

5. Applicazione ai dati sperimentali, un modello di previsione delle temperature

5.1 Ricostruzione dello spazio delle fasi

L'utilizzo del teorema di embedding per ricostruire lo spazio delle fasi relativo alla serie di temperature di Udine-Castello descritta nel capitolo 4, richiede la determinazione del *time delay* e della dimensione di *embedding* (cfr. par. 2.1.). A tal fine sono stati utilizzati i metodi di analisi esposti nel capitolo 2.

5.1.1 Mutua informazione media

Nelle figure 5.1 e 5.2 è rappresentata la mutua informazione media (2.8) per la serie dei dati sperimentali e per la serie *denoised* (cfr. par. 4.2.2.). Come si vede, entrambe le curve non presentano minimi, tuttavia la mutua informazione media della serie *denoised* diminuisce più lentamente e quindi il *time delay* è stato posto uguale a 1 per la serie dei dati sperimentali e uguale a 2 per quella *denoised*.

5.1 Ricostruzione dello spazio delle fasi

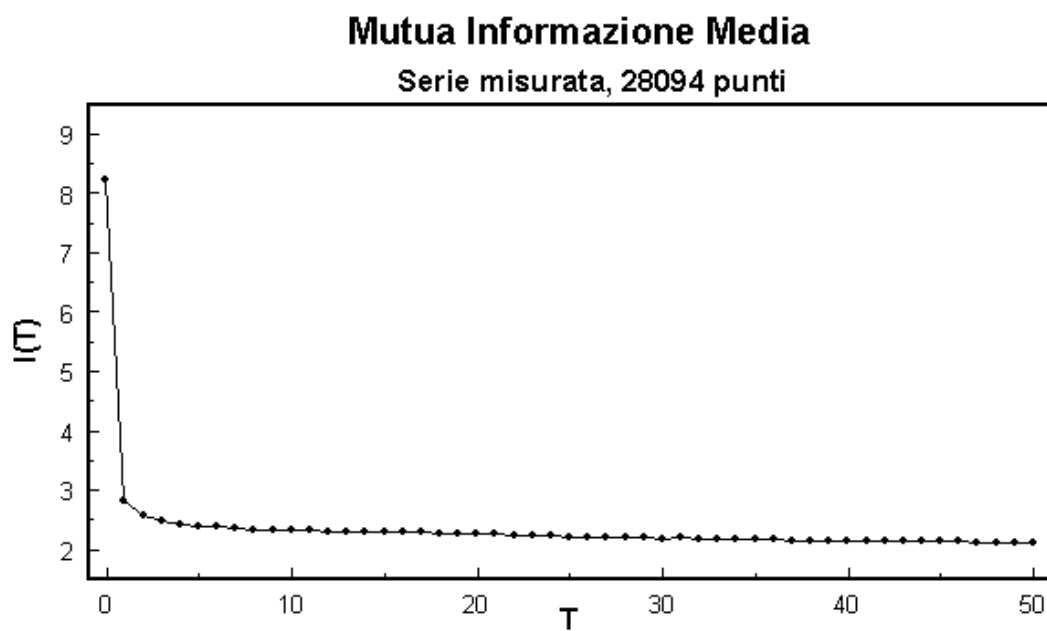


FIGURA 5.1 Mutua informazione media per la serie misurata.

5.1.2 Falsi vicini

Il risultato del calcolo dei falsi vicini per la determinazione della dimensione di embedding è rappresentato in Figura 5.3 per la serie misurata e in Figura 5.4 per quella *denoised*.

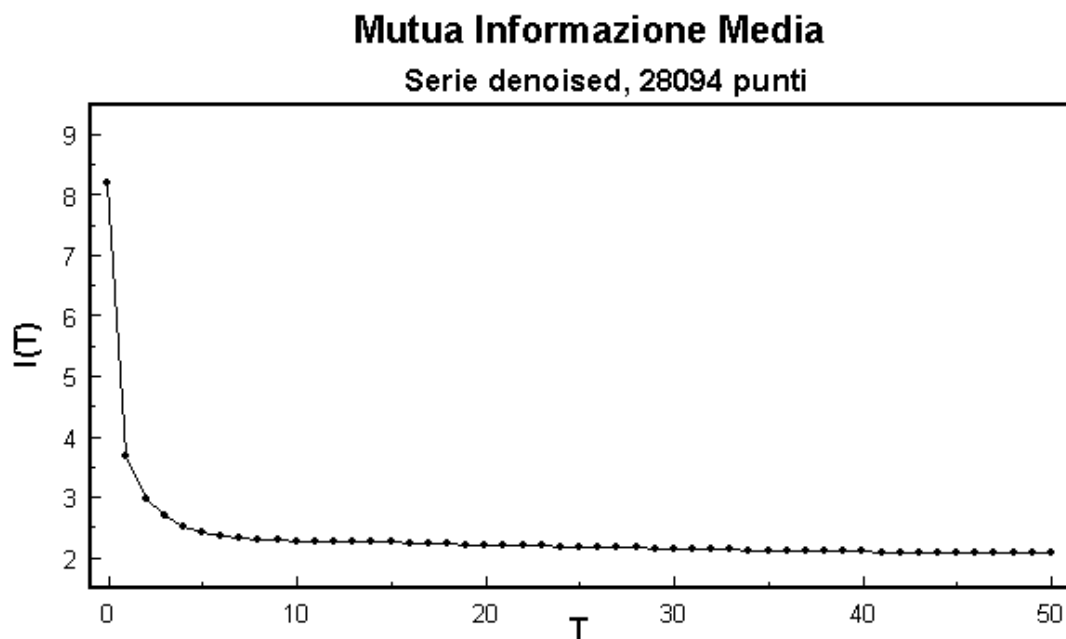


FIGURA 5.2 Mutua informazione media per la serie *denoised*.

5. Applicazione ai dati sperimentali, un modello di previsione delle temperature

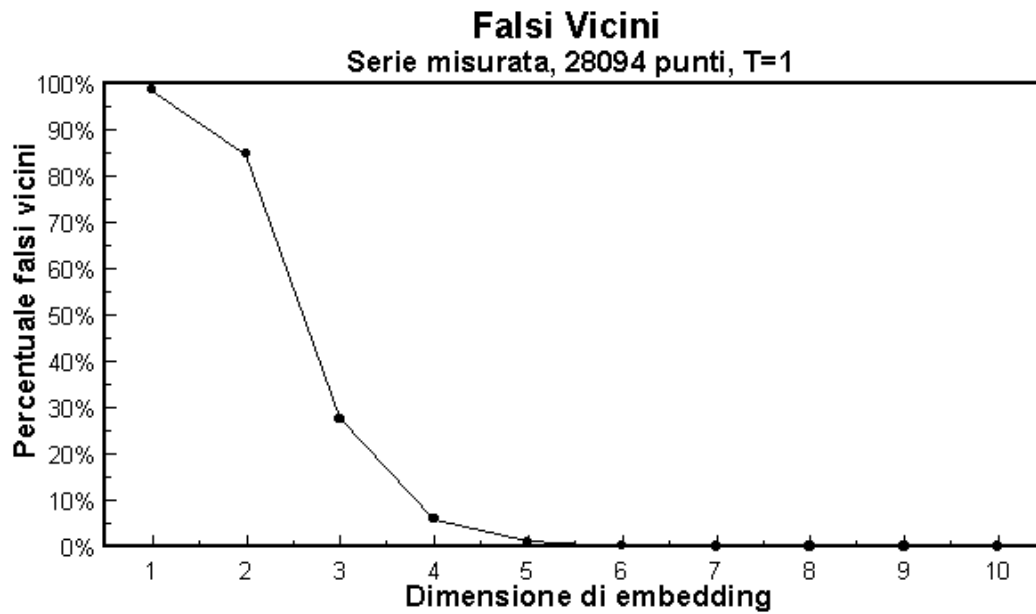


FIGURA 5.3 Falsi vicini per la serie misurata.

Come si vede dalle figure, in entrambi i casi il metodo individua una dimensione di *embedding* pari a 5 anche se, a causa del rumore, la percentuale di falsi vicini è sempre leggermente maggiore nel caso della serie misurata. Questo

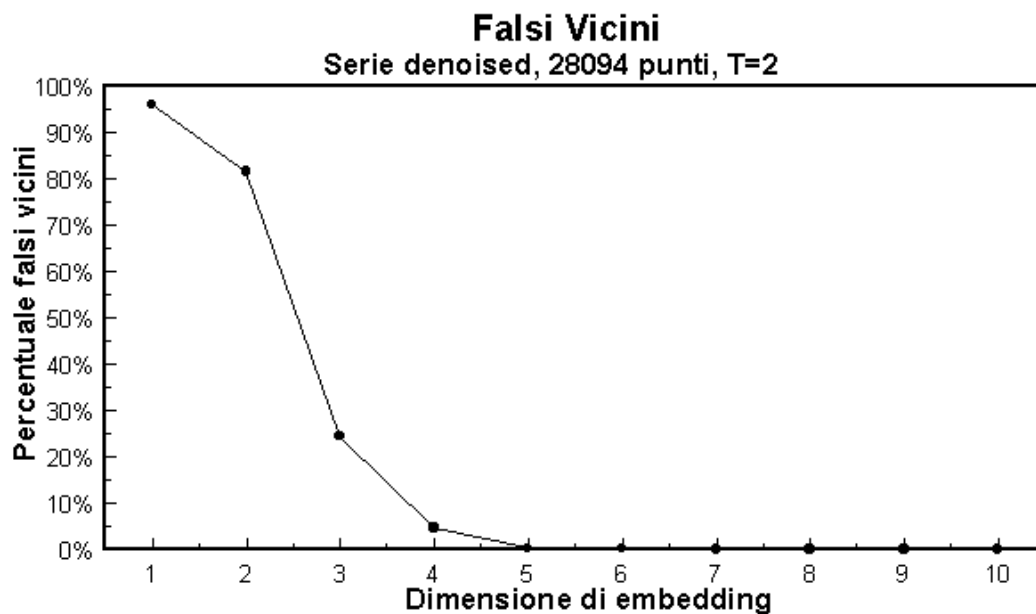


FIGURA 5.4 Falsi vicini per la serie *denoised*.

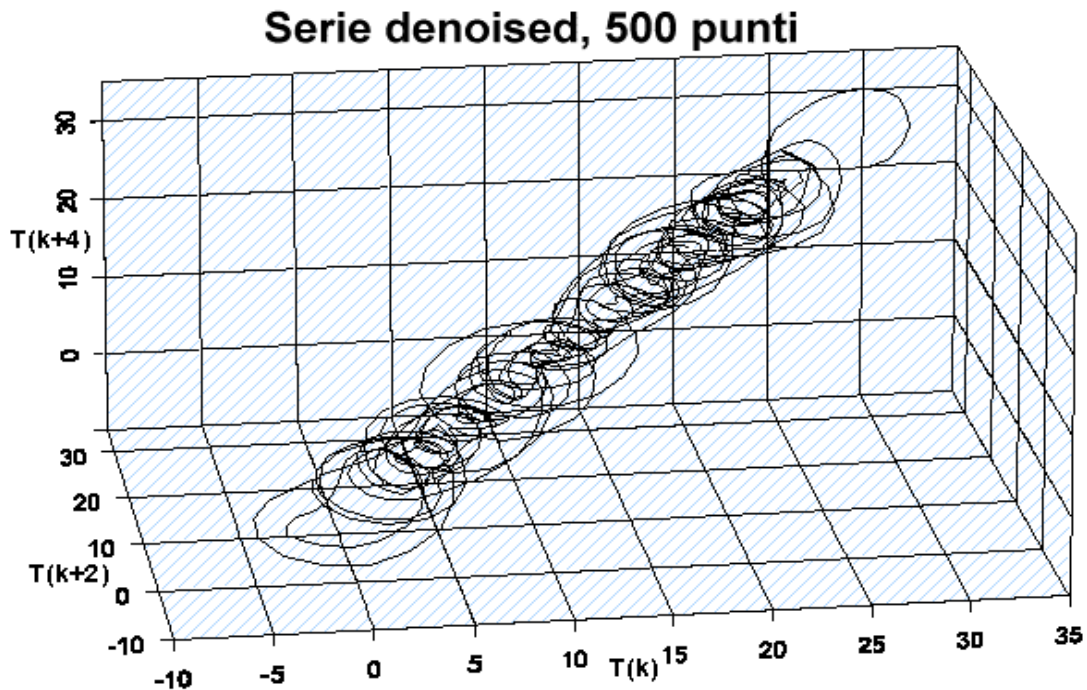


FIGURA 5.5 Sezione tridimensionale dell'attrattore della serie *denoised*.

risultato mostra che il metodo dei falsi vicini è efficace anche se applicato direttamente alla serie rumorosa.

Nella Figura 5.5, infine, è rappresentata una sezione tridimensionale dell'attrattore della serie *denoised* nello spazio delle fasi, ricostruito usando terne di valori della temperatura della forma: (T_k, T_{k+2}, T_{k+4}) .

5.2 Invarianti della dinamica

Come già fatto per i sistemi di Lorenz ed Hènon, anche alla serie dei dati delle temperature giornaliere di Udine sono stati applicati gli algoritmi di Grassberger-Procaccia ed è stata effettuata la ricerca del primo esponente di Lyapunov.

5.2.1 Algoritmo di Grassberger-Procaccia

L'applicazione dell'algoritmo di Grassberger-Procaccia sia alla serie misurata che a quella *denoised* ha prodotto risultati ambigui a causa dei problemi illustrati nel paragrafo 3.2. poiché non è stato possibile determinare con sicurezza la zona lineare delle curve. Ciò è probabilmente dovuto a due cause principali: in primo luogo

5. Applicazione ai dati sperimentali, un modello di previsione delle temperature

L'algoritmo richiede il calcolo delle distanze tra tutti i punti dell'attrattore e quindi l'allocazione di un vettore avente un numero di componenti proporzionale a $N^2/2$ ove N è il numero dei punti; le dimensioni di tale vettore sono però limitate dalla memoria a disposizione e quindi nel calcolo della funzione di correlazione non è stato possibile utilizzare più di 10000 punti. L'altra causa è rappresentata dal rumore presente nella serie misurata e da quello residuo della serie *denoised*, evidentemente ancora sufficiente a rendere inefficace l'algoritmo.

D'altro canto l'uso del metodo di Grassberger-Procaccia per determinare la caoticità di una serie non è consigliato (Theiler, 1986; Osborne e Provenzale, 1989; Provenzale et al., 1992) essendo molto più attendibile il metodo dei falsi vicini o il calcolo del primo esponente di Lyapunov; inoltre la dimensione di correlazione, unica informazione aggiuntiva fornita dall'algoritmo di Grassberger-Procaccia, non è necessaria agli scopi del presente lavoro.

5.2.2. Esponente di Lyapunov

I risultati dell'applicazione del metodo di Rosenstein (par. 3.5.) alla serie *denoised* (l'unica che verrà utilizzata di qui in avanti) sono rappresentati nella Figura

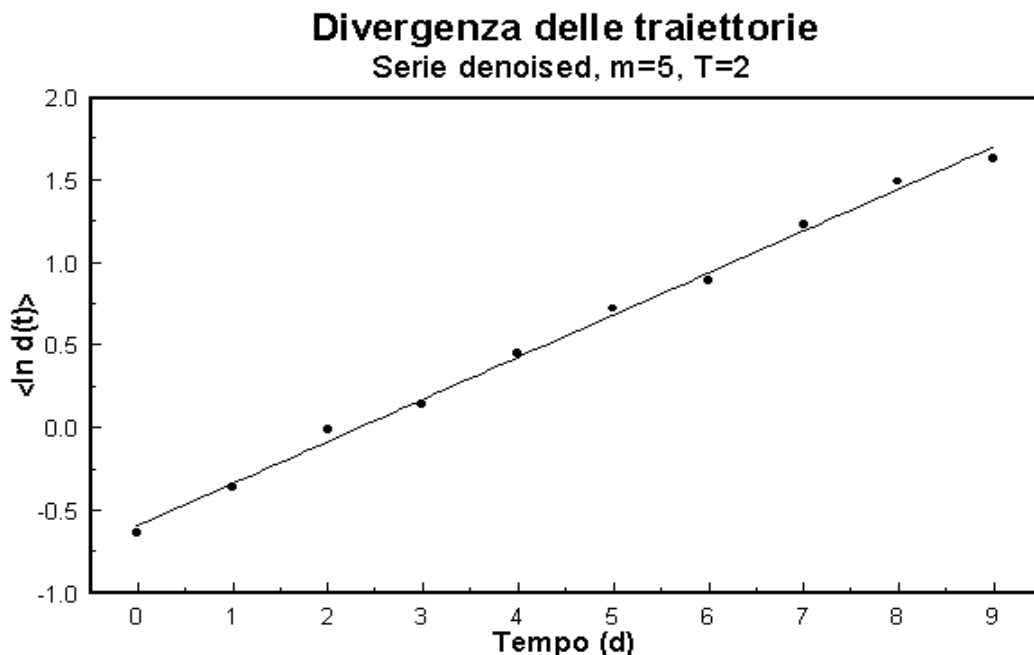


FIGURA 5.6 Divergenza media delle traiettorie per la serie delle temperature.

5.3 Un modello di previsione delle temperature

5.6. Il valore dell'esponente di Lyapunov maggiore risulta essere $\lambda_1 = 0.25$ il che conferma ulteriormente la caoticità della serie esaminata.

5.3 Un modello di previsione delle temperature

Dopo aver determinato la *time delay* e la dimensione di *embedding* è possibile ricostruire l'attrattore del sistema il quale, a sua volta, può essere usato per ottenere un modello della dinamica locale per predire l'evoluzione di un punto dello spazio delle fasi all'interno del bacino di attrazione nel quale tale punto si trova.

Il modello qui proposto si basa su di un'idea molto semplice: due punti vicini dell'attrattore tendono a separarsi a causa della divergenza delle traiettorie, ma per un tempo limitato (che dipende dall'esponente di Lyapunov) essi rimarranno abbastanza vicini in modo tale che l'evoluzione di uno segua quella dell'altro. Se dunque uno dei due punti è l'ultimo della serie, l'evoluzione temporale dell'altro può essere usata per ottenere una previsione dell'evoluzione del primo.

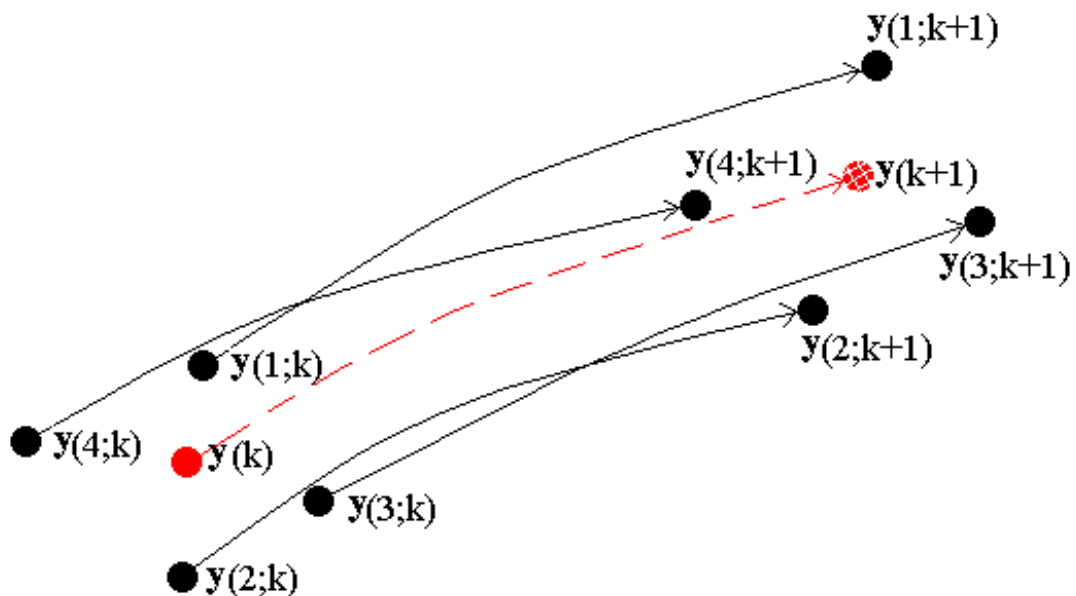


Figura 5.7 Rappresentazione schematica del modello previsionale: $y(k)$ rappresenta il punto del quale si vuole prevedere l'evoluzione $y(k+1)$, i punti neri rappresentano i primi vicini di $y(k)$ con le rispettive evoluzioni.

5. Applicazione ai dati sperimentali, un modello di previsione delle temperature

La previsione può essere resa più accurata se, anziché considerare solo il primo vicino del punto del quale si vuole ottenere l'evoluzione, si calcola il punto previsto come una media dell'evoluzione di un certo numero N_B di punti vicini al punto dato. Nella Figura 5.7 è rappresentato schematicamente il modello.

Sia dunque $\mathbf{y}(k)$ il punto del quale si vuole predire l'evoluzione $\mathbf{y}(k+1)$; dati i primi N_B vicini di $\mathbf{y}(k)$: $\mathbf{y}(r;k)$; $r=1,2,\dots,N_B$, si assume che $\mathbf{y}(k+1)$ sia dato dalla media pesata dei punti $\mathbf{y}(r;k+1)$, successivi dei punti $\mathbf{y}(r;k)$, con i pesi inversamente proporzionali alle distanze di tali punti da $\mathbf{y}(k)$:

$$\mathbf{y}(k+1) = \sum_{r=1}^{N_B} w_r \cdot \mathbf{y}(r;k+1) \quad (5.1)$$

ove

$$w_r = \frac{d_r}{d}, \quad d_r = \frac{1}{|\mathbf{y}(r;k) - \mathbf{y}(k)|}, \quad d = \sum_{r=1}^{N_B} \frac{1}{d_r}.$$

Con una generalizzazione immediata il modello (5.1) può essere utilizzato per prevedere l'evoluzione di $\mathbf{y}(k)$ dopo n passi temporali:

$$\mathbf{y}(k+n) = \sum_{r=1}^{N_B} w_r \cdot \mathbf{y}(r;k+n). \quad (5.2)$$

Ovviamente la qualità della previsione subirà un deterioramento all'aumentare di n a causa della divergenza delle traiettorie; infatti, se E_i è l'errore che si compie nella determinazione di $\mathbf{y}(k+i)$, in condizioni ideali l'errore nella previsione del passo temporale successivo sarà dato da $E_{i+1} = e^{\tau\lambda} E_i$ ove τ è il periodo di campionamento e λ il primo esponente di Lyapunov del sistema; essendo quest'ultimo positivo, l'errore tende necessariamente ad aumentare. Di fatto poiché il modello, approssimando un punto con una media pesata di punti vicini, inevitabilmente introduce degli errori indipendenti dalla divergenza delle traiettorie, la qualità della previsione sarà sempre peggiore di quella ideale.

5.3 Un modello di previsione delle temperature

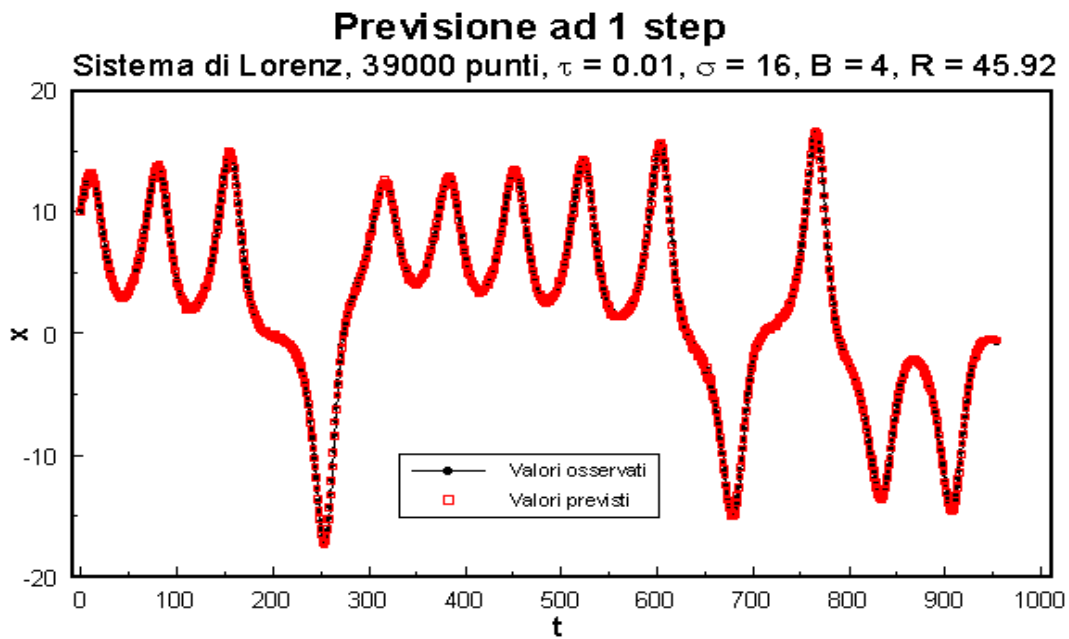


FIGURA 5.8 Confronto dei valori osservati con quelli previsti ad uno step d'integrazione per la variabile x del sistema di Lorenz.

Nelle Figure 5.8 e 5.9 sono rappresentate le previsioni, rispettivamente ad uno e tre passi temporali, relative alla variabile x del sistema di Lorenz: l'errore medio compiuto sulla previsione ad un passo è $E_1 = 1.03$, quello sulla previsione a

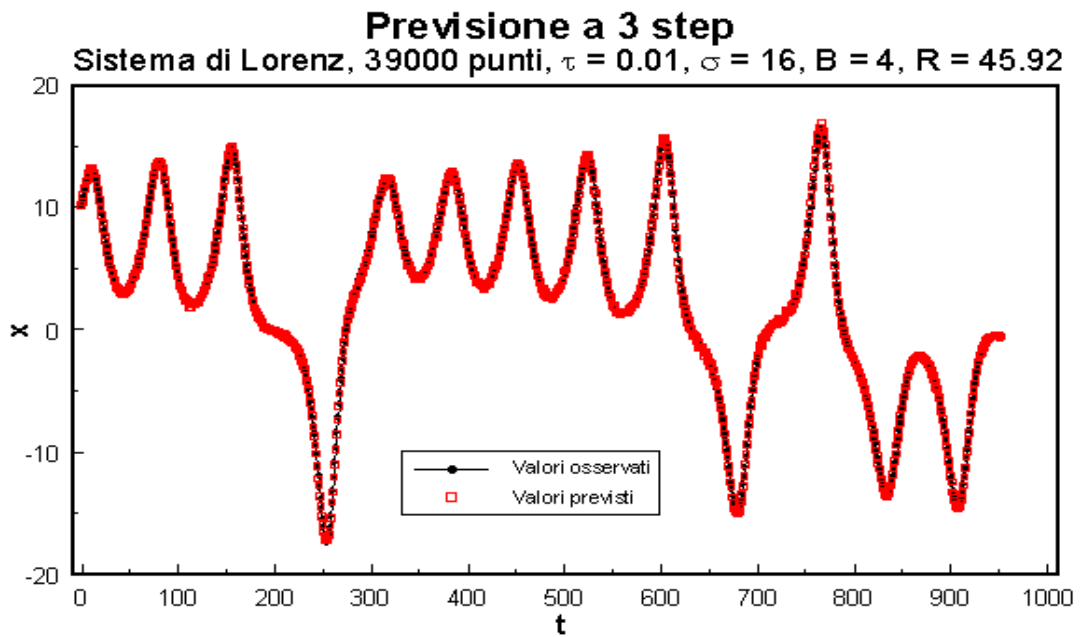


FIGURA 5.9 Confronto dei valori osservati con quelli previsti a tre step d'integrazione per la variabile x del sistema di Lorenz.

5. Applicazione ai dati sperimentali, un modello di previsione delle temperature

tre passi è $E_3 = 3.07$; se si confronta quest'ultimo con l'errore indotto dalla divergenza delle traiettorie: $E_3 = e^{2\tau\lambda} E_1 = 1.06$, si vede come l'errore indotto dal modello sia preponderante.

Nelle Figure 5.10 e 5.11 sono rappresentate le previsioni ad uno, due e tre giorni della serie delle temperature. Come si vede, nonostante gli errori introdotti, il modello consente di ottenere previsioni abbastanza buone.

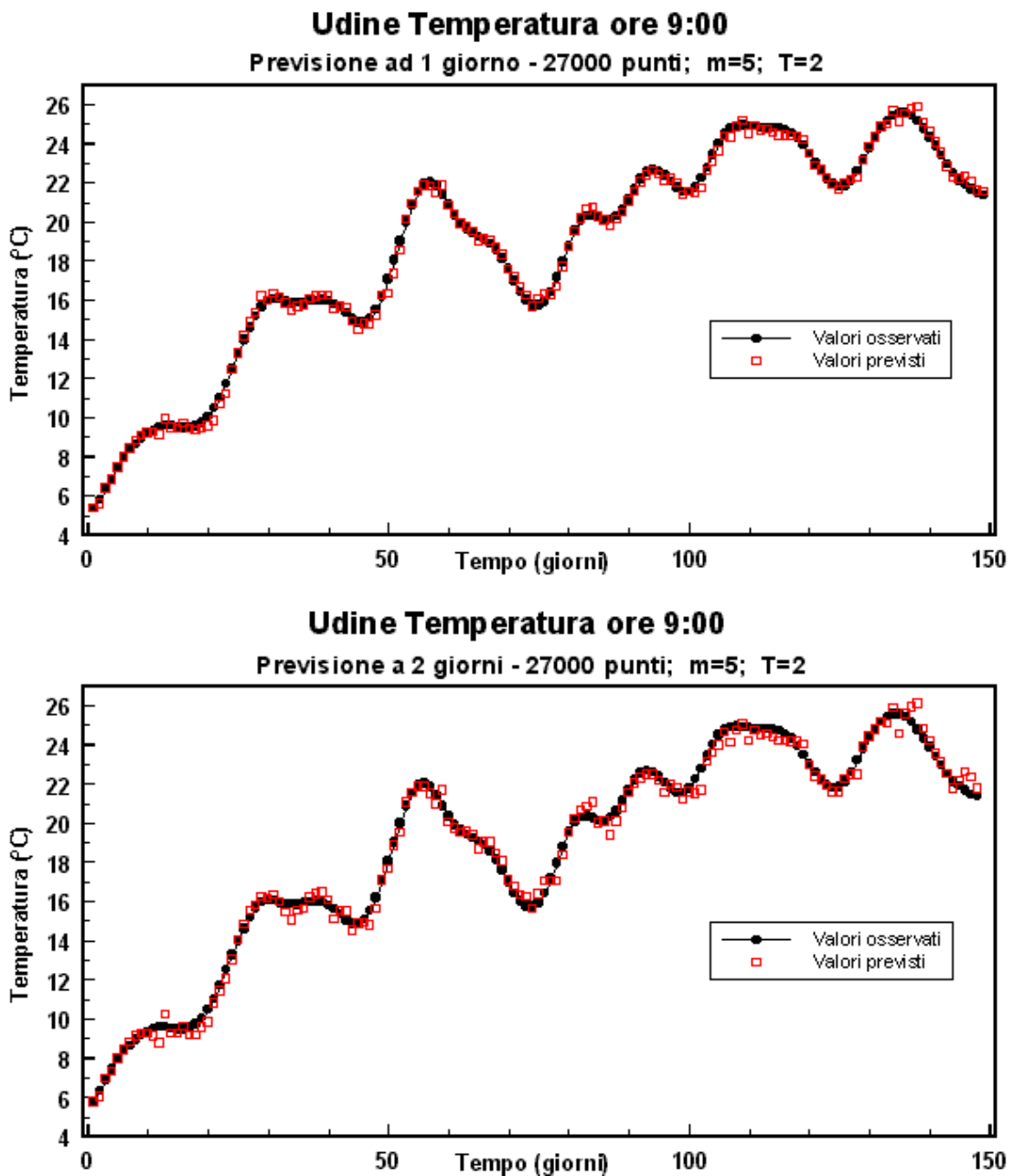


FIGURA 5.10 Confronto dei valori osservati con le previsioni ad uno e due giorni per la serie di temperature di Udine.

5.4 Risultati e conclusioni

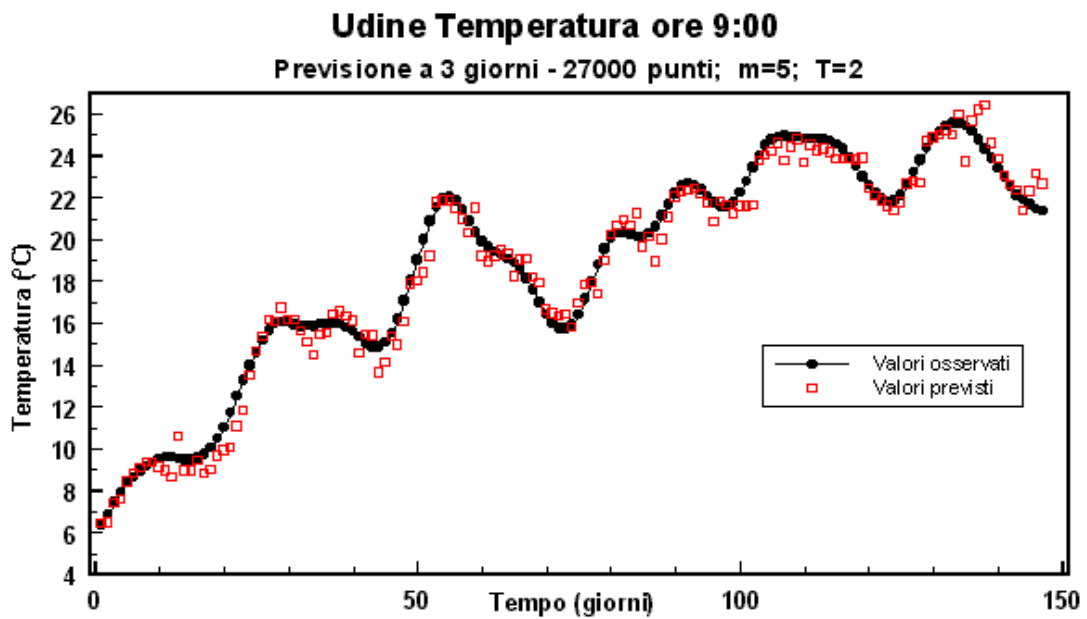


FIGURA 5.11 Confronto dei valori osservati con le previsioni a tre giorni per la serie di temperature di Udine.

Nella Tabella 5.1 sono riportati gli errori minimi, massimi e medi compiuti nelle previsioni ad uno, due e tre giorni delle Figure 5.10 e 5.9:

Ampiezza previsione (giorni)	Errore minimo (°C)	Errore massimo (°C)	Errore medio (°C)
1	0.00	0.75	0.20
2	0.01	1.32	0.33
3	0.02	2.06	0.58

TABELLA 5.1 Errori compiuti nelle previsioni del modello.

5.4 Risultati e conclusioni

L'applicazione del metodo dei falsi vicini ha permesso di dimostrare che la temperatura dell'aria di Udine-Castello è governata, almeno localmente, da un sistema caotico a 5 gradi di libertà e non da un sistema stocastico. Questo risultato è

5. Applicazione ai dati sperimentali, un modello di previsione delle temperature

stato confermato dal calcolo dell'esponente di Lyapunov maggiore. L'attrattore di tale sistema caotico nello spazio delle fasi a 5 dimensioni è stato ricostruito utilizzando il teorema di Takens ed avendo determinato il valore del *time delay* con il metodo della mutua informazione media.

Le informazioni sull'evoluzione temporale di un punto nello spazio delle fasi contenute nell'attrattore di un sistema caotico sono state utilizzate per sviluppare un modello locale di previsione della serie studiata. Tale modello consente di ottenere previsioni abbastanza buone, almeno entro un limite pari a tre volte il periodo di campionamento della serie in esame.

Le tecniche qui descritte hanno carattere generale nel senso che possono essere applicate ad una serie temporale a priori qualsiasi; e, qualora questa risultasse generata da un sistema caotico a basso numero di gradi di libertà, il modello proposto nel presente lavoro potrà essere utilizzato per ottenere delle previsioni.