

6.2 Modelli per il monossido di carbonio

La serie temporale scelta per l'addestramento è quella rilevata in via Piave poiché tra le stazioni idonee, per legge, al controllo di questo inquinante essa presenta un numero maggiore di eventi acuti.

Risultati piuttosto soddisfacenti sono stati ottenuti nella previsione a passo temporale di un'ora introducendo in input - oltre alla consueta componente autoregressiva dell'inquinante, alle variabili chimiche ad esso maggiormente correlate ed alle informazioni meteorologiche - il traffico registrato in Via Piave ed altre due variabili costruite per indicare il giorno della settimana e l'ora del giorno corrispondenti all'output. Tale modello (rete '1COt' di tab. 6.5) è dunque rappresentabile tramite l'espressione:

$$\text{CO}(t+1) = f_{[11,8,1]} \left(\text{CO}(t), \text{CO}(t-1), \text{CO}(t-2), \text{CO}(t-3), \text{NO}_x(t), \right. \\ \left. \text{T}(t), \text{VV}(t), \mathbf{g}(t), \text{Tr}(t), \text{d}(t+1), \text{h}(t+1); \mathbf{W} \right)$$

in cui Tr, d e h sono rispettivamente le variabili 'traffico', 'giorno settimanale' ed 'ora'.

Proprio l'introduzione di queste tre variabili ha permesso di migliorare la previsione già piuttosto buona della più semplice rete '1CO' descritta in tabella 6.5.

Il monossido di carbonio è difatti un inquinante primario la cui sorgente principale è costituita dal traffico veicolare. L'input raffigurante il numero di veicoli in transito apporta dunque al sistema un contenuto informativo importante, anche se non si tratta del numero assoluto di veicoli le cui emissioni raggiungono effettivamente gli analizzatori¹². Le variabili che indicano l'ora del giorno e il giorno della settimana sono state introdotte per pesare il potenziale accumulo dell'inquinante, la cui fenomenologia è, appunto, correlata all'ora del giorno (*cfr.* Cap. V). Trattandosi di funzioni perfettamente determinate il valore in input è quello contemporaneo al valore di output.

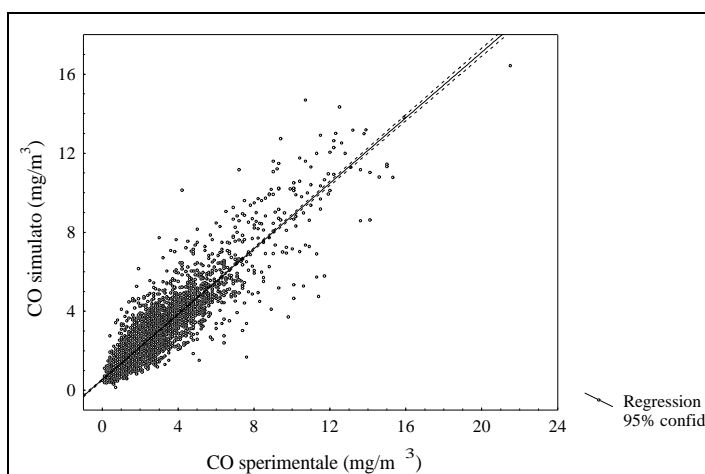
¹² Si ricorda che il conteggio degli autoveicoli è limitato ad un solo senso di marcia e che la sezione di misura è posta a un centinaio di metri dalla stazione.

Tab. 6.5. Reti neurali per la previsione del monossido di carbonio a passo temporale di un'ora.

Rete	Tipo	N°patterns	vettore di input	ME (mg/m ³)
1CO	f-8-6-1	4386	NO _x -1; T-1; VV-1; γ-1; CO-4 CO-3 CO-2 CO-1	± 0.99
1COt	f-11-8-1	4386	NO _x -1; T-1; VV-1; γ-1; Tr d h CO-4 CO-3 CO-2 CO-1	± 0.85

L'errore medio di previsione raggiunto dalla rete '1COt' (ME di tab.6.5) è di ± 0.85 mg/m³, ed è inferiore al 6% dei 15 mg/m³ a cui è fissato il *livello di attenzione*. La correlazione CO sperimentale - CO simulato è 0.91, a cui corrisponde una varianza spiegata dell'82% (fig. 6.12).

Fig. 6.12. Analisi dei risultati della rete '1COt': regressione CO sperimentale-CO simulato.



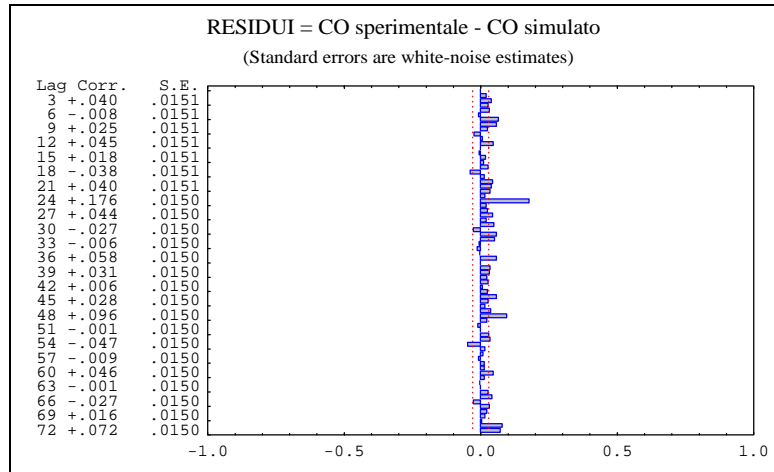
In tabella 6.6 sono presentate le statistiche descrittive della serie del CO sperimentale e quella del CO simulato dallo strumento neurale. Il confronto mette in mostra la buona prestazioni della rete '1COt'.

Tab. 6.6. Statistiche descrittive per il confronto tra CO sperimentale e CO simulato dalla rete '1COt'.

	Media	Int. Confid. 95%	Mediana	Min	Max	Dev. Std.	Err. Std
CO sperimentale	2.47	2.41 ; 2.53	1.90	0.00	21.50	2.03	.03
CO simulato	2.59	2.53 ; 2.64	2.10	.16	16.43	1.86	.03
RESIDUI	-.12	-.14 ; -.09	-.14	-5.94	6.65	.86	.01

Il correlogramma dei residui (fig. 6.13) conferma il giudizio positivo sulla previsione effettuata dalla rete '1Cot': solo in corrispondenza dell'intervallo di tempo di 24 ore si nota un valore di correlazione significativamente diverso da quello attribuibile ad un processo generatore di tipo *white noise*.

Fig. 6.13. Funzione di autocorrelazione dei residui del modello della rete '1Cot'



Per quanto riguarda la previsione dei picchi di concentrazione, dei 5 valori al di sopra del *livello di attenzione* (15 mg/m³) presenti nel *training set* la rete '1Cot' riesce a segnalarne uno solo; per i rimanenti, però, l'output previsto non si allontana troppo dal valore soglia. Estendendo l'analisi a tutti i valori al di sopra dei 12 mg/m³ (in tutto solo 21), si ottiene una previsione corretta nel 53.4% degli eventi acuti sperimentali.

Nelle figure 6.14 e 6.15 sono presentati, rispettivamente, il diagramma completo ed un dettaglio dell'uscita fornita dallo strumento neuronale sovrapposta alla serie sperimentale dell'inquinante. Si osserva che in alcuni casi l'output del modello eccede il valore sperimentale.

Fig. 6.14. Confronto dati sperimentali-simulati dalla rete '1Cot'

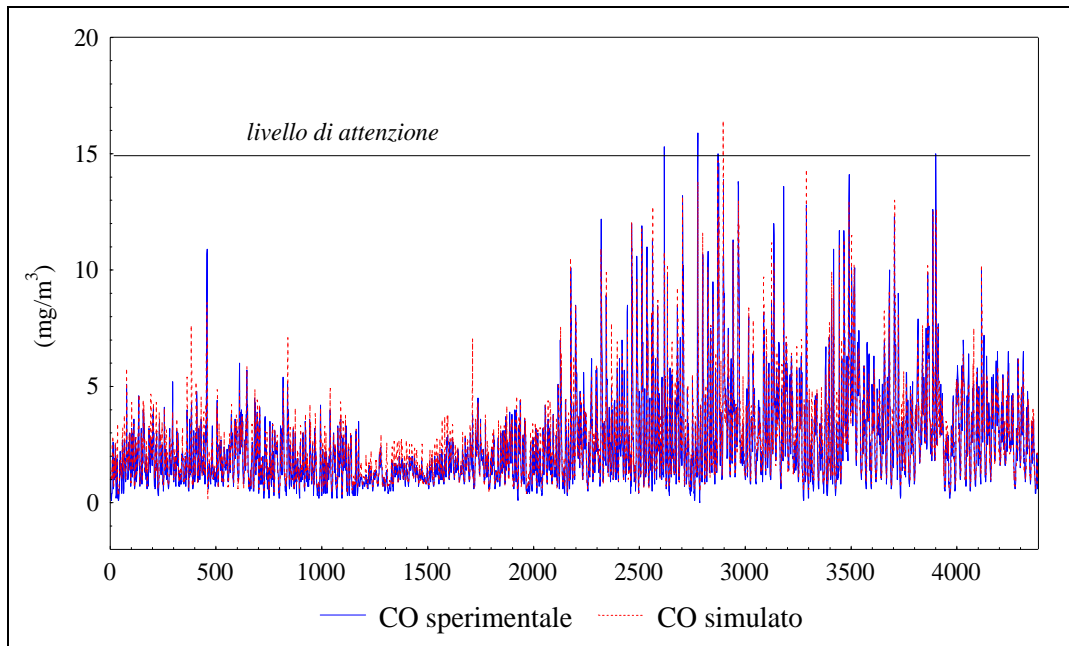
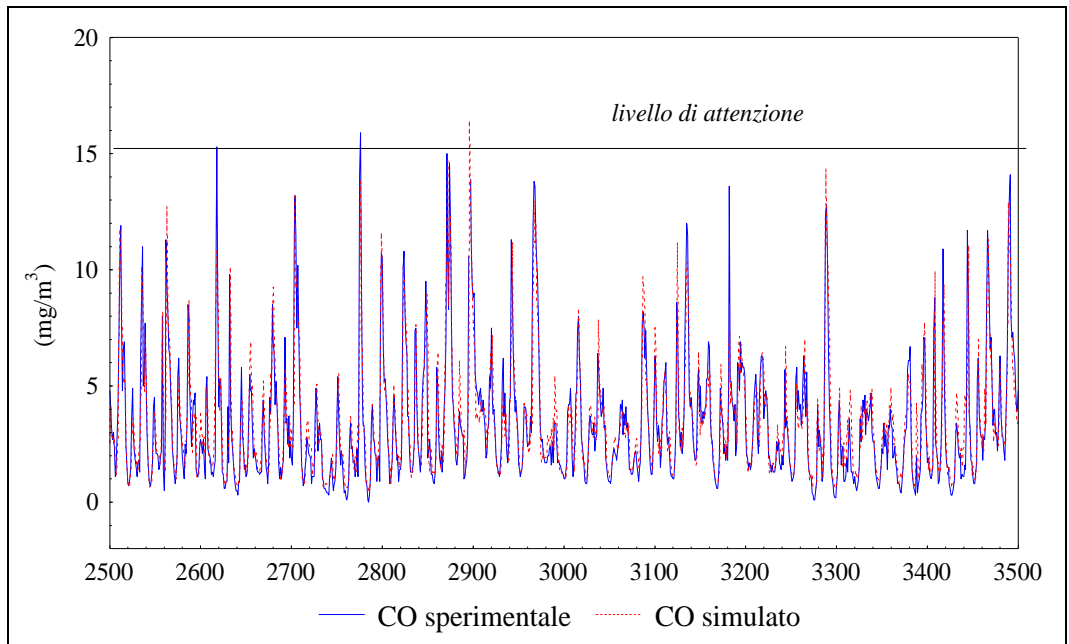
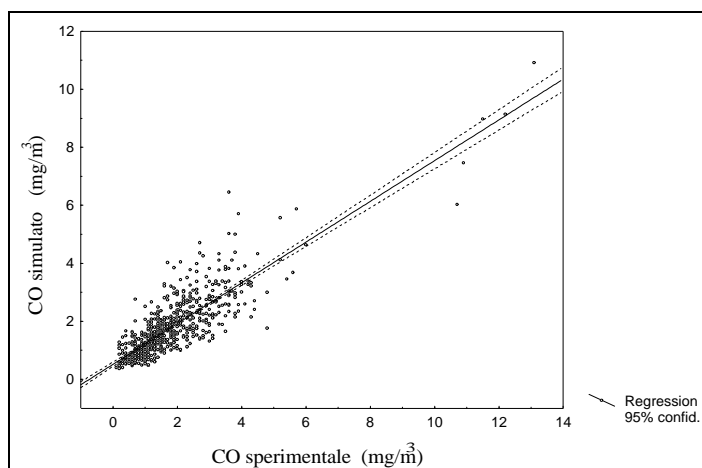


Fig. 6.15. Confronto dati sperimentali-simulati dalla rete '1COt': dettaglio



Il test di verifica del modello è stato eseguito su di un insieme di 664 *pattern*. L'errore medio nel *validation set* è leggermente migliore di quello ottenuto sui *pattern* di addestramento: $\pm 0.77 \text{ mg/m}^3$; la correlazione tra valore sperimentale e valore simulato è invece scesa a 0.85 (fig. 6.16), un valore comunque soddisfacente.

Fig. 6.16. Test della rete '1COt': regressione dati sperimentali-previsti.



Nel *validation set* gli eventi acuti sono in numero estremamente ridotto: solo 3 concentrazioni al di sopra dei 12 mg/m³, per i quali la rete non riesce a fornire un valore al di sopra di tale soglia, con un errore comunque non eccessivo.

Per la **previsione a passo temporale di 3 ore** è stato ‘dilatato’ il modello della rete ‘1COt’. La rete ‘3COt’ così ottenuta può essere descritta tramite l’espressione:

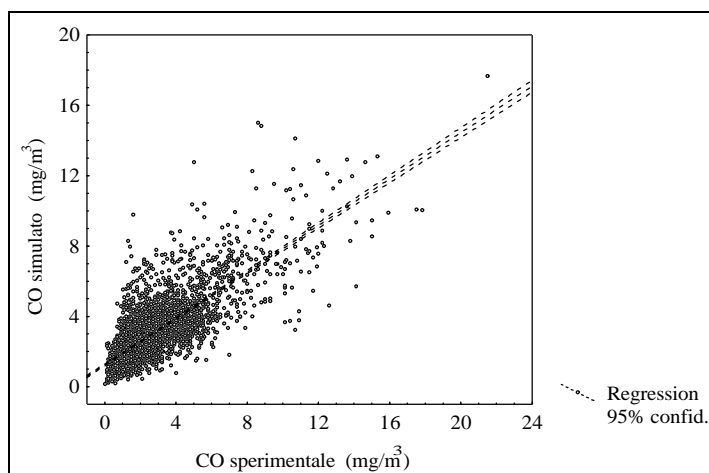
$$CO_{(t+3)} = f_{[11,8,1]} \left(\begin{array}{l} CO_{(t)}, CO_{(t-3)}, CO_{(t-6)}, CO_{(t-9)}, NO_X(t), \\ T_{(t)}, VV_{(t)}, \mathbf{g}_{(t)}, Tr_{(t)}, d_{(t+3)}, h_{(t+3)}; \mathbf{W} \end{array} \right)$$

La struttura è dunque rimasta invariata: si tratta ancora di un modello neuronale *feed-forward*, con 11 neuroni di input, 8 nel *layer* nascosto ed un solo output, costituito dal valore del monossido di carbonio al tempo $(t+3)$, essendo t l’ultimo istante per il quale si dispone di informazioni¹³.

L’errore medio raggiunto alla fine dell’addestramento (su 4339 *pattern*) è ± 1.22 mg/m³, circa il 8.1% del valore di attenzione. La correlazione tra CO sperimentale e CO previsto è 0.78 (fig. 6.17); la varianza spiegata dal modello è perciò il 60%, un valore piuttosto basso rispetto a quanto ottenuto per la previsione dell’ozono con il medesimo passo temporale.

Fig. 6.17. *Analisi dei risultati della rete ‘3COt’: regressione CO sperimentale-CO simulato.*

¹³ Fanno eccezione le variabili ‘giorno settimanale’, d , e ‘ora’, h , poiché funzioni note.

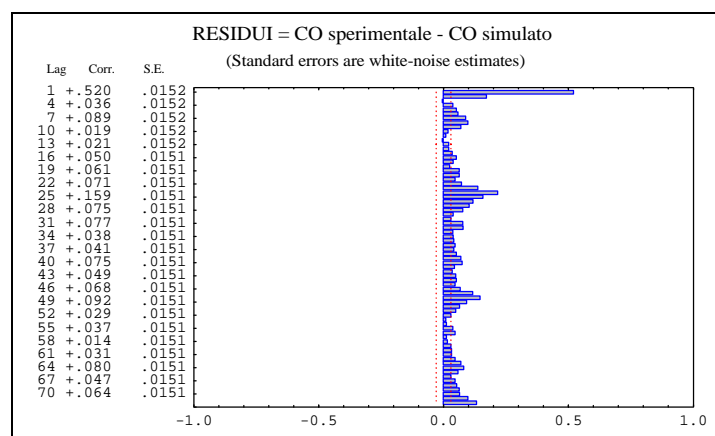


Le statistiche descrittive per il confronto tra la serie sperimentale e quella simulata sono riportate in tabella 6.7.

Tab. 6.7. *Statistiche descrittive per il confronto tra CO sperimentale e CO simulato dalla rete '3COt'.*

	<i>Media</i>	<i>Int. Confid. 95%</i>	<i>Mediana</i>	<i>Min</i>	<i>Max</i>	<i>Dev. Std.</i>	<i>Err. Std.</i>
CO sperimentale	2.48	2.42 ; 2.54	1.90	0.00	21.50	2.04	.03
CO simulato	2.88	2.83 ; 2.94	2.54	.16	17.66	1.73	.03
RESIDUI	-.40	-.44 ; -.37	-.46	-8.20	8.36	1.29	.02

Fig. 6.18. *Funzione di autocorrelazione dei residui del modello della rete '3COt'.*



Il correlogramma dei residui (fig. 6.18) testimonia una simulazione insoddisfacente da parte del modello.

L'analisi in dettaglio sulla previsione degli eventi acuti conferma il giudizio negativo sulla rete

'3COt': dei 5 valori di concentrazione al di sopra del *livello di attenzione* essa riesce a segnalare uno solo; anche estendendo l'analisi a tutti i valori eccedenti i 12 mg/m^3 , si ottiene una previsione corretta solo nel 33% dei casi.

L'andamento dell'output della rete è osservabile nelle figure 6.19 e 6.20.

Fig. 6.19. Confronto dati sperimentali-simulati dalla rete '3COt'.

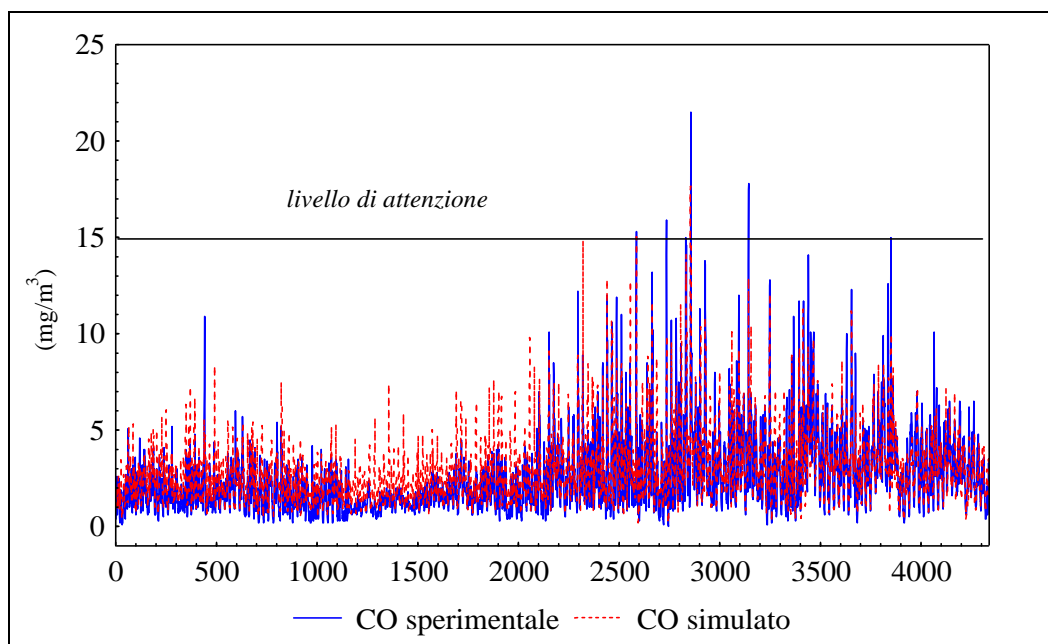
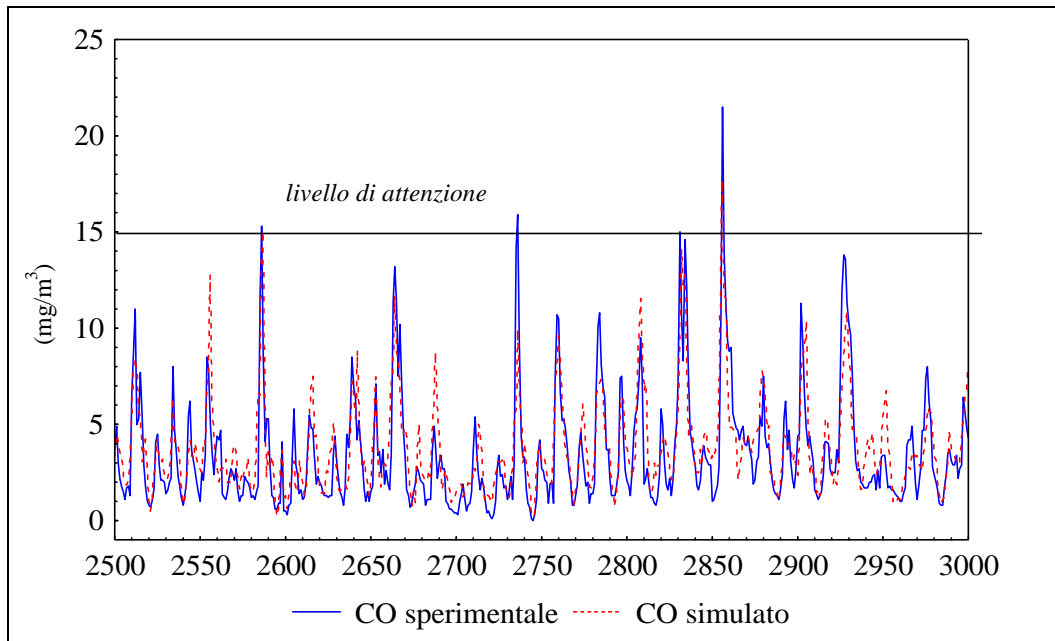


Fig. 6.20. Confronto dati sperimentali-simulati dalla rete '3COt': dettaglio.



I risultati piuttosto deludenti per la previsione del monossido di carbonio hanno indotto a cambiare metodo di addestramento. Nelle prove fin qui esposte il *training set* conteneva una mole d'informazione irrilevante ai fini della segnalazione degli eventi acuti. Non è infatti interessante che la rete impari a riprodurre in dettaglio l'andamento della concentrazione in *range* di valori lontani dal limite, ma piuttosto che segnali con precisione la possibilità di giungere ad uno *stato di allarme* o di *attenzione*.

In quest'ottica è stato sperimentato il seguente approccio: Si pensi ad una serie temporale $\{C_i\}$ delle concentrazioni di un inquinante e si immagini che il valore C_i misurato all'istante t_i sia il valore più probabile delle concentrazioni x misurabili all'istante t_i (*ipotesi di rappresentatività del campione sottoposto a misura*).

Per semplicità si supponga che la distribuzione delle x sia *normale* con *deviazione standard* σ : $p(x, \mathbf{m}, \mathbf{s})$.

In conseguenza delle ipotesi fatte lo sperimentatore che misura il valore C_i (in una distribuzione *normale* il valore più probabile è il valor medio μ , quindi $\mu = C_i$) se potesse ripetere la misura avrebbe la probabilità $P_i = 1 - \int_0^{C_A} p(x, C_i, \mathbf{s}) dx$ di incappare in una concentrazione superiore a C_A .

Con un tale ragionamento è possibile associare al valore C_i la probabilità di superare un limite prefissato C_A (ad esempio il valore di attenzione, come rappresentato in figura 6.21). In questo modo si può quindi trasformare una serie temporale di concentrazioni in una serie temporale di probabilità di superamento di una concentrazione prefissata, purché sia nota la densità di probabilità $p(x, \mathbf{m}, \mathbf{s})$. Quest'ultima può essere ottenuta dalla distribuzione di frequenza delle C_i (*ipotesi ergodica*: la media di insieme coincide con la media eseguita nel tempo).

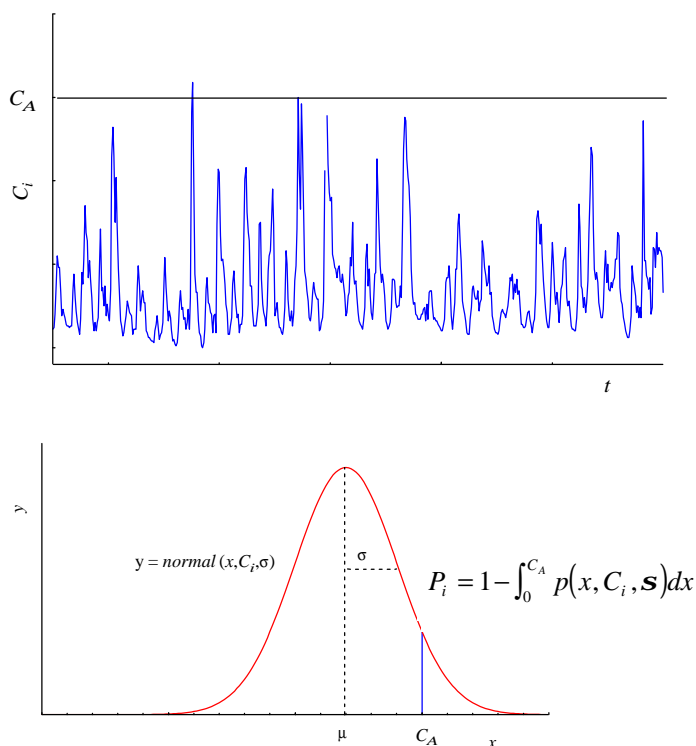


Fig. 6.21. Probabilità di superamento del limite C_A associata alla concentrazione C_i

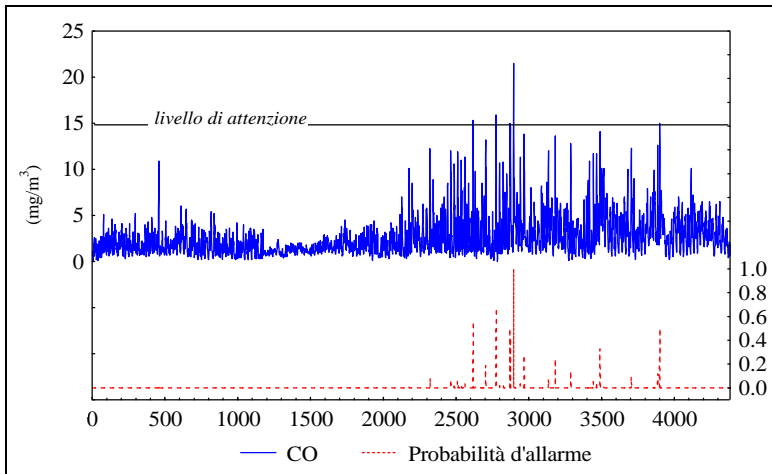


Fig. 6.22. Probabilità di allarme associata alla serie delle concentrazioni

In figura 6.22 sono state affiancate la serie temporale del monossido di carbonio e quella ritraente la probabilità di superamento del *livello di attenzione*. Si osserva dunque come l'andamento della *'probabilità d'allarme'* sia estremamente semplificato. Il calcolo

della probabilità già per concentrazioni di 12 mg/m^3 fornisce infatti valori molto prossimi a zero¹⁴. Al valore di concentrazione pari al limite C_A (identificato in questo caso nei 15 mg/m^3 del *livello di attenzione*), corrisponde invece un valore di probabilità di 0.5. Probabilità 1 è infine associata a valori di concentrazione pari al *livello di allarme*, essendo quest'ultimo esattamente il doppio di quello di *attenzione*.

I successivi modelli neurali sono dunque stati costruiti per prevedere la *probabilità d'allarme*.

La rete '1COtAP' per la previsione a passo temporale di un'ora è stata addestrata secondo il vettore di input di seguito specificato:

$$AP_{(t+1)} = f_{[11.8,1]} \left(AP_{(t)}, AP_{(t-1)}, AP_{(t-2)}, AP_{(t-3)}, NO_X_{(t)}, T_{(t)}, VV_{(t)}, \mathbf{g}_{(t)}, Tr_{(t)}, \mathbf{d}_{(t+1)}, \mathbf{h}_{(t+1)}; \mathbf{W} \right)$$

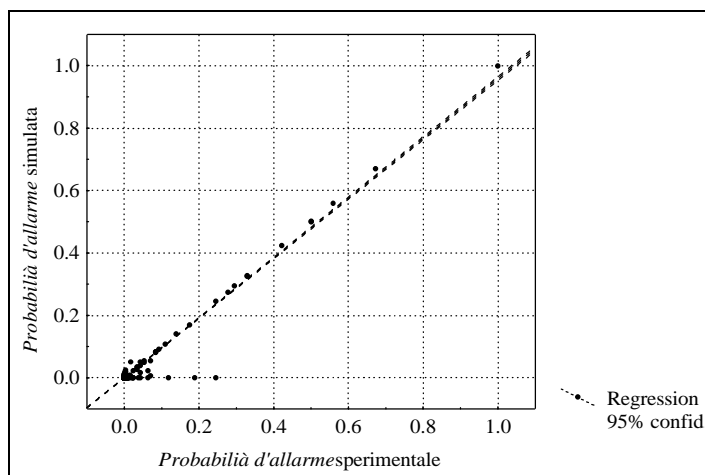
in cui AP è la variabile *'probabilità d'allarme'*.

L'errore medio a cui si è assestata la risposta neuronale è di ± 0.0055 , valore, questa volta, adimensionale perché riferito alla scala 0-1 delle probabilità.

¹⁴ Con $C_i = 12$ e $\sigma = 2.03$ (valore della deviazione standard della serie sperimentale del CO)
 $P_i = 1 - \int_0^{C_A} p(x, C_i, \mathbf{s}) dx = 0.07$.

La correlazione tra la *probabilità d'allarme* sperimentale e quella simulata è estremamente elevata: 0.98 (fig. 6.23); a tale valore corrisponde una varianza spiegata dal modello pari al 96%.

Fig. 6.23. *Analisi dei risultati della rete 'ICOtAP': regressione AP sperimentale - AP simulata.*



Le statistiche descrittive riportate in tabella 6.9 testimoniano una simulazione eccellente. Il valore massimo presentato dalla *probabilità d'allarme* calcolata dalla serie sperimentale delle concentrazioni è questa volta raggiunto dalla simulazione neuronale. Sembra dunque risolta la difficoltà, riscontrata in altri addestramenti, di far assumere al neurone di output valori posti al limite superiore della propria scala di attivazione.

Tab. 6.9. *Statistiche descrittive per il confronto tra 'probabilità d'allarme' sperimentale e 'probabilità d'allarme' simulata.*

	<i>Media</i>	<i>Int. Confid 95%</i>	<i>Mediana</i>	<i>Min</i>	<i>Max</i>	<i>Dev. Std.</i>	<i>Err. Std</i>
AP sperimentale	.00172	.00094 ; .00251	0.00	0.0000	.9993	.0266	.0004
AP simulata	.00152	.00074 ; .00229	0.00	0.0000	.9994	.0261	.0004
RESIDUI	.00021	.00005 ; .00037	0.00	-.0344	.2450	.0055	.00008

Il confronto tra l'output del modello e quello desiderato è riportato nelle figure 6.24 e 6.25.

Fig. 6.24. *Confronto dati sperimentali-simulati dalla rete 'ICOtAP'.*

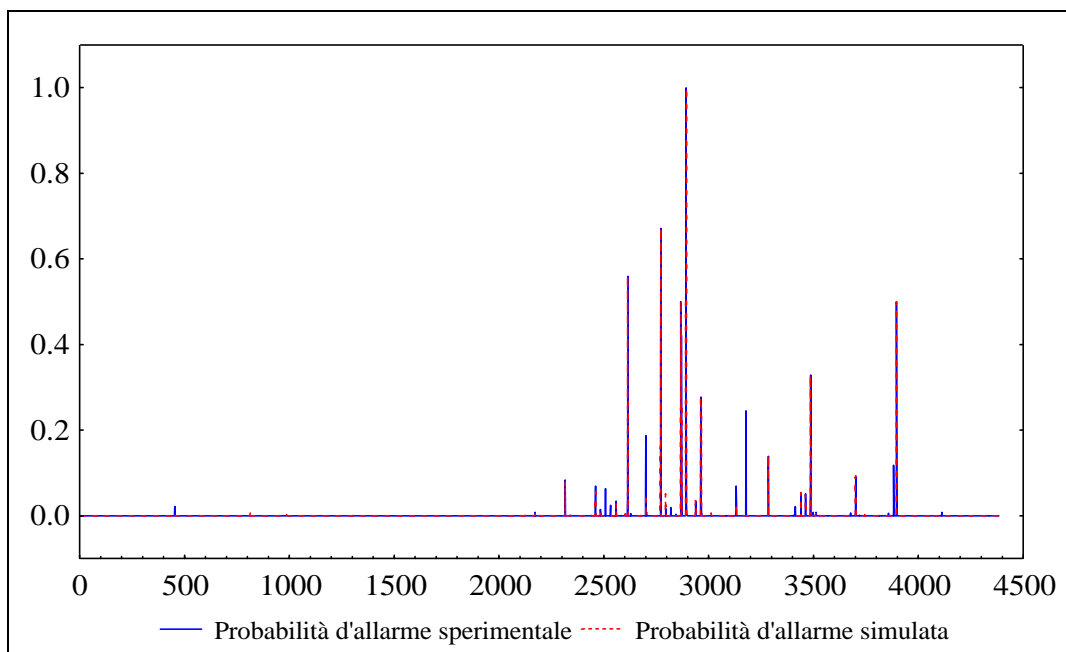
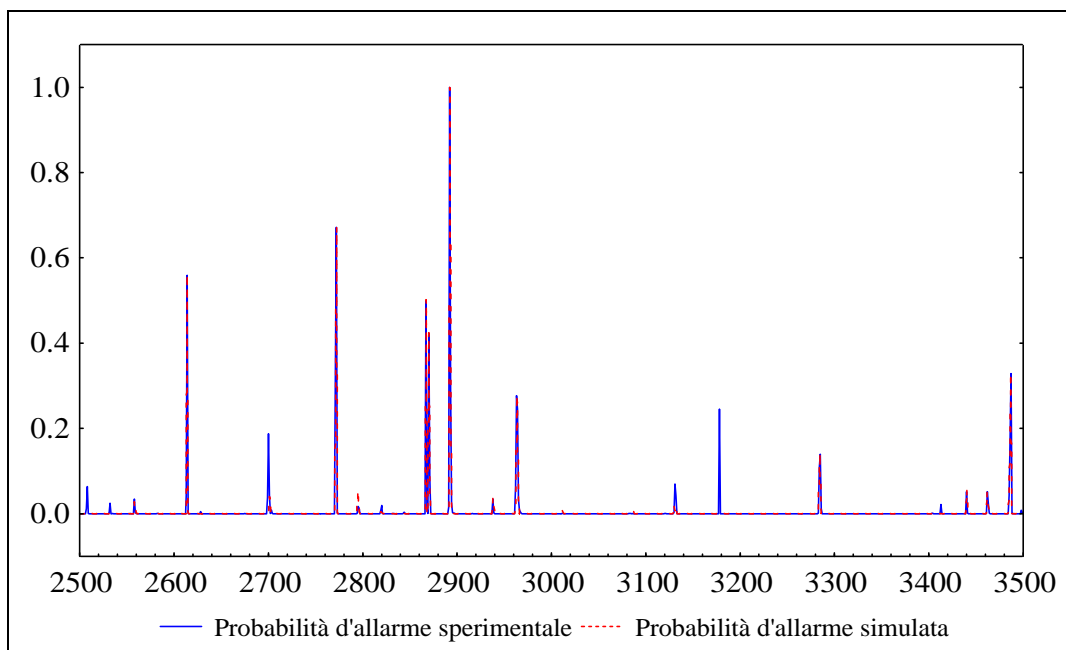
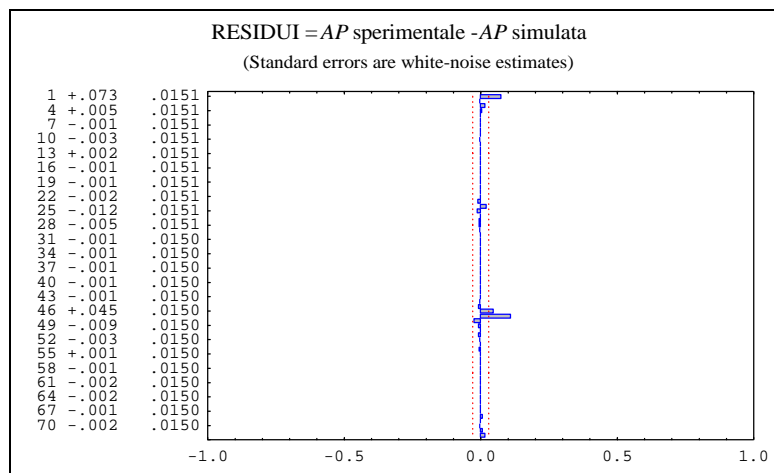


Fig. 6.25. *Confronto dati sperimentali-simulati dalla rete 'ICOtAP': dettaglio*



Il correlogramma dei residui (fig. 6.26) conferma il giudizio positivo sulla simulazione effettuata dalla rete 'ICOtAP'.

Fig. 6.26. *Funzione di autocorrelazione dei residui del modello della rete 'ICOtAP'.*



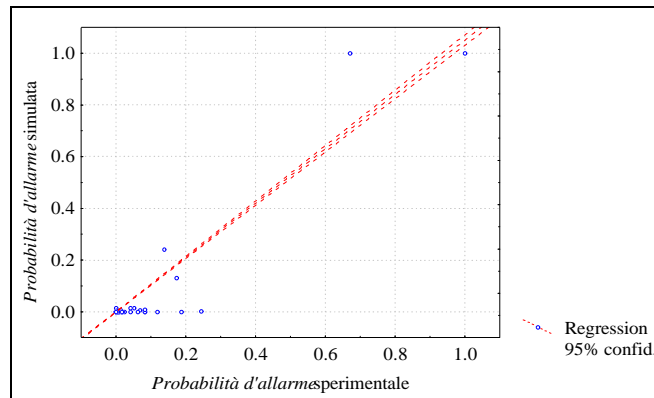
L'analisi di dettaglio sulla previsione dei picchi presentati dalla funzione *probabilità d'allarme* conferma l'ottimo comportamento dello strumento neuronale. I cinque *eventi acuti*¹⁵ del *training set* sono tutti correttamente previsti dalla rete neuronale. Estendendo l'analisi a tutti gli episodi corrispondenti a concentrazioni maggiori di 12 mg/m³ (cioè a valori di AP maggiori di 0.07), la previsione risulta corretta nel 87% dei casi.

Dato il numero piuttosto ridotto di dati 'significativi' all'interno della serie della *probabilità d'allarme* è particolarmente importante eseguire il test di verifica del modello. Gli ottimi risultati, dato il numero esiguo di valori non nulli e data la struttura piuttosto complessa della rete, potrebbero essere dovuti ad un apprendimento 'a memoria' degli output corretti. Viene infatti indicato in letteratura un rapporto massimo di 6 *pattern* per connessione al di sotto del quale una rete neuronale può imparare 'a memoria' l'output corretto associato al vettore di input. In questo caso il numero totale di *pattern* per connessione (99 in tutto) è notevolmente maggiore ma se si considerano i soli *pattern* 'significativi' il rapporto diviene inferiore a 5.

Il test di validazione è stato eseguito su un insieme di 1400 *pattern*. L'errore medio è risultato pari a ± 0.014 e la correlazione a 0.94 (fig. 6.27).

Fig. 6.27. Test della rete '1COtAP': regressione dati sperimentali-simulati.

¹⁵ Ad essi è associata un valore della AP maggiore di 0.5.



Nel *validation set* gli eventi acuti associati a concentrazioni maggiori di 12 mg/m^3 sono solo 10, di questi la rete neurale ne segnala correttamente 4. Le prestazioni in fase di test sono dunque inferiori a quelle raggiunte al termine dell'addestramento, ma comunque piuttosto soddisfacenti.

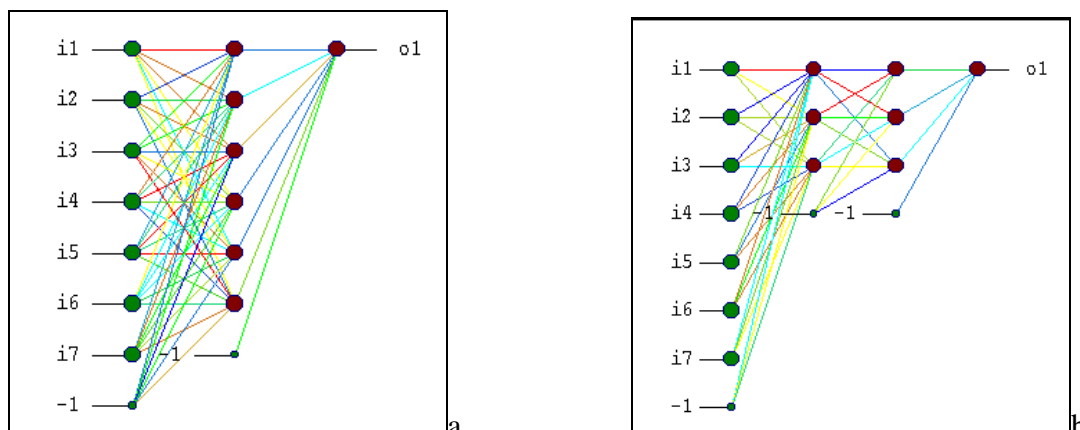
Pur con qualche difficoltà nel reperire un set di dati sufficientemente ricco di eventi acuti, questo tipo di approccio al problema della previsione degli episodi di alte concentrazioni sembra essere una delle strade percorribili.

6.3 Modelli per il biossido di azoto

Per il biossido di azoto si presentano solo alcune delle prove eseguite per confrontare diverse architetture neurali. Le conclusioni a cui si è giunti in questa fase di studio hanno orientato la scelta dei modelli anche per gli altri inquinanti.

Una prima prova è stata dedicata a verificare se l'utilizzo di due strati nascosti invece di uno migliorasse le previsioni. Si è dunque addestrato, con la medesima struttura del vettore di input e con il medesimo *training set*, due reti costituite da un numero di neuroni uguale ma diversamente distribuito: la rete '1NO₂' (tab. 6.11 e fig. 6.28 a), una *feed-forward* a 7 neuroni di input, 6 nell'unico strato nascosto ed un output, e la rete '1NO₂b' (tab. 6.11 e fig. 6.28b), sempre una *feed-forward* con i medesimi 7 neuroni di input, 3 neuroni nel primo strato nascosto, 3 nel secondo ed il consueto neurone di output.

Fig. 6.28. Struttura delle reti '1NO₂' e '1NO₂b'.



L'espressione formale dei due modelli di previsione del biossido di azoto dell'ora successiva, è, per la rete '1NO₂':

$$\text{NO}_2(t+1) = f_{[7,6,1]}(\text{NO}_2(t), \text{NO}_2(t-2), \text{NO}_2(t-3), \text{NO}_2(t-4), T(t), \text{VV}(t), \mathbf{g}(t); \mathbf{W})$$

mentre per la rete '1NO₂b' cambia solo la disposizione dei neuroni nascosti:

$$\text{NO}_2(t+1) = f_{[7,3,3,1]}(\text{NO}_2(t), \text{NO}_2(t-2), \text{NO}_2(t-3), \text{NO}_2(t-4), T(t), \text{VV}(t), \mathbf{g}(t); \mathbf{W})$$

I risultati raggiunti dai due modelli sono molto simili, con una prestazione leggermente migliore per la rete ad un solo strato nascosto (errore medio leggermente inferiore e correlazione dati sperimentali - simulati lievemente maggiore, come riportato in tab. 6.11). Questa verifica ha orientato l'implementazione di modelli strutturati in un solo strato nascosto.

Tab. 6.11. Reti neurali a confronto per la previsione del biossido di azoto.

Rete	Tipo	N° patterns	vettore di input	ME (mg/m ³)	r sper-sim
1NO ₂	f-7-6-1	7975	T-1; VV-1; γ-1; NO ₂ -4; NO ₂ -3; NO ₂ -2; NO ₂ -1	± 11.5	.91
1NO ₂ a	f-7-6-1	7975	T-1; VV-1; γ-1; d34; d23; d12; NO ₂ -1	± 11.6	.90
1NO ₂ b	f-7-3-3-1	7975	T-1; VV-1; γ-1; NO ₂ -4; NO ₂ -3; NO ₂ -2; NO ₂ -1	± 11.7	.90

Un secondo approfondimento ha riguardato la ricerca della 'forma' migliore con cui inserire nel vettore di input la componente autoregressiva dell'inquinante. In particolare è

stato verificato se fornire le derivate del valore dell'inquinante piuttosto che i valori assoluti dell'inquinante stesso desse migliori risultati.

Così, mentre nella rete '1NO₂' la componente autoregressiva è rappresentata dai valori assunti dall'inquinante nelle quattro ore precedenti al valore che si vuole prevedere, nella rete '1NO_{2a}' vengono inseriti in input gli ultimi tre valori della derivata prima, affiancati dal valore assoluto dell'inquinante solo per l'ora precedente l'output:

$$\text{NO}_2(t+1) = f_{[7,6,1]}(\text{NO}_2(t), d(t), d(t-1), d(t-2), T(t), VV(t), \mathbf{g}(t); \mathbf{W})$$

dove d indica appunto la derivata¹⁶.

Il confronto tra i risultati dei due modelli ha indicato una previsione lievemente migliore per la rete addestrata con i valori assoluti dell'inquinante. Anche questa conclusione è stata tenuta presente nei successivi sviluppi modellistici.

¹⁶ Per il calcolo della derivata prima è stato utilizzato il metodo *backward*: $f' = \frac{f_t - f_{(t-1)}}{\Delta t} + o(\Delta t)$, in cui l'ultimo termine è l'errore di ordine primo in δt . Il termine $d_{(t)}$ ad esempio è:

$$d_{(t)} = \frac{\text{NO}_2(t) - \text{NO}_2(t-1)}{\Delta t}.$$