritenere che l'intero bacino subisca vicende comuni. Infatti, mediante analisi di sensitività dei risultati, è possibile individuare i

punti del territorio più esposti e quelli che maggiormente determinano l'inquinamento degli altri.

Per quanto riguarda gli eventi acuti è inoltre importante adottare dispositivi in grado di fornire previsioni a scala di bacino con anticipo. L'anticipo per le previsioni è a due scale temporali: a breve termine (*nowcasting*) e a lungo termine (*forecasting*). Il primo è utile soprattutto per intervenire, ad esempio, nel traffico automobilistico indirizzandolo tramite azioni su semafori segnaletici a messaggio variabile; il secondo, la cui scala temporale è dell'ordine dei giorni, potrebbe permettere di intervenire, impartendo le disposizioni previste dalla legge (limitazione del traffico, ecc.), preventivamente, in tempo utile perché non si raggiungano livelli d'inquinamento pericolosi per la popolazione.

Nel breve termine le previsioni a scala di bacino hanno fornito risultati soddisfacenti, mentre, nel lungo termine, lasciano ancora notevoli spazi d'incertezza.

Nel seguito saranno descritti i modelli neurali addestrati per la previsione di O_3 in ciascun sito di monitoraggio, tenendo conto che le concentrazioni di ozono sono un fenomeno tipicamente estivo. Per questo motivo i dati utilizzati riguardano il periodo 1 giugno-30 settembre 1996.

Una notevole quantità di lavoro è stato dedicato alla scelta dell'architettura neuronale più adatta, che è risultata essere costituita da un solo *vettore di input*, da un unico *strato nascosto* ed un solo neurone di output. In *input*, sono stati provati vettori di dimensioni diverse, mentre in *output* si è sempre considerata la concentrazione dell'ozono da prevedere.

Prima di tutto, si è costruito un unico modello neuronale per ciascun sito di monitoraggio (Bassano del G., Montecchio M., Schio, Valdagno e Vicenza Ovest), addestrandolo a simulare l'andamento dell'ozono con informazioni dello stesso sito (Liguori, 1996; Liguori et Al., 1997 a; 1997 b; Liguori et Al., 1998). Lo scopo di tali modelli era effettuare un *nowcasting* della qualità dell'aria. La struttura del modello è sintetizzabile nella seguente espressione:

$$O_3(t) = f_{[11,8,1]} \left\{ \begin{array}{l} NO_2(t-1), T(t-1), VV(t-1), DV(t-1), UR(t-1), RST(t-2), \\ O_3(t-4), O_3(t-3), O_3(t-2), O_3(t-1); W \end{array} \right\}$$

dove, nell'indice tra parentesi quadra è dato il numero di neuroni negli strati rispettivamente di input, nascosto e di output, f è la trasformazione effettuata dalla rete sulle variabili di input, indicate tra parentesi graffa, mediante i pesi delle connessioni W .

Successivamente, si è studiata la struttura spaziale dei dati analizzando l'O₃ rilevato in ciascuna stazione, mettendolo in relazione con i valori di O₃ rilevati nelle altre. L'aggiunta di informazioni di tipo "spaziale" ha generalmente migliorato le previsioni rispetto a quelle ottenute con sole informazioni locali.

Per quanto riguarda il numero di unità nascoste utilizzate, non è stato trovato finora alcun criterio che indichi il numero delle unità più conveniente, o il numero di strati su cui distribuirle, né a tali numeri sono connessi significati fisici particolari. In generale, una rete *feed-forward*, con un solo strato nascosto, è ritenuta capace di approssimazione universale, quindi il problema è limitato alla sola ricerca del più adeguato numero di unità nascoste (Hornik. K., 1991). Bisogna tener presente che, all'aumentare del numero delle connessioni (che dipende dal numero di unità nascoste), la rete, mentre acquista una maggiore capacità di riprodurre il *training set*, perde in flessibilità, cioè in capacità di generalizzazione. Per questo motivo, nelle reti utilizzate nel presente lavoro, il numero di unità nascoste è sempre inferiore al numero degli elementi di *input* ed in particolare è stato orientato da molteplici prove.

Prima di dare una rassegna dei modelli utilizzati, dei relativi risultati e dei metodi di valutazione della qualità di questi ultimi, si ritiene utile descrivere le modalità utilizzate per la costruzione dei modelli e per la preparazione dei dati.

Una volta definito il vettore di input si preparano due set di dati:

- il *training set*, da utilizzare per l'addestramento della rete neuronale;
- il *testing set*, da utilizzare per la validazione della rete neuronale addestrata.

Il *training set* ed il *testing set* si ottengono mediante estrazione casuale di *patterns* (vettori d'input), senza ripetizione, da un set originario costituito da una matrice le cui righe sono tutti i *patterns* disponibili, ordinati in sequenza temporale. L'estrazione casuale senza ripetizione si effettua sulle righe di tale matrice (e quindi sui *patterns*) ed ha lo scopo di creare *file* rappresentativi della stessa popolazione. Inoltre, l'estrazione viene applicata in modo che il *training set* ed il *testing set* siano tra loro complementari, condizione necessaria per un efficace addestramento e per una corretta validazione (Boznar M., 1993).

Generalmente, il *training set* presenta un numero di *patterns* maggiore rispetto a quello del *testing set*, in quanto l'addestramento richiede un notevole contingente di esempi, al fine del raggiungimento della configurazione dei pesi delle connessioni corrispondente all'apprendimento del *problema*.

I dati dei *file*, di addestramento e verifica, vengono normalizzati affinché risultino tra loro confrontabili ed esercitino la medesima influenza sulla rete neuronale.

La normalizzazione è stata effettuata secondo la formula:

$$\bar{x}_i = 0.2 + 0.6 \cdot \frac{(O_i - O_{\min})}{(O_{\max} - O_{\min})}$$

dove O_i è la i -esima osservazione della variabile O , O_{\min} è il valore minimo assunto dalla variabile O e O_{\max} è il valore massimo della stessa variabile.

La conversione ha consentito di ottenere valori compresi nell'intervallo 0.2-0.8, con l'intento di evitare le code della funzione di trasferimento *sigmoide* (Comrie A., 1997).

Una volta addestrata la rete si è provveduto alla conversione inversa dei valori di output ed alla loro analisi qualitativa.

I criteri di valutazione dei risultati sono molteplici. Una prima indicazione della bontà dei risultati è fornita dal valore del coefficiente di correlazione tra valori sperimentali e valori simulati. La valutazione delle capacità previsionali dei modelli, prevede anche dei test sull'abilità degli stessi a riprodurre gli eventi acuti d'inquinamento da ozono. Data l'assenza di valori sperimentali che eccedano i *livelli d'allarme*, il test è stato applicato solamente alla previsione di eventi di superamento delle soglie di concentrazione di O_3 , stabilite per il *livello d'attenzione*. Tale test, consiste in una serie di indici statistici calcolati sulla base di parametri ricavati da tabelle di contingenza¹.

5.1. Modelli per l'ozono a Bassano del Grappa

La prima rete neuronale per la previsione a scala di bacino, è stata addestrata sulla concentrazione attuale di ozono di Bassano del Grappa, intregando l'informazione in ingresso costituita dai dati rilevati a Montecchio Maggiore, Schio, Vicenza Ovest e

¹ Per ulteriori delucidazioni si rimanda al cap. 4, paragrafo 4.4.