

ANALISI DI GUASTO IN TEMPO REALE PER UN PROCESSO DI DEPURAZIONE BIOLOGICA

F. Baggiani*, S. Marsili-Libelli**

Sommario – La rivelazione automatica dei guasti riveste un'importanza sempre maggiore nella conduzione degli impianti di depurazione a causa degli elevati standard di trattamento richiesti e della necessità di proteggere il capitale investito dalle conseguenze di un guasto prolungato che si può propagare nell'impianto con effetti a valanga. In questo lavoro si descrive lo sviluppo e l'applicazione di un sistema di analisi automatica dei guasti (FDI) basato su un'estensione dinamica dell'analisi alle componenti principali (ACP). Dopo aver brevemente richiamato la struttura dell'impianto oggetto dell'applicazione e la teoria alla base della rilevazione dei guasti, si è sviluppato un sistema di Fault Detection and Isolation (FDI) in grado di rilevare la presenza e la natura di malfunzionamenti. Il sistema è basato su un'estensione in tempo reale dell'analisi alle componenti principali (ACP) attraverso la quale si confronta il funzionamento corrente del processo con un "modello di buon funzionamento" definito precedentemente sulla base di dati che corrispondono ad un funzionamento certamente soddisfacente. Una volta messo a punto l'algoritmo, se ne sono valutate le prestazioni del sistema utilizzando un insieme di dati rilevati con cadenza di un minuto compresi fra Agosto 2007 e Maggio 2008 presso l'impianto di depurazione di Pagnana (Empoli) gestito da Acque SpA. Il sistema può effettuare due tipi di analisi: di breve e di lungo periodo, finalizzate al riconoscimento di malfunzionamenti di diversa natura. Il sistema è inoltre in grado di distinguere diverse modalità di malfunzionamento: dal guasto al sensore, al disturbo momentaneo (spike) alla presenza di dati anomali. La versatilità del sistema è dovuta all'impiego di più statistiche, che aggiungono robustezza alla diagnosi e si rivelano particolarmente utili nel funzionamento con orizzonte lungo.

REAL-TIME FAULT DETECTION FOR A WASTEWATER TREATMENT PLANT

Summary – Automatic fault detection is becoming increasingly important in wastewater treatment plant operation, given the stringent treatment standards and the need to protect the investment costs from the potential damage of an unchecked fault propagating through the plant. This paper describes the development of a real-time Fault Detection and Isolation (FDI) system based on an adaptive Principal Component Analysis (PCA) algorithm, used to compare the current plant operation with a "good behaviour" model based on a preliminary set of data. The algorithm was developed in the LabView 8.20 (National Instruments, Austin, TX, USA) platform for real-time operation in the compact Field Point, a Programmable Automation Controller by National Instruments supervising the plant operation. The FDI was tested with a large set of

operational plant data with 1 hour sampling time from August 2007 through May 2008. Two time horizons were used in the analysis: a short term monthly horizon proved very reliable in isolating sensor failures and short duration disturbances such as spikes, whereas the long term horizon provided accurate detection of long-term drifts. The system robustness is enhanced by the use of multiple statistics, not only control charts but also contribution plots, which proved instrumental in discriminating among the various causes of malfunctioning.

Parole chiave: analisi di guasto, analisi alle componenti principali, depurazione biologica, automazione, controllo di processo.

Keywords: fault detection, principal component analysis, wastewater treatment, programmable automation controllers, process control.

1. INTRODUZIONE

Il monitoraggio degli impianti di trattamento riveste un'importanza sempre maggiore nella conduzione di impianti di depurazione per ottenere un effluente in accordo con la normativa vigente in materia di scarichi (D.Lgs 152/06) e garantire una continua efficienza dell'impianto. Questo studio descrive lo sviluppo di uno strumento di analisi e diagnosi dei sensori operanti nell'impianto di trattamento delle acque reflue di Pagnana nel comune di Empoli e gestito da Acque S.p.A.

I moderni impianti di depurazione sono dotati di sensori in grado di raccogliere informazioni su un ampio numero di variabili. Le misure raccolte sono utilizzate per monitorare il processo fornendo indicazioni per azioni manuali o per attuare delle azioni di controllo in retroazione mediante sistemi automatici di controllo. Questo presuppone in ogni caso l'affidabilità dei sensori e la bontà delle misure disponibili, mentre la presenza di guasti o di misure non veritiere è incompatibile con l'automazione e porta ad una perdita di fiducia nel sistema da parte degli operatori. Proprio per evitare tutto ciò è importante che il sistema di supervisione del processo comprenda anche, oltre ad un programma di manutenzione delle sonde, un sistema di rivelazione dei guasti che evidenzia subito la presenza di un guasto o di una variazione anomala nel processo (Schraa *et al.*, 2006).

I dati raccolti dai sensori negli impianti di depurazione sono numerosi, non stazionari, possiedono non linearità non sempre perfettamente compensate e sono spesso autocorrelati, quindi è necessario utilizzare metodi multivariati in grado di trattare simultaneamente tutte le informazioni provenienti dal campo ed estrarre il nucleo dell'informazione, presentando solo le informazioni essenziali all'operatore o al sistema esperto di supervisione. Questo lavoro descrive appunto un tale sistema, che permette di migliorare l'affidabilità generale dell'impianto, sia in gestione manuale che automatica.

* Ing. Francesco Baggiani; Università degli studi di Firenze, Dipartimento di Sistemi e Informatica – Via di S. Marta, 3 – 50139, Firenze – E-mail: francesco.baggiani@gmail.com

** Prof. Stefano Marsili Libelli; Università degli studi di Firenze, Dipartimento di Sistemi e Informatica – Via di S. Marta, 3 – 50139, Firenze – E-mail: marsili@dsi.unifi.it.

2. CARATTERISTICHE DELL'IMPIANTO DI DEPURAZIONE

2.1 L'impianto di Pagnana

L'impianto di depurazione preso in esame è quello di Pagnana, nel Comune di Empoli. Si tratta di un impianto di depurazione biologica a fanghi attivi con una potenzialità di 88600 AE. I liquami trattati sono essenzialmente di due tipologie: reflui di origine mista (civile e meteorologica) ed extraflussi (spurghi da fosse settiche, in quantità minore anche percolati di discarica, fanghi industriali, liquami alimentari e soluzioni di lavaggio delle cantine).

La filiera di trattamento della linea acque dell'impianto in esame è strutturata come in Fig. 1 ed è composta da due stadi: 1) pre-trattamenti e trattamenti primari di tipo fisico; 2) trattamenti secondari biologici.

I reflui in arrivo dal collettore fognario subiscono prima una grigliatura grossolana, allo scopo di permettere l'eliminazione di elementi solidi (come pezzi di legno o plastiche di varia natura) che potrebbero danneggiare le apparecchiature meccaniche o ostruire le condotte. Dopo la grigliatura grossolana, il refluo viene sollevato mediante delle coclee ed è sottoposto ad una grigliatura più fine, arrivando così al comparto di disabbatura/disoleatura. Dopo il sollevamento con le coclee, il liquame transita attraverso l'impianto per gravità. Successivamente il refluo viene convogliato nella vasca di sedimentazione primaria e da qui ai trattamenti biologici. Il comparto biologico dell'impianto di depurazione di Pagnana è costituito da tre vasche anossiche di pre-denitrificazione in serie, da tre vasche aerobiche di ossidazione funzionanti in parallelo e da tre sedimentatori secondari, anch'essi operanti in parallelo (Fig. 1).

Infine il liquido chiarificato, uscito dai sedimentatori secondari, può subire un ulteriore processo di disinfezione mediante clorazione, qualora il livello di patogeni nell'effluente sia tale da richiederne l'applicazione.

Sull'impianto di Pagnana è attualmente operante il sistema di acquisizione dati e controllo compact Field Point (cFP)

(National Instruments, Austin, TX, USA) che acquisisce le misure provenienti dal campo e precisamente:

- concentrazione di ossigeno disciolto in mezzo e in uscita dalle tre vasche aerobiche;
- concentrazione di nitrati in uscita sia dalla vasca anossica che da quella aerobica;
- concentrazione di ammonio in uscita dalla vasca aerobica;
- misura dei solidi sospesi in ossidazione e nella linea di ricircolo;
- potenza assorbita dagli inverter delle sei turbine di aerazione;
- portate di ingresso e uscita del refluo;
- portate di ricircolo della miscela aerata e di ricircolo dei fanghi;
- misure chimico-fisiche: pH, potenziale RedOx e temperatura.

Il sistema di monitoraggio dell'impianto, realizzato attorno al compact Field Point (cFP) di National Instruments, consente l'acquisizione di tutti i parametri di processo disponibili (Figura 2); la frequenza di campionamento dal campo è di 1 secondo, mentre l'archiviazione, che avviene dopo il filtraggio, ha la cadenza di 1 minuto. Tutti i dati raccolti nell'arco delle 24 ore vengono salvati su un foglio di calcolo Excel denominato con la data corrente, che viene chiuso ogni giorno alle ore 24 e contestualmente ne viene aperto un altro con la data del giorno successivo. Il cFP opera in tempo reale e può funzionare in modo autonomo, dopo che sia stato scaricato a bordo il software di controllo. Il sistema è visibile via internet in quanto capace di pubblicare la vista del pannello di controllo come pagina web, a cui si può accedere attraverso la rete aziendale nella quale è contrassegnato da uno specifico indirizzo IP.

3. ANALISI IN COMPONENTI PRINCIPALI (ACP)

L'Analisi in Componenti Principali (ACP) è un semplice metodo non parametrico per estrarre l'informazione contenuta in un insieme di dati apparentemente confusi per ridondanza, rumore, o inadatto riferimento. Mediante questo metodo è

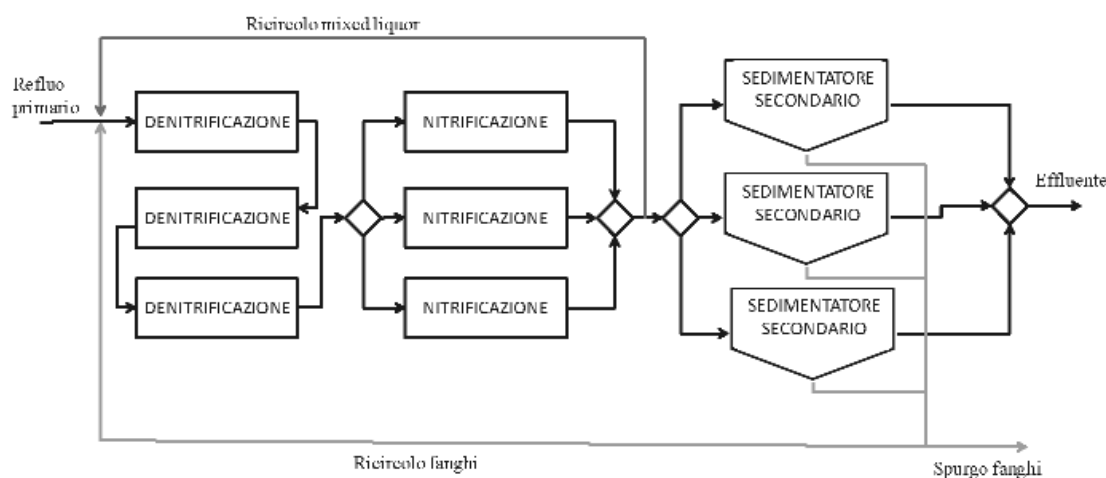


Fig. 1 – Schema del comparto biologico dell'impianto di Pagnana

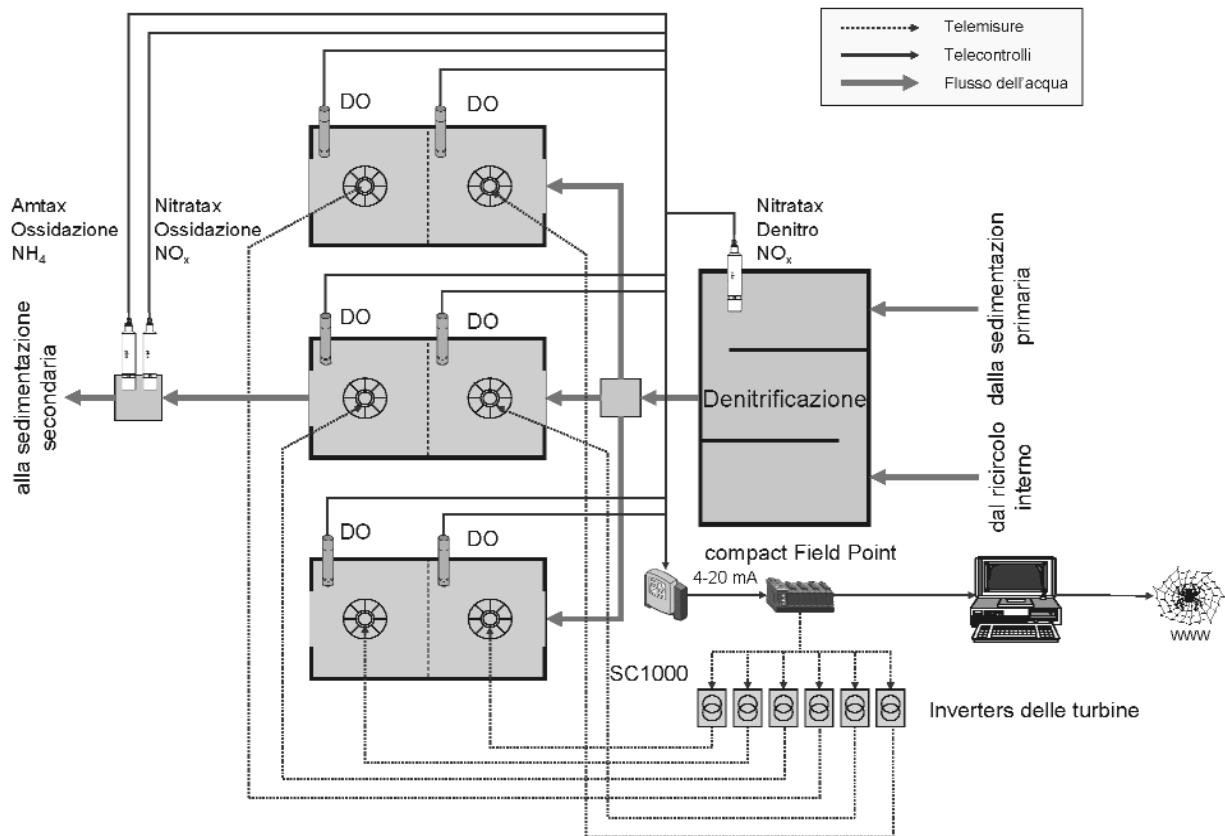


Fig. 2 – Sistema di acquisizione e monitoraggio dell'impianto di Pagnana

anche possibile ridurre un insieme complesso di dati ad una dimensione inferiore, rivelando così strutture informative più semplici che la complessità nascondeva. Infatti può accadere che ci si accorga che la base di rappresentazione originale è ridondante; in questo caso l'ACP permette di determinare una base di dimensionalità inferiore che spiega quasi altrettanto bene i dati, rivelando quali dinamiche sono più importanti, quali ridondanti e quali sono semplicemente rumore.

L'ACP è dunque un'operazione lineare effettuata sui dati allo scopo di rimuovere le tre principali fonti di disturbo: rumore, riferimento e ridondanza. Fra le infinite trasformazioni lineari che realizzano un cambio di base nello spazio di riferimento dei dati, l'ACP è quella che rende massima la varianza dei dati attraverso un cambio di base. Se una volta fatta questa operazione poi si vedrà che la maggior parte di variazione è contenuta in un sottoinsieme di dimensione inferiore si sarà anche risolto il problema della ridondanza, escludendo quelle componenti che contengono una parte limitata di variabilità dei dati. Supponendo che i dati da analizzare provengano da n sensori e che per ciascuno di essi si abbiano a disposizione p campioni, i dati possono essere raggruppati in forma compatta nella matrice $\mathbf{X} \in \mathbb{R}^{n \times p}$. Dopo averli depurati dal loro valor medio, se ne calcola la matrice di covarianza $\mathbf{C}_X \in \mathbb{R}^{n \times n}$ come:

$$\mathbf{C}_X = \frac{1}{p-1} \mathbf{X}\mathbf{X}^T.$$

La trasformazione lineare cercata dovrà perciò minimizzare la ridondanza, misurata dagli elementi non diagonali e massi-

mizzare la varianza, rappresentata dagli elementi diagonali. Dato che ciascun termine di \mathbf{C}_X è positivo, ciò significa rendere diagonale la matrice di covarianza. La trasformazione che rende \mathbf{C}_X simmetrica è fornita dalla matrice dei suoi autovettori, che formano una base ortonormale essendo \mathbf{C}_X simmetrica. Perciò nella nuova base i dati trasformati, detti *scores*, saranno espressi come

$$\mathbf{Z} = \mathbf{X}\mathbf{T} \quad \text{con} \quad \mathbf{C}_Z = \frac{1}{p-1} \mathbf{Z}\mathbf{Z}^T \text{ diagonale.} \quad (1)$$

La matrice di trasformazione $\mathbf{T} = [\mathbf{t}_1 \quad \mathbf{t}_2 \quad \dots \quad \mathbf{t}_n]$ rappresenta la base ortonormale ($\mathbf{T}^T \mathbf{T} = \mathbf{T} \mathbf{T}^T = \mathbf{I}$) composta dalle *componenti principali* cercate, dette anche *vettori di loading*, mentre i corrispondenti autovalori $\{\lambda_i, i = 1, \dots, n\}$ presentano l'importanza relativa di ciascuna componente principale nel rappresentare i dati e sono. Inoltre, è possibile classificare le dimensioni trasformate ordinando le colonne di \mathbf{T} partendo dall'autovettore \mathbf{t}_1 che corrisponde al più grande autovalore, λ_1 , e che spiega la massima frazione della varianza dei dati e proseguendo con gli autovalori che spiegano una percentuale sempre minore della variabilità dei dati. Si può perciò definire una successione di percentuali di variabilità spiegata da un certo numero di componenti principali ordinate a partire dalla più significativa (la numero 1), come:

$$p_1 = 100 \frac{\lambda_1}{\lambda_1 + \lambda_2 + \dots + \lambda_n} \quad p_2 = 100 \frac{\lambda_1 + \lambda_2}{\lambda_1 + \lambda_2 + \dots + \lambda_n} \quad \dots \text{etc.} \quad (2)$$

L'ACP può essere utilizzata non solo per ridurre la dimensio-

nalità di un problema, ma anche per ridurre il livello di rumore presente nei dati. È infatti possibile eliminare parte della variabilità residua, e quindi anche parte del rumore che accompagna l'informazione rilevante, prendendo in considerazione solo un numero a di componenti principali minore del numero delle variabili originali n . Maggiore è la correlazione tra le variabili iniziali, maggiore sarà la possibilità di ridurre la dimensionalità del problema in esame, riducendone la ridondanza intrinseca senza eccessiva perdita di informazione. La suddivisione fra la quota dei dati trattenuta e quella esclusa è evidente se si riesprimono i dati originali in funzione delle componenti principali come $\mathbf{X}=\mathbf{Z}\mathbf{T}$, tenendo presente che essendo \mathbf{T} ortonormale è $\mathbf{T}^T=\mathbf{T}^{-1}$. Nel caso di riduzione di dimensionalità $a < n$, l'ACP scompone la matrice dei dati \mathbf{X} in:

$$\mathbf{X} = \mathbf{Z}_a \mathbf{T}_a + \mathbf{Z}_{n-a} \mathbf{T}_{n-a} = \mathbf{Z}_a \mathbf{T}_a + \mathbf{E} \quad (3)$$

dove \mathbf{Z}_a è formata dalle prime a righe di \mathbf{Z} e \mathbf{T}_a la matrice trattiene solamente le prime a colonne di \mathbf{T} corrispondenti agli a autovalori più grandi $\{\lambda_1 > \lambda_2 > \dots > \lambda_a\}$. La matrice \mathbf{E} , detta dei *residui*, contiene le componenti dei dati corrispondenti agli autovalori meno significativi e che non sono spiegate dalle componenti principali trattenute.

3.1 Applicazione di ACP all'individuazione ed isolamento del guasto

I sistemi di individuazione ed isolamento del guasto (*fault detection and isolation*, FDI) rivelano in tempo reale il guasto nel sistema e il sensore che lo ha provocato. Sono anche in grado di segnalare la variazione dallo stato attuale di processo dovuta ad un evento temporaneo, consentendo così un migliore monitoraggio del processo e della prestazione della strumentazione impiegata rispetto alla semplice osservazione dell'andamento temporale delle diverse variabili misurate.

L'utilizzo della ACP in questo contesto consiste nel riconoscere se il dato corrente di funzionamento rientra in una partizione precedentemente definita come *normale* sulla base di dati sicuramente corrispondenti a questa condizione. Se al contrario il dato corrente se ne discosta, verrà segnalata un'anomalia al sistema di supervisione. Questo approccio consente il monitoraggio simultaneo di più variabili di processo acquisite nello stesso istante. La realizzazione di sistemi di diagnostica FDI tramite ACP si articola in varie fasi (Li e Rong, 2006; Aguado e Rosen, 2007; Detroja *et al.*, 2007; Mina e Verde, 2007):

1. Individuazione di un periodo di buon funzionamento dei sensori da utilizzare come modello del sistema in esame;
2. Applicazione della ACP al periodo di riferimento e scelta delle componenti principali da trattenere;
3. Determinazione delle soglie di buon funzionamento del processo tramite l'utilizzo di statistiche appropriate, che verranno introdotte fra poco;
4. Monitoraggio in tempo reale delle nuove misure e calcolo delle statistiche da paragonare con i limiti precedenti in modo da individuare la presenza e la natura dell'eventuale malfunzionamento.

3.2 Selezione del periodo di riferimento

Lo sviluppo di un sistema di monitoraggio in tempo reale necessita di un set di dati preliminari, individuati nei dati storici, in cui sia accertato il corretto funzionamento dei sensori. Questa prima fase è necessaria per poi definire i limiti delle statistiche che individuano il guasto. La grandezza del set di dati di riferimento è definita in base alla quantità di misure valide disponibili e in modo tale da mantenere una velocità computazionale adeguata.

3.3 ACP in prima fase e scelta del numero di componenti principali trattenute

Data una matrice di dati di riferimento \mathbf{X}_r con n variabili e m campioni, si applica l'ACP e si decide la dimensione delle a componenti principali da trattenere considerando le grandezze relative dei vari autovalori.

3.4 Statistiche T^2 e Q e loro limiti

Due tipi di statistiche sono utilizzate normalmente per estrarre informazioni dalle componenti principali: la T^2 di Hotelling e la Q , derivata dagli studi di *Statistical Process Control* (Oakland e Porter, 1995). La statistica T^2 , introdotta da Hotelling nel 1947, coglie la variazione delle componenti all'interno del modello di riferimento (o spazio di processo) ed è definita come

$$T^2(k) = \mathbf{x}^T(k) \mathbf{T}_k \mathbf{\Lambda}^{-1} \mathbf{T}_k^T \mathbf{x}(k) \quad (4)$$

dove \mathbf{T}_k è la matrice dei *loading* all'istante k ; $\mathbf{\Lambda} = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$ è la matrice diagonale degli autovalori della matrice di covarianza \mathbf{C}_X formata dai dati raccolti in una finestra di opportuna lunghezza fino all'istante corrente k . A questa statistica si affianca di solito la statistica Q , costruita come somma dei quadrati dei residui delle componenti principali (cioè le componenti non trattenute); essa è capace di misurare la variazione non considerata dal modello preso in esame con la ACP (spazio dei residui).

$$Q(k) = \mathbf{x}^T(k) (\mathbf{I} - \mathbf{T}\mathbf{T}^T) \mathbf{x}(k) \quad (5)$$

Entrambe le statistiche, in accordo alla teoria delle carte di controllo multivariate, presentano dei limiti superiori di controllo:

$$T_{lim}^2 = \frac{a(p-1)}{p-a} F_{a,p-a,\alpha} \quad (6)$$

dove a sono le componenti principali trattenute;

p sono i campioni esaminati dalla ACP;

$F_{a,p-a,\alpha}$ è l' α -percentile della distribuzione F di Fisher.

$$Q_{lim} = \theta_1 \left[\frac{h_0 c_\alpha \sqrt{2\theta_2}}{\theta_1} + 1 + \frac{\theta_2 h_0 (h_0 - 1)}{\theta_1^2} \right]^{(1/\theta_3)} \quad (7)$$

dove c_α è il quantile della distribuzione normale a livello $(1-\alpha)$ e

$$h_0 = 1 - 2 \frac{\theta_1 \theta_2}{3\theta_3^2} \quad \theta_1 = \sum_{i=a+1}^n \lambda_i \quad \theta_2 = \sum_{i=a+1}^n \lambda_i^2 \quad \theta_3 = \sum_{i=a+1}^n \lambda_i^3$$

con λ autovalori determinati precedentemente.

La procedura di controllo sul T^2 viene effettuata sugli *score* dei dati di riferimento. Se si riscontrano valori superiori alla soglia si deve verificare la loro attendibilità, distinguendo gli outliers, che devono essere eliminati, dai campioni anomali ma possibili, che devono essere trattenuti per migliorare la performance futura dell'algoritmo (Rosen, 1998).

3.5 Monitoraggio dei malfunzionamenti in tempo reale

In seguito ogni campione acquisito \mathbf{x}_{new} delle n variabili viene proiettato nello spazio di processo assunto come riferimento; così facendo è possibile ricavare la sua proiezione nello spazio delle componenti principali (*scores*) come

$$\mathbf{z}_{new} = \mathbf{x}_{new} \mathbf{T}_k \quad (8)$$

È importante notare che anche i nuovi dati devono essere standardizzati così come viene fatto per i dati utilizzati per l'identificazione del modello iniziale. Dagli *score* e dai *loading* trattenuti dei nuovi dati è possibile ricavare il valore delle due statistiche T^2 e Q .

Il passo conclusivo dell'algoritmo FDI è quello di evidenziare quale delle variabili originali provochi la maggior parte della variazione del valore di T^2 e Q , cioè sapere quale dei sensori è la causa principale del malfunzionamento (*isolation*). Ciò si realizza introducendo il concetto di "contributo della variabile" CV come descritto da Lee *et al.* (2004). Ogni variabile originale contribuisce secondo la formula:

$$CV(k) = \mathbf{z}(k) \mathbf{A}^{-1/2} \mathbf{T}_k^T = \mathbf{x}(k) \mathbf{T}_k \mathbf{A}^{-1/2} \mathbf{T}_k^T \quad (9)$$

dove \mathbf{A} è la matrice diagonale che ha gli elementi diagonali uguali agli autovalori determinati precedentemente, $\mathbf{z}(k)$ sono gli *score* al tempo k , \mathbf{T}_k è la matrice corrente dei *loading* e $\mathbf{x}(k)$ sono le misure originali standardizzate acquisite al tempo k .

3.6 Aggiornamento dello spazio di processo

Le misure provenienti dagli impianti di trattamento di acque reflue sono abitualmente numerose, ridondanti e occasionalmente rumorose e mancanti. Inoltre le caratteristiche del pro-

cesso sono dinamiche e non stazionarie; tutto ciò rende il monitoraggio e il sistema di *fault detection* più complicato da realizzare in quanto non possono essere applicate con successo le consolidate tecniche statiche come la ACP tradizionale (Rosen, 2001).

Proprio per questo tipo di problema si è dovuti ricorrere all'implementazione di una ACP dinamica che tenga conto delle autocorrelazioni e delle correlazioni tra variabili (Ku *et al.*, 1995; Lee C. *et al.*, 2004; Lee C. *et al.*, 2006).

La ACP è resa dinamica variando la natura della finestra di riferimento da statica a mobile, come mostrato in Fig. 3. In questo modo è possibile tener conto del valore appena acquisito e delle correlazioni rispetto agli istanti precedenti. Infatti la finestra mobile, di ampiezza m , scorre di un passo lungo l'asse del tempo ogni volta che è acquisito un nuovo campione di misure e può decidere a seconda di soglie fissate se tenerlo in memoria eliminando l'elemento più vecchio o scartarlo, perciò la finestra mobile può contenere dati non contigui nel tempo. Inoltre ad ogni passo la matrice dei dati deve essere riscalata rispetto alle nuove medie e varianze.

Il nuovo campione viene proiettato nello spazio di processo in memoria, ottenendo il valore delle due statistiche T^2 e Q . La finestra di riferimento mobile trattiene al suo interno il nuovo campione solo e soltanto se esso ha valori di T^2 e Q inferiori alle soglie di aggiornamento fissate come segue:

$$T_{aggiornamento}^2 \cong \frac{1}{8} \div \frac{1}{10} T_{lim}^2 \quad (10)$$

$$Q_{aggiornamento} \cong Q_{lim} \quad (11)$$

In quest'ultimo caso viene riapplicata la ACP e ricalcolata la nuova matrice dei *loading* in modo da aggiornare lo spazio di processo, nel caso contrario lo spazio di processo rimane inalterato.

3.7 Carte di controllo

Il monitoraggio in tempo reale ha lo scopo di analizzare e interpretare continuamente le misure acquisite in modo tale da

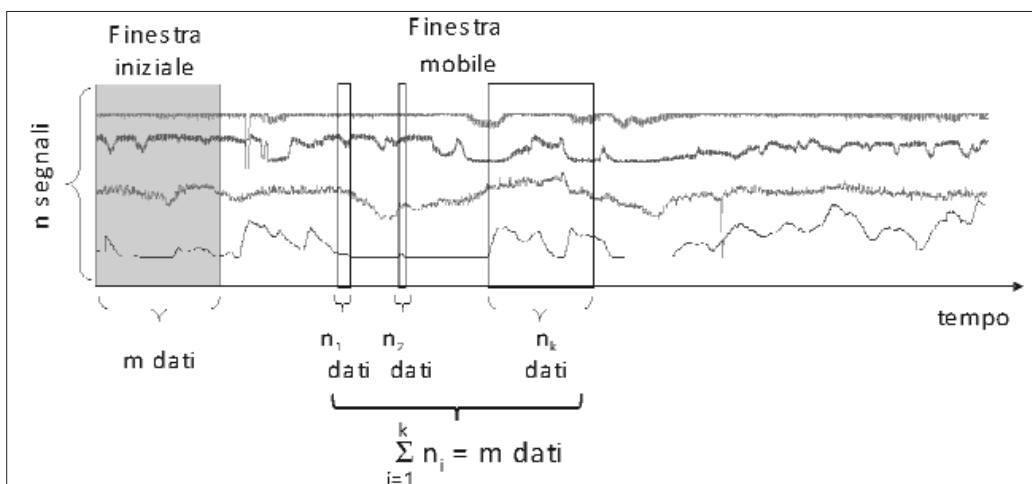


Fig. 3 – Avanzamento della finestra mobile per ACP dinamica nell'algoritmo FDI

individuare il malfunzionamento o guasto e scoprire la causa. È importante quindi avere strumenti di analisi ed interpretazione che possano essere facilmente codificate nell'algoritmo di analisi automatica, in modo da presentare all'operatore un risultato rapidamente interpretabile.

- *Grafici di T^2 e Q*

Gli andamenti temporali delle due statistiche riassumono la deviazione provocata da una o più variabili rispetto al modello di riferimento. Si possono fissare soglie di attenzione e di allarme all'aumentare del valore delle due statistiche.

- *Grafici degli score*

Ogni punto è identificato dai valori degli *score* trattenuti al tempo di campionamento corrente. Sotto condizioni di normale funzionamento il centro di gravità dell'insieme degli *scores* si troverà vicino all'origine, data la standardizzazione iniziale. I punti si allontaneranno dall'origine alla presenza di un disturbo.

È molto utile visualizzare l' i -esimo limite di confidenza $\delta_{i,\alpha}$ degli *score* per l' i -esima componente principale per avere un'idea della regione di corretto funzionamento

$$\delta_{i,\alpha} = \sqrt{\lambda_i \bar{t}_{p-1, \alpha/2}} \quad (12)$$

dove λ_i è l'autovalore i -esimo, $\bar{t}_{p-1, \alpha/2}$ è il percentile della distribuzione *t di Student* a livello α e p è il numero di campioni.

- *Contribution plot*

Un altro grafico solitamente utilizzato è quello che visualizza il contributo delle variabili (CV) al malfunzionamento, definito dall'eq. (9). Sull'asse delle ascisse si riportano le variabili originali acquisite, mentre sulle ordinate il contributo al malfunzionamento. Maggiore è l'altezza della barra, maggiore sarà il suo contributo.

3.8 Struttura dell'algoritmo FDI

L'algoritmo diagnostico di *Fault Detection and Isolation* (FDI) realizzato per il depuratore di Pagnana è stato inizialmente concepito per il comparto Azoto, segnalando ogni anomalia di funzionamento dei relativi sensori. Il software, inizialmente sviluppato in ambiente Matlab 7 è stato poi trasportato in ambiente LabView 8.20 (National Instruments, Austin, TX, USA) per il funzionamento in tempo reale a bordo del cFP che attualmente controlla l'impianto. Esso sfrutta la tecnica della ACP dinamica precedentemente descritta, appoggiandosi alla Advanced Signal Processing Toolkit, un pacchetto aggiuntivo di LabView 8.20 che ne accresce le funzionalità di analisi dei segnali. L'algoritmo FDI è strutturato complessivamente in tre blocchi, come mostrato in Fig. 4.

1. Acquisizione ed elaborazione dei dati del periodo di riferimento (fase preliminare).
2. Calcolo dello spazio di riferimento iniziale e dei limiti delle statistiche di buon funzionamento (fase I).

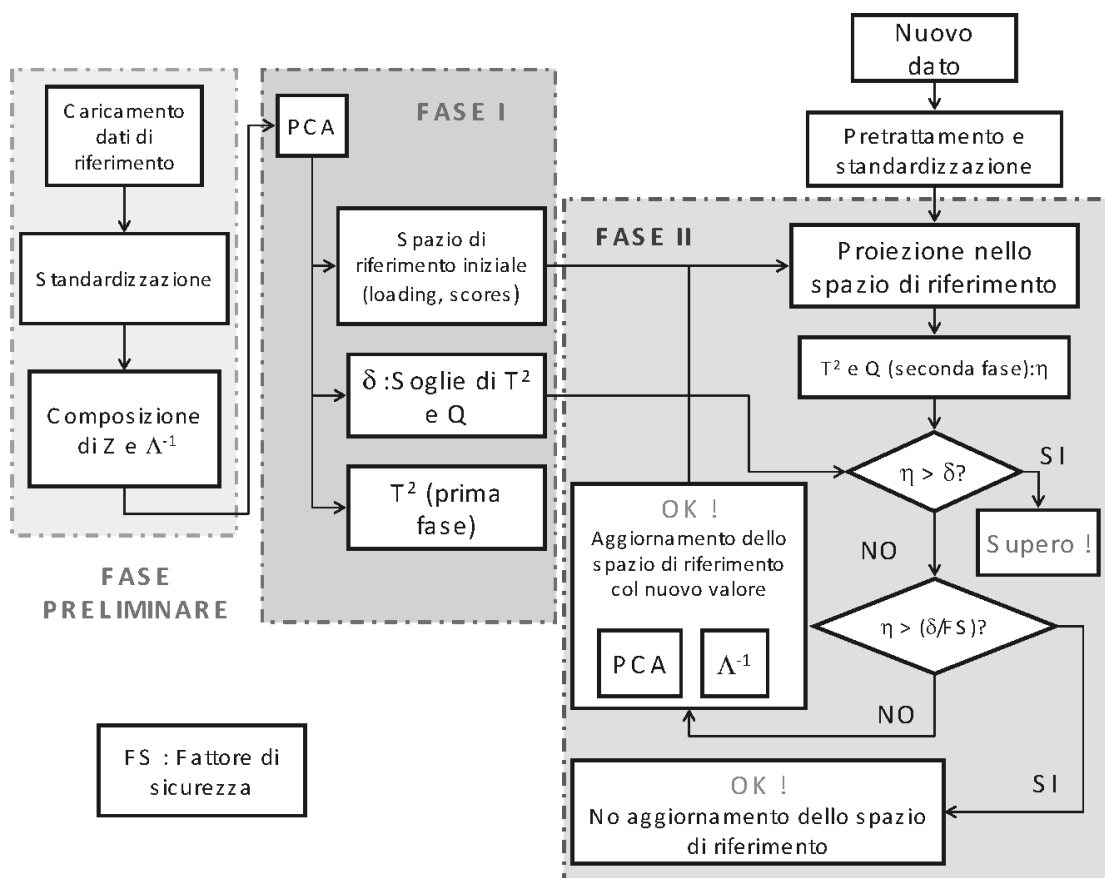


Fig. 4 – Schema dell'algoritmo di controllo

3. Monitoraggio in tempo reale e aggiornamento della finestra di riferimento (fase II).

Nella prima parte il software carica i dati della finestra scelta come riferimento, li standardizza, compone la matrice su cui applicare la ACP e ricava la matrice Λ^{-1} necessaria per le operazioni successive. Nel secondo blocco si ha l'utilizzo del toolkit Advanced Signal Processing che fornisce in uscita i *loading*, gli *score* e le varianze di ogni componente principale del periodo di riferimento. Con questi valori è possibile calcolare, una volta deciso il numero di componenti da trattene-re, la statistica T^2 della prima fase e le soglie di controllo per T^2 e Q , tramite le equazioni viste precedentemente. I limiti ricavati nella fase 1 e la costruzione dello spazio di riferimento iniziale consentono di effettuare il successivo monitoraggio in tempo reale (fase 2). Infine nella fase 3 si costruisce la finestra mobile necessaria a rendere dinamica la ACP: ogni nuovo set di dati acquisito viene standardizzato e proiettato nello spazio di riferimento attuale in modo tale da ricavare i valori degli *score* e quindi i valori delle due statistiche. A questo punto confrontando il valore delle statistiche con le soglie precedentemente ricavate si è in grado di valutare la presenza di anomalie. Nel caso di valori inferiori alle soglie è stato realizzato un sistema di aggiornamento del periodo di riferimento in modo tale da includere l'ultimo set di dati nella finestra ed eliminare il set di dati più vecchio presente nella matrice se il valore delle statistiche è inferiore alle soglie divise per un fattore di sicurezza empirico, in caso contrario il periodo di riferimento rimane invariato.

L'aggiornamento, che comporta una nuova applicazione della ACP, dipende dal confronto tra valore istantaneo e soglia di aggiornamento della statistica: se i dati sono inferiori alla soglia, si effettua l'aggiornamento e si riapplica l'ACP ai dati aggiornati, in caso contrario (se il valore supera la soglia di aggiornamento) tutte le variabili vengono analizzate, ma non contribuiscono all'aggiornamento della base dei dati "buoni".

4. RISULTATI

I sensori presi in esame sono quelli riguardanti l'azoto e precisamente:

- 1) NH_4 in uscita dalla vasca di ossidazione;
- 2) NO_x in uscita dalla vasca anossica ($\text{NO}_{x \text{ DEN}}$);
- 3) NO_x in uscita dalla vasca di ossidazione ($\text{NO}_{x \text{ OX}}$).

Le elaborazioni qui presentate sono basate sui dati provenienti dall'archivio storico nel periodo Agosto 2007 – Maggio 2008 e sono state eseguite in modalità *real-time*, nel senso che i dati venivano elaborati via via che venivano acquisiti dal campo. Si è assunto come finestra di riferimento un campione di 19830 dati, che copre approssimativamente un periodo di 14 giorni, acquisiti nel mese di Dicembre 2007. Questo campione è stato realizzato concatenando due diversi periodi dal 1 al 7 Dicembre circa e dal 13 al 18 Dicembre circa (Fig. 5).

Applicando l'ACP a questi dati di riferimento si è scelto un modello che conserva due componenti principali che spiega circa l'84 % della variabilità osservata, come mostra la Tabella 1. Questo modello viene assunto come riferimento per l'identificazione di anomalie nel sistema.

Dai dati di riferimento di Fig. 5 si sono anche calcolati i valori limite delle statistiche T^2 e Q , ponendo α pari a 0.05 e ottenendo i seguenti valori di soglia $T_{lim}^2 = 6.0$ e $Q_{lim} = 3.2$. Sono state effettuate due tipi d'analisi: una nel breve periodo allo

Tab. 1– Varianza spiegata da ogni componente di ACP

Componente principale	Autovalore	Percentuale di varianza	Percentuale cumulativa
1	1.685	56.2	56.2
2	0.832	27.7	83.9
3	0.484	16.1	100.0

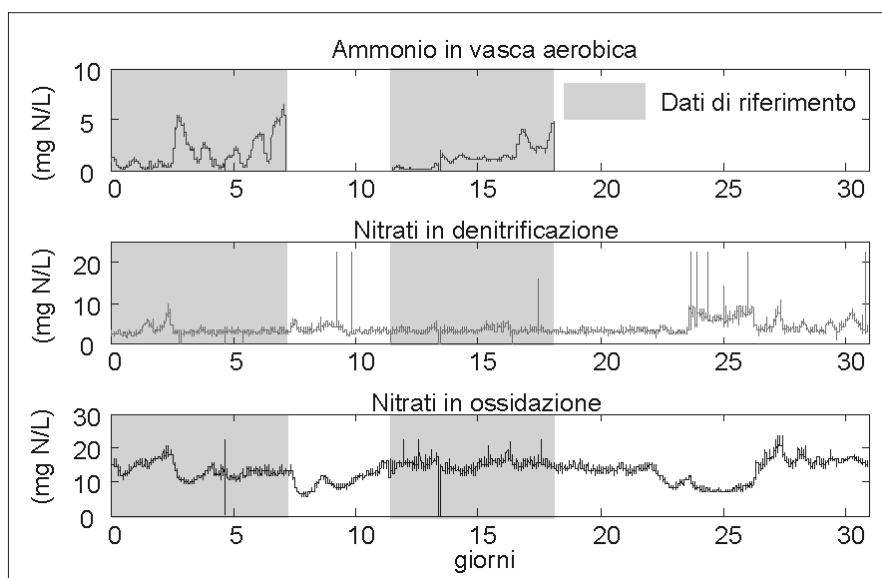


Fig. 5 – Scelta dei dati di riferimento (riquadri grigi) ricavati dai dati rilevati nel mese di Dicembre 2007 e ritenuti rappresentativi di buon funzionamento

scopo di verificare che l'algoritmo rilevasse ogni singolo malfunzionamento, ed una di lungo periodo, con l'obiettivo di controllare se il sistema diagnostico rimanesse sempre efficiente nonostante la mancata applicazione delle correzioni in seguito al susseguirsi di malfunzionamenti nel tempo. Infatti lavorando su dati d'archivio, dove non è possibile provvedere ad azioni correttive alle anomalie, è importante sincerarsi che la mancata applicazione di correzioni nel tempo non porti ad una diagnosi di guasto che peggiori nel tempo.

4.1 Analisi di breve periodo

Nel caso di buon funzionamento dei sensori le due statistiche T^2 e Q assumono valori molto bassi, inferiori alla metà del valore delle soglie, e rimangono tali finché perdura la bontà della misura, come mostra la Fig. 6.

I grafici degli *score* per i periodi monitorati di buon funzionamento confermano i risultati della teoria della ACP in quanto mostrano che tutti i punti sono vicini o interni all'ellissi di confidenza (Fig. 7), che dà un'idea della regione di corretto funzionamento delle sonde. L'ellisse di confidenza è stata ricavata ponendo α pari a 0.01.

4.1a Guasto di un sensore

Si esaminano ora alcuni casi particolari di diagnostica di guasto. In Fig. 8 è riportato un esempio di malfunzionamento della sonda dell'ammonio nel Dicembre 2007: l'algoritmo percepisce il netto salto che si verifica in caso di guasto ed entrambe le statistiche assumono immediatamente valori molto al di sopra della soglia. Col passare del tempo i valori diminuiscono, ma non si torna alla situazione antecedente il guasto se il sensore non fornisce più misure plausibili e la diagnosi di guasto viene mantenuta. Inoltre, il *contribution plot* riportato in basso nel grafico mostra come l'algoritmo

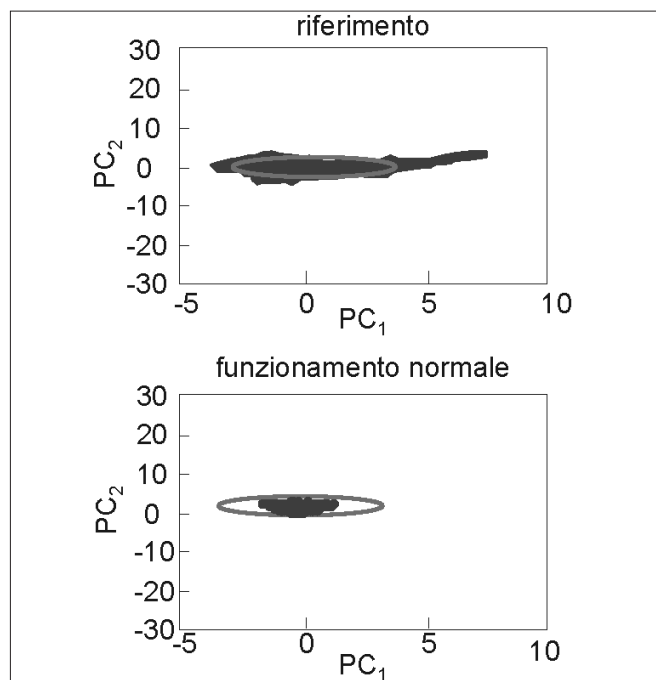


Fig. 7 – Grafico degli *score* del periodo di riferimento iniziale e di alcuni giorni di funzionamento normale (28 – 30 Novembre 2007) proiettati nello spazio delle prime due componenti principali ($\alpha=0.01$)

individui correttamente il sensore guasto, in quanto esso porta il contributo maggiore alle statistiche.

4.1b Spikes

Si indica con questo termine uno o più picchi di brevissima durata con valori non compatibili con il normale andamento

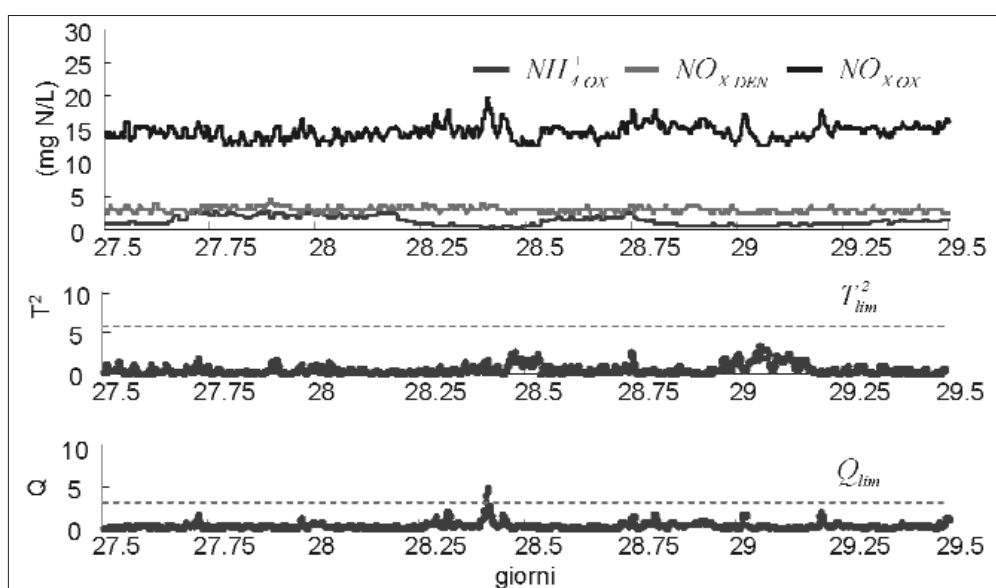


Fig. 6 – Andamento delle concentrazioni di NH_4 e NO_x nel periodo 28 – 30 Novembre 2007 e relativi grafici di controllo T^2 e Q . Gli indicatori T^2 e Q mostrano il buon funzionamento dei sensori, salvo un piccolo picco di Q_{lim} di scarso significato

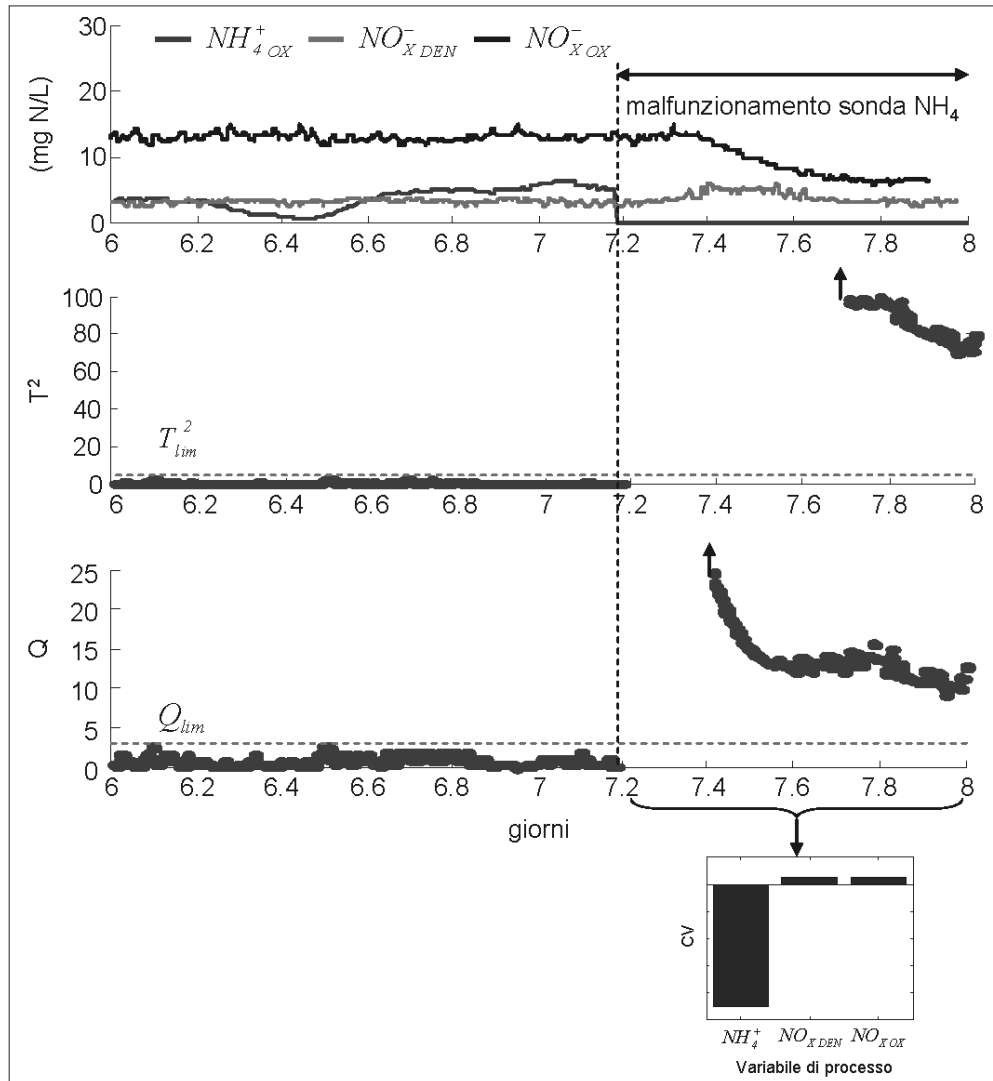


Fig. 8 – Diagnosi di guasto del sensore di ammonio: andamento delle concentrazioni e statistiche T^2 e Q per i giorni 7–8 Dicembre 2007. La diagnosi è confermata dal contribution plot (in basso) che indica il sensore di ammonio come responsabile del malfunzionamento

delle misure correnti. In Fig. 9 si mostrano due *spikes* provenienti dalla sonda dei nitrati in vasca di denitrificazione che temporaneamente assume valori negativi per banali motivi di sbilanciamento elettrico della linea di trasmissione; il sistema diagnostico rileva questo tipo di malfunzionamento attraverso l'utilizzo di carte di controllo basate sui superamenti delle soglie; infatti gli *spikes* sono difficilmente rilevabili dai comuni grafici del T^2 e Q perché producono valori delle due statistiche superiori di almeno due ordini di grandezza rispetto ai dati ordinari. Al contrario, l'utilizzo dei *contribution plots* permette di individuare subito quale sia il sensore che provoca l'anomalia.

4.1c Anomalia di processo

Il sistema diagnostico FDI, utilizzando la ACP, permette di esercitare anche la funzione di controllo statistico di processo, cioè di segnalare situazioni anomale se vengono rilevate misure molto diverse da quelle nello spazio di riferi-

mento anche se ciò non è causato dal malfunzionamento di un sensore. Al fine di chiarire questo aspetto si riporta come esempio un periodo di misure riguardante Gennaio 2008 (Fig. 10) nella quale si ha sia un'anomalia di processo dovuta ad una concentrazione esageratamente bassa degli NO_x nella vasca di ossidazione (primo evento a sinistra), sia un malfunzionamento del sensore di ammonio (secondo evento a destra).

La statistica T^2 in questo caso fornisce un contributo significativo in quanto per come è costruita coglie la variazione delle componenti all'interno del modello di riferimento (spazio di processo). Di conseguenza l'aumento del valore di questa statistica può non essere necessariamente causata da un guasto che metta fuori uso il sensore, ma anche ad una situazione non presente nello spazio di processo ritenuto "normale". Perciò ad un aumento della statistica T^2 corrisponde un allontanamento dal riferimento voluto. Nella figura 15 si notano chiaramente due superamenti della soglia: il secondo

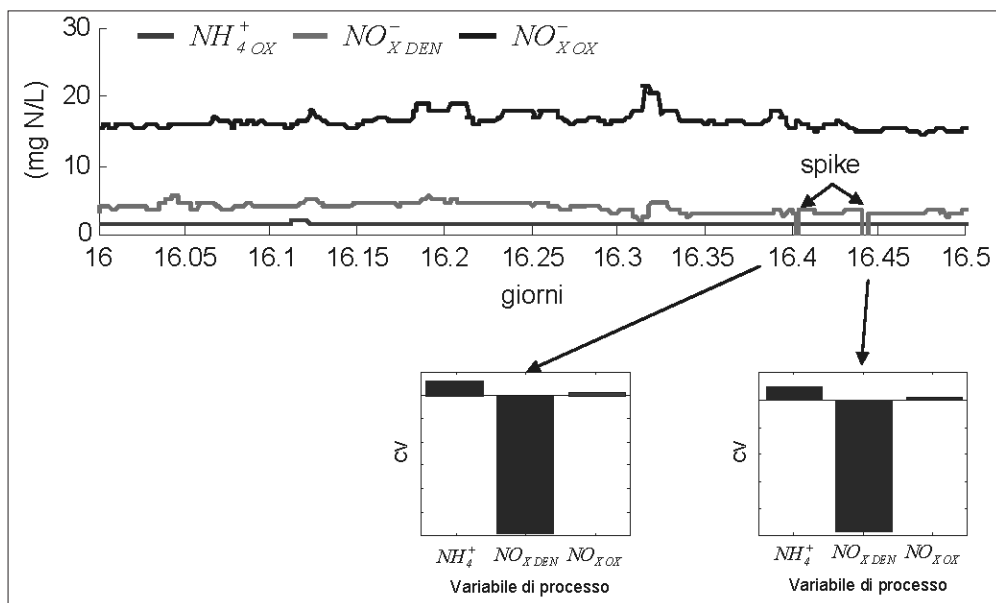


Fig. 9 – Analisi di due spikes della sonda nitrati in denitrificazione grazie ai contribution plots, che individuano il sensore che ha causato il malfunzionamento transitorio. In questo caso le statistiche T^2 e Q non forniscono informazioni significative

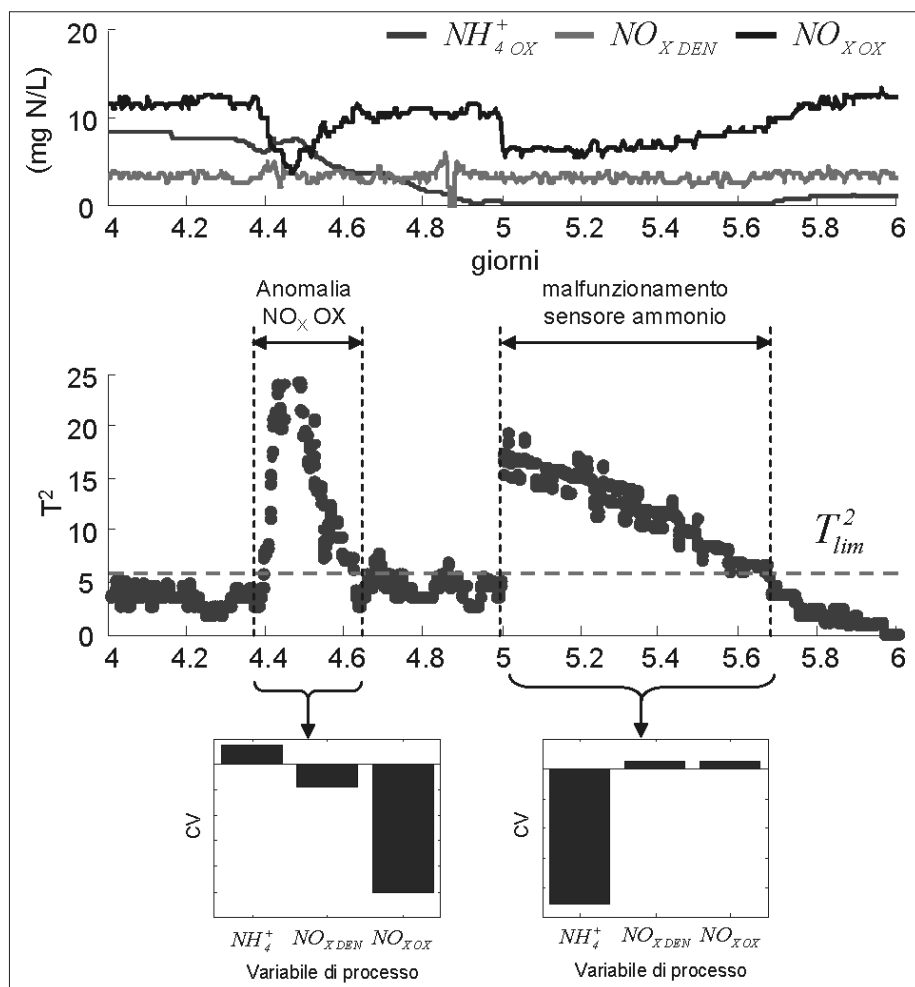


Fig. 10 – Andamento delle concentrazioni e del T^2 per i giorno 5-6 Gennaio 2008. I contribution plots in basso sono fondamentali per discriminare le cause dei malfunzionamenti rilevati

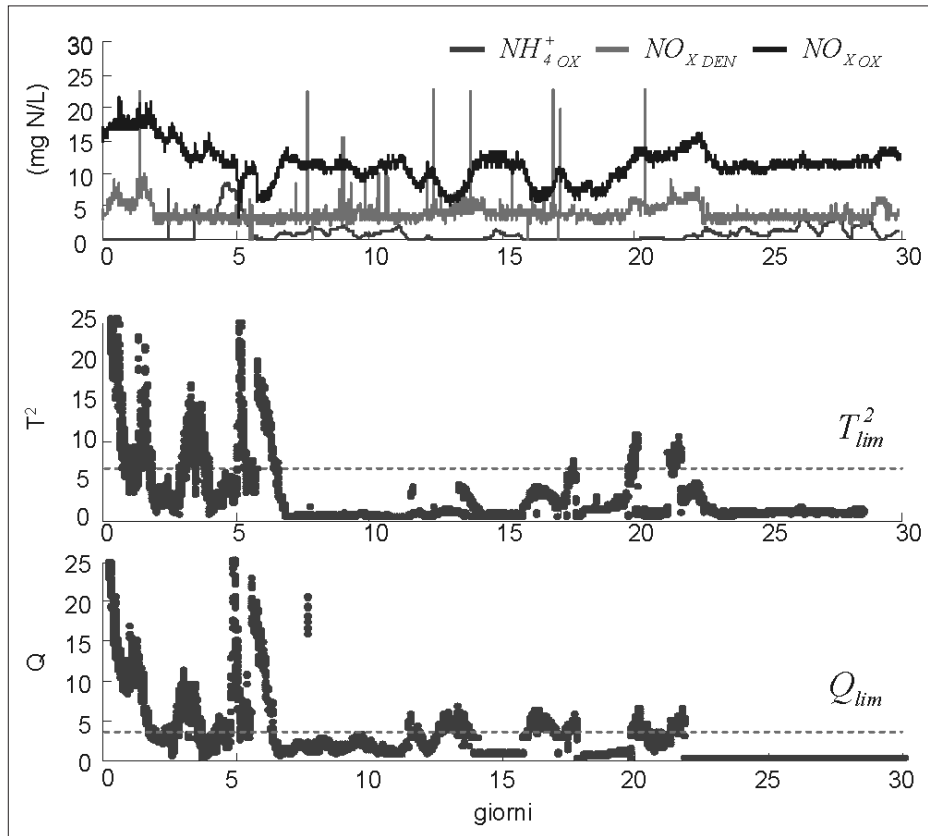


Fig. 11– Esempio di monitoraggio di lungo periodo: concentrazione delle varie specie azotate e relative statistiche T^2 e Q per il mese di Gennaio 2008

durante il giorno 6 è dovuto ad un guasto del sensore dell'ammonio, ma per il primo caso non è possibile individuare un guasto di alcun sensore esaminando le serie temporali nel giorno 4. Infatti il primo picco nel grafico del T^2 è dovuto alla notevole diminuzione della concentrazione dei nitrati in vasca di ossidazione e alla stabilità degli altri due segnali. In questo caso l'algoritmo FDI ci segnala grazie alla ACP che sta accendendo qualcosa che sta allontanando il processo dal proprio spazio di riferimento. L'analisi dei *contribution plots*, mostrati in basso nella Fig. 10 avvalorava quanto detto poiché il contributo principale al T^2 è dato dal sensore dei nitrati in vasca di ossidazione.

4.2 Analisi di lungo periodo

Si è esaminato il comportamento del software sul lungo periodo, prendendo in considerazione un intervallo di tempo di un paio di mesi in modo tale da verificare la risposta dell'algoritmo nel caso in cui non venga effettuata alcuna azione di manutenzione dopo un allarme. Anche in questo caso il funzionamento dell'algoritmo è buono, anche se leggermente peggiore, in quanto riesce sì ad individuare tutte le tipologie di anomalia già illustrate nel breve periodo, ma la rilevazione è meno efficiente se basata sulle statistiche T^2 e Q (Fig. 11) mentre il grafico comparativo degli *scores* (Fig. 12) mostra chiaramente l'allontanamento del processo dalle condizioni normali di funzionamento.

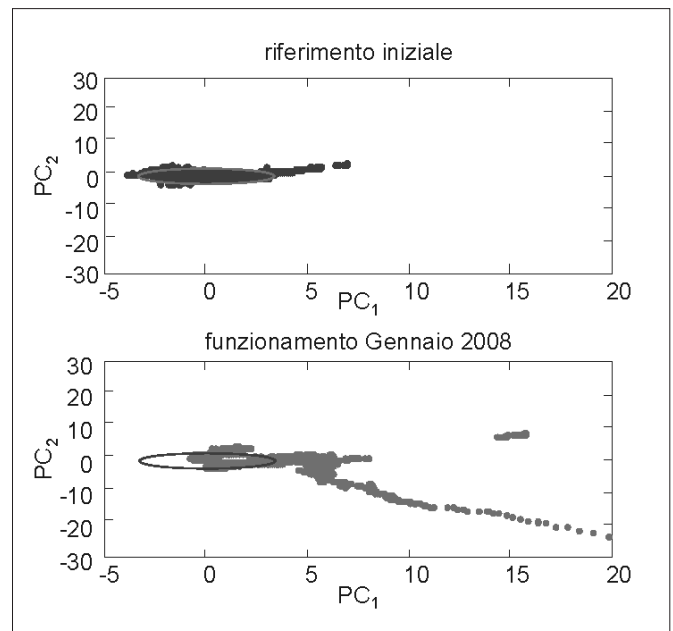


Fig. 12 – Grafico degli *scores* del periodo di riferimento iniziale e del mese di Gennaio 2008, durante il quale si nota un notevole allontanamento dal funzionamento normale (le ellissi di confidenza si riferiscono al livello $\alpha=0.01$)

Tab. 2 – Valutazione delle prestazioni sul lungo periodo: confronto fra malfunzionamenti effettivi e rilevati

			Fermi sonda	Spike	Outliers	Anomalie di processo	
2007	Agosto	Malfunzionamenti effettivamente osservati	2	4	2	1	
		Malfunzionamenti rilevati dall'algoritmo	Finestra corta	2	4	2	1
			Finestra lunga	2	3	1	1
	Settembre	Malfunzionamenti effettivamente osservati	4	0	0	1	
		Malfunzionamenti rilevati dall'algoritmo	Finestra corta	4	0	0	1
			Finestra lunga	3	0	0	1
	Ottobre	Malfunzionamenti effettivamente osservati	4	6	2	3	
		Malfunzionamenti rilevati dall'algoritmo	Finestra corta	4	6	2	3
			Finestra lunga	4	4	2	2
	Novembre	Malfunzionamenti effettivamente osservati	1	4	1	0	
		Malfunzionamenti rilevati dall'algoritmo	Finestra corta	1	4	1	0
			Finestra lunga	1	3	1	0
Dicembre	Malfunzionamenti effettivamente osservati	2	20	7	2		
	Malfunzionamenti rilevati dall'algoritmo	Finestra corta	2	20	7	2	
		Finestra lunga	2	15	6	1	
2008	Gennaio	Malfunzionamenti effettivamente osservati	6	20	10	1	
		Malfunzionamenti rilevati dall'algoritmo	Finestra corta	6	20	10	1
			Finestra lunga	5	16	9	1
	Marzo	Malfunzionamenti effettivamente osservati	2	55	2	1	
		Malfunzionamenti rilevati dall'algoritmo	Finestra corta	2	55	2	1
			Finestra lunga	2	46	1	1
	Aprile	Malfunzionamenti effettivamente osservati	3	4	1	0	
		Malfunzionamenti rilevati dall'algoritmo	Finestra corta	3	4	1	0
			Finestra lunga	2	3	1	0
	Maggio	Malfunzionamenti effettivamente osservati	5	0	1	0	
		Malfunzionamenti rilevati dall'algoritmo	Finestra corta	5	0	1	0
			Finestra lunga	4	0	1	0

4.3 Valutazione complessiva delle prestazioni

Le prestazioni dell'algoritmo sono riepilogate nella Tabella 2, dove per ciascun mese si riporta il numero di malfunzionamenti effettivamente riscontrati e le rilevazioni dell'algoritmo con i due diversi orizzonti di funzionamento. Utilizzando la finestra corta si analizza il mese corrente suddividendolo in parti contenenti pochi malfunzionamenti allo scopo di verificare l'efficacia del sistema diagnostico per ogni singolo guasto, mentre tramite la finestra lunga si verifica se i mancati interventi di manutenzione influenzino il comportamento dell'algoritmo nel tempo.

Inoltre si deve precisare che in presenza di un guasto prolungato l'algoritmo è costretto a non inserire i dati acquisiti più recentemente fra quelli considerati di "buon funzionamento"; così facendo si ha la perdita di una notevole quantità di dati potenzialmente validi e di conseguenza non si ha un aggiornamento efficiente della cosiddetta "memoria a breve termine" del processo con cui effettuare il controllo statistico di processo. In altre parole, il funzionamento *off-line* qui simu-

lato rappresenta il caso peggiore in cui l'algoritmo FDI può trovarsi ad operare.

Invece, dall'analisi nel breve periodo, dove il numero di anomalie e malfunzionamenti analizzati per volta è limitato, si può rilevare che se il sistema dei sensori è accompagnato da una manutenzione regolare o se in seguito all'allarme dato dal sistema FDI si apportano le necessarie correzioni, le prestazioni dell'algoritmo sono decisamente migliori.

5. CONCLUSIONI

In questo lavoro si è descritto lo sviluppo e l'applicazione di un sistema di analisi automatica dei guasti (FDI) basato su un'estensione dinamica dell'analisi alle componenti principali (ACP). Dopo aver brevemente richiamato la struttura dell'impianto su cui il sistema è stato impiegato e la teoria alla base della rilevazione dei guasti, si sono esaminate le prestazioni del sistema utilizzando un insieme di dati rilevati con cadenza di un minuto compresi fra Agosto 2007 e

Maggio 2008 presso l'impianto di depurazione di Pagnana (Empoli) gestito da Acque SpA. Le condizioni in cui il sistema è stato fatto funzionare sono le più severe, in quanto operando in modalità *off-line*, l'eventuale rilevazione del guasto non provocava un intervento di ripristino. In tal modo i dati successivi non potevano essere utilizzati per ricondizionare l'algoritmo rinforzando la base di dati di "buon funzionamento". Il sistema è in grado di effettuare due tipi di analisi: di breve e di lungo periodo. Nel primo caso, l'orizzonte corto (un mese) porta ad un riconoscimento praticamente certo del guasto, mentre sull'orizzonte lungo si ha un lieve peggioramento delle prestazioni, principalmente dovuto alla non rimozione del guasto e conseguente mancanza di aggiornamento della base di dati. Una volta operante in modalità *on-line* ci si può attendere che le prestazioni migliorino. Il sistema è inoltre in grado di distinguere diverse modalità di malfunzionamento, dal guasto al sensore, al disturbo momentaneo (*spike*) alla presenza di dati anomali. La versatilità del sistema è dovuta all'impiego di più statistiche, non solamente gli indicatori T^2 e Q , ma anche i *contribution plots*, che si rivelano particolarmente utili nel funzionamento con orizzonte lungo.

Nell'applicazione finale il sistema è destinato a operare *in linea* nel compact Field Point usato per il controllo dell'impianto, aggiungendo alle attuali funzionalità una serie di allarmi che indichino all'operatore la natura del malfunzionamento.

RINGRAZIAMENTI

Nello svolgimento di questa ricerca si riconosce il contributo e la collaborazione di Acque SpA ed Acque Ingegneria srl per il supporto prestato attraverso un'apposita convenzione di ricerca, lo svolgimento di un periodo di tirocinio e l'accesso alle strutture aziendali.

BIBLIOGRAFIA

- Aguado D., Rosen C.**, 2007. Multivariate statistical monitoring of continuous wastewater treatment plants. *Engineering Applications of Artificial Intelligence*, **21**, 1080–1091.
- Detroja K.P., Gudi R.D., Patwardhan S.C.**, 2007. Plant-wide detection and diagnosis using correspondence analysis. *Control Engineering Practice*, **15**, 1468–1483.
- Dillon W.R., Goldstein M.**, 1984. *Multivariate Analysis: Methods and Applications*, J. Wiley and Sons, New York, USA, pp. 587.
- Ku W., Storer R.H., Georgakis C.**, 1995. Disturbance detection and isolation by dynamic principal component analysis. *Chemometrics and Intelligent Laboratory Systems*, **30**, 179–196.
- Lee C., Choi S.W., Lee I.B.**, 2004. Sensor fault identification based on time-lagged ACP in dynamic processes. *Chemometrics and Intelligent Laboratory Systems*, **70**, 165–178.
- Lee C., Choi S.W., Lee I.B.**, 2006. Sensor fault diagnosis in a wastewater treatment process. *Water Science & Technology*, **53** (1), 251–257.
- Lee J.M., Yoo C.K., Lee I.B.**, 2004. Statistical process monitoring with independent component analysis. *Journal of Process Control*, **14**, 467–485.
- Li R., Rong G.**, 2006. Dynamic Process Fault Isolation by Partial DACP. *Chemical and Biochemical Engineering Quarterly*, **20** (1), 69–77.
- Mina J., Verde C.**, 2007. Fault Detection for large scale systems using Dynamic Principal Components Analysis with adaptation. *International Journal of Computers, Communications & Control*, **2** (2), 185–194.
- Oakland, J.S., Porter L.J.**, 1995. *Total quality management*. Oxford Butterworth Heinemann, pp. 370.
- Rosen C.**, 1998. *Monitoring wastewater treatment systems*, Lic. thesis, Department of Industrial Electrical Engineering and Automation, University of Lund (S), pp. 190.
- Rosen C.**, 2001. *A chemometric approach to process monitoring and control with applications to wastewater treatment operation*. Doctoral Dissertation in Industrial Automation, Department of Industrial Electrical Engineering and Automation, University of Lund (S), pp. 278.
- Schraa O., Tole B., Copp J.B.**, 2006. Fault detection for control of wastewater treatment plants. *Water Science & Technology*, **53** (4–5), 375–382.

CURRICULA

Francesco Baggiani – Si è laureato con lode in Ingegneria per l'Ambiente ed il Territorio presso l'Università di Firenze nell'anno 2008 discutendo la tesi dal titolo "Analisi di guasto in tempo reale per un processo di depurazione biologica". Durante lo svolgimento della tesi ha effettuato attività di tirocinio presso Acque SpA. Attualmente collabora con il Dipartimento di Sistemi e Informatica per ricerche sulla diagnosi dei guasti negli impianti di depurazione. Svolge anche attività di libero professionista nello stesso settore.

Stefano Marsili-Libelli – Laureato con lode in Ingegneria Elettronica presso l'Università di Pisa, è attualmente Professore ordinario di Automatica (ING/INF04) ed afferisce al Dipartimento di Sistemi e Informatica dell'Università di Firenze dove insegna Modellistica dei Sistemi Ambientali e Automazione dei Processi di Trattamento delle Acque. Gli interessi di ricerca comprendono la modellistica dei ecosistemi acquatici e dei processi di trattamento delle acque. Partecipa a programmi finanziati dalla Comunità Europea, fra cui i COST 682 e COST 624 sull'automazione degli impianti di depurazione. È stato consulente della Regione Toscana e del Tribunale di Firenze in materia di inquinamento ambientale, membro del comitato tecnico dell'Autorità di Bacino del fiume Arno. Ha collaborato alla stesura del Piano di Tutela delle Acque (D.Lgs 152/99) della Regione Emilia-Romagna per la parte informatica e modellistica. È autore di oltre 150 pubblicazioni scientifiche recensite su riviste e congressi internazionali. Fa parte del comitato editoriale della rivista internazionale *ISI Environmental Modeling & Software*. Ha fatto parte di diversi dottorati di ricerca, fra cui il Dottorato in Ingegneria dell'Automazione (consorzio delle Università di Firenze, Bologna e Padova) e del Dottorato di Ricerca in Idrodinamica e Modellistica Ambientale, con sede amministrativa presso l'Università di Padova. Attualmente fa parte del collegio dei docenti del dottorato in Ingegneria Sanitaria della seconda Università di Roma Tor Vergata.