Università degli Studi di Cagliari

DOTTORATO DI RICERCA IN DIFESA E CONSERVAZIONE DEL SUOLO, VULNERABILITÀ AMBIENTALE E PROTEZIONE IDROGEOLOGICA

Ciclo XXIII

TESI IN COTUTELA con l'Université de Strasbourg

DOCTORAT SCIENCES DE LA TERRE DE L'UNIVERS ET DE L'ENVIRONNEMENT DE STRASBOURG (STUES)

APPLICATION OF ARTIFICIAL NEURAL NETWORKS IN HYDROGEOLOGY : IDENTIFICATION OF UNKNOWN POLLUTION SOURCES IN CONTAMINATED AQUIFERS

GEO/05

Presentata da:	Ing. Maria Laura Foddis
Coordinatore Dottorato	Prof. Felice Di Gregorio
Tutor	Prof. Ing.Gabriele Uras Prof. Philippe Ackerer

Esame finale anno accademico 2009 - 2010

Université de Strasbourg

THESE

Présentée en cotutelle avec l'Université de Cagliari
en vue de l'obtention du grade de
Docteur de l'Université de Strasbourg
et Docteur de l'Université de Cagliari

DOCTORAT EN SCIENCES DE LA TERRE DE L'UNIVERS ET DE L'ENVIRONNEMENT DE STRASBOURG (STUES)

par

Maria Laura Foddis

APPLICATION OF ARTIFICIAL NEURAL NETWORKS IN HYDROGEOLOGY : IDENTIFICATION OF UNKNOWN POLLUTION SOURCES IN CONTAMINATED AQUIFERS

Soutenue le 25 mars 2011 devant le jury constitué de :

Directeur de Thèse	Philippe ACKERER
Codirecteur de Thèse	Gabriele URAS
Rapporteur interne	Gerhard SCHÄFER
Rapporteur externe	Alessandra FANNI
Rapporteur externe	Linda SEE

ACKNOWLEDGMENTS

It was a pleasure for me to work with all the wonderful people in the *Dipartimento di Ingegneria del Territorio* and the *Dipartimento di Ingegneria Elettrica ed Elettronica* of the University of Cagliari and the *Laboratoire d'Hydrologie et de Géochimie de Strasbourg (LHyGeS)* of the University of Strasbourg. I enjoyed the atmosphere, their friendship, and their support.

I really appreciate the continuous support and useful inputs of my two thesis supervisors Gabriele URAS and Philippe ACKERER: their ideas had a major influence on this thesis. I enjoyed their interest in my research as well as the fruitful discussions. In particular, I would like to thank Philippe ACKERER for reviewing and proof-reading my thesis. I am happy to have such a supportive supervisor.

I wish to express my gratitude to Augusto MONTISCI - *Dipartimento di Ingegneria Elettrica ed Elettronica* of the University of Cagliari - that spent a lot of time helping and supporting me designing artificial neural networks. Without his precious contribution, I would not have achieved all the results presented in this work.

It is an honor for me that Alessandra FANNI - *Dipartimento di Ingegneria Elettrica ed Elettronica* of the University of Cagliari - and Gerhard SCHÄFER - *Laboratoire d'Hydrologie et de Géochimie de Strasbourg (LHyGeS)* of the University of Strasbourg – took part to the jury during the PhD defence.

I am grateful that Linda SEE - *School of Geography* of the University of Leeds - accepted to be external jury member for awarding the PhD doctor diploma of the University of Strasbourg and external Referee for awarding the label of *Doctor Europeaus* of the University of Cagliari.

I also want to thank Robert MOSE - *Ecole Nationale du Génie de l'Eau et de l'Environnement* of the University of Strasbourg - for accepting to be External Referee for awarding the label of *Doctor Europeaus* of the University of Cagliari.

A special thanks to my husband Martin HEIBEL for proof-reading the entire thesis. I know it has been a very hard work.

Last but not least, I wish to thank my family an my friends who have always supported me in this challenging and demanding work.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS

ACP : Analyse en Composantes Principales

ANN : Artificial Neural Networks

ATSDR: Agency for Toxic Substances and Disease Registry

BC(S): Boundary Condition(s)

BRGM : Bureau de Recherche Géologique et Minière

CCl_3: trichloromethyl radical

CCl_4: carbon tetrachloride- tétrachlorure de carbone-tetracloruro di carbonio

CEREG: Centre d'Etudes et de Recherches Eco-Géographiques

CIENPPA: Commission Interministérielle d'Etude de la Nappe Phréatique de la Plaine d'Alsace

CO_2: carbon dioxide

CPHF: California Public Health Foundation

DCM: DiChloroMethane

DDA: Direction Départmentale de l'Agriculture

DG: Discontinuous Galerkin

DIREN: Direction Régionale de l'ENvironnement

DNAPL(S): Dense Non-Aqueous Phase Liquid(S)

EBP: Error Back Propagatio

EFMH: Eléments Finis Mixtes Hybrides

EPA: Environmental Protection Agency

FD: Finite Difference

FV: Finite Volume

GAC: Granular Activated Carbon

HSG: Health and Safety Guide

HTMP: Hydrodynamique et Transfert en Milieu Poreux

IMFS: Institut de Mécaniques des Fluids et Solids

IPCS: International Programme on Chemical Safety

LM: Levenberg-Marquardt

LOO: Leave One Out

MCL: Maximum Contaminant Level

MCLG: Maximum Contaminant Level Goals

MLP: Multi Layer Perceptron

MHFE: Mixed Hybrid Finite Element

MSE: Mean Square Error

NIOSH: National Institute for Occupational Safety and Health

PE : Processing Element

PIREN: Programme Interdisciplinaire de Recherche en ENvironnement du CNRS

RNA: Reti Neurali Artificiali - Réseaux de Neurones Artificiels

SEMA: Service des Eaux et des Milieux Aquatiques de la DIREN

SGAL: Service Géologique d'Alsace-Lorraine

T: duration of the pollution source activity

TRACES: Transport of RadioACtive Elements in the Subsurface

UNEP: United Nations Environment Programme

VOC: Volatile Organic Chemical

WHO: World Health Organization

X,Y: spatially coordinates

1D, 2D, 3D: one-, two-, three-dimensional

RIASSUNTO

Introduzione

Lo studio della vulnerabilità degli acquiferi rappresenta uno degli aspetti di più stretta attualità nell'ambito della protezione del territorio.

Il lavoro presentato, in questa tesi di ricerca, verte sullo studio della contaminazione delle risorse idriche sotterranee. Le acque sotterranee rappresentano la maggior parte delle riserve d'acqua potabile sulla terra. Generalmente la loro qualità è maggiore di quella delle acque superficiali, grazie alle proprietà filtranti del sottosuolo. Tuttavia, le acque sotterranee sono esposte a inquinamenti generati dall'uomo che rendono questa importante risorsa sempre più vulnerabile.

Quando le acque sotterranee sono inquinate, il ripristino della loro qualità e la rimozione degli inquinanti richiede tempi molto lunghi, ed è, talvolta, un processo praticamente irrealizzabile.

Nel campo della contaminazione delle risorse idriche sotterranee è opportuno sottolineare che, in alcuni casi, i fenomeni di inquinamento possono derivare da contaminazioni le cui origini provengono da eventi generati in luoghi e tempi diversi da quelli in cui si è riscontrata l'effettiva contaminazione. Tali situazioni rendono necessaria la ricerca di tecniche che permettano l'individuazione delle coordinate spazio-temporali di sorgenti contaminanti incognite.

In generale, l'individuazione e la delimitazione della fonte di un pennacchio contaminante appare di notevole importanza ai fini della definizione delle migliori politiche di gestione del sito contaminato e della pianificazione degli opportuni interventi di bonifica del sottosuolo.

La determinazione delle condizioni iniziali dell'inquinamento appare di notevole interesse anche ai fini dell'applicazione della Direttiva dell'Unione Europea 2004/35/CE in materia di responsabilità ambientale, di prevenzione e risarcimento dei danni ambientali basata sull'affermazione del principio secondo cui: "chi inquina paga". La Direttiva 2004/35/CE è stata recepita nella normativa italiana con il D. Lgs 152/2006 (Parte VI - "Norme in materia di tutela risarcitoria contro i danni all'ambiente") e nella normativa francese con la legge 2008 – 757 (Titre 1er – "Dispositions relatives a la prévention et a la réparation de certains dommages causes a l'environnement").

Alcune caratteristiche idrogeologiche e qualitative delle acque sotterranee, in molti casi, non sono direttamente misurabili e fisicamente devono essere valutate in funzione di altri parametri direttamente misurabili. Il problema di determinare i parametri sconosciuti del modello è abitualmente definito come "problema inverso". La risoluzione del problema inverso per la modellazione del flusso e del trasporto dei contaminanti nelle acque sotterranee è l'obiettivo principale di questo lavoro di ricerca.

In riguardo alla risoluzione del problema inverso, in questo lavoro, ci si pone l'obiettivo di individuare una metodologia atta a identificare le caratteristiche spazio temporali di sorgenti di contaminazione incognite. In questo caso il problema inverso è risolto sulla base delle misurazioni di concentrazione del contaminante nei pozzi di monitoraggio situati nel dominio di interesse. Noto l'effetto generato da un determinato fenomeno si cerca di ricostruirne la causa.

La ricerca, si è sviluppata sui seguenti temi:

- Modellazione della contaminazione delle acque sotterranee mediante l'utilizzo di un software non commerciale per la modellazione del flusso e trasporto del contaminante nei mezzi porosi (TRACES - Transport or RadioActiver Elements in the Subsurface) sviluppato da Hoteit e Ackerer (2003). Il software TRACES combina gli elementi finiti ibridi misti ad elementi finiti discontinui per la risoluzione dello stato idrodinamico e del trasferimento di massa nei mezzi porosi.
- Modellazione delle relazioni causa-effetto nella contaminazione delle acque sotterranee mediante le Reti Neurali Artificiali (RNA). Le RNA sono realizzate utilizzando il Neural Network Toolbox di Matlab 7.1.
- Applicazione delle RNA per la risoluzione del problema inverso su due casi di contaminazione delle falde acquifere.

Negli ultimi decenni, le RNA sono diventate sempre più popolari come strumento di soluzione dei problemi e di previsione in molte discipline.

Una RNA è costituita dall'interconnessione di una serie di processori elementari chiamati neuroni (Perceptron), i quali, sono logicamente disposti in due o più strati ed interagiscono tra loro attraverso connessioni pesate. In particolare, mediante le reti di tipo Multi Layer Perceptron (MLP), utilizzate in questo lavoro, è possibile creare il modello di un sistema semplicemente sulla base di un opportuno insieme di coppie

input/output di pattern di esempi. Le caratteristiche delle RNA sviluppate in questo lavoro dipendono dalla natura dei problemi analizzati e non esistono linee guida teoriche per stabilire il modo migliore di operare. Il modello è specifico per il sistema in esame e non può essere costruito a priori.

L'addestramento di una RNA consiste in una regola di apprendimento che modifica i pesi delle connessioni in base alla differenza tra l'uscita calcolata della rete e il modello desiderato (target). L'obiettivo, di questa fase, è quello di rendere la RNA in grado di generalizzare le informazioni acquisite: fornendo alla rete un input, proveniente da un esempio non incluso nell'insieme dei pattern di addestramento, la rete deve determinare l'output corretto.

Per sviluppare la metodologia precedentemente citata, in questo lavoro, sono stati considerati due casi diversi di sorgenti di contaminazione puntuali e continue: un caso di inquinamento di un acquifero teorico e un caso reale (sorgente di inquinamento incognita dell'acquifero alsaziano in Francia).

Nel caso teorico, esaminato in questo studio, ci si è proposti di definire le coordinate spazio-temporali (X,Y,T) della sorgente contaminante incognita sulla base di poche misure di concentrazione del contaminante acquisite nei pozzi di monitoraggio in un certo tempo *t*.

Nel caso reale ci si è proposti di definire il comportamento di una sorgente di inquinamento incognita che, a causa di un incidente nel 1970, ha inquinato con tetracloruro di carbonio (CCl_4), uno dei più grandi acquiferi dell'Europa Occidentale: la falda acquifera Alsaziana (Regione Alsazia - Francia). Il comportamento della sorgente di inquinamento nel luogo dell'incidente non è noto. L'obiettivo è stato quello di individuare, sulla base delle simulazioni delle concentrazioni del contaminante nei pozzi, le caratteristiche della sorgente di inquinamento sconosciuta in termini di variazione spazio-temporale del tasso di iniezione del contaminante nel luogo dell'incidente.

Per risolvere il problema inverso, in entrambi i due casi in esame, è stata messa a punto una metodologia basata sull'applicazione della tecnologia delle RNA. In particolare sono state addestrate diverse RNA per la risoluzione del problema diretto. Lo scopo è stato quello di associare i pattern di ingresso (rappresentati dalle caratteristiche della sorgente) con i pattern di uscita (rappresentati dalle concentrazioni del contaminante nei pozzi di monitoraggio). Al fine di risolvere il problema inverso e di

identificare le caratteristiche incognite della sorgente di inquinamento, la RNA addestrata è stata invertita. Fissando il pattern di uscita si è potuto ricostruire l'ingresso corrispondente.

Ai fini della formazione del consistente pattern di esempi necessari all'addestramento, alla validazione ed al test della RNA, sono stati generati vari scenari idrogeologici caratterizzati da differenti sorgenti di contaminazione.

La tesi è organizzata come segue: nel primo capitolo sono riportate le equazioni classiche dell'idrodinamica e del trasporto dei contaminanti per convezione, diffusione e dispersione nel saturo. Nel secondo capitolo vengono descritte le reti neurali artificiali ed i modelli matematici alla base delle architetture utilizzate nella parte sperimentale. Il terzo capitolo è dedicato alla parte sperimentale inerente la contaminazione di un acquifero teorico: è descritto l'iter della metodologia adottato e risultati ottenuti. Il quarto capitolo è rivolto alla parte sperimentale ed ai risultati conseguiti per il caso reale di contaminazione dell'acquifero alsaziano. L'ultimo capitolo è dedicato alle conclusioni.

Capitolo 1: modellazione del flusso e del trasporto dei contaminanti

In questo capitolo è presentato il modello matematico utilizzato per la modellazione del flusso e trasporto dei contaminanti nelle falde acquifere.

Il moto delle acque attraverso i mezzi porosi, come possono essere assimilati gli acquiferi, viene governato, essenzialmente, da due parametri fondamentali: la conducibilità idraulica che rappresenta la quantità d'acqua che attraversa l'unità di superficie ad un gradiente unitario e dal gradiente idraulico che rappresenta la variazione del carico idraulico per unità di distanza percorsa.

La migrazione del contaminante negli acquiferi è espressa dalle equazioni di flusso e trasporto. L'equazione di flusso è data dalla legge di Darcy e l'equazione di continuità. La legge di Darcy descrive il moto di un fluido in un mezzo poroso. L'equazione di continuità esprime il principio della conservazione della massa di un fluido in movimento in un dato volume di controllo.

I meccanismi legati al trasporto di un contaminante in un mezzo poroso sono collegabili ai fenomeni di convezione, diffusione e dispersione.

Capitolo 2: Reti Neurali Artificiali

In questo capitolo viene descritto il concetto di rete neurale artificiale e delle sue più importanti caratteristiche che rendono questa tecnologia attraente nella ricerca idrogeologica. Sono dettagliate le componenti, la struttura e l'architettura delle RNA utilizzate in questo lavoro di ricerca. In particolare viene descritto l'iter per la progettazione di una rete Multilayer Perceptron e il problema dell'apprendimento. In fine, un paragrafo è dedicato al raffronto tra l'approccio classico e l'approccio neurale alla modellazione.

Le RNA sono dei modelli capaci di elaborare le informazioni in modo simile al cervello umano, in quanto, esse hanno la capacità di adattarsi a situazioni nuove utilizzando la conoscenza su situazioni simili. La caratteristica peculiare delle RNA risiede nella loro capacità di apprendimento, di generalizzazione e di approssimazione.

In generale l'approccio di studio ha visto il susseguirsi delle seguenti fasi:

- **Pre-processing:** analisi descrittiva dei dati, loro trasformazione e codifica a seconda delle esigenze del modello neurale, eventuale riduzione delle variabili attraverso diverse tecniche.
- **Selezione dell'architettura, delle regole di apprendimento e della funzione di errore:** scelta del numero di unità e degli strati nascosti sulla base dell'individuazione del tasso di errore più basso fra le varie architetture individuate. Questa fase è fondamentale in quanto ad essa è legata la capacità o meno di generalizzazione delle RNA elaborate. In questa fase vengono, anche, scelte della funzione di errore e l'algoritmo di ottimizzazione.
- **Addestramento:** fase relativa al problema della stima dei parametri, assegnato un certo pattern di esempi. In questo lavoro si è utilizzato esclusivamente l'addestramento supervisionato, dove l'apprendimento viene guidato dall'esterno fornendo alla rete l'insieme di pattern di esempi costituiti dall'ingresso e l'uscita del sistema. La rete, in questa fase, attribuisce le migliori coppie di pesi in base ai dati sperimentali in modo da ricostruire la relazione ingresso/uscita del sistema e ottenere partendo dall'input un output il più vicino possibile all'output esibito.
- **Verifica della stabilità:** fase necessaria per valutare il funzionamento della rete, la sua capacità di generalizzazione e, quindi, verificare la stabilità dei

risultati ottenuti. La stabilità della rete è influenzata da diversi fattori, quali: l'architettura selezionata, l'inizializzazione dei pesi, il campione utilizzato per l'apprendimento (pattern di esempi).

- **Post-processing:** fase dedicata alla valutazione dell'interpretabilità dei risultati ottenuti. I dati di partenza che a seconda delle esigenze del modello neurale erano stati ridotti e trasformati, in questa fase vengono riportati attraverso tecniche di inversione allo stato originario.

Capitolo 3: Applicazione delle RNA allo studio di un acquifero inquinato

Questo capitolo è dedicato alla definizione di una metodologia atta all'identificazione delle coordinate spazio-temporali di sorgenti contaminanti incognite sulla base di pochi valori della concentrazione del contaminante rilevati nei pozzi di monitoraggio in un certo tempo *t*. In particolare, l'obiettivo è quello di valutare la capacità delle Reti Neurali Artificiali di ricostruire le condizioni iniziali e, quindi, di localizzare, nello spazio e nel tempo, le sorgenti di fenomeni di contaminazione dell'acquifero esaminato.

In una prima fase, si è andati ad individuare un bacino idrogeologico da utilizzare come caso di prova e conseguentemente si è creato un modello di flusso e di trasporto dello stesso. È stata, quindi, studiata la contaminazione di un acquifero omogeneo isotropo confinato con un unico strato, per il quale si è simulato, mediante l'applicazione del software TRACES, l'andamento della piezometrica e delle isoconcentrazioni in una situazione stazionaria di partenza, in presenza di un pozzo di assorbimento. Il fenomeno è stato studiato supponendo che i parametri iniziali del modello non variassero nell'intervallo temporale della simulazione. Inoltre, è stata considerata l'ipotesi restrittiva che la contaminazione avvenisse con un solo generico contaminante a partire da una sola cella del modello a rilascio costante durante tutto il periodo della simulazione.

Il pattern di esempi è stato costruito considerando 40 sorgenti di contaminazione posizionate in diverse zone del dominio con tempi di attività variabili di 10, 20 e 30 anni. In totale sono stati considerati 120 scenari idrogeologici. I campioni ricavati dalla simulazione sono rappresentati dalle matrici delle concentrazioni del contaminante acquisite in 50 maglie distribuite uniformemente nel dominio.

Nello studio di un acquifero contaminato le curve di concentrazione del contaminante misurate nei pozzi di monitoraggio possono essere utilizzate per identificare le caratteristiche della sorgente contaminante che le ha determinate. Tuttavia, l'identificazione di sorgenti di contaminazione incognite diventa più difficile nel caso della mancanza di serie storiche delle curve di concentrazione del contaminante nel dominio. Per questi motivi in questo studio ci si è posti nella situazione peggiore, prendendo in considerazione, in particolare, il caso in cui si abbia la totale assenza di serie storiche della contaminazione, bensì un solo valore di concentrazione del contaminante per ogni pozzo di monitoraggio. In particolare si considera il caso della scoperta di una contaminazione per la prima volta in un dato dominio.

Sono stati presi in considerazione solo gli ultimi valori delle curve di concentrazione del contaminante ottenute attraverso le simulazioni per le 50 maglie del dominio. Queste 50 maglie corrispondevano a 50 ipotetici pozzi di monitoraggio. Una metodologia basata sull'applicazione delle reti neurali artificiali è stata sviluppata al fine di ridurre il numero di queste maglie e scegliere, tra le 50 posizioni delle maglie, quella più conveniente per l'eventuale realizzazione dei pozzi di monitoraggio. Alla fine della procedura, delle 50 maglie di partenza, sono state tenute solo 8 maglie.

A questo punto si è potuto valutare l'architettura della rete più adatta alla soluzione del problema dell'individuazione delle coordinate spazio-temporali delle sorgenti contaminanti incognite.

Le reti elaborate sono di tipo Multi Layer Perceptron, con un unico strato nascosto. L'algoritmo di apprendimento scelto è stato quello di Levenberg-Marquardt, in quanto produce i migliori risultati in termini di errore e di velocità di addestramento.

Il modello neurale prevedeva: 3 neuroni nello strato in ingresso, un unico strato nascosto di 8 neuroni e 8 neuroni nello strato in uscita. I neuroni nello strato in ingresso erano: 2 per le coordinate spaziali (X,Y) e 1 per il tempo di attività della sorgente e poteva avere valore di 10, 20 and 30 anni. I neuroni in uscita rappresentavano gli 8 valori della concentrazione del contaminante per gli 8 pozzi di monitoraggio.

I risultati ottenuti nelle prime elaborazioni neurali erano influenzati dalla non elevata numerosità del campione, per cui si è dimostrato necessario l'utilizzo della regola del "Leave One Out Cross Validation"(LOO). In questo caso il data set non viene più suddiviso in training, validation e test set, bensì un esempio alla volta viene escluso dal training per essere utilizzato come test set. In questo modo training set e test set sono

sempre diversi e a rotazione tutti i casi fanno parte del test. Sulla base di questa regola sono state addestrate 120 reti.

In questo caso la regola del LOO non è utilizzata per addestrare una rete che verrà utilizzata per risolvere un particolare caso, ma solo per stimare la capacità di estrapolare l'informazione delle 120 reti applicandola all'esempio, di volta in volta, rimasto fuori. Nel caso in cui si voglia individuare una nuova sorgente non compresa nelle 120, tutti i 120 pattern verranno usati per l'addestramento e la nuova sorgente fungerà da test. La metodologia sviluppata ci permette di avere la ragionevole presunzione che l'errore sul caso in esame non sarà maggiore di quello riscontrato nelle reti precedentemente addestrate col LOO.

Le 120 reti sono state inizialmente addestrate per la risoluzione del problema diretto. Una volta completata la fase di addestramento le reti sono state invertite per la risoluzione del problema inverso. Note le misurazioni delle concentrazioni del contaminante nei pozzi di monitoraggio le coordinate spazio-temporali delle sorgenti sono state determinate.

In generale, i risultati mostrano una buona capacità della rete nella localizzazione della sorgente. Nella maggior parte dei casi l'errore nell'identificazione delle coordinate spaziali è stato minore della dimensione di una maglia e l'errore massimo commesso è minore della dimensione di due maglie del dominio.

Meno soddisfacente è il risultato ottenuto per l'individuazione della durata dell'attività della sorgente contaminante. Solo il 76% delle reti sono state in grado di fornire una risposta al 100% corretta. In particolare, le reti in cui l'attività della sorgente era di 10 e 30 anni si è avuto rispettivamente su 40 casi un solo errore. Per le reti che consideravano un'attività della sorgente di 20 anni si sono avuti, su un totale di 40, 26 casi in cui in cui l'errore minimo era di 6 mesi e l'errore massimo era di 5 anni e 3 mesi.

Capitolo 4: RNA per la stima della sorgente inquinante dell'acquifero Alsaziano

Questo capitolo è dedicato all'elaborazione, sulla base dell'applicazione delle Reti Neurali Artificiali (RNA), di una metodologia innovativa per l'identificazione dell'evoluzione temporale della sorgente contaminate incognita dell'acquifero alsaziano (Benfeld -Francia).

Nella prima parte del capitolo è stata introdotta la problematica inerente la contaminazione della falda acquifera alsaziana e il modello di acquifero utilizzato per la realizzazione della RNA.

Nel 1970 a Benfeld (villaggio a 35km da Strasburgo), a seguito di un incidente, un autobotte riversò una quantità non nota di tetracloruro di carbonio (CCl_4) e contaminò parte del più grande acquifero dell'Europa occidentale. Noti i dati di concentrazione del contaminante nei pozzi di monitoraggio si è cercato di individuare le curve di concentrazione nei quattro fronti stratigrafici della sorgente contaminante incognita che li ha determinati.

Lo studio dell'inquinamento dell'acquifero alsaziano è di difficile realizzazione, in quanto, la quantità di tetracloruro di carbonio infiltrata al momento dell'incidente è incognita: una parte di volume di tetracloruro di carbonio contenuta nella cisterna del camion è stata recupera, una parte è evaporata e la restante si è infiltrata nell'insaturo per poi raggiungere l'acquifero. Il comportamento della sorgente nello spazio e nel tempo è incognito: per via della sua bassa solubilità in acqua il composto si comporta da sorgente di contaminazione a rilascio continuo, le cui dinamiche possono essere solo ipotizzate.

Oggigiorno, l'acquifero alsaziano, nonostante i continui trattamenti di bonifica, verte in una situazione di contaminazione ancora elevata ed essendo incognita la quantità del contaminate diffusa nel sottosuolo non si possono attuare stime sui tempi necessari alla decontaminazione. Conoscere le reali caratteristiche della sorgente è di fondamentale importanza ai fini della progettazione e della scelta delle più adatte tecniche di bonifica da mettere in opera.

Le fasi operative, in questa parte della ricerca, sono state le seguenti:

- realizzazione del pattern di esempi rappresentati dalla simulazione di diversi scenari di attività della sorgente sulla base di un modello numerico di flusso e trasporto della contaminazione dell'acquifero alsaziano derivante da un precedente lavoro (Aswed, 2008). Il pattern di esempi, in totale, è costituito da 104 diversi scenari della sorgente contaminante,
- pre-processing dei dati, implementazione e addestramento della RNA,
- inversione della RNA addestrata al fine della risoluzione del problema inverso.

A seguito di varie prove la tipologia di rete neurale scelta, in quanto meglio si adatta al problema in esame, è di tipo: Multi Layer Perceptron (MLP) implementata in back-propagation mediante l'algoritmo di Levenberg-Marquardt.

I campioni ottenuti dallo studio del modello erano costituiti da due matrici in cui le colonne erano riferite alle curve di concentrazione del contaminante rispettivamente nei 4 fronti stratigrafici della sorgente e in 45 pozzi di monitoraggio diversamente distribuiti nel dominio. In entrambe le matrici le righe rappresentavano il tempo. Queste matrici avevano dimensioni troppo elevate per poter essere successivamente elaborate tramite le RNA, richiedendo un numero troppo grande di esempi di input e, quindi, una rete di grandi dimensioni di difficile gestione, perdendo, in tal modo, una peculiarità dell'applicazione delle RNA consistente nella velocità di calcolo. Per questi motivi si è resa necessaria una fase di pre-processing finalizzata alla riduzione della dimensione dei dati, pratica, del resto, comunemente utilizzata nell'applicazione delle RNA.

I due gruppi di matrici di ingresso (sorgenti) e uscita (pozzi), durate la fase di pre-processing, sono stati considerati separatamente. Ad ogni matrice sono state applicate *Trasformate di Fourier del secondo ordine (FTT-2D)*, in questo modo, il dato veniva traslato dal dominio del tempo al dominio della frequenza. Le matrici, in seguito, sono state convertite in vettori, dove: la prima metà era rappresentata dalle componenti relative alle ampiezze e la seconda metà dalle componenti relative alle fasi. I vettori sono stati, in seguito, riuniti a formare un'unica matrice dove ogni colonna rappresentava un esempio. In totale si avevano 2 matrici: una per gli ingressi e una per le uscite della RNA.

Il secondo passo è rappresentato dalla riduzione attraverso l'*Analisi delle Componenti Principali (ACP)*. Questa metodologia presenta il vantaggio di eliminare le componenti principali che contribuiscono a meno di un valore predefinito λ, espresso come percentuale della totale variazione nell'insieme di dati a disposizione, in questo modo risulta possibile definire a priori l'ordine di approssimazione dovuta alla perdita di informazione. Questa applicazione presenta, tuttavia, un inconveniente legato alla numerosità dei campioni, infatti il numero dei campioni relativi a ciascun esempio deve essere minore o uguale al numero di esempi. Come conseguenza dell'applicazione di questa metodologia era necessario un numero di esempi molto elevato.

Per questo motivo, si è proceduto ad un'ulteriore riduzione della dimensione delle matrici sulla base di una soglia delle sole ampiezze in modo da tenere solo le

componenti significative. Le componenti delle ampiezze al di fuori della soglia sono state poste uguali a zero e con esse anche le rispettive fasi. Alle nuove matrici è stata applicata la *ACP* preceduta dalla normalizzazione nell'intervallo [-1,+1].

La rete elaborata è composta da 11 neuroni nello strato di ingresso, un unico strato nascosto da 11 neuroni ed uno strato di uscita da 36 neuroni. L'interruzione della fase di allenamento era basata sulla metodologia "Stopped training", per cui pattern di 104 esempi è stato suddiviso in: un training set di 74 esempi, un validation set di 19 ed un test set di 11 esempi.

Durante la fase di training i pesi delle connessioni sono modificati in modo da minimizzare l'errore tra l'output calcolato e il target (output desiderato). L'obiettivo è quello di ricostruire la relazione ingresso/uscita del sistema e ottenere partendo dall'input un output il più vicino possibile all'output esibito. Allo stesso tempo, viene calcolato l'errore negli esempi del validation set e quando questo inizia a crescere l'addestramento viene interrotto. Il test set viene utilizzato esclusivamente per valutare la capacità di generalizzazione della rete su esempi non noti che non hanno partecipato all'addestramento.

In questo modo, la rete è stata addestrata per la risoluzione del problema diretto. Al termine dell'addestramento i pesi vengono congelati e la rete è stata invertita per la risoluzione del problema inverso.

Conoscendo le misure di concentrazione nei pozzi di monitoraggio le corrispondenti curve di concentrazione del contaminante della sorgente per i quattro fronti stratigrafici sono state determinate.

Conclusioni

Nel lavoro di ricerca presentato in questa tesi si è considerato un nuovo approccio, basato sull'applicazione delle RNA, per la risoluzione del problema inverso nel caso di acquiferi contaminati.

Per quanto riguarda l'applicazione descritta nel capitolo tre, discende che la metodologia applicata può risultare utile non solo per l'identificazione delle coordinate spazio temporali delle sorgenti incognite, bensì, anche come metodica per circoscrivere delle aree in cui effettuare delle analisi più approfondite, in modo da minimizzare i costi di eventuali sondaggi nei domini contaminati.

Sulla base dell'applicazione descritta nel capitolo quattro appare evidente che le RNA rappresentano una nuova tecnologia il cui potenziale per la risoluzione di problemi non lineari come quello studiato nel caso della contaminazione dell'acquifero alsaziano.

È chiaro che le RNA rappresentano una tecnologia emergente grazie alla loro principale proprietà rappresentata dalla capacità di essere approssimatori universali. Appare ovvio che il pieno potenziale delle RNA per la risoluzione di problemi non lineari, tenendo in considerazione l'assenza e l'effetto delle incertezze nei parametri, deve essere maggiormente esplorato.

La metodologia sviluppata potrebbe offrire, in tempi di elaborazione relativamente brevi e costi bassi, soluzioni concrete atte a proteggere le risorse idriche sotterranee. A riguardo, si ritiene, a seguito del presente lavoro di ricerca, che tecniche di indagine basate sull'applicazione delle RNA dovrebbero essere ulteriormente esaminate, in quanto, in grado di offrire un valido contributo al campo delle soluzioni esistenti in materia di inquinamento delle acque sotterranee.

RESUME

Introduction

L'étude de vulnérabilité des eaux souterraines est l'un des problèmes les plus actuels de la protection du territoire.

Le travail présenté dans cette thèse, porte sur l'étude de la contamination des ressources hydriques souterraines. Les eaux souterraines constituent la plus grande partie des réserves d'eau potable de la Terre. En règle générale, leur qualité est supérieure à celle des eaux de surface, grâce aux propriétés filtrantes du sous-sol. Cependant, les eaux souterraines sont exposées à des pollutions générées par l'Homme qui rendent cette importante ressource de plus en plus périssable.

Lorsque les eaux souterraines sont contaminées, la restauration de leur qualité et l'élimination des polluants requièrent beaucoup de temps, et, parfois, il s'agit des processus quasi-impossibles.

Dans le domaine de la contamination des eaux souterraines, dans certains cas, la pollution peut résulter d'une contamination dont la localisation et les origines sont inconnues. De telles situations nécessitent la recherche de techniques qui permettent l'identification des caractéristiques de ces sources contaminantes inconnues.

En général, l'identification et la délimitation de la source d'un panache de contamination est d'une grande importance dans la définition des politiques appropriées pour la gestion des sites contaminés et la planification de l'assainissement du sous-sol.

La détermination des conditions initiales de la pollution est, encore, d'un intérêt considérable dans l'application de la Directive Européenne 2004/35/CE sur la responsabilité, la prévention et l'indemnisation des dommages environnementaux. La Directive est fondée sur l'affirmation du principe «pollueur-payeur» et a été transposée en droit italien par le Décret Législatif 152/2006 (Partie VI - «Norme in materia di tutela risarcitoria contro i danni all'ambiente») et en droit français par la loi 2008-757 (Titre 1er – "Dispositions relatives à la prévention et à la réparation de certains dommages causés à l'environnement").

Certaines caractéristiques concernant la qualité et l'hydrogéologie des eaux souterraines, dans de nombreux cas, ne sont pas directement mesurables et doivent être évaluées en fonction d'autres paramètres directement mesurables. Le problème de la

détermination des paramètres inconnus du modèle est généralement dénommé "problème inverse".

La résolution du problème inverse pour la modélisation de l'écoulement et le transport des contaminants dans les eaux souterraines est l'objectif principal de ce travail de recherche.

Quant à la résolution du problème inverse, dans le présent document, nous avons pour objectif la définition d'une méthodologie qui permette l'identification des caractéristiques dans l'espace et le temps des sources inconnues de contaminations. Dans ce travail de recherche, le problème inverse est résolu sur la base de mesures de concentrations du contaminant dans les puits de surveillance situés dans un domaine d'intérêt. Une fois connu l'effet d'un certain phénomène, nous cherchons à reconstruire la cause qui l'a généré.

Ainsi, la recherche a-t-elle été élaborée selon les points suivants :

- Modélisation de la contamination des eaux souterraines par l'utilisation d'un logiciel non-commercial pour la modélisation des flux et le transport des contaminants dans les milieux poreux (TRACES - Transport dans le sous-sol ou RadioActiver Elements - développé par Hoteit Ackerer (2003)).
- Modélisation des relations cause-effet de la contamination des eaux souterraines par les Réseaux de Neurones Artificiels (RNA). Les RNA ont été créés en utilisant le Neural Network Toolbox de Matlab 7.1.
- Application des RNA pour la résolution du problème inverse dans deux cas de contamination des eaux souterraines étudiés.

Le model numérique TRACES est un code numérique (2D-3D) développé au sein du LHyGes qui permet de simuler l'écoulement et le transport réactif dans un milieu poreux saturé.

Depuis les dernières décennies, les RNA sont devenus de plus en plus utilisés comme outil de résolution de problèmes et de prévision dans de nombreuses disciplines.

Un RNA est réalisé par l'interconnexion d'un nombre de processeurs élémentaires appelés neurones (Perceptron). Les neurones sont logiquement disposées en deux ou en plusieurs couches et peuvent interagir les uns avec les autres par des connexions. En particulier, à travers les réseaux de neurones Multi Layer Perceptron (MLPS), utilisés

dans ce travail, il est possible de créer un modèle d'un système uniquement sur la base d'un ensemble approprié de couples d'exemples d'entrées/sorties du système étudié. Les caractéristiques des RNA, développées dans ce travail, sont fonction de la nature des problèmes analysés. Il n'existe pas de lignes directrices théoriques pour déterminer la meilleure approche pour la création des RNA. En règle générale, le modèle est spécifique pour le système étudié et ne peut être construit de manière générale.

L'apprentissage d'un RNA est constitué d'une règle qui modifie les poids de connexions sur la base de la différence entre la sortie calculée par le réseau et la sortie réelle du système (objectif). Le but de la formation est de permettre au RNA de généraliser les informations obtenues au cours de l'entrainement et de fournir la sortie correcte pour des exemples non compris dans l'ensemble des exemples utilisés pendant l'apprentissage.

Dans ce travail, afin d'élaborer la méthodologie mentionnée ci-dessus, deux cas différents de sources ponctuelles et continues de la pollution ont été considérés, en particulier : un cas de pollution d'un aquifère théorique et un cas réel (source inconnue de la pollution de l'aquifère d'Alsace en France).

Dans le cas théorique, l'objectif était de définir les coordonnées spatio-temporelles (X, Y, T) de la source contaminant inconnue sur la base de mesures de la concentration du contaminant acquis dans le puits de surveillance à un certain moment *t*.

Dans le cas réel, le but est de définir le comportement d'une source de pollution inconnue qui suite à un d'un accident en 1970, a contaminé par le tétrachlorure de carbone (CCl_4), l'un des plus grands aquifères en Europe occidentale. Le comportement de la source de la pollution dans le lieu de l'accident n'est pas connu. L'objectif est d'identifier, à partir de simulations des concentrations du contaminant, obtenues dans des puits, les caractéristiques de la source de la pollution inconnue en termes de variations spatiales et temporelles (lieu et évolution des flux).

Une méthodologie fondée sur l'application de la technologie des RNA a été mise au point pour les deux cas étudiés. En particulier, des RNA ont été créés pour résoudre le problème direct. L'objectif était d'associer les entrées au système (représentées par les caractéristiques de la source) avec les sorties du système (représentées par les mesures acquises dans les puits distribués dans le domaine). Pour résoudre le problème inverse et identifier les caractéristiques de la source de la pollution inconnue, les RNA formés

ont été inversés. Fixant les caractéristiques des sorties du système, les entrées correspondantes peuvent être reconstruites.

Afin de construire des exemples nécessaires à l'apprentissage du RNA, plusieurs scénarios caractérisés par différentes sources de contamination ont été calculés à l'aide du logiciel TRACES.

Le travail est organisé de la manière suivante : dans le premier chapitre, nous rappelons les équations classiques de l'hydrodynamique et du transport par convection et dispersion d'un soluté en milieu poreux saturé. Le deuxième chapitre décrit les réseaux de neurones artificiels ainsi que les modèles mathématiques de base utilisés dans la partie expérimentale. Le troisième chapitre est consacré à la partie expérimentale concernant de contamination réelle de la nappe d'Alsace et les résultats ainsi obtenus. Le quatrième chapitre s'attache, quant à lui, à la partie expérimentale de la contamination d'un l'aquifère théorique.

Chapitre 1: Modélation du flux et du transport du contaminant

Dans ce chapitre, le modèle mathématique utilisé pour modéliser le mouvement contaminant dans les eaux souterraines est présenté. La migration du contaminant est décrite par l'équation de l'écoulement et l'équation de transport. L'équation de l'écoulement est régie par deux équations principales qui sont la loi de Darcy et l'équation de continuité. La loi de Darcy exprime la vitesse de filtration en fonction du gradient de charge. L'équation de continuité exprime le principe de la conservation de la masse d'un fluide en mouvement. Dans un volume élémentaire, la masse du fluide prélevé ou injecté est égale à la somme de la variation de la masse du fluide durant un intervalle de temps élémentaire et des flux massiques traversant la surface de ce volume. Le transport de polluant est décrit par l'équation de convection-dispersion.

Chapitre 2: Réseaux de Neurones Artificiels

Dans ce chapitre, le concept de réseaux de neurones artificiels et ses caractéristiques les plus importantes sont est décrits. Les composants, la structure et l'architecture des RNA utilisés dans cette recherche sont également détaillés. Plus particulièrement dans cette partie sont exposés le processus pour la construction d'un réseau multicouche Perceptron et le problème de l'apprentissage. Un paragraphe est

consacré à la comparaison entre l'approche classique et l'approche neuronale de la modélisation d'un phénomène.

Les RNA sont capables de traiter les informations d'une manière similaire au cerveau humain, car ils ont la capacité de s'adapter aux nouvelles situations en utilisant la connaissance des situations similaires. Le trait distinctif du RNA réside dans leur capacité de généralisation, d'apprentissage et d'approximation.

En général, l'approche fondée sur des réseaux neuronaux artificiels utilisé suivi dans ce travail de recherche, se compose des phases suivantes :

- Pré-traitement des données : analyse descriptive des données, leur traitement et le codage en fonction des besoins du modèle neuronal parmi la réduction des variables à travers différentes techniques.
- Sélection de l'architecture, les règles d'apprentissage et des fonctions d'erreur : choix du nombre d'unités élémentaires (neurones) et des couches cachées, sur la base du plus bas taux d'erreur calculé parmi les différentes architectures. La capacité de généralisation du RNA est liée strictement à cette phase. Dans cette étape sont aussi choisis l'algorithme d'optimisation et la fonction d'erreur.
- Apprentissage: cette phase concerne l'estimation des poids des connections du réseau. Dans ce travail, nous n'avons utilisé que la formation supervisée, où l'apprentissage est piloté de l'extérieur du réseau en fournissant un ensemble d'exemples d'entrée et de sortie du système étudié. Le réseau attribue les meilleures paires de poids sur la base des données expérimentales, en vue de reconstruire la relation entrée/sortie du système et d'obtenir, en fonction de l'entrée, une sortie la plus proche possible de la sortie désirée.
- Évaluation de la stabilité: étape nécessaire pour évaluer le fonctionnement du réseau, sa capacité de généralisation et, par conséquent, d'assurer la stabilité des résultats. La stabilité du réseau est influencée par plusieurs facteurs, notamment: l'architecture sélectionnée, l'initialisation des poids et l'échantillon utilisé pour l'apprentissage (couples d'exemples entrées/sorties du système).
- Post-traitement: cette phase est consacrée à l'évaluation des résultats. Les données, qui avaient été réduites et transformées selon les besoins du réseau de neurones, sont signalées par des techniques d'inversion dans le même état.

Chapitre 3: Application des RNA à l'étude d'un aquifère pollué

Ce chapitre est consacré à la définition d'une méthodologie pour identifier la position et la durée de l'activité des sources polluantes inconnues en utilisant quelques valeurs de concentration du polluant mesurées dans les puits à un certain moment *t*. L'objectif est d'évaluer la capacité des RNA à identifier les caractéristiques mentionnées ci-dessus d'une source contaminant. Dans la première phase de cette partie de la recherche un cas théorique de pollution est etudié et un modèle d'écoulement et de transport a été créé à l'aide du logiciel TRACES. L'aquifère étudié est un aquifère mono couche confiné homogène et isotrope. De plus, une hypothèse restrictive d'une source contaminant générique diffusée par une seule maille du domaine a été étudiée.

Les exemples utiles pour les RNA ont été construits en prenant en compte 40 sources de contamination situées dans les différentes parties du domaine. Le temps d'activité des 40 sources sont de 10, 20 et 30 ans. Ainsi, 120 différents scénarios hydrogéologiques ont été calculés. Les échantillons sortant de TRACES sont représentés par des matrices de la concentration des contaminants obtenue dans 50 mailles uniformément réparties dans le domaine.

Dans l'étude d'un aquifère contaminé, les courbes de concentration des contaminants mesurées dans les puits peuvent être utilisées pour identifier les caractéristiques de la source contaminant. Toutefois, l'identification des sources de contamination devient plus difficile eu l'absence de courbes de séries chronologiques de la concentration du contaminant dans le domaine. Pour ces raisons, cette étude a pris en compte la situation la plus extrême, en particulier le cas d'une absence totale de séries chronologiques de la concentration du contaminant : en fait, le cas de la découverte de la contamination pour la première fois dans un domaine donné.

Dans cette optique, seule les dernières valeurs de concentration obtenues avec les simulations pour les 50 mailles du domaine ont été prises en compte.

Une méthodologie basée sur des réseaux neuronaux artificiels a été développée pour réduire le nombre et choisir parmi ces 50 mailles les plus appropriées pour la réalisation éventuelle du réseau suivi. A l'issu de la procédure, seulement 8 des 50 mailles ont été retenues.

À ce stade, l'architecture du réseau le mieux adapté pour résoudre le problème de l'identification des coordonnées spatio-temporelles des sources polluantes inconnues a été choisie.

Les réseaux constitués sont de type multi-couches Perceptron avec une couche cachée, l'algorithme d'apprentissage retenu est celui de Levenberg-Marquardt

Le modèle neuronal inclus : 3 neurones dans la couche d'entrée, une couche cachée de 8 neurones et 8 neurones dans la couche de sortie. Les neurones de la couche d'entrée sont les suivants : 2 pour les coordonnées spatiales (X, Y) et 1 pour le temps d'activité de la source pour 10, 20 et 30 années. Les neurones de sortie représentent les 8 valeurs de la concentration du contaminant pour les 8 puits de surveillance.

Le nombre réduit d'exemples après divers essais a imposé l'utilisation de la règle d'apprentissage Leave-One-Out Cross Validation (LOO). Dans ce cas, l'ensemble des données est divisé en deux : un ensemble pour l'apprentissage composé par 119 exemples et un ensemble pour le test composé d'un exemple. 120 RNA ont été formés avec différents ensembles d'apprentissage; les 120 exemples ont tous été utilisés un par un dans l'ensemble test.

De cette façon, les ensembles d'apprentissage et de test sont toujours différents. Cette règle permet d'acquérir le plus possible d'informations des couples d'exemples.

Dans ce cas, LOO n'est pas utilisé pour la formation d'un réseau qui sera utilisé pour résoudre un cas particulier, mais uniquement pour évaluer l'habileté du réseau à retrouver les caractéristiques de la source inconnue pour l'exemple test.

Les 120 réseaux ont été formés pour résoudre le problème direct. Après la phase de formation, les réseaux ont été inversés pour résoudre le problème inverse. Quand on veut trouver une nouvelle source qui n'est pas incluse dans les 120, les 120 modèles peuvent être utilisés pour l'apprentissage et la nouvelle source servira de test. La méthodologie développée nous permet d'avoir une présomption raisonnable un l'erreur de la nouvelle RNA qui ne sera pas plus élevée que celle trouvée avec les 120 réseaux formés avec le LOO.

En général, les résultats montrent une bonne capacité du réseau à localiser la source. Dans la plupart des cas, l'erreur sur l'identification des coordonnées spatiales a été inférieure à la taille d'une maille et l'erreur maximale commise est inférieure à la taille de deux mailles du domaine.

Moins satisfaisant est le résultat obtenu par l'évaluation de la durée de la source de contamination. Seulement 71% des réseaux ont été en mesure de fournir à 100% de réponses correctes. En particulier, les réseaux pour lesquels la durée de l'activité de la source était de 10 et 30 ans respectivement, il y a eu une seule erreur sur les 40 cas. Quant aux réseaux pour une durée d'activité de 20 ans, 21 sont erronés, avec une erreur minimum de 6 mois et un maximum de 5 ans, sur un total de 40 cas.

Chapitre 4: RNA pour l'estimation de la source polluante de l'aquifère alsacien

Ce chapitre est consacré à l'élaboration d'une méthodologie innovante pour l'identification de l'évolution temporelle de la source contaminant inconnue de l'aquifère Alsacien fondée sur l'application des réseaux neuronaux artificiels.

Dans la première partie du chapitre est introduite la problématique liée à la contamination de la nappe alsacienne et le modèle de l'aquifère alsacien utilisé pour la construction du RNA.

En 1970, à BENFELD (village situé à 35 km de STRASBOURG), suite à un accident, un camion-citerne a renversé une quantité inconnue de tétrachlorure de carbone (CCl_4) et contaminé une partie du plus grand aquifère en Europe occidentale.

L'étude de la pollution de l'aquifère alsacien est compliquée, car la quantité de tétrachlorure de carbone infiltré au moment de l'accident est inconnue. En particulier le volume de tétrachlorure de carbone contenue dans le camion-citerne est estimé à 4000l dont une partie a été récupérée, une autre partie s'est évaporée et le reste s'est infiltré dans le sous-sol avant d'atteindre la nappe. Le comportement de la source dans l'espace et le temps est inconnu : en raison de sa faible solubilité dans l'eau, ce contaminant se comporte comme une source de contamination continue, dont la dynamique ne peut être que supposée.

Aujourd'hui, malgré les traitements de dépollution, l'aquifère d'Alsace reste dans une situation de contamination. Connaître les caractéristiques précises de la source est d'une importance fondamentale pour l'estimation des temps et des techniques de dépollution à mettre en ouvre.

Les étapes opérationnelles, pour cette partie du travail, sont décrites ci-dessous :

- le modèle numérique tridimensionnel de la pollution de la nappe, développé par Aswed (2008), a été utilisé comme base pour créer les exemples nécessaires à l'apprentissage du RNA. Au total, 104 scénarios différents de la source de contamination ont été pris en compte.
- pré-traitement destiné à réduire la taille des données sortant de TRACES pour les rendre utiles à l'apprentissage du RNA,
- inversion du RNA formé en vue de la résolution du problème inverse.

Après différents tests, le RNA qui s'adapte le mieux au problème étudié est le multi-couches Perceptron mise en œuvre avec l'algorithme rétro-propagation utilisant l'algorithme de Levenberg-Marquardt.

Les échantillons sortant du modèle numérique se composait de deux matrices :

- l'une composée de quatre colonnes correspondant aux quatre niveaux de contaminations (mailles sur quatre couches du modèle numérique),
- l'autre composée de 45 colonnes correspondant aux 45 courbes des concentrations du contaminant dans les 45 puits.

Dans ces deux matrices les lignes représentent le temps.

Ces matrices étant trop grandes pour être traitées avec les RNA avec un temps de calcul acceptable, elles ont été réduites.

Pendant le pré-traitement les deux groupes des matrices d'entrée (source) et de sortie (puits) ont été examinés séparément. A chaque matrice ont été appliquées les transformées de Fourier du second ordre (FFT-2D). De cette façon, les données ont été transférées du domaine du temps au domaine de la fréquence. Les matrices ont été réduites pour former des vecteurs où la première moitié représente les composantes relatives aux amplitudes et la seconde moitié les composants relatives aux phases. Les vecteurs concernant les 104 entrées et les 104 sorties ont ensuite été réunis pour former deux matrices où chaque colonne représente un exemple. Les deux matrices étaient encore trop grandes, une ultime réduction a été réalisée par la technique de l'analyse en composantes principales (ACP).

L'application de cette technique, cependant, comporte un problème lié au nombre d'échantillons; en effet, le nombre d'échantillons pour chaque exemple doit être inférieur ou égal au nombre d'exemples. Dans ce cas, les matrices ont été normalisées dans

l'intervalle [-1,1] avant d'appliquer la dernière réduction basée sur l'application de l'ACP. Les matrices ainsi réduites sont prêtes à être utilisées pour l'apprentissage du RNA.

Le RNA créé est composé de 11 neurones dans la couche d'entrée, une couche cachée de 11 neurones et une couche de sortie de 36 neurones. Pour la formation, les deux matrices entrée/sortie composées des 104 exemples ont été divisées en : 74 exemples pour l'ensemble d'apprentissage, 19 pour l'ensemble de validation et 11 exemples pour l'ensemble de test.

Au cours de la phase de formation, les poids des connexions sont modifiés de façon à minimiser l'erreur entre la valeur calculée et la sortie désirée. Durant l'apprentissage, l'erreur est calculée sur les exemples de l'ensemble d'apprentissage. Dans le même temps, l'erreur est calculée avec les exemples de l'ensemble de validation. Lorsque cette erreur commence à augmenter, la formation est arrêtée. L'ensemble test est utilisé exclusivement pour évaluer la capacité de généralisation du réseau sur des exemples qui n'ont pas été utilisés lors de l'apprentissage.

Le RNA a été formé pour résoudre le problème direct. À la fin de la formation, les poids sont congelés et le réseau a été inversé pour résoudre le problème inverse.

Conclusions

Le travail de recherche présenté est fondé sur l'application de l'approche du RNA pour la résolution du problème inverse dans le cas des aquifères contaminés.

En ce qui concerne l'application théorique décrite dans le chapitre 3, il semble que la méthode peut être utile non seulement pour l'identification de l'espace-temps coordonnées des sources inconnues, mais aussi comme une méthode pour identifier les zones dans lesquelles une analyse plus approfondie est nécessaire, cette technique peut permettre une importante réduction des coûts pour la sourveillance.

Sur la base de l'application décrite dans le chapitre 4, il est clair que le RNA représente une nouvelle technologie dont le potentiel pour résoudre les problèmes non linéaires (comme celle étudiée dans le cas de la contamination de l'aquifère d'Alsace), devrait être approfondie.

Il s'agit là d'une technologie émergente qui pourrait apporter avec un temps de traitement relativement court et de faible coût, des solutions pratiques pour protéger les

ressources en eaux souterraines. À cet égard, d'après les résultats de ce travail de recherche, il apparait que l'application des RNA peut être une contribution précieuse à l'ensemble des solutions existantes dans le domaine de la pollution des eaux souterraines.

INTRODUCTION

Only a small percentage of water present on earth is useful for human use and 98% of this water is represented by water reserves contained in aquifers. Therefore groundwater represents an important resource for the production of drinking water. Groundwater has generally a higher quality than surface water because of the filtering properties of the ground.

However, groundwater is exposed to man-made pollution. One of the major issues for groundwater specialists is the effective management of the groundwater quality because contamination of groundwater may prevent its use for drinking as well as for other domestic, industrial and agricultural purposes. Due to increased pollution phenomena, groundwater has become increasingly vulnerable and its sustainable management is nowadays extremely important to protect global health.

When groundwater is polluted, the restoration of quality and removal of pollutants is a very slow, hence, lengthy, and, sometimes, a practically impossible process. In consequence a management aimed at protecting the groundwater quality and at safeguarding the groundwater resources has consequently a vital importance for life support systems.

This work focuses on groundwater resources contaminations. In this field, it should be underlined that in some cases, pollution may result from contaminations whose origins are generated in different times and places where these contaminations have been actually found. To address such situations of pollution, it is necessary to develop specific techniques that allow to identify in time and space the behaviour of unknown contaminant sources. In general, the identification and delineation of the source of a contaminant plume is of utmost importance regarding both the improvement of management policies and the planning of subsurface remediation in the polluted site.

The determination of the initial conditions of pollution is of considerable interest in the framework of the implementation of the European Union Directive 2004/35/EC: this directive concerns environmental liability with regard to the prevention and compensation of environmental damages, based on the affirmation of the principle of polluter-payer. Directive 2004/35/EC is transposed into Italian law by the “Legislative Decree 152/2006” (Parte VI – “Norme in materia di tutela risarcitoria contro i danni all’ambiente”) and into French legislation by “Loi numéro 2008-757 (Titre 1er –

“Dispositions relatives à la prévention et à la réparation de certains dommages causés à l'environnement”).

State of groundwater pollution implies the necessary development of an effective protection and monitoring of key zones, especially in those areas where the geological characteristics of the soils strata allow relatively easy penetration of anthropogenic pollution into groundwater. To protect the aquifers, the quantification of parameters that governs the groundwater flow has become imperative an key requirement. Due to the complexity of the hydrogeological processes, attention should be paid to the variations of the domain characteristics in space and time, the interaction between ground and water and a large number of physical parameters, mainly hydrodynamic and structural. Finally one should underline that in certain formation, pollutant may travel long distances in an aquifer without being attenuated. Representing all these parameters in a model can be very useful in terms of development and management of water resources for environmental studies. Unfortunately, in some cases, the model parameters are highly uncertain. These uncertainties regarding the parameters must be taken into account to ensure a better modeling of the aquifer pollution.

In many cases, some hydrogeological and groundwater quality characteristics, are not directly measurable and must be physically assessed in function of directly measurable parameters. The problem of determining the unknown model parameters is usually identified as "inverse problem". Solving the inverse problem is the main goal of modeling groundwater flow and contaminant transport. The validity of an aquifer forecasting model is closely related to the reliability and accuracy of the parameters assessment. With respect to the resolution of the inverse problem, this work aims at defining a methodology that allows to identify the features in space and time of unknown contamination sources. In our case, the inverse problem is solved on the basis of measurements of contaminant concentrations in monitoring wells located in the studied areas.

In the framework of this thesis, the research is developed under the following themes:

- groundwater contamination modeling using a non-commercial software for the flux and transport model in porous media (TRACES - Transport or RadioActiver Elements in the Subsurface);

- modeling of the cause and effect relationships in groundwater contamination with Artificial Neural Networks (ANN) technology;
- application of ANN to solve the inverse problem in two cases of groundwater contamination.

Over the past decades, Artificial Neural Networks (ANN) have become increasingly popular as a problem solving tool and have been extensively used as a forecasting tool in many disciplines.

An ANN consists of a number of interconnected processing elements (Perceptrons) called neurons, which are logically arranged in two or more layers and interact with each other through weighted connections. In particular, the networks Multi Layer Perceptrons (MLPs) used in this work can create a model of a system only on the basis of a suitable set of input/output pairs of example patterns. Several algorithms have been proposed in the literature, allowing one to obtain the desired accuracy of the model in any kind of engineering problem. The features of the developed ANN depend on the nature of the problems analyzed and there are no theoretical guidelines for determining the best way out. The model is specific to the system under consideration and cannot be built a priori.

The training of the ANN consists in a learning rule that modifies the weights of the connections on the basis of the difference between the calculated output of the network and the desired pattern. The aim of the training is to make the ANN able to generalize the acquired information, i.e. to give the correct output even for examples not included in the patterns of the training set. This aspect is crucial for the application described in this work, because the assumption is to reconstruct the input by inverting the trained ANN. In practice, the network is trained through an input-output relationship. After that, the network is inverted to solve the inverse problem by reconstructing the output-input relationship. In other words, once the output of the system is known, the input is reconstructed by inverting the trained ANN.

To develop the methodology previously cited, this work consider two different cases of continuous point contamination sources: a theoretical case and a real case (unknown pollution source of the Alsatian aquifer in France).

In the theoretical case addressed in this study, we aim to define the time-space coordinates (X, Y, T) of the unknown contaminant source based on the measures of contaminant concentration acquired in the monitoring wells at a certain time t.

In the real case considered in this study, we aim to define the behaviour of an unknown pollution source that – following an accident in 1970 - has polluted with carbon tetrachloride (CCl_4) one of the largest aquifers in Western Europe and the main sources of drinking water in the Alsace Region (France), i.e. the Alsatian aquifer. The pollution source behaviour at the accident location is unknown. The objective has been to identify this unknown pollution source in terms of temporal variations, injection rates and duration of the source activity based on the measures of the contaminant concentration curves acquired in the monitoring wells.

To solve the inverse problem for both studied cases, an innovative methodology based on the application of ANN technology has been developed. In particular, different ANNs were trained to solve the direct problem. The objective was to combine the input patterns (which represent the pollution source characteristics) with the output patterns (which represent the measures acquired in the monitoring wells). In order to solve the inverse problem and to identify the unknown pollution source characteristics, the trained ANN has been inverted. By fixing the output pattern, the ANN has been able to reconstruct the corresponding input. The inverse problem solution method developed during this research allow us to solve the problem

Various scenarios of the two contamination sources behaviour have been generated in order to form a consistent pattern of examples necessary for training, validation and testing of the ANN.

The patterns have been constructed using TRACES that combines the mixed-hybrid finite element and discontinuous finite element to solve the hydrodynamic state and mass transfer in the porous media.

For making the problem handy for the ANNs applications, for both the two studied cases, different feature extraction pre-processing techniques have been applied. The input/output patterns dimension have been drastically reduced to a very manageable size in order to limit the number of free parameters of the neural networks.

Data reduction and ANN implementation have been carried out with the Neural Network toolbox of Matlab 7.1.

The results of the research work that are described in this thesis show how ANNs can be used as efficient tools to describe unknown pollution sources characteristics on the basis of the contaminant concentration measures acquired in the monitoring wells.

STATE OF THE ART

The identification and remediation of polluted aquifers represent nowadays an important challenge in groundwater resource management. In order to efficiently manage the groundwater quality, it is fundamental to know pollution source characteristics such as location, magnitude, duration of the activity. Inaccuracies/inadequacies in determining the pollution sources may result an inefficient or unsuccessful management/remediation efforts. Information regarding the pollution sources is also necessary and useful for addressing the judicial issues of responsibility and compensation for environmental damage.

In the field of hydrogeological studies, the literature, starting from the late '80s, but especially in the mid '90s, contains many examples of implementation of different applications of ANN related to various issues. The last decade, there has been seen a significant activity in ANN applied to various hydrogeological problems such as groundwater modeling, modeling of hydrogeological parameters, modeling of various kinds of aquifers contamination, water quality modeling. It is clear that ANNs represent an emerging new technology and their full potential for solving hydrogeological problems must be further explored. This is due to the main properties of ANN, represented by the ability to be universal approximators. Several studies are dedicated to the development of different models for solving the inverse problem, however works using the ANN approach are less popular.

In the following paragraph, some works that address an approach to solve the inverse problem in hydrogeology are presented, although not all of these are based on ANN application. Each technology requires a groundwater flow and contaminant transport model to simulate the physical processes in the aquifer system.

Examples of inverse problem solution

Rizzo and Dougherty (1994), developed a method for pattern completion based on the application of ANNs possessing many operational objectives of the neural kriging approach. A neural kriging (NK) network is implemented in a parallelized algorithm, and applied to develop maps of discrete spatially distributed fields (e.g., log hydraulic conductivity).

Zio (1997) investigated the feasibility of solving the inverse problem by using artificial neural networks. He considered a simple analytic model of contaminant transport due to a point source in stationary flow field to generate simulated concentration histories for various values of the dispersion coefficient. The simulated observations have been used to train the ANN to identify the value of the associated dispersion coefficient. The approach seems to offer a versatile and efficient tool for parameter identification. However no high expectations should be attached to this method in regards of the instability and non unique of the inverse problem solution, since these inherent difficulties are mainly due to the insufficiency and inaccuracy of the available observations and this method will suffer from it as well as the other methods.

Schwarz et al. (1998) proposed a new investigation approach for the assessment of groundwater contamination based on the inversion of concentration time series measured within pumping wells. Using the inversion approach, it is possible to investigate the mean pollutant concentration and the concentration distribution over a control plane perpendicular to the groundwater flow direction downstream of a pollutant source, as well as the mass flux over this control plane.

Holder et al. (1998) tested the method proposed by Schwarz (1998) in the abandoned industrial site Neckar Valley in Stuttgart. They try to reconstruct the best description of the catchment area with the lowest number of wells and shortest pumping time.

Gümrah et al. (1999) described an ANN approach that can be used to forecast the future pollutant concentrations and hydraulic heads of a groundwater source. In order to check the validity of the approach, a hypothetical field data as well as a case study were produced by using groundwater simulator with the method of characteristics (MOC). Hydraulic heads and chlorine concentrations were obtained from groundwater simulations. ANN was trained by using the historical data of the last two years. The chlorine concentrations and hydraulic heads were estimated by applying both the long-term and the short-term ANN predictions. In long-term predictions, the chlorine concentrations and heads were estimated for the future eight years by using the historical data of two years for each well. The short-term approach was applied for the wells where the higher errors have been obtained during the application of long-term predictions (in total two wells).

Mahar & Datta (2000) proposed a methodology based on nonlinear optimization model for estimating unknown magnitude, location and duration of groundwater pollution sources by using measured values of pollutant concentration at selected locations. The performance of the developed model is evaluated for a transient flow and transport of a conservative pollutant in an hypothetical confined homogeneous and isotropic aquifer system. Different scenarios are considered, such as: concentration measurement errors, missing measurements, location of observation wells vis-à-vis actual source locations, and non-uniqueness of solutions in terms of local or global optimal solutions. In total, 8 monitoring wells were considerate. The total study time of five years is divided into twenty equal time steps of three months each. It is assumed that the potential sources are active only during the first four time steps of the solution time horizon. The nonlinear programming formulation minimizes the weighted sum of the squares of the differences between the model estimated and observed concentrations. The results were validated by comparing the results of the proposed model with those obtained by using the USGS-Method of characteristics (USGS-MOC) computer code.

Bockelmann et al. (2001) proposed a new integral groundwater investigation approach to quantify natural attenuation rates at field scale. In this approach, pumping wells positioned along two control planes downstream of a contaminant source. Flux through these plans is calculated using an analytical solution derived by Schwarz (1998). With this technique the spatial-temporal distribution of the contaminant may be reconstructed according to a few concentration data.

Fanni et al. (2002) proposed the use of the neural network paradigm in an innovative way. A neural network is trained to capture the functional relationship between geometrical and chemical properties of the contaminants and the hydrological map of the basin in prefixed measurement points. The network is then use to solve the inverse problem of locating the source of the pollutant, and how many time steps before the event occurred under the restricted hypothesis of groundwater contaminated by a single pollutant injected in a single point.

Rajanayaka et al (2002) utilize a hybrid approach based on a combination of two types of ANN models to solve the groundwater inverse problem. Supervised Multi Layer Perceptron (MLP) ANN and Self-Organising Network (SON) were amalgamated to estimate parameters reasonably accurately by using solute concentration observations

that were obtained from a two-dimensional groundwater transport numerical model. A three layers MLP network was used to find the complex relationship of output, K, and the associated concentration values.

Singh and Datta (2006) proposed an ANN based methodology to identify unknown groundwater pollution sources, when a portion of the concentration observation data is missing. The source characteristics and the corresponding concentration measurements at time steps for which it is not missing, constitute a pattern for training the ANN. A groundwater flow and transport numerical simulation model is utilized to generate the necessary patterns for training the ANN. Performance evaluation results show that the back-propagation based ANN model is essentially capable of extracting hidden relationship between patterns of available concentration measurement values, and the corresponding sources characteristics, resulting in identification of unknown groundwater pollution sources. The performance of the methodology is also evaluated for different levels of noise (or measurement errors) in concentration measurement data at available time steps.

Zhiqiang et al. (2006) presented a novel technique utilizing ANN to backtrack source location and earlier plume concentrations from recent plume information. For proof-of-concept, two tracer tests (a non-point-source and a point-source) were performed in a large-scale (10'×14'×6') groundwater physical model. The physics-based flow and transport model (MODFLOW 2000 and MT3DMS) was calibrated using the data from the non-point-source tracer test and validated using the point source tracer test data. ANN was trained using the calibrated model predictions and compared to actual data. The ANN developed is capable of approximating results quickly, which is important for real-time modeling and long-time monitoring optimization design. The ANN can also back out the earlier plumes to better identify the contaminant source.

Bashi-Azghadi et al. (2010) presented a new methodology for estimating location and amount of leakage from an unknown pollution source using groundwater quality monitoring data. The proposed methodology includes a multi-objective optimization model, namely Non-dominated Sorting Genetic Algorithm-II (NSGA-II) which is linked with MODFLOW and MT3D groundwater quantity and quality simulation models. The main characteristics of an unknown groundwater pollution source are estimated using two probabilistic simulation models, namely Probabilistic Support Vector Machines (PSVMs) and Probabilistic Neural Networks (PNNs). In real-time

groundwater monitoring, these trained probabilistic simulation models can present the probability mass function of an unknown pollution source location and the relative error in estimating the amount of leakage based on the observed concentrations of water quality indicator at the monitoring wells.

Aswed (2008) worked on an accident happened in north-eastern of France on 1970. The goal was to model and simulate the transfer in the aquifer of contaminant (chlorinated solvent). They determined the source term at the accident location. To estimate the contaminant concentration at the source, the travel time between the source and measurement-wells is calculated by the method of moments.

Conclusion and comments

In this chapter presents an overview of the studies in the research area treated in this work.

THESIS'S STRUCTURE

Chapter 1: modeling groundwater flow and contaminant transport

In chapter 1, we briefly present the basic definitions of the porous-medium properties, the governing equations of single-phase flow and transport of solute in saturated porous media. The water flow is described by Darcy's law and the continuity equation that governs the volumetric balance equations. The solute transport is described by a convection-diffusion-dispersion equation.

Chapter 2: Artificial Neural Networks

In chapter 2, the concept of artificial neural networks and the most important features that make this technology attractive in hydrogeology research are explained. Structure, components and architecture models of artificial neural networks are detailed. In particular, the supervising learning problems and the useful tools are discussed. In addition, the classical modeling approach is compared with artificial neural network modeling. The design and training of a Multilayer Perceptron Network is also addressed.

Chapter 3: ANN applied to study a polluted aquifer

Chapter 3 is dedicated to the inverse problem in order to identify the spatial location (X,Y) and the duration of the activity (T) of a theoretical case of groundwater pollution. This chapter provides a description of the theoretical aquifer and its numerical model. Various source scenarios are applied to the theoretical aquifer hydrogeological model in order to generate the examples patterns used for training and testing of the ANN. The patterns elaboration and reduction are detailed. Finally, the ANN development and the inversion procedure are explained.

Chapter 4: ANN for estimating Alsatian aquifer pollution source

Chapter 4 proposes a new methodology that aims at solving the inverse problem in order to reconstruct the behaviour in time and space of the carbon tetrachloride unknown pollution source of the Alsatian aquifer (France). The chapter provides a description of the Alsatian aquifer, the carbon tetrachloride tanker accident occurred in

1970, as well as of the physical characteristics of this dangerous chemical. The numerical model of the Alsatian aquifer is described in order to explain the model produced example patterns for the ANN. Various source scenarios are applied in order to generate the examples patterns used for training, validating and testing the ANN. The patterns elaboration and reduction are detailed. Finally, the ANN development and the inversion procedure are explained.

1 MODELING GROUNDWATER FLOW AND CONTAMINANT TRANSPORT

In this chapter, we briefly present the basic definitions of the porous-medium properties, the governing equations of single-phase flow and transport of solute in saturated porous media. The water flow is described by Darcy's law and the continuity equation that governs the volumetric balance equations. The solute transport is described by a convection-diffusion-dispersion equation.

1.1 Flux and transport phenomena

1.1.1 Properties of saturated porous media

A porous medium consists of a solid matrix with interconnected void spaces. The void spaces can be completely or partly filled with water and/or other fluids like oil or gas. The pores are the spaces that are not filled by solid material. In this thesis, only water saturated porous media are considered. The saturated zone is, therefore, formed of porous material whose pores are filled with water. The main parameters that characterize a porous medium are the porosity and permeability.

Porosity

The Porosity of the medium is represented by the void space distributed in the solid matrix. It is a dimensionless parameter expressed by the volume of void spaces per unit volume of the aquifer material. Since the isolated or disconnected pores do not account for the flow, the concept of effective porosity is introduced, which is the ratio of volume of the interconnected pores to the total volume of the soil or the rock. In granular porous media, such as the alluvial aquifer, the effective porosity is typically almost equal to the total or bulk porosity.

Permeability

The Permeability is a measure of the medium ability to transmit a fluid flow under the influence of a driving pressure. This parameter depends on the size, shape and interconnectness of pores spaces. Finer-grained material exhibits low permeability, while coarser-grained material generally exhibits higher permeability. The intrinsic permeability is expressed, as follows:

$$k = \frac{\mu}{\rho g}\frac{Q}{A}\frac{1}{(\Delta p / \Delta s)} \qquad (1.1)$$

where,

Q: the volumetric flow rate, [L^3T^{-1}];

μ : the dynamic viscosity of the fluid, [$ML^{-1}T^{-1}$];

A: the cross–sectional area of the flowing fluid, [L^2];

$\Delta p / \Delta s$: the applied head difference across the length, [-].

The permeability is a function of the medium. For a medium saturated with water, it is customary to define the hydraulic conductivity. Unlike the permeability, the hydraulic conductivity takes into account the particular fluid that is present in the medium. The hydraulic conductivity is defined as:

$$K = \frac{k\rho g}{\mu} \qquad (1.2)$$

where,

K: is the hydraulic conductivity, [LT^{-1}];

k: is the intrinsic permeability of the medium, [L^2];

ρ: is the fluid density, [ML^{-3}];

g: is the acceleration of gravity, [LT^{-2}];

μ: is the fluid viscosity, [$ML^{-1}T^{-1}$].

It is clear from equation (1.2) that K incorporates the medium permeability k, and the fluid properties ρ and μ.

Many geological formations are anisotropic where the permeability in the direction of the geological layers is greater than in the perpendicular direction. Moreover, in heterogeneous media, the permeability varies in space. The permeability in natural soils may vary from 10^{-8} m^2 for very conducting to 10^{-16} m^2 for poorly conducting aquifers [Bear, 1988]. The permeability depends on the micro scale geometry of the medium, i.e., the grain sizes and the interconnectedness and the orientation of pores. Several empirical and theoretical relationships relate the permeability to the porosity, the effective grain diameter, and other medium parameters.

1.1.2 Groundwater flow equations

In order to model groundwater flow in complicated large scale media, the governing equations are solved numerically and given by Darcy's law and the conservation equations.

Darcy's law

The fundamental law of fluid flow in a porous medium is the Darcy's law. The basic concept is that the groundwater flows from levels of higher energy to the levels of lower energy. This energy is essentially the results of the height and the pressure. Darcy's law in the porous media expresses the filtration velocity in a steady state or transient state as a function of the pressure gradient and the gravity. The Darcy's law is written by the general formula [Bear, 1979], as follows:

$$q = -\frac{k}{\mu}\left(\nabla p + \rho g \nabla z\right) \qquad (1.3)$$

where,

q: is the Darcy's velocity or specific discharge, [LT^{-1}];

μ: is the dynamic fluid viscosity, [$ML^{-1}T^{-1}$];

p: is the fluid pressure, [$ML^{-1}T^{-2}$];

z: is the elevation above some arbitrary datum (is the vertical coordinate), [L].

If the density is assumed to be constant, the Darcy's law (1.3) can then be simplified as:

$$q = -\frac{k\rho g}{\mu}\nabla\left(\frac{p}{\rho g} + z\right) = -K\nabla h \qquad (1.4)$$

where,

$h = \frac{p}{\rho g} + z$ represents the groundwater head (the piezometeric head), [L];

$K = \frac{k\rho g}{\mu}$ is the hydraulic conductivity coefficient or the permeability, [LT^{-1}].

Remark: q is not the true velocity, as it assumes flow through an open pipe and does not take into account the fact that water is only able to flow through the pores

between solid grains. To find the actual groundwater velocity (average velocity), the Darcy velocity is divided by effective porosity w:

$$u = \frac{q}{w}$$

In an isotropic medium, the hydraulic conductivity K, or similarly the intrinsic permeability k is a scalar. However, if the porous medium in three-dimensional space is anisotropic, the hydraulic conductivity is defined as a symmetric tensor of the form:

$$K = \begin{pmatrix} K_{xx} & K_{xy} & K_{xz} \\ K_{xy} & K_{yy} & K_{yz} \\ K_{xz} & K_{yz} & K_{zz} \end{pmatrix}$$

The hydraulic conductivity tensor can be diagonalized introducing three mutually orthogonal axes called *principal directions* of anisotropy. In the following, we suppose that the principal axes are aligned with the *x, y, z* directions. The tensor K is therefore diagonal, that is:

$$K = \begin{pmatrix} K_{xx} & 0 & 0 \\ 0 & K_{yy} & 0 \\ 0 & 0 & K_{zz} \end{pmatrix}$$

In practice, two permeabilities are distinguished: the vertical permeability K_{zz} and the horizontal permeability $K_{xx}=K_{yy}$ [De Marsily, 1981].

The continuity equation

The continuity equation is based on the principle of the conservation of mass of water. In a control volume, the mass flux due to the sources and sinks is equal to the temporal change of mass and the mass flux across the volume boundaries [Bear, 1979]:

$$\frac{\partial(\rho\omega)}{\partial t} + \nabla \cdot (\rho q) = \rho f \qquad (1.5)$$

where, f represents the sink/source term for the fluid, [T^{-1}].

The porosity is generally slightly pressure-dependent [Kinzelbach, 1986]. However, this aspect is neglected in this work, i.e., the matrix is considered incompressible. The density ρ depends only on the pressure p at a constant temperature. One can then write:

$$\frac{\partial(\rho\omega)}{\partial t} = \frac{\partial(\rho\omega)}{\partial p}\frac{\partial p}{\partial t} = \frac{s}{g}\frac{\partial p}{\partial t} \qquad (1.6)$$

where, $s = g\dfrac{\partial(\rho\omega)}{\partial p}$, [$L^{-1}$], is the specific storage coefficient which gives the mass of fluid added to storage (or released from it) in a unit volume of porous medium per unit rise (or decline) of the pressure head $p/(\rho g)$.

By substituting Eq. (1.6) into Eq. (1.5), the continuity equation is obtained in terms of the pressure. If the spatial variation of density is negligible, the continuity equation becomes:

$$\left(\frac{s}{\rho g}\right)\frac{\partial p}{\partial t} + \nabla \cdot q = f \qquad (1.7)$$

The relation between the hydraulic head and the pressure is given by [Bear, 1979]:

$$p = \rho g\left(h - z\right)$$

Then, one can write:

$$\frac{\partial p}{\partial t} = g(h - z)\frac{\partial \rho}{\partial t} + \rho g\frac{\partial h}{\partial t}$$

By using, we obtain:

$$\frac{\partial p}{\partial t} = \left(\frac{(h - z)s}{\omega}\right)\frac{\partial p}{\partial t} + \rho g\frac{\partial h}{\partial t}$$

It follows that,

$$\frac{\partial p}{\partial t} = \left(\frac{1}{1 - (h - z)s/\omega}\right)\rho g\frac{\partial h}{\partial t} \qquad (1.8)$$

In particle, the quantity $\left((h - z)s/\omega\right)$ is negligible with respect to 1 [Banton & Bangoy 1997]. Then:

$$\frac{\partial p}{\partial t} \approx \rho g\frac{\partial h}{\partial t} \qquad (1.9)$$

By replacing Eq. (1.9) in Eq. (1.7), the mass balance equation of an incompressible fluid in the non-deformable porous medium is written in the general form:

$$s\frac{\partial h}{\partial t} + \nabla \cdot (K\nabla h) = h \qquad (1.10)$$

In steady-state, the piezometric head is constant over time. Equation (1.10) reduces to the following form:

$$\nabla \cdot q = f \qquad (1.11)$$

Initial and boundary conditions

The transitory flow problem described by the continuity equation and Darcy's law (1.4) requires knowledge of the initial and boundary conditions. Initial conditions provide the necessary set of primary variables in the computational domain at the beginning of the simulation. Additionally boundary conditions (BC) have to be supplied at the margins of the model domain. These boundary conditions represent the interaction between the domain and the surrounding environment. Various types of boundary conditions are the following:

- Dirichlet-BC: has fixed value of the head at the boundary of the domain.

$$h(x,t) = h^D(x,t)$$

where h^D is a known function.

In the steady state, this type of boundary conditions is necessary to guarantee the uniqueness of the solution. The conditions of prescribed head value can be, for example, the contact of the aquifer with a river, rivers/lakes, etc.

- Neumann-BC: describes the flux of a quantity perpendicular to the boundary of the domain. It is expressed by:

$$q \cdot n = -K\frac{\partial h}{\partial n}(x,t) = q^N(x,t)$$

where, n is the outward normal vector on the boundary and q^N is a known function.

A condition for prescribed flux can be the impermeable boundaries where the flux is zero, inflow or outflow through the boundaries.

- Cauchy or Fourier-BC: is a combination of Dirichlet and Neumann boundary conditions. The flow across the boundary is calculated from a given value of the head, such that:

$$-K\frac{\partial h}{\partial n}(x,t) = g^F(x,t)h + f^F(x,t)$$

where f^F and g^F are known functions.

An example of this case is the interactions of an aquifer with a river.

1.1.3 Porous medium transport equations

Water, in its movement, can carry materials in dissolved form. The transport of such pollutants is a process which takes into account several physical mechanisms such as convection, hydrodynamic dispersion, molecular diffusion, and chemicals mechanisms such as adsorption, radioactive decay/filiations, and precipitation/dissolution. The fluid-medium interaction may fasten or reduce the spreading of the pollutant in the porous medium.

Convection

The convection is the movement of the pollutant dissolved in the groundwater in the direction of the flow. The convection is derived by the mean velocity of the groundwater. Thus, an increase in groundwater velocity will result in farther travel of the contaminant.

The convection is, generally, the dominant mass transport process in groundwater flow system [Domenico et al., 1990]. The migration of a contaminant owing to convection is significantly influenced by the aquifer hydraulic conductivity, effective porosity, and hydraulic gradient [Wiedemeier et al. 1999].

In a uniform porous media, water will travel vertically downward until it hits the water table and then move in the down gradient direction of the aquifer [Wolfe et al., 1997].

The convection equation is given by:

$$\frac{\partial C}{\partial t} = -\nabla \cdot (Cu) \qquad (1.12)$$

where,

C: is the concentration of solute, [ML^{-3}];

u: is the actual groundwater velocity, [LT^{-1}].

Dispersion and diffusion

These mechanisms may lead the contaminant to spread in directions different from the water flow paths. The molecular diffusion is due to concentration gradients within the liquid phase. This mechanism is independent of the flow velocity. It produces a flux of particles from region of high contaminant concentration to regions of low concentration. The mechanical dispersion is a phenomenon of spreading caused by fluctuations in the velocity field and heterogeneities at the microscopic scale. Reactive and non-reactive solutes may spread due to dispersion both along and perpendicular to the groundwater flow. The dispersion increases in heterogeneous material due to non-uniform groundwater flow paths.

The equation of dispersion-diffusion is given by:

$$\frac{\partial C}{\partial t} = \nabla \cdot (D \cdot \nabla C) \qquad (1.13)$$

where D is the dispersion-diffusion tensor which represents the contribution of the mechanical dispersion and of molecular diffusion in porous media. This tensor, in the three-dimensional space, takes the form:

$$D = D_c + D_m$$

where D is the dispersion-diffusion tensor which represents the contribution of the mechanical dispersion and of molecular diffusion in porous media. This tensor, in the three-dimensional space, takes the form:

where,

Dc: is the mechanical dispersion tensor [Bear, 1979]:

$$D_c = \|u\|(\alpha_l E(u) + \alpha_t (I - E(u)))$$

with,

$$E_{i,j}(u) = \frac{u_i u_j}{\|u\|^2} \qquad i,j=1,\ldots.3;$$

Dm: is the diagonal tensor of the molecular diffusion in porous medium [L^2T^{-1}];

α_l : is the longitudinal dispersivity, [L];

α_t :is the transversal dispersivity, [L].

Unlike the dispersion, the diffusion can occur both in the absence or presence of convective flow. It is generally less significant than dispersion in most groundwater flow problems.

The three mechanisms mentioned above (convection, dispersion, and diffusion) cause the contaminant to spread in the direction of flow both longitudinally and transversally. The combined processes of advection and dispersion result in a reduced concentration of the dissolved solute (dilution) as well as plume spreading. Dispersion generally causes contaminants to migrate 10 to 20 percent further than migration created by advection alone. The processes of advection, dispersion, and diffusion control the movement of the contaminant [Clement et al. 2004].

The equation convection-diffusion-dispersion

In the case of conservative and non-reactive transport, the integration processes of convection, molecular diffusion and mechanical dispersion are given by the following equation:

$$\frac{\partial C}{\partial t} = -\nabla \cdot (Cu) + \nabla \cdot (D \cdot \nabla C) \qquad (1.14)$$

In the presence of an instantaneous and linear adsorption, the relation between the solute concentration C and the sorbed concentration in the solid C_s, is given by Eq. (1.15). The total mass of solute per unity of volume can be written as:

$$\omega C + \rho_s (1-\omega) C_s \qquad (1.15)$$

Where ρ_s represents the density of solid, [ML^{-3}].

Assuming instantaneous linear adsorption, and constant porosity the retardation factor *R* is defined by:

$$\omega \frac{\partial C}{\partial t} + \rho_s (1-\omega) \frac{\partial C_s}{\partial t} = \omega \left(1 + \rho_s \frac{(1-\omega)}{\omega} K_d \right) \frac{\partial C}{\partial t} = \omega R \frac{\partial C}{\partial t}$$

where,

$$R = \left(1 + \rho_s \frac{(1-\omega)}{\omega} \right) K_d$$

By taking into account the spatial immobility of sorbed solute due to the convection or to the dispersion, the equation of transport becomes:

$$R\frac{\partial C}{\partial t} = \nabla \cdot (D \cdot \nabla C) - \nabla \cdot (Cu) \qquad (1.16)$$

With, $R \geq 1$, this term decreases the transport velocity of the solute with respect to the velocity of the groundwater.

In the case of a degradation mechanism of first order (decrease radioactive) and in the present of source/sink function, the equation of transport becomes:

$$R\left(\frac{\partial C}{\partial t} + \lambda C\right) = \nabla \cdot (D \cdot \nabla C - Cu) + f_c \qquad (1.17)$$

where,

λ: is the degradation coefficient, [T^{-1}];

f_c: is the sink/source term which describes the outlet/inlet in the domain,[$ML^{-3}T^{-1}$].

In addition to the initial conditions given for C at $t=0$, the boundary conditions, related to the transport problem, can be:

- Dirichlet type: prescribed concentration: $C(x,t) = C^D(x,t)$,
- Neumann type: prescribed hydraulic head value: $-D_c \frac{\partial C}{\partial n}(x,t) = q^N(x,t)$,
- Fourier conditions: a combination of hydraulic head and concentration;

1.2 Summary

In this first chapter, we have provide a definition of the porous-medium properties and the governing equations of single-phase flow and transport of solute in saturated porous media has also been given. The water flow is described by the Darcy law and the continuity equation that governs the volumetric balance equations. The solute transport is described by a convection-diffusion-dispersion equation.

2 ARTIFICIAL NEURAL NETWORKS

Before discussing the artificial neural network application developed to solve the hydrogeological inverse problem, this chapter aims at presenting the key features of artificial neural networks technologies as well as the biological origin of artificial neural networks research.

2.1 Introduction

The first studies on Artificial Neural Networks (ANNs) were prompted by a desire to have computers mimic human learning. The initial attempt was to reproduce the neural structure of the brain tissue on computational tools. ANNs possess the unique attribute of universal approximation, ability to learn from examples without the need of explicit physics, and the capability of processing large data volumes at high speeds.

The basic notion of ANNs was first formalized by McCulloch and Pitts (1943) in their model of an artificial neuron. Attention to research in this field remained somewhat dormant in their early years due to the unknown/undiscovered capabilities of this method and of its potential use. However, interest in this area picked up momentum in a dramatic fashion with the work of Hopfield (1982) and Rumelhart (1986). Not only did these studies place ANN on a firmer mathematical footing, but also opened the door to a host of potential applications for this computational tool. Consequently, ANN computing has progressed rapidly on different fronts such as theoretical development of different learning algorithms and computing capabilities [Govindaraju et al., 2000].

Over the past decade, ANNs have become increasingly popular in many disciplines as a problem solving tool. This technology was developed in many different fields always characterized by a high degree of interdisciplinary. ANNs have the ability to solve extremely complex problems with highly non-linear relationships. ANNs have a flexible structure that are capable of approximating almost any input-output relationships.

Complex and heterogeneous hydrogeology systems are extremely difficult to model. However, it has been proved that ANN's flexible structure can provide simple and reasonable solutions to various problems in hydrogeology. ANNs have been successfully employed in hydrogeology research [Morshed et al., 1998; ASCE Task Committee on Application of ANN in Hydrology, 2000; Maier et al., 2000].

In this work, the ANN has been developed with version 7.1 of the Neural Networks Toolbox of Matlab.

2.2 Natural Neural Network

Before describing the technical side of ANN, it would be useful to briefly discuss the Natural Neural Networks (NNN) and the cognition of living organisms. This brief summary explained the few elements of biological neural networks we want to take over into the ANN.

A natural neuron is comparable to a switch with information input and output. In the switch, called soma, weighted information are accumulated. It is activated if there are enough stimuli from other neurons hitting the information input. Incoming signals, from other neurons, are received by neuron dendrites and then transferred to the neuron by special connections: the synapses. In the soma, as soon as the accumulated signal exceeds a certain value, an electrical pulse is activated. The inputs are summarized to a pulse according to the chemical change. The pulse represents the output information that is transmitted to the others connected neurons. Outgoing pulses are transferred by the axons. The axon is a long slender extension of the soma and it is electrically isolated in order to achieve a better conduction of the electrical signal.

Depending on how the neuron is stimulated by the cumulated input, the neuron itself emits a pulse or not. The output is non-linear and not proportional to the cumulated input.

Input and output of an Artificial Neuron (AN) may be vectors or scalars. In the ANN, the inputs are multiplied by a number: the weight. The set of such weights represents the information storage of the ANN. The weights of the inputs are variable, similar to the chemical processes at the synaptic cleft. This adds a great dynamic to the network because a large part of the ”knowledge” of an ANN is saved in the weights.

2.3 Structure and components of Artificial Neural Networks

2.3.1 Definition of Artificial Neural Network

There are many definitions of artificial neural networks (ANNs). We will use a pragmatic definition that emphasizes the key features of the technology. ANNs are

distributed, adaptive, generally nonlinear learning machines built from many different Processing Elements (PEs) [Principe, 2002].

The ANN is symbolized like a graph where patterns are represented in terms of numerical values attached to the nodes of the graph (PEs or neurons), and transformations between patterns achieved via simple message passing algorithms (Figure 2.1).

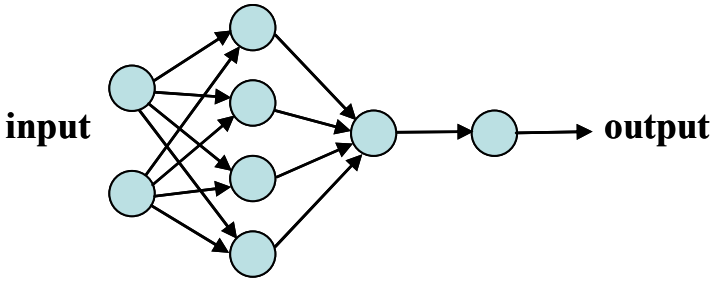


Figure 2.1: graph scheme of an Artificial Neural Network.

PEs, also called AN, are logically arranged in two or more layers and interact with each other through the weighted connections The interconnectivity of the AN defines the ANN topology.

The input value and the corresponding set of desired target (output value used to train, validate and test an ANN) is called pattern. ANN may be trained to perform a particular function by adjusting the values of the connections (weights) between elements. Commonly ANN are adjusted, or trained, so that a particular input leads to a specific target output. Such a situation is shown in Figure 2.2 below.

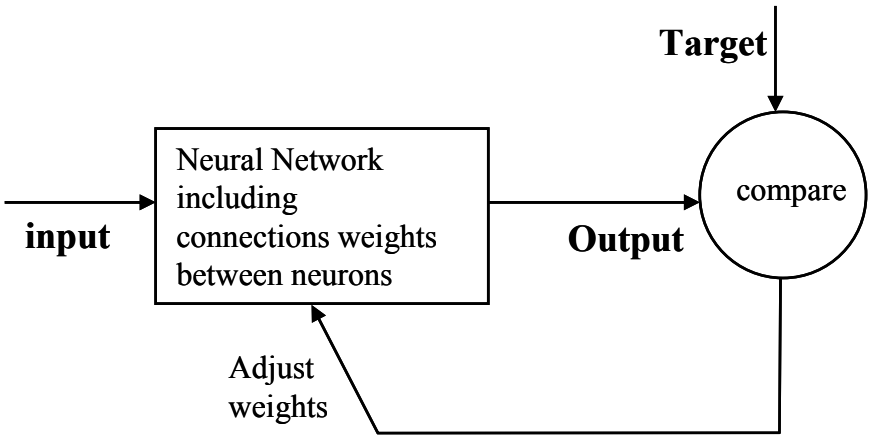


Figure 2.2: Adjust weight between neurons.

The network weights are adjusted, based on a comparison of the output (output calculated by the network) and the target (desired output that correspond to the output pattern), until the network output matches the target. [The mathworks, 2005].

2.3.2 Mc-Culloch-Pitts Processing Element

The Mc-Culloch-Pitts (M-P) PE make simply a sum-of-products followed by a threshold nonlinearity (Figure 2.3). Its input-output equation is:

$$y = f(net) = f\left(\sum_{i=1}^{D} W_i x_i + b\right) \qquad (2.1)$$

Where D is the number of inputs, x_i are the inputs, W_i are the input connection weights, x_0 is the bias, W_0 is bias weight and y is the output. The activation function f is a threshold function defined by:

$$f(net) = \begin{cases} 1 \ \ for\ net \geq 0 \\ -1\ for\ net < 0 \end{cases} \qquad (2.2)$$

which is commonly called the signum function.

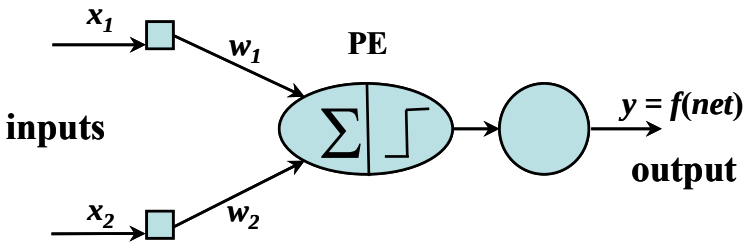


Figure 2.3: Two inputs, one output McCulloch-Pitts PE.

The McCulloch-Pitts PE is created by the concatenation of a synapse and an axon. The synapse contains the weights W_i, and performs the sum-of-products. The synapse shows that the element has 2 inputs and one output. The number of inputs x_i is set by the input axon.

2.3.3 The Perceptron

The perceptron is a pattern-recognition machine composed by multiple inputs fully connected to an output layer with multiple McCulloch-Pitts PEs. The Perceptron produce a simply sum of the product between inputs and connection weights (2.3). Each input x_j is multiplied by an adjustable constant W_{ij} (the weight) before being fed to the ith processing element in the output layer, yielding [Principe, 2000].

Perceptron input-output equation is:

$$y = f\left(\sum_{i=1}^{D} W_i x_i + x_0 W_0\right) \qquad (2.3)$$

Where D is the number of inputs, x_i are the inputs, W_i are the input connection weights, x_0 is the bias, W_0 is bias weight and y is the output [Principe, 2000].

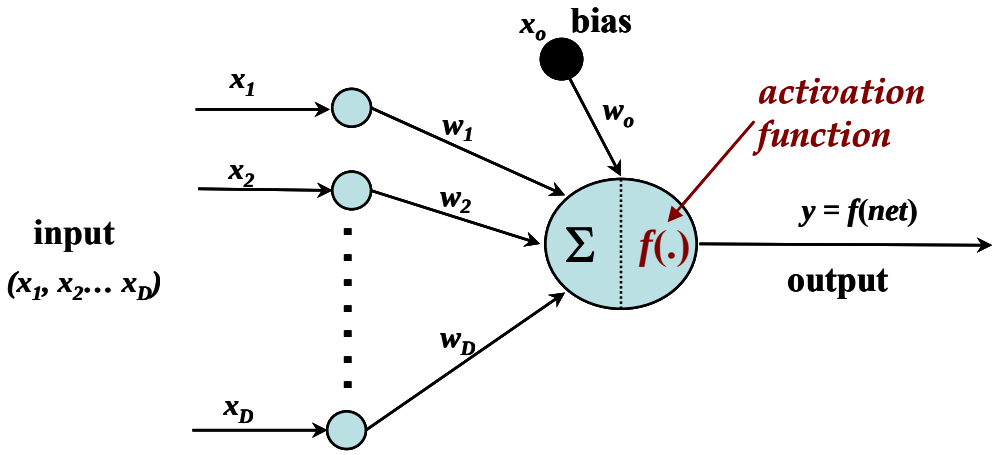


Figure 2.4: Perceptron with D inputs and 1 output.

The PEs sum all these contributions and produce an output that is a nonlinear (static) function of the sum. The PEs' outputs become either system outputs or are sent to the same or other PEs depending to the ANN architecture. Each PE in the ANN receives connections from other PEs and/or itself. The interconnectivity defines the topology. The signals flowing on the connections are scaled by adjustable parameters called weights, w_{ij}. A weight is associated with every connection [Principe, 2000].

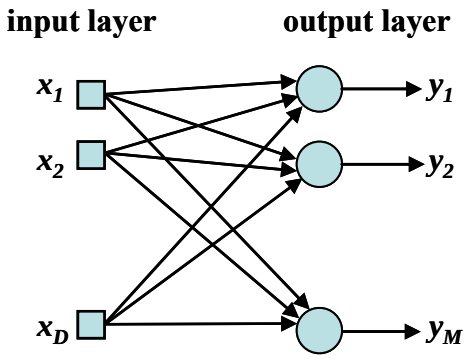


Figure 2.5: Two layer perceptron with D inputs and M outputs (D-M).

The number of outputs is normally determined by the number of classes in the data (Figure 2.5). These PEs add the individual scaled contributions and respond to the entire input space.

There are different Perceptron models and the following ones can be considered as the most significant and consolidated :

- Hopfield model, based on associative memory;

- Rumelhart model, based on Backpropagation networks;
- Kohonen model, based on self-organizing networks.

Connection weights and transfer function

The signals flowing in the connections are scaled by adjustable connection weights W_{ij}. A connection weight is associated with every connection and can have:

- positive value: excitatory connections;
- negative value: inhibitory connections;
- zero: no connection.

Following, taking into consideration an ANN composed by N interconnected neurons, connection weights and activation state of an ANN are briefly described. In the ANN, the neuron i receives input, from the neuron j. This input is composed by the output O_j (a real number) multiplied the correspondent connection weight $W_{ij.}$

Each neuron receives simultaneously a series of signals which can activate or not the neuron producing an output signal. The signal received by the neuron i is $WijOj$. In particular, depending on the value of the signal we can distinguish three situations:

- $0 < W_{ij}O_j < 1$ attenuate function,
- $W_{ij}O_j > 1$ amplifier function,
- $W_{ij}O_j < 0$ inhibitor function.

Considering all the neurons, the total input signal for the neuron i is the sum of the product of each input associated weight by its output:

$$I_{ij} = \sum_j W_{ij} O_j$$

The correspondent neuron i output O_i is influenced by the transfer function T, that describes the output behaviour of the neuron [Mazzetti A, 1991]:

$$O_i = T(I_i)$$

The weights can be implemented in a square weight matrix or in a weight vector.

Activation function

A trained ANN has to be able to generalise, in other terms, it should produce the correct output for given inputs, that belong to the same class, but were not used for training. The input-output relationship is reconstructed by the activation function A. The input-output relationship can be described as follows:

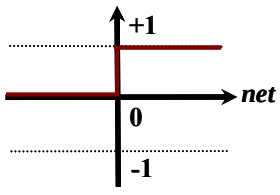
$$Y = A(X) \qquad (2.4)$$

Where X and Y are respectively the input and output network and A is the activation function.

Before the training of the network is tested, the neuron i is not activated and its output O_i is zero. During the training phase, the network is submitted to the presentation of a casual sequence of the training patterns. One complete presentation of the pattern to the network is called epoch. In this phase, connection weights W_{ij} are modified. Each neuron is then activate in different ways and can propagate signals to other neurons. The output Y correspondent to the input X is finally determined by the activation function A of the network [Mazzetti A, 1991].

The activation function controls the amplitude of the output of the neuron. The choice of the activation function can considerably change the behaviour of the network. The most popular activation functions for perceptron are described below:

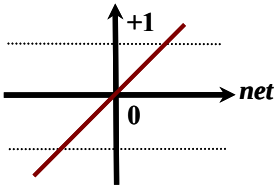
- *Hard Limit activation function*: it limits the output of the neuron to either 0, if the net input argument n is less than 0; or 1, if n is greater or equal to 0. This function is used in Perceptrons, to create neurons that make classification decisions. In the Neural Networks Toolbox of Matlab 7.1, the *hardlim* function realizes the mathematical hard-limit activation function. Graph (Figure 2.6) and algorithm (2.5) are expressed below:



$$f(net) = \begin{cases} 1 \ for \ net \geq 0 \\ 0 \ for \ net < 0 \end{cases} \qquad (2.5)$$

Figure 2.6: Hard Limit activation function.

- *Linear activation function*: the neurons activated with a linear activation function are used as linear approximators in Linear Filters. In the Neural Networks Toolbox of Matlab 7.1, the *purelin* function realizes the mathematical Linear activation function. Graph (Figure 2.7) and algorithm (2.6) are expressed below:

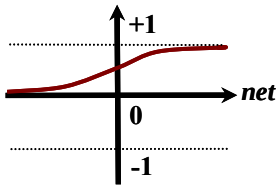


$$f(net) = \alpha \cdot net \qquad (2.6)$$

Figure 2.7: Linear activation function.

The linear activation function is used in the framework of this work.

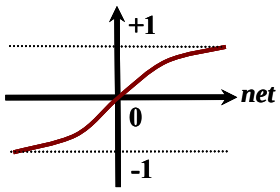
- *Logarithmic sigmoid activation function*: this function takes the input, which can have any value between plus and minus infinity, and squashes the output into the range 0 to 1. This activation function is commonly used in Backpropagation networks because it is differentiable and continue. In the Neural Networks Toolbox of Matlab 7.1, the *logsig* function realizes the mathematical logarithmic sigmoid activation function. Graph (Figure 2.8) and algorithm (2.7) are expressed below:



$$f(net) = \frac{1}{1 + e^{-\alpha \cdot net}} \qquad (2.7)$$

Figure 2.8: Logarithmic sigmoid activation function.

- *Hyperbolic tangent sigmoid activation function*: it takes the input, which can have any value between plus and minus infinity, and squashes the output into the range -1 to +1. This activation function is commonly used in Backpropagation networks because it is differentiable and continue. In the Neural Networks Toolbox of Matlab 7.1, the *tansig* function realizes the mathematical tangent sigmoid activation function. Graph (Figure 2.9) and algorithm (2.8) are expressed below:



$$f(net) = \tanh(\alpha \cdot net) \tag{2.8}$$

Figure 2.9: Hyperbolic tangent sigmoid activation function.

The hyperbolic tangent sigmoid activation function is used in this work, where the range of activation function values was between [−1, 1]. In fact, input and output data were normalized in order to take in consideration such limits. This procedure allows the control of the signals transmitted into the ANN.

2.4 Multi Layer Perceptron (MLP)

The single layer perceptron can only classify linearly separable problems. For non-separable problems, it is necessary to use more layers. The Multi-Layer Perceptron neural network is the most widely known and used neural network architecture in hydrogeology problems. Conceptually, a Multi-Layer Perceptron (MLP) is a Feedforward neural network with is a cascade of Perceptrons.

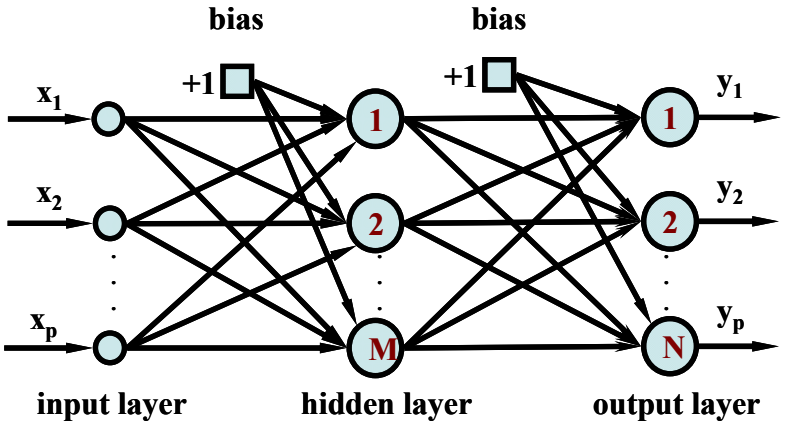


Figure 2.10: Multi layer perceptron (MLP).

A Feedforward MLP is a kind of neural network where neuron can be numbered, in such a way that each neuron has weighted connections. The neurons are partitioned into layers and those can be numbered in such a way that the nodes in each layer are connected only to nodes in the next layer.

The partition of layers consists in three or more layers: an input layer, an output layer and one or more hidden layers with non linear PEs. The hidden layers are not directly connected to the outside world.

2.4.1 How does a Multi Layer Perceptron work?

This paragraph explains the functioning of a MLP. The MLP networks are used in this work.

Taken into consideration N interconnected neurons, X and Y respectively input and output of the examples patterns (composed by real numbers). The best set of connection weights are those that ensure a low error between output calculated and output desired. In this work, the cost function used is the mean squared error performance function (2.9).

$$E = \frac{1}{N}\sum_{i=1}^{P}(y_i - t_i)^2 \qquad (2.9)$$

Where:

- $Y=(y_1, y_2, \dots, y_P)$ are the calculated output;
- $T=(t_1, t_2, \dots, t_P)$ are the desired output called target.

From a mathematical point of view, ae MLP network realizes a non-linear combination of the components of the inputs vector:

$$y_k(x) = f\left(\sum_{i=1}^{m} w_{ki}\, g\left(\sum_{j=1}^{n} v_{ij}x_j + \vartheta_i\right) + \vartheta_k\right) = f\left(net_k\right)$$

Where:

- y_k are the kth value of the neuron in the output layer;
- m = number of the neurons in the hidden layer;
- n = number of the neurons in the input layer;
- x_j: = jth component of the input vector;
- w_{ki} are the connection weight in the hidden layer;
- v_{ij} are the connection weight in the input layer;
- θ_i and θ_k are the bias values;
- f and g are the activation functions.

The construction of a network can be made on the basis of Cybenko's theorem [Cybenko, 1989].

The Cybenko's universal approximation theorem states that a MLP network (composed of single hidden layer containing finite number of hidden neurons with sigmoid activation functions) is a universal approximator. In other words, a MLP can approximate any continuous function arbitrarily if the number of hidden neurons is sufficiently large. One hidden layer MLP can create local regions in the input space. Another layer with several PEs can be thought of as combining bumps in disjoint regions of the space. This is a very important property, because in the theory of function approximation there are well established theorems that state that a linear combination of localized bumps can approximate any reasonable function. Therefore an MLP with two hidden layers is also a universal approximator, which means that it can realize any input-output map, just as the one hidden layer MLP can [Principe, 2000].

2.5 Artificial Neural Network learning rules

Learning methods can be described in two different categories:

- supervised learning;
- unsupervised learning.

In the case of supervised learning, the network is trained by providing it with couples of input and target patterns. This input-target pairs can be provided by an external teacher, or by the system which contains the neural network (self-supervised). This kind of learning is also called learning with a teacher, since a control process knows the correct answer for the set of selected input patterns. During the supervised train, the connection weights are changed in order to minimize the error between the target and the network output. Figure 2.11 represents the scheme of supervised learning.

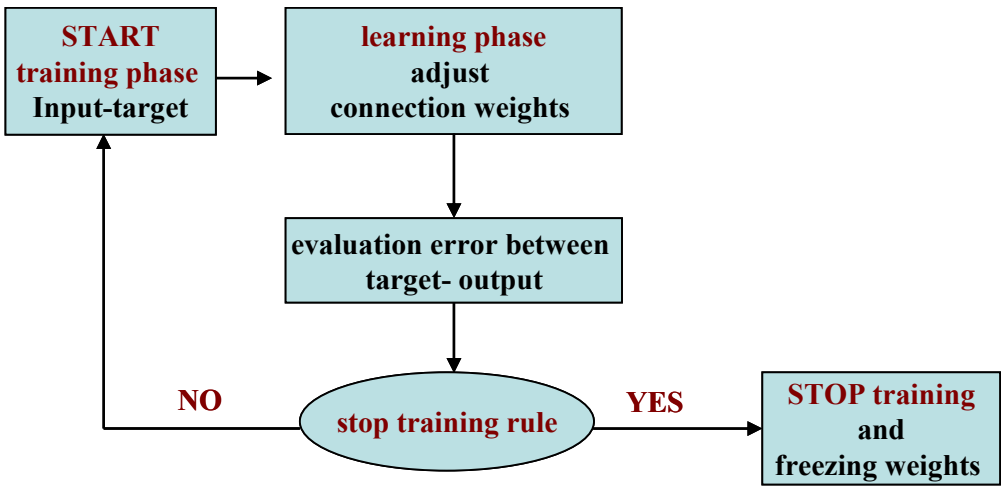


Figure 2.11: scheme of supervised learning.

Unsupervised learning is used when, for a given input, the exact numerical output that the network should produce is unknown [Rojas, 1996]. In unsupervised learning an output unit is trained to respond to clusters of input patterns. In this paradigm the system is supposed to discover statistically salient features of the input population. Unlike the supervised learning paradigm, there is no *a priori* set of categories into which the patterns are to be classified; on the contrary the system must develop its own representation of the input stimuli. This kind of network is usually used when it is not possible to apply supervised learning. Figure 2.12 shows the scheme of unsupervised learning.

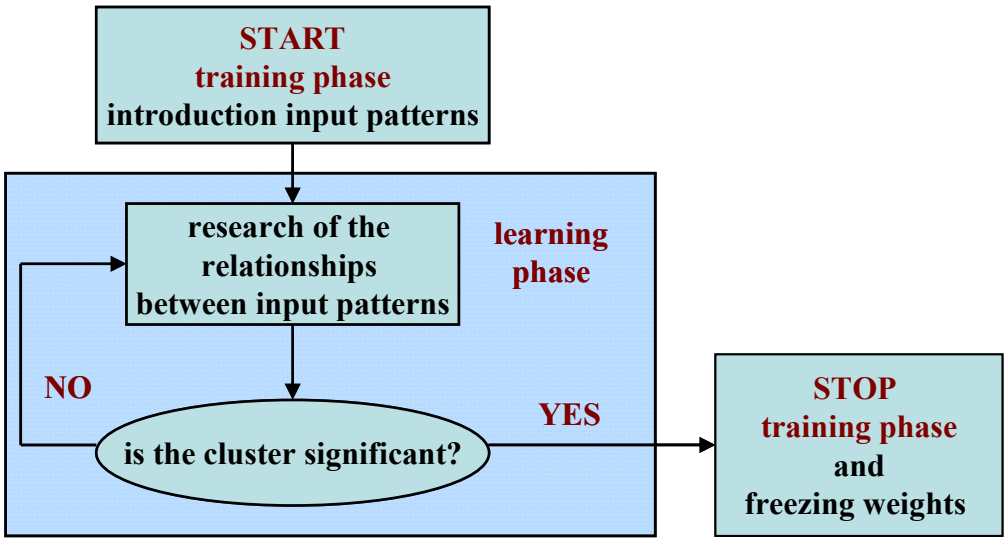


Figure 2.12: scheme of supervised learning.

The main difference between these methods is that in the unsupervised learning the network finds itself the relationship between the input, while in the case of supervised learning, the network is trained to learn the input-output relationship.

2.6 The supervised learning problem

The learning problem consists of finding the optimal combination of weights accordingly that the network function f approximates a given function F as closely as possible. However, we are not given the function F explicitly but only implicitly through some couples of input-output examples [Rojas, 1996].

To consider the learning problem, the attention have to focus on the choice of [Ingrassia S. and Davino C., 2002]:

- *the error function:* called also the cost function, it measures how far away a particular solution is from an optimal solution to the problem to be solved.

- *the learning algorithm*: it is the algorithm that minimizes the error function. Learning algorithms search through the solution space to find a function that has the smallest possible cost.

In this work have been chosen the mean squared error performance function cost and the Levenberg-Marquardt learning algorithm.

2.6.1 Supervised learning algorithms

Error Back Propagation (EBP), or propagation of error, is the most common learning algorithm for training MLP networks. In this work the Error Back Propagation algorithm is used.

The EBP algorithm looks for the minimum of the error function in weight space, that is used to find a local minimum of the error function. The combination of weights which minimizes the error function is considered to be a solution of the supervised learning problem. Since this method requires computation of the gradient of the error function at each iteration step, the continuity and differentiability of the error function must be guaranteed. The network is initialized with randomly chosen weights. The gradient of the error function is computed and used to correct the initial weights (connection weights are adjustable in the training phase) [Rojas, 1996].

Every time the network makes an error, during the training phase, this error is propagated through the synaptic connections and summed for each unit from which they receive the signal (from input to output layer by layer). In EBP algorithm practice, the error between the calculated output and the desired output, for a particular input state, is propagated backwards through the weights of the hidden layers, until it reaches the input layer (from output to input). The goal is to isolate the influence of each connection weight in the error between calculated output and target (desired output).

This approach allows high efficiency in the achievement of results and a generalization of the solution for unknown new examples. This algorithm may be used for MLP network, with any number of connections and different structures.

In order to minimize the error function, different algorithms may be used:

- the Gradient Descent method;
- the Levenberg-Marquardt method.

Gradient Descent method

The Gradient Descent method is a first-order optimization algorithm. Following is explained the detailed features [Ingrassia, 2002].

Supposed to have an MLP network with logarithmic sigmoid activation function

$$f(net) = \frac{1}{1 + e^{-\alpha \cdot net}}$$

Where α is a positive constant.

If x_k^μ is the input network pattern. h_j^μ is the output in the hidden layer(s):

$$h_j^\mu = f\left(\sum_{k=o}^{n} v_{jk} x_k^\mu \right)$$

and y_i^μ is the output in the output layer:

$$y_i^\mu = f\left(\sum_{j=0}^{m} w_{ij} h_j^\mu \right)$$

The network goal is to minimize the error function E_w:

$$E_w = \frac{1}{2} \sum_\mu \sum_i \left(t_i^\mu - y_i^\mu \right)^2$$

According to the Gradient Descent method, in order to improve the network performance, the parameters vary along the surface of the error in the direction of gradient maximum negative slope.

$$\Delta w_{ij} = -\eta \frac{\partial E}{\partial w_{ij}} = \eta \sum_\mu \left(t_i^\mu - y_i^\mu \right) \cdot f\left(net_i^\mu \right) h_j^\mu \qquad (2.10)$$

Where η is the learning rate that determines the speed by which the network learns. High values of learning rate speed up the learning process but, at the same time, may provoke convergence problems and generate instability.

The 2.10 can be further simplified by placing the product of the sum equal to:

$$\vartheta_i^\mu = \left(t_i^\mu - y_i^\mu \right) \cdot f\left(net_i^\mu \right) \qquad \text{that becomes:} \quad \Delta w_{ij} = \eta \sum_\mu \vartheta_i^\mu h_j^\mu$$

For the lower connections, the variation between intern units must be considered with regard to the synaptic weights :

$$\Delta v_{jk} = -\eta \frac{\partial E}{\partial v_{jk}} = -\eta \frac{\partial E}{\partial h_j^{\mu}} \frac{\partial h_j^{\mu}}{\partial v_{jk}}$$

that becomes:

$$\Delta v_{ik} = \eta \sum_{\mu} \sum_{i} \vartheta_i^{\mu} w_{ij} f\left(net_j^{\mu}\right) x_k^{\mu}$$

Observing these relationships, it is possible to remark that in the activation function of each unit h_i, the sum of products between the upper layer elements ϑ_i^{μ} and the corresponding synaptic weights w_{ij} is required. In other words, the error is propagated backwards from a layer to the previous layer.

The disadvantage of this the method is due to the fact that it is not very fast, and that it requires a lot of computing power. In addition, the nonlinear activation of the nodes (as described above, a logarithmic sigmoid activation function is used) determines a complicated error function. The surface is associated to the error function can have several local minima (that can trap the net function during the training phases) and flat surfaces (that can slow down the speed of learning).

Levenberg-Marquardt method

Levenberg-Marquardt algorithm is used in this work because it is better adapt at the studied cases.

Levenberg-Marquardt is a method that provides a numerical solution to the problem of minimizing a function, generally nonlinear, over a space of parameters of the function. These minimization problems arise especially in least squares curve fitting and nonlinear programming, therefore a sufficient low number of iterations is required [Ingrassia, 2002].

The performance index that has to be optimized is represented by the following equation:

$$F\left(w_N^T\right) = \sum_{p=1}^{P} \left[\sum_{k=1}^{K} \left(d_{kp} - o_{kp}\right)^2 \right]$$

Where w is the vector of all the N weights of the network, d_{kp} and o_{kp} are respectively the desired output value and the real output value for the kth output and for the pth input pattern, P and K represent respectively the number of output network patterns and the number of the input network patterns of the ANN. In matrix notation it can be express by the equation as:

$$F(w) = E^T E$$

Where E represents the matrix of the error for all the patterns.

The weights are commutated through the following equation:

$$w_{t+1} = w_t - \left(J_t^T J_t + \lambda_t I\right)^{-1} J_t^T E_t$$

Where $\lambda_.$ represents a parameter that governs the step size, J is the Jacobian matrix of the m errors in regard to the n network weights. The Levenberg-Marquardt algorithm consequently requires the calculation of the Jacobian matrix of errors at each iteration step.

From a practical point of view, one selection criteria of the learning algorithms can be based on the number of parameters of the error function. Levenberg-Marquardt algorithm is useful if the number of the error function parameters is low. Otherwise expressed algorithms are related to conjugate gradient. In the opposite case, gradient descended algorithms are preferred.

2.6.2 The Overfitting problem and training stop techniques

The critical issue in developing a neural network is *generalization*. In fact, the loss of the generalization capacity rend an artificial neural network unusable. Such situation may be generated by an exasperate training. To avoid such risks, it is useful to find rules that permit the evaluation of the best duration of the training phase.

Learning rules that allow to establish a limit of the training phase assume a key role in order to safeguard the generalization capacity. Typically a threshold value of the error that determines once reached the training stop is established.

Depending on the choice of the threshold, the overfitting event may arise. The overfitting determines an excessive specialization of the network on training examples creating a disadvantage in the generalization ability.

Three useful methods for the training interruption are represented by:

- the *cross validation*;
- the L*eave-One-Out cross-validation* (LOO);
- the *stopped training*.

In this work, the Leave-One-Out cross-validation is used to train an ANN specialized in finding the spatial and temporal coordinates of an unknown source of contamination. The stopped training rule is used for the training of the ANN used to reconstruct the Alsatian aquifer pollution source behaviour.

Each method is explained below.

Cross validation

In the *cross validation* method, the patterns are divided in two different parts. The first part is used to train and build the artificial neural network. The second part is used to verify the performance of the network and for error evaluation. When the error on the second set begins to increase, the training process is interrupted. During the training phase, the training is stopped even if the number of epochs is not reached.

Leave-one-out cross validation

The simplest and commonly used method of cross validation is the *LOO* method. This method is applied especially when the available set of patterns are not very numerous. In the Leave-one-out procedure, the examples pattern is divided in p sets, where p is the number of the couple input/output.

Each set is divided by two subsets:

- one set composed by p-1 examples: used for training the network;
- one set composed by 1 example: used for validating or testing the network.

This subdivision is repeated during p steps: in each steps one different example is left out of the training set (p-1) until all the p are used both for training and validation or test.

If the LOO is used for training and validation during training phase, the training is stopped although the number of epochs provided is not reached. During the training, the error is calculated on the training set; at the same time the error is calculated also on the

validation set, independently from the training set. When the validation error begins to rise, the training process is interrupted.

In this work, as it will be explained in chapter 4, the LOO rule is used for training and testing. As a consequence, during the training phase, the set number of epochs is totally reached. In this case, the LOO is used because the examples pattern is too small. This method allows to extract a maximum of information from all the patterns.

Stopped training

The "Stopped training" method derives from the need of finding a compromise between the attempt to identify as accurately as possible all the examples presented and the generalization ability of the network. This method is used, especially when a large number of examples is available.

The initial pattern is divided in three sets:

- *Training set*: it is a set of data used in the training phase. During this phase, connection weights are adjusted in order to minimize the error function. The network is trained by the interactive presentation of the couples input/output.
- *Validation set*: it is a set of data used in the validation phase. During this phase the connection weights are already adjusted. Consequently, the validation set are used to calculate the error function between calculated output and desired output. The network is validated by the interactive presentation of the examples of the couples input/output. When the validation error begins to rise, the training process is interrupted;
- *Test set*: it is a set of data used in the test phase. During this phase, the connection weights are already adjusted and validated. The test set is used to evaluate the error function value between the calculated output and the desired output. The performance of the network is tested by the presentation of new input examples: the difference between the calculate output and the target is assessed. The goal is to evaluate the generalization capacity of the network considering new examples.

In other words, during the training phase, the connection weights are modified and the error is calculated in order to minimize the error on the training set, but at the same

time the error is calculated also on a validation set, independently from the training set, and as the validation error begins to rise, the training process is interrupted.

In this work, as it will explained in chapter 3, the stopped training rule is used for training, validating and testing. During the training phase, when the validation error begun to rise, although the number of epochs provided were not reached, the training was stopped. After that, the weights were frozen and the network was tested with the unknown example.

2.6.3 Convergence criteria

There are no general convergence criteria to assess the learning capacity of the network. However two simple rules may be adopted:

- evaluate the learning rate for the training patterns by considering the overfitting phenomena, the connection weights and the average error committed;
- evaluate the learning rate for the validation patterns by considering the average error committed for the validation patterns.

2.7 Artificial Neural Network architecture models

Neural networks are composed by a number of interconnected neurons. Based on the network architecture that is structured and on the type of connections, different models may be found. A list of the most popular models is available below:

- *Recurrent networks*: some of the outputs are taken and feed them back to the inputs or to hidden layer neurones;
- *Non-Recurrent networks*: the connection weights have an unique direction from the input to the output (Figure 2.13 - a);
- *completely connected networks*: each neuron is connected with all the over neurons (Figure 2.13 - b);

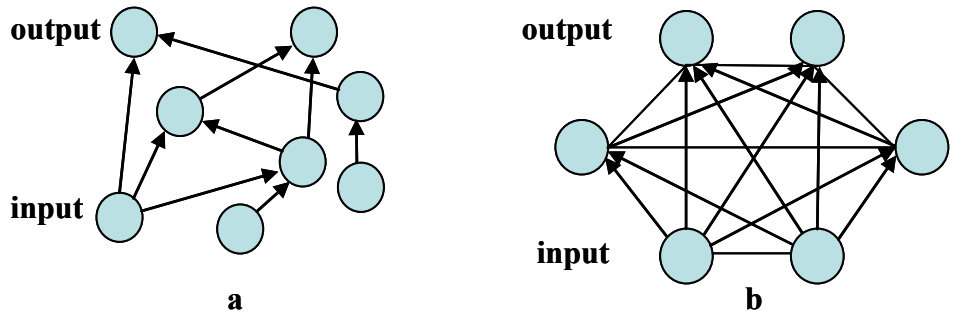


Figure 2.13: Non-Recurrent network (a), completely connected network (b).

- *layered networks*: neurons are organized on speared layers (Figure 2.14 - a);
- *Symmetric networks*: the connection between each two neurons is the same in the two senses (Figure 2.14 - b);
- *auto-associative networks*: input units and output units are the same (Figure 2.14 –c);

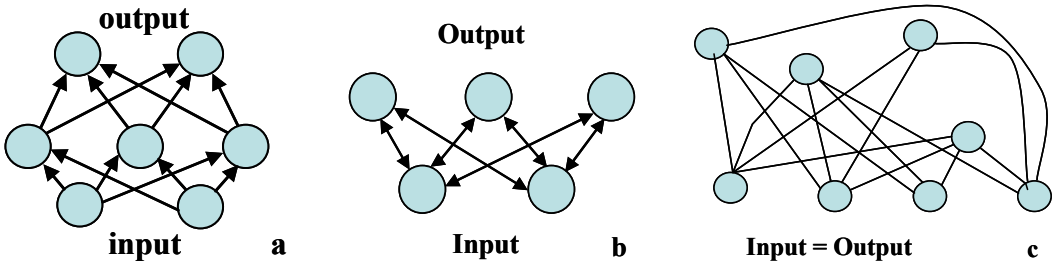


Figure 2.14: layered network (a); symmetric network (b); auto-associative network (c).

- *stochastic networks*: into the network are introduced random variations either by giving the network's neurons stochastic transfer functions, or by giving them stochastic weights. There is some probability that a unit is not activated even when it receives a stimulus;
- *asynchrony networks*: neurons are activated casually one by one.

2.8 ANN modeling approach

In general, in the classical modeling approach, a model is defined as a simplified version of a real system and phenomena. The system is a "transparent box" where internal components and their operation are known. The constitutive laws are used to implement the model of the system and to solve the problem. The classical approach needs precise governing equations and the assumption of hypothetical simplification.

A complex model may yield more accurate results if it describes the real situation thoroughly. However, a complex model usually involves more computation time, makes the analysis more difficult, and often requires more information for its construction. Such situations require more simplifying assumption in order to obtain a manageable solution of the problem. Simplified assumptions and not available governing equations may, in some cases, render the model not reliable.

In other words, ANN modeling approach allows to model a complex environmental phenomena without precise governing equations and with no simplifying assumption. The model device is considered as a "black box" where the internal structure is unknown or inaccessible. Only experimental data are used to derive the system model.

Disadvantages associated to ANN modeling is that the model is specific for the system under consideration and cannot be built *a priori*.

ANNs allow to analyze a physical system where mathematical models are complex or not existent. ANN dissociates itself from the system physical model, i.e. the physical meaning of parameters is totally lost. Based on external data, ANNs are able to build simple algebraic equations that can reproduce the cause-effect relationship of the studied phenomena.

In practice, the main feature of an ANN is the "Generalization Propriety". The ANN, based on a given data set (examples of the phenomena), called training set, tries to build a statistical model able to reproduce the process that has generated the data set. This kind of technology is not programmed but just trained. During the training, the ANN learns to build inside the knowledge necessary to perform the requested task: reproduce the studied phenomena. In this way, the information is processed in a distributed manner by the elementary units, reaching a resistance to noise. In fact, the network is able to operate despite the presence of uncertain data, incomplete or slightly erroneous.

The learning ability is a function of various factors, including the topology of the network itself, the number of neurons, the learning rules and the training patterns.

It is possible to summarize the ANN actions in basic steps:

- learning: phase in which connections weights are calculated based on training patterns (examples of the system or phenomena that we want to model);

- execution: phase during which the ANN trained network calculates the output, taking into account the values of inputs and weights obtained in the previous phase. The ANN is executed with new examples that are not used during the training in order to reproduce the modelled phenomena.

The networks can be implemented in both hardware and software level, by means of special mathematical and statistical computational models. The ANNs models used in this work are developed with the Neural Network Toolbox of MATLAB 7.1.

In the framework of ANN technologies, information are processed in parallel, and input is distributed in a number of different neurons, which contribute at the same time to output production. This parallelism reduces the risk to have irreparable damage in the case of neuron loss when we have architecture with many neutrons and complex ANNs. It may seem strange to simulate parallel networks on sequential computer, however, the use of parallel simulators appear advantageous in terms of computation time during the training phase.

2.9 Designing and training of a Multilayer Perceptron Network

The issues involved in designing and training a MLP are:

- *the problem of "mapping":* it deals with the identification of the *n* input variables;
- *problem of the "threshold":* it concerns the identification of the *m* output variables;

In order to design a MLP architecture, it is necessary to define the following parameters:

- *choose the architecture model*;
- *decide how many neurons should be use in input and output layer* on the basis of the *n* input variables and *m* output variables;
- *choose the number of hidden layers to be use in the network:* for almost all problems, one hidden layer is sufficient. Two hidden layers are required for modeling data with discontinuities. Using two hidden layers rarely improves the model, and it may introduce a greater risk of converging to a local minima. There is no theoretical reason for using more than two hidden layers;

- *decide how many neurons should be used in each hidden layer*: one of the most important features of a perceptron network is the number of neurons in the hidden layer(s). If an inadequate number of neurons is used, the network will be unable to model complex data, and the resulting fit will be poor. If too many neurons are used, the training time may become excessively long, and the network may over fit the data. When overfitting occurs, the network will begin to model random noise in the data. The result is that the model fits the training data extremely well, but it generalizes poorly new and unseen data. Validation must be used to test this;
- *choose the activation function for each layer*: the activation function controls the amplitude of the neuron output. The choice of the activation function may significantly impact the performance of the ANN network;
- *choose the learning rules;*
- *choose the training stop technique and convergence criteria.*

The goal of the training process is to find the set of weight values that reconstruct the input-output relationship. In particular, the weights that may generate the output from the neural network have to match the target values as closely as possible.

During the training phases, it is important to find and converge in an optimal solution that avoids local minima in a reasonable period of time. The last steps before "freezing the weights" (save the trained network) consist in validating the neural network and test it regarding overfitting.

2.10 Summary

In this chapter, the concept of artificial neural networks and the most important features that render this technology attractive in hydrogeology research have been explained. Structure, components and architecture models of artificial neural networks have been detailed. In particular, the supervising learning problems and the useful tools to solve or mitigate them have been discussed. Classical modeling approach is compared with artificial neural network modeling. The designing and training of a Multilayer Perceptron Network has also been addressed.

3 ANN APPLIED TO STUDY A POLLUTED AQUIFER

This chapter is dedicated to a theoretical case of pollution. In particular, this work aims to define a method to identify the spatial location (X,Y) and the duration of the activity (T) for an unknown pollution source based on the measures of contaminant concentration acquired in the monitoring wells at a certain time *t,* which represents the current time.

Groundwater contaminations, in some cases, may result from pollutions whose origins are at times and places different than where the contaminations have been actually found. Such situations require the development of techniques that allow the identification of these unknown pollution sources.

In this chapter we propose the use of the neural network paradigm in an innovative way in order to identify unknown pollution sources. Artificial neural networks are used as a tool to locate the source of a contamination process in a homogeneous and isotropic two dimensional domains. This case takes under consideration the restricted hypothesis of groundwater contaminated by a single pollutant injected in a single point. Training patterns are constructed by simulating hydrogeological scenarios through the use of a non commercial software for flux and contamination transport modelling.

The huge amount of data carried out by each time step of simulation of the domain is not suitable to be inputted in an artificial neural network. Feature extraction techniques have been therefore implemented to reduce data dimensionality. A neural network is trained to capture the functional relationship between the contaminant concentration measured in pumping wells and the position and duration of the pollution source that generated these contaminations. The network is then used to solve the inverse problem consisting in locating the pollutant source, and how many time steps before the event occurred. Results of the study have shown the suitability of the neural approach for extracting hidden relationship between patterns and monitor groundwater resources.

This part of the research is carried out to improve in a innovative way the results of two previous works [Scintu, 2004; Fanni et al., 2002], aimed at analysing if the

ANNs were able to identify the geological domain, the position of a contaminant source and known concentrations of contaminants in the monitoring network.

In Fanni (2002) and Scintu (2004) works, different traditional feedforward, MultiLayer Perceptron (MLP) networks are trained to predict the coordinates of the pollutant source and the time the pollution occurred. In particular three ANNs were trained: one for the time step concerning the duration of the activity (T) and two for geometrical coordinates (X,Y). The inverse problem approach, applied in both works, consists of presenting as input to the neural model the knowledge of the distribution map of the contaminant at different time steps and associating them, as desired output, the geometrical coordinates of the injection wells and the time step before the injection occurred. Therefore the inverse problem is directly solved using the ANNs. These previous works showed very good performances in locating the pollutant source, however worse results have been obtained with time step prediction network (see paragraph 3.5).

In this work, in order to improve these results, a new approach is applied. One ANN is trained to solve the direct problem: presenting as input to the neural model the spatial location (X,Y) and the duration of the activity (T) for the unknown pollution sources and associating them, as desired output, to the measures of contaminant concentration acquired in the monitoring wells at the final time *t* of the numerical simulation, which represents the current time. After the training phases, the trained ANN is inverted in order to solve the inverse problem. Starting from the contaminant concentration in monitoring wells, the unknown contaminant source characteristics have been found. The approach, presented in this chapter, thanks to a drastically reduction of the input/output data to a very manageable size allows a strong computational time reduction. Moreover the implemented method is useful not only to identify the location and duration activity of unknown pollution sources, but also to bound the study area defining best location of the monitoring wells in the domain and optimize the investigation costs.

3.1 Introduction to the implemented methodology

Groundwater pollution sources are characterized by varying spatial location, injection rates and duration of activity. Concentration measurement data from monitoring wells may be utilized to identify these unknown pollution sources in terms

of spatial location and duration of the activity. Therefore, the identification of an unknown groundwater pollution source becomes more difficult in the lack of complete breakthrough curves of historical concentration. If concentration observations are missing over a length of time after an unknown source has become active, it is even more difficult to correctly identify the unknown pollution source.

An artificial neural networks based methodology has been developed to solve the inverse problem for such a missing data scenario when concentration measurement data is not available for the entire duration of the pollutant source activity. In the studied case the worst situation is taken into consideration: in particular, one single value of concentration measurement at the current time t is available.

Data for training the ANN are simulated using a groundwater flow and contaminant transport numerical simulation model. A generic conservative pollutant is considered. Contamination observed in monitoring wells result from a single source with a constant injection rate and different release periods.

The model implemented has been chosen to be very simple: the purpose of this research is to explore the potentialities of the artificial neural network methodology for solving the inverse problem of unknown pollution source position estimation and not to actually apply this methodology to address a given real problem.

In a first step, an artificial neural network was trained to solve the direct problem. In this part of the application, the networks were trained, by means of examples, to recognize the pollution sources position and duration of activity corresponding to the output contaminant concentration in monitoring wells. The input patterns were made of the pollution source features in terms of spatial position and activity duration. The output patterns were contaminant concentration observation data at given monitoring wells. After the training, the network's generalization properties can be exploited to estimate the contaminant concentration data in monitoring wells corresponding to new pollution sources.

In a second time, the trained network was inverted in order to solve the inverse problem. On the basis of the known contaminant concentration data in monitoring wells, the pollution sources position and the duration of the activity can be identified.

In order to train the ANNs, it is necessary to generate a consistent data set of patterns for training and test. To generate the patterns, the numerical flux and

contaminant transport simulation software TRACES (Transport of Radio ACtive Elements in the Subsurface) has been used.

As documented by Hoteit et al.(2004), TRACES is a computer program for the simulation of flow and reactive transport in saturated porous media. It is written in FORTRAN 95 and is portable to different platforms. TRACES handles transient and steady state computations in 2D and 3D heterogeneous domains. It is based on mixed and discontinuous finite element methods for solving the hydrodynamic state and mass transfer problems. The code is flexible in describing complicated geometries by using triangles or quadrangles in 2D, and tetrahedrons, prisms or hexahedrons in 3D. Boundary conditions and almost all parameter values can vary in space. A material property index is assigned to each grid element. Boundary conditions, source terms, fluid and porous matrix properties can change with time, based on a user-specified tabular function.

ANN patterns were constructed through a coherent number of hydrogeological scenarios, based on a simple 2D geometry.

The data provided by the model were treated by feature extraction techniques in order to significantly reduce the size so that they could be processed by the ANN.

3.2 Flux an transport model of the studied aquifer

To train and test the ANN, it is necessary to construct a consistent set of training patterns. These training patterns are generated by simulating hydrogeological scenarios using flow and contaminant transport model.

The first step in the modeling procedure consisted in defining the appropriate conceptual model. In a second time, the conceptual model has been expressed in the form of a mathematical model by using TRACES.

The modeling procedure is explained in the following paragraphs.

3.2.1 Conceptual model of the theoretical aquifer

To study the phenomena generated in aquifers contaminations, it has been necessary to build a conceptual model to organize field data and assess how these data are translated into a physical or mathematical model. The construction of the conceptual model implies a number of simplifying assumptions such as the definition of the basin

boundaries, initial conditions, boundary conditions, recharge and discharge sources, pollution sources, hydraulic heads and the hydrochemical patterns.

The theoretical hydrogeological basin and its principal features have been defined as reported in Table 3.1.

THEORETICAL AQUIFER	
Aquifer type: confined and isotropic aquifer system composed by one horizontal layer characterized by only one stratigraphic unit whit a constant thickness. It is delimited by no-flow boundaries on the North and South sides.	
Domain dimension	1000* 1000m^2
Hydraulic head on the west boundary	9 m
Hydraulic head on the east boundary	8 m
Horizontal hydraulic conductivity [k_o]	0.0001 m/s
Effective porosity	10%
Pumping well	1
Pumping rate of the well (the pumping starts from the beginning of the simulation)	0.0012 m^3/s
Constant punctual pollution source concentration	100 $\mu g/m^3$.

Table 3.1: theoretical aquifer features.

No recharge rate is applied to the aquifer. The initial contaminant concentration, in the domain, is assumed equal to zero.

Totally, 40 constant punctual pollution sources with a concentration of 100 $\mu g/m^3$ were considered. Each pollution source has different position in the domain.

3.2.2 Numerical model

In order to solve the partial differential equation by the numerical model TRACES, a grid is superimposed over the studied area.

The grid consists of a set of consecutive cells, for which it is necessary to specify the input parameters. The size of the grid can be different to allow a less approximate study of the most significant resolutions and more detailed in areas where greater accuracy is required. The development of the grid, therefore, represents one of the most critical part of modeling because it strongly impacts the numerical solutions obtained. In

this case, a regular quadrangular two-dimensional mesh is imposed in the whole domain (Figure 3.1). The adopted model discretization may provide a sufficient geometric configuration to reproduce the best possible hydrodynamic conditions of the aquifer. In Table 3.2, the model discretization has been summed up.

Model discretization	
Number of X cells	50
Number of Y cells	50
Dimensions of the X and Y cells	20*20m^2

Table 3.2: model discretization.

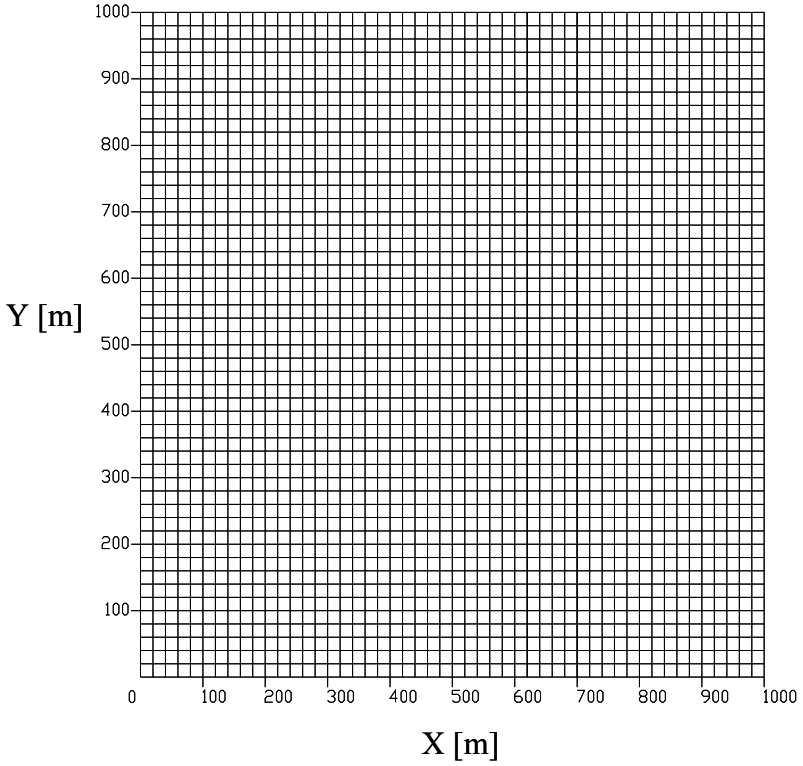


Figure 3.1: two dimensional quadrangular mesh.

Thanks to TRACES software, it is possible to assign in the numerical model the aquifer input features to each cell or cell edge of the grid. Mixed and discontinuous finite elements methods, used by TRACES, ensure exact local mass balance, handle high parameter discontinuities between adjacent elements, and treat full tensors without approximation. The mathematical model describing the flow in the porous material is solved by the mixed hybrid finite element method. The transport equation is split in 2 parts, where the advective part is solved by discontinuous finite element and the rest by

the mixed hybrid finite element method. Discontinuous Galerkin finite element can solve advective dominant transport without oscillations and with very limited numerical diffusion [Hoteit et al, 2004].

Through the application of the code, the trend of the piezometric head and contaminant concentration in the domain in stationary state was developed. The phenomenon has been studied assuming that the contamination happened from a single cell (pollutant source) with a single generic contaminant. It is also assumed the presence of a pumping well with a constant pumping rate and no variation of the initial parameters of the model during the simulation time.

Figure 3.2 shows hydraulic head and contaminant concentration distribution for a generic pollutant source after 10 years activity at the top of the aquifer domain.

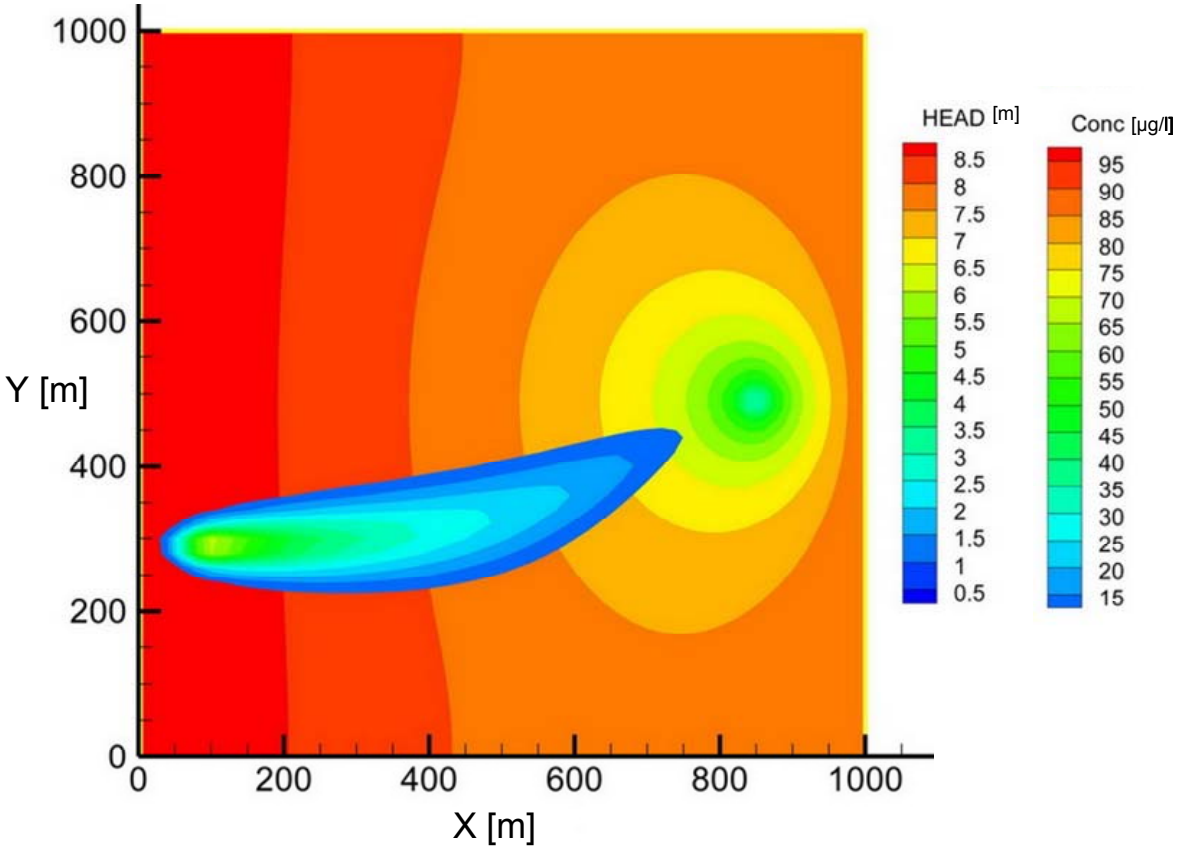


Figure 3.2: hydraulic head and solute concentration in the domain.

3.3 ANN pattern construction: elaboration and reduction

The main source of information about the contaminant movement and accumulation in aquifers comes from measurements of solute concentrations in

monitoring wells. In fact, these are the parameters used in this work to solve the inverse problem in order to identify the unknown pollution sources.

The ANN patterns have been constructed by the simulation of 40 different hydrogeological scenarios. In order to uniformly cover the entire basin area, the 40 sources were located at different positions of the domain (Figure 3.3). For all 40 sources, three activity source duration have been considered. In particular the timing of activity of the sources were 10, 20 and 30 years. So for each of the 40 scenarios, three different durations of the activity have been considered.

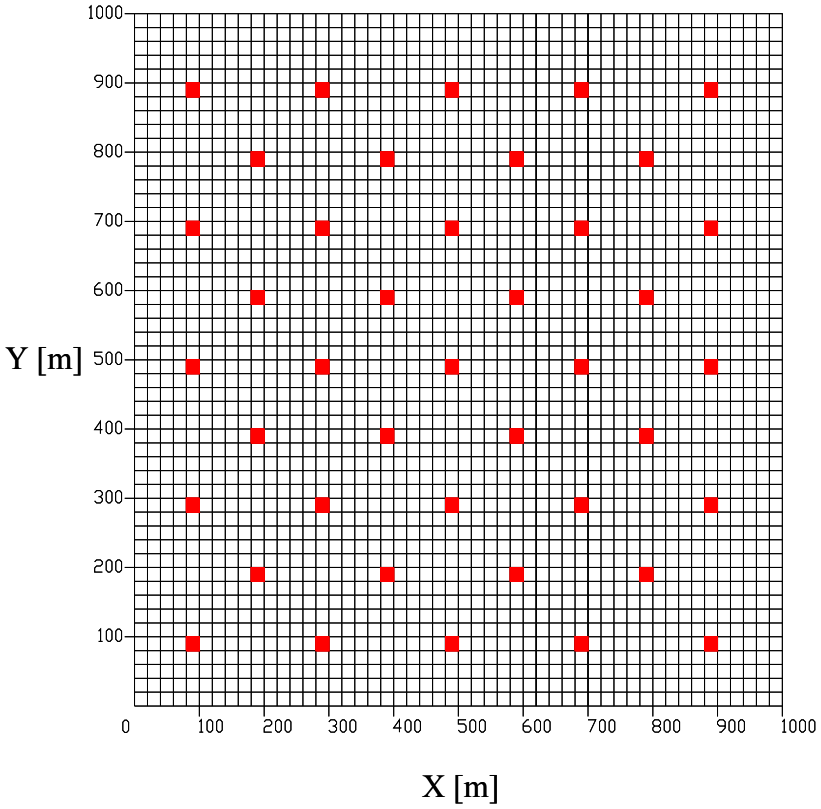


Figure 3.3: distribution of the 40 sources in the domain.

In total, 40 different initial source locations for the 3 durations were considered, resulting in 40*3=120 sample maps of contaminant distributions. Each source of pollution has been assigned 100 $\mu g/m^3$ of contaminant concentration.

The samples obtained from the simulation model are the matrix of contaminant concentration for 50 cells distributed in order to cover the entire basin area in the domain (Figure 3.4).

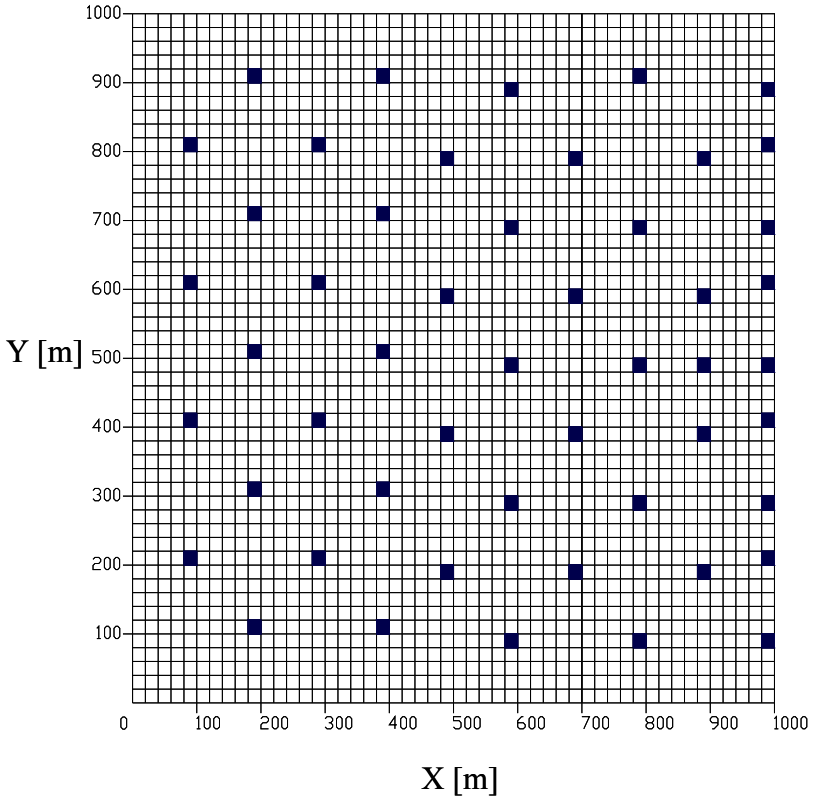


Figure 3.4: distribution of the 50 cells for the contaminant concentration in the domain.

In the matrices, rows represent the time and columns represent the value of concentration in the 50 cells.

Matrices were too large to be processed through ANN, requiring too many examples of inputs and thus a large network quite difficult to handle. For these reasons, it has been necessary to perform a data pre-processing aimed at reducing the matrices size. Feature extraction is a practice commonly used in ANN applications.

Different neural models have been developed depending on data reduction used in order to choose the best feature extraction procedure.

In the following paragraph, the feature extraction techniques needed to reduce the data dimensionality and to limit the number of free parameters of the ANN are detailed.

3.3.1 Input data reduction

The input data for the ANN were the positions (X,Y) and the duration of the activity (T) of the pollution sources for the 120 hydrological scenarios developed. In particular, 40 couples of coordinates for a total of 120 simulations. The three input parameter (X,Y,T) have been pre-processed by normalizing so that they fall in the

interval [-1,+1]. In the Matlab neural network toolbox, input data are normalized through the command *premnmx*. The algorithm is presented in Equation 3.1.

$$pn = \frac{2 \cdot (p - \min p)}{(\max p - \min p) - 1} \quad 3.1$$

Where:

pn is the matrix of normalized input vectors,

p is the matrix of input (column) vectors,

max p is the vector containing maximums for each p ,

min p is the vector containing minimums for each p .

The pre-processed input pattern matrix had size 3*120.

3.3.2 Output data reduction

TRACES simulation samples have been saved as an ASCII matrix file. We have one file for each examined scenario. Each matrix component, $a_{i,j}$, represents j concentration value of the contaminant for time i.

In total, the output data consist of 120 matrices regarding the concentrations of the contaminant in the domain associated with 120 different sources. The 120 matrices had size [m,n], where the rows represent the time and columns represent the value of concentration in the 50 cells. In order to make these data useful for the ANN, it is necessary to reorganize the 120 matrices to form one big unique matrix for all the ANN output.

Several data pre-processing may be used in such reduction cases. The choice is strictly linked to the chosen ANN approach. In our case, the total absence of complete breakthrough curves of concentration time series at all the time steps was hypothesized. So, for each one of the 50 cells, only one concentration observation data are taken into consideration, in particular the concentration of the final time t was taken.

The following scheme (Figure 3.5) present the pre-processing to select the most significant components for the monitoring wells selection procedure used for the output pattern matrix construction.

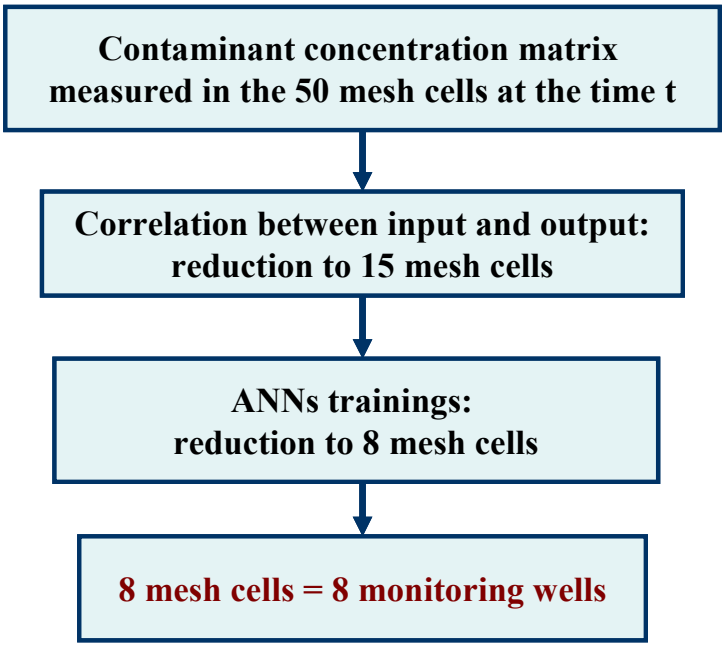


Figure 3.5: schema of the monitoring wells selection procedure.

For the output data reduction, the restricted hypothesis of the total absence of historical data concerning the aquifer pollution was taken into consideration, in particular, the case of a contamination detection for the first time in a generic domain.

Only the last value of contaminant concentration in the time, obtained through simulations, were considered for the 50 cells. These 50 cells may correspond to 50 hypothetical monitoring wells. At this point, the matrices became vectors composed by 50 elements. The 120 vectors were joined to make an unique matrix of output patterns. This matrix had dimension 50*120.

However, these vectors were too large to be subsequently processed through the ANN, requiring too many examples and a large network with a lot of hidden neurons. In this way, the ANN becomes too big and it may loose its specific feature consisting in the calculation speed. Moreover, the number of 50 hypothetical wells was too large for a small domain such as that taken into account.

To reduce the number of hypothetical wells, a procedure has been developed that permits to identify the minimum number of monitoring wells needed to solve the problem directly using the ANN.

Firstly, it was determined which of the 50 cells/wells were less related to the sources. In practice, it was calculated the correlation coefficient between the 120 sources and the corresponding 120 concentrations vectors for the 50 cells. The 120 cases were considered one by one and the matrix of correlation coefficients between the

sources and the 50 cells were calculated. For each source a classification have been drawn and just the first five cells better correlated for each sources were kept into consideration.

Based on this initial reduction, 15 cells were kept (Figure 3.6). So the output patterns matrix size becomes 15*120.

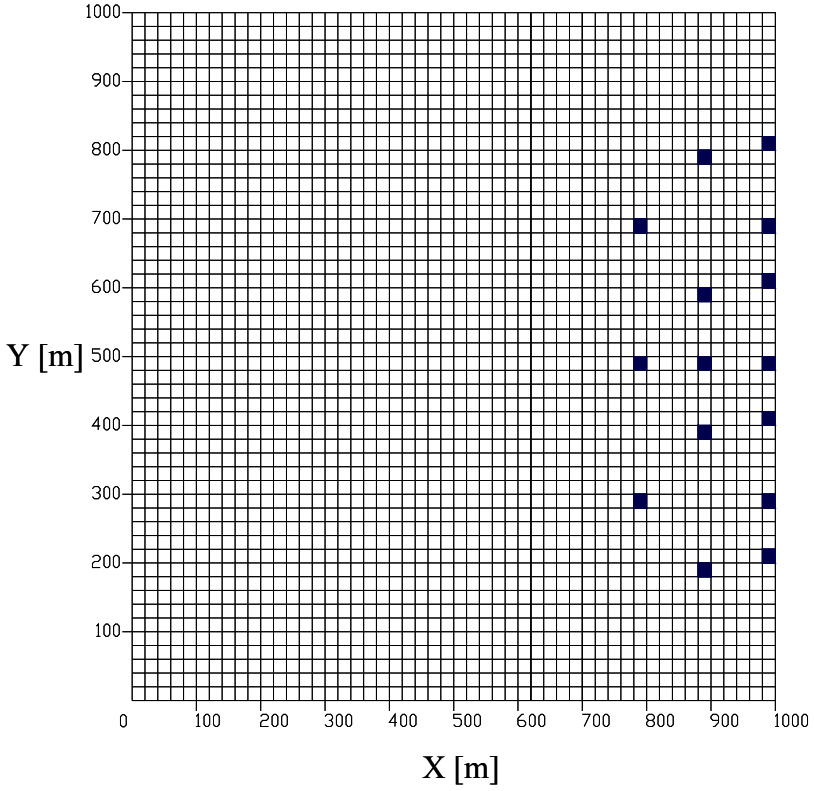


Figure 3.6: distribution of the 15 cells or hypothetical monitoring wells.

The number of wells was still too large determining high investigation costs. In order to further reduce this number, an interactive procedure based on the application of a ANN was developed.

ANNs have been over-trained with the sole training set composed by the all patterns. After each session, the cell (hypothetical monitoring well) that was less linked to the sources was removed from the output patterns. So the matrix became smaller.

The neural network trained for the additional reduction of the number of wells was a traditional feedforward Multi Layer Perceptron (MLP) composed by 3 layers:

- input layer composed of 3 neurons corresponding to coordinates XYT,
- hidden layer composed of 15 neurons

- output layer composed of 15 neurons corresponding to the 15 cells.

The number of the hidden neurons is determined by means of a trial and error procedure performing several trainings assuming a growing number of neurons in the hidden layer.

Once defined the network features, it was over-trained in order to reduce the error between calculated output and desired output, evaluating at the same time the gradient evolution (using the method of gradient descent paying attention to local minima). The network is trained with the iterative presentation of a whole set of patterns called "epoch" (input and output pairs). During the training the consecutive steps of 100 epochs were exercised until the error stop to decrease. At the end of the training phase, the calculated outputs and desired outputs have been compared. The worst results were removed from the output pattern matrix.

For each training phase, the output pattern matrix became smaller, reducing the number of hidden and output neurons. At the same time, the number of hidden neuron is reduced for keeping the same number of hidden and output neurons.

The ANN was trained in order to assess and preserve the learning capacity of the network. This procedure was carried out until the minimum number of monitoring wells necessary to solve the direct problem was found.

Finally, the minimum number of monitoring wells has been set to 8 (Figure 3.7). A further reduction of the number may cause a deterioration of learning ability and generalization capacity of the network.

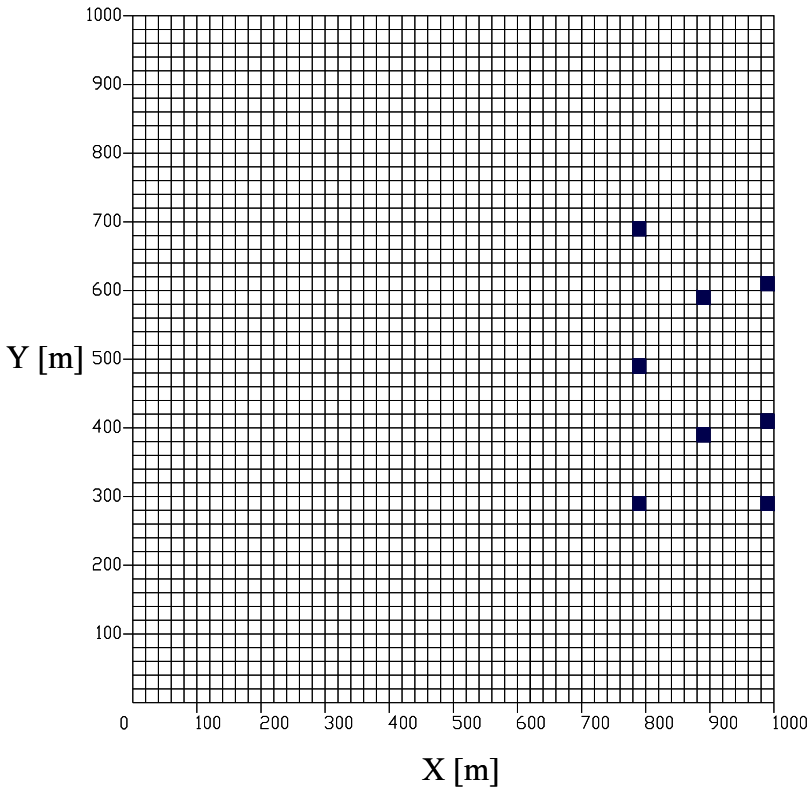


Figure 3.7: distribution of the 8 cells or monitoring wells.

As a consequence, the output pattern matrix is composed of contaminant concentrations at 8 monitoring wells for the 120 simulations. This matrix has been pre-processes by normalizing so that the components fall in the interval [-1,+1] (see Equation 3.1). The pre-processed output pattern matrix has size 8x120.

3.4 MLP networks development and inverse problem solution

The ANN was initially trained to solve the direct problem: starting from the time-space coordinates of the pollutant source, the value of the contaminants concentration in monitoring wells has been reconstruct. The trained network was subsequently inverted to solve the inverse problem: starting from the measurement of contaminants concentration in monitoring wells, the time-space coordinates of the unknown pollutant source have been found. Figure 3.8 shows the scheme of the ANN implementation.

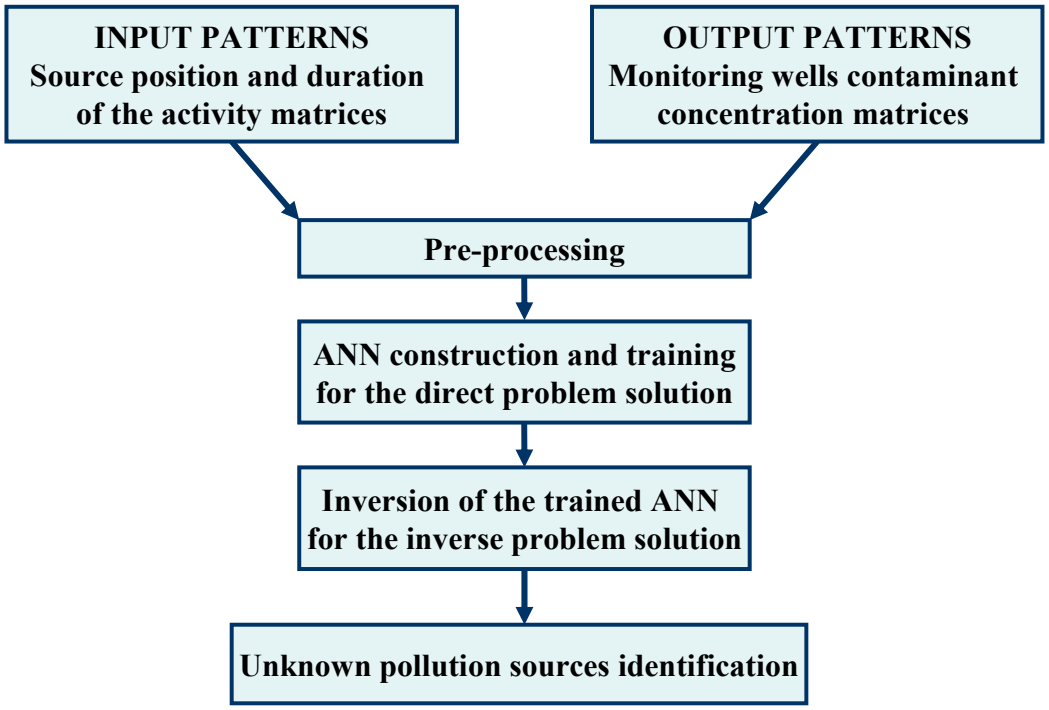


Figure 3.8: schema of the applied procedure.

3.4.1 MLP networks development

The most important features that should be defined for the artificial neural networks construction are: the learning method, the network architecture, the activation functions for each layer, the learning algorithm and the learning rule.

A 3 Layer Perceptron (MLP) architecture was trained with the supervised learning by furnishing couples of input and target patterns (desired output). The network was formed of one input layer, one hidden layer and one output layer. The input layer is composed of 3 neurons corresponding to the two spatial coordinates X, Y and time T. The output layer is composed of 8 neurons corresponding to the contaminant concentrations in the 8 monitoring wells. Once the input and output layer dimension has been defined, it is necessary to set the number of hidden neurons. Usually, this number is determined by means of a trial and error procedure, so that several trainings are performed assuming a growing number of neurons in the hidden layer. In our case, we have taken the same number of neurons for the output layer in the hidden layer. As a result, the hidden layer is made of 8 neurons.

The functional dependence between input and output is defined by hyperbolic tangent activation function for the hidden layer and linear activation function for the output layer.

The learning algorithm chosen is the Error Back Propagation (EBP) algorithm optimized by the Levenberg-Marquardt (LM) algorithm. The LM algorithm gave the best performances in terms of Mean Square Error (MSE).

To ensure satisfying sample generalization performances, we have used a Leave one Out Cross Validation (LOO) learning rule.

The training of the ANN is a critical part of the proposed process. In fact, a special attention has been dedicated to train the ANN in such a way that it is able to generalize the information contained in the training set. To this end, during the training phase the connection weights were modified in order to minimize the error of the training set.

Once the above-mentioned key features of the artificial neural network are selected, it may be trained. During supervised training, in our case, the connection weights are changed in order to minimize the error between the target and the network output. It is proceeded with the presentation of a whole set of patterns (input and output pairs). The set of patterns is called “epoch”. The training is based on the iterative presentation of the epochs according to a random sequence of the patterns.

During the training, thanks to the EBP algorithm, the error between the calculated output and the desired output, for a particular input state, is propagated backwards through the weights of the hidden layers, until it reaches the input layer (from output to input). The goal is to isolate the influence of each connection weight in the error between calculated and target.

On the basis of the results obtained during all the training trials, the appropriate number of training iterations (epochs) is assumed at 100.

The LOO learning rule is generally applied when the available number of examples patterns is limited. In this procedure, the examples pattern is divided in p sets, where p is the number of the couple input/output patterns. Each set is divided in two subsets: one composed of p-1 examples is used for the training and the other one composed of 1 example that may be used for validation or test. In our case, the training set is made of 119 couples of input-output patterns and the test set is composed of one couple of input-output patterns. One by one, each couple is part of the training set and of the test set. In this procedure, the validation set was not considered because we

wanted to exclude the stop training process during the training procedure. This method allows to extract a maximum of information from all the patterns.

A consequence of the validation set absence is that a number of epochs have to be set. The selected epochs are totally reached during the training phase. Accordingly, in a preliminary training, an increasing number of epochs have been considered by taking steps of 100 epochs for each trial. These trials allowed to understand that 100 epochs were enough. An additional number of epochs did not improve the ANN. In addition, this arrangement was successful in terms of reducing computation time. A further increase of the epochs generated a negative cost-benefit ratio, in fact, the overall improvement of the results was not sufficient to justify an increase of computing time. Totally, 100 epochs have been set for the training phase.

A consequence of using the LOO learning rule is the need of training a number of networks equal to the number of examples couples, allowing a coherent comparison between the results.

In total 120 networks, based on the application of the LOO learning rule, were trained with different training set and test set. Each network was trained with 120 examples and tested with one example (kept out from the original training set). The test example allows us to estimate the network generalization capacity.

In our case, the LOO procedure is not used to train the network that will be used in a particular case, but only to estimate the generalization ability of the 120 trained networks. If one wants to consider a new source not included in the 120 patterns, all the patterns will be used for the training set and the new case will be used for the test set.

The developed methodology allows us to reach the reasonable presumption that the error for the new case will not be greater than the errors experienced in the 120 networks already trained.

3.4.2 Solution of the inverse problem with the MLP networks

The networks were initially trained to solve the direct problem, as explained before. The connection weights have been frozen after the end of each training session. For each case, during the training the mean square error and the gradient progress were considered.

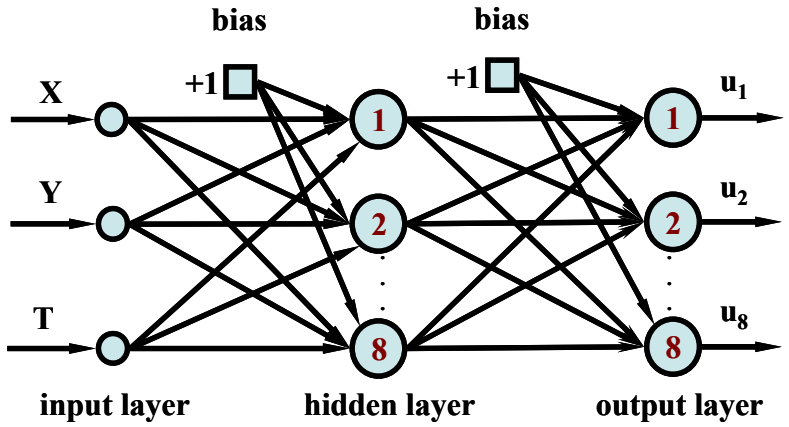


Figure 3.9: generic artificial neural network trained with the LOO.

Figure 3.9 show one or the ANNs trained that calculate the a relationship between input and output patterns.

The ANNs trained demonstrate a relationship between input and output patterns described by the following algebraic equations systems:

$$\begin{cases} \text{Input layer} & \underline{\underline{W}}_1 \cdot \underline{x} + \underline{b}_1 = \underline{y} \\ \text{Hidden layer} & \underline{h} = \sigma(\underline{y}) \\ \text{Output layer} & \underline{\underline{W}}_2 \cdot \underline{h} + \underline{b}_2 = \underline{u} \end{cases} \tag{3.2}$$

Where:

$\underline{x}$ is the input of the network,

$\underline{\underline{W}}_1$ is the weights matrix of the input layer,

$\underline{b}_1$ is the bias vector of the input layer,

$\underline{y}$ is the input of the hidden layer,

$\underline{h}$ is the output of the hidden layer,

$\sigma(\cdot)$ is the hidden neurons logistic activation function,

$\underline{u}$ is the output of the network,

$\underline{\underline{W}}_2$ is the weights matrix of the output layer,

$\underline{b}_2$ is the bias vector of the output layer.

Once the training phase is completed, meaning that all the weights have been determined, the inversion of the network can be performed. The trained networks were inverted to solve the inverse problem. On the basis of the known output of the system, which derives from a set of measurements in the monitoring wells at a certain time t, the corresponding input can be calculated exploiting the method described in [Carcangiu et al, 2007; Fanni et al, 2003].

During the inversion process, explained below, the difference between the calculated input and the desired input was considered.

On the basis of the third equation described in (3.2), starting from the output u, the vector h can be determined.

Provided that the matrix $\underline{\underline{W}}_2$ is full rank, the solution corresponding to the minimum sum squared error can be found as:

$$\underline{h} = \underline{\underline{W}}_2^{-1} \cdot \left(\underline{u} - \underline{b}_2\right) \qquad (3.3)$$

The second equation in (3.2) states a biunivocal relation between y and h, therefore the vector y can be calculated as:

$$\underline{y} = \sigma^{-1}\left(\underline{h}\right) \qquad (3.4)$$

Finally, provided that the matrix $\underline{\underline{W}}_1$ is full rank, the input pattern x can be calculated as:

$$\underline{x} = \underline{\underline{W}}_1^{-1} \cdot (\underline{y} - \underline{b}_1) \qquad (3.5)$$

The desired source position and duration of activity have been obtained by applying backward the pre-processing of the input data obtained from the inversion of the ANN. In the following paragraph results of the ANN inversion are explained.

3.5 Performance evaluation and conclusions

In general, the results showed very good performances in locating the pollutant source. Less satisfying results have been obtained with time step prediction.

The error made by all the networks has been calculated considering the difference between the real position and the simulated position of the pollutant sources.

Figure 3.10, Figure 3.11 and Figure 3.12 show the hydrogeological domain with spatial coordinates X and Y corresponding to the 40 pollution sources positions. The blue circle represents the input patterns corresponding to the pollution sources position used during the training and the red x represents the simulated input patterns corresponding to the pollution sources position after the inversion of the 120 trained ANNs. Each of the three figures shows the results of the inversion of the 40 networks for the three time steps (10, 20 and 30 years).

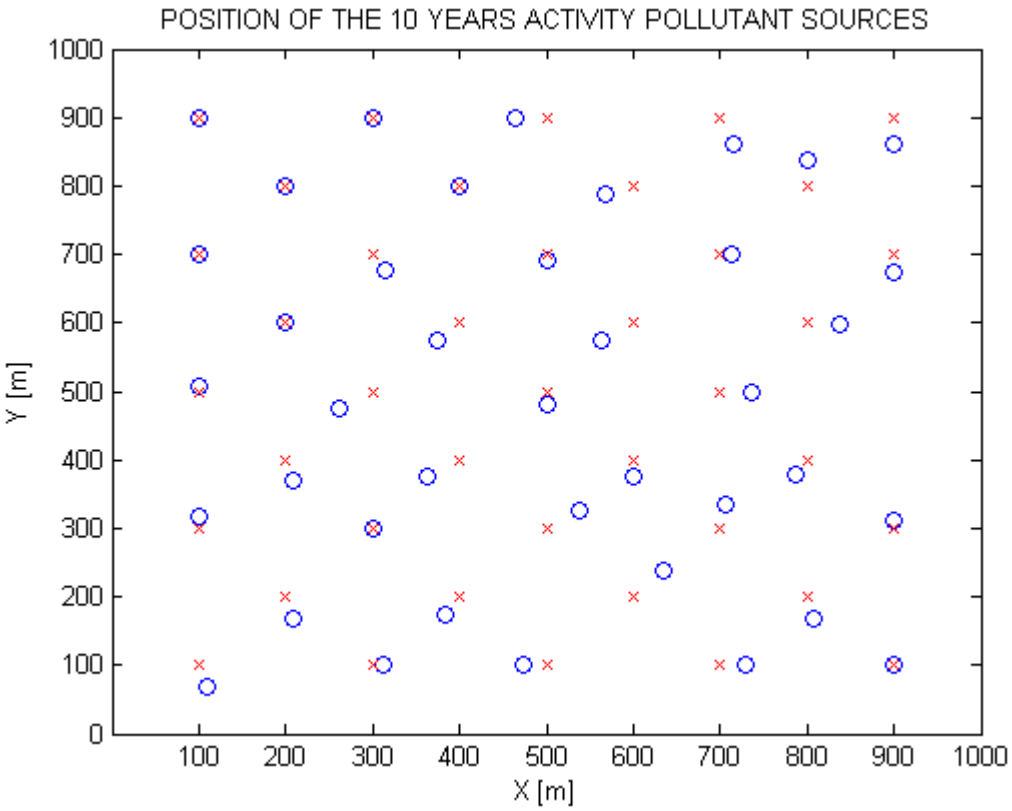


Figure 3.10: real (red x) and simulated position (blue circle) of the 10 years activity pollutant sources.

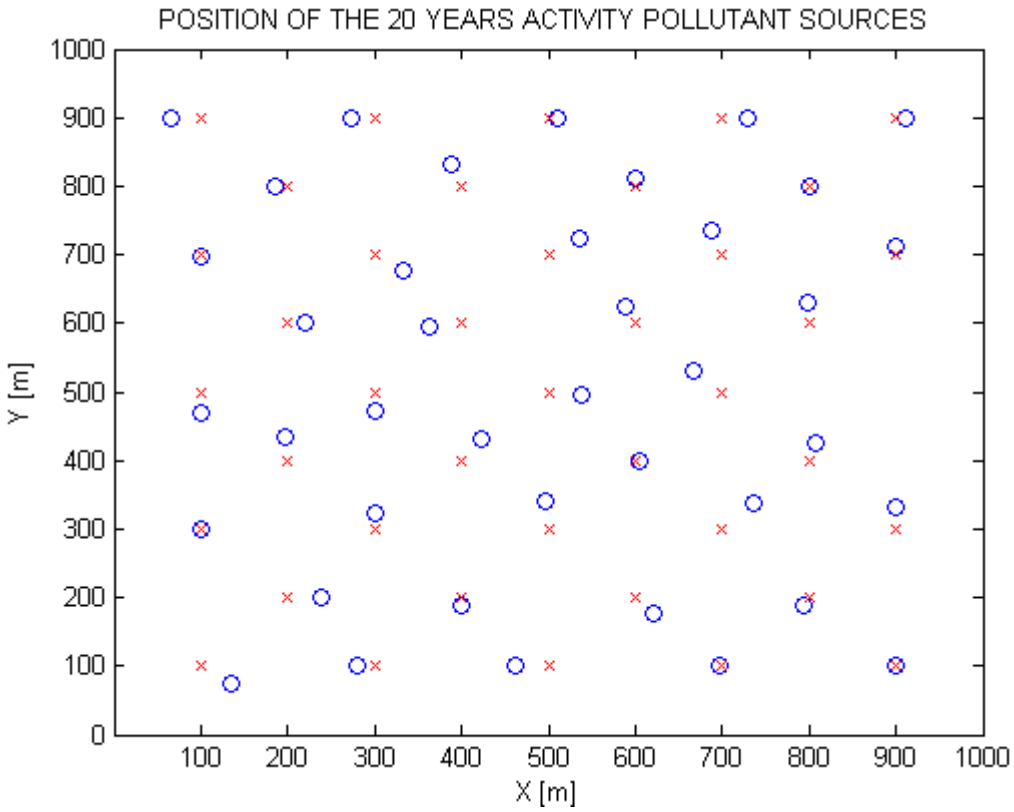


Figure 3.11: real (red x) and simulated position (blue circle) of the 20 years activity pollutant sources.

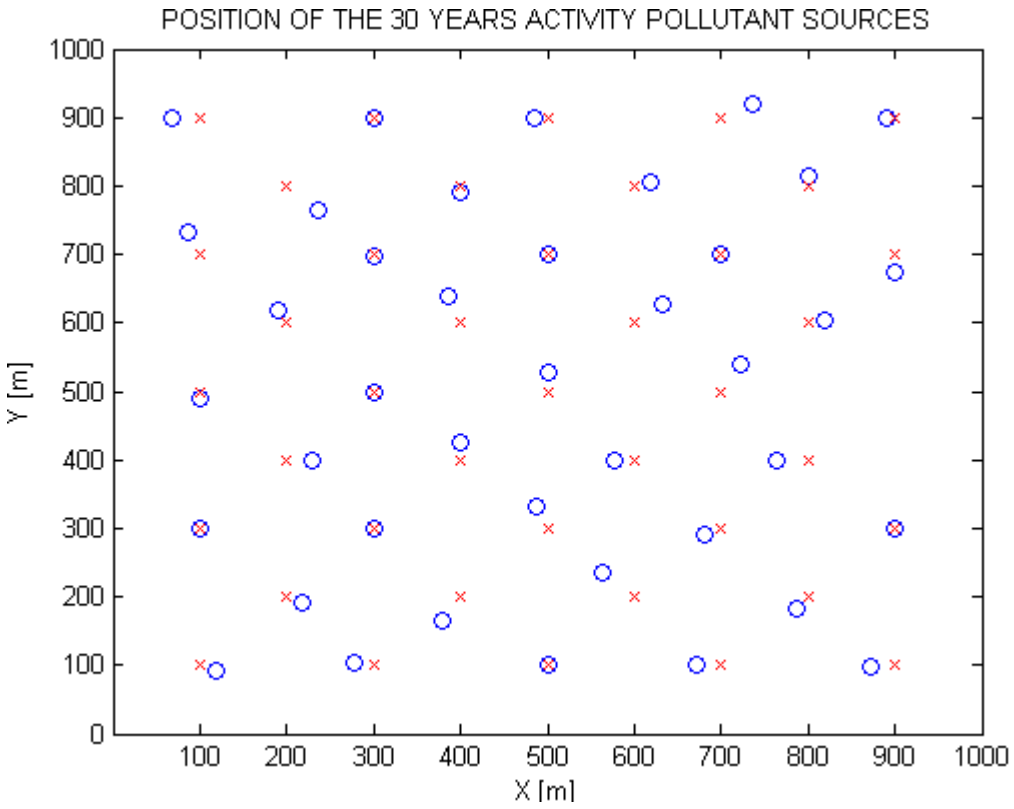


Figure 3.12: real (red x) and simulated position (blue circle) of the 30 years activity pollutant sources.

Table 3.3 illustrates the percentages of success of identifying the unknown pollution sources recognition of the artificial neural networks. Concerning the pollution sources position we decided to divide the results between the error smaller and largest than one domain cell size. For the time of activity of the pollution sources we consider the case where time was 100% correct and the cases where it was smaller than 6 years, because 5,26 years was the maximum error committed in time estimation.

Patterns results examples	%
X,Y,T 100% correct	14
T 100% correct / X,Y error < 20m	40
T 100% correct / X,Y error > 20m	17
T error < 6 years / X,Y error < 20m	23
T error < 6 years / X,Y error > 20m	6

Table 3.3: performance of the inversion of the artificial neural networks.

The final results as per Table 3.4 are expressed not as a percentage of error, but as the average mean and maximum error for the results. This way is considered more interesting to show how the network could fail in the pollution sources features approximation.

	X [m]	Y [m]	Time [years]
Em – mean error	14.19	14.33	0.70
EM – maximum error	39.17	39.82	5.26

Table 3.4: results related to the identification of the pollution sources features.

Concerning the positions of the unknown sources, the results show that most of the time the identification error is less than the size of one cell, in fact the cell size is equal to 20*20 m^2. At the same time the maximum error, which represents the worst case, is less than the size of two cells (Table 3.4).

In Table 3.5, the results obtained in this part of the research have been compared with the results of Scintu (2004) and Fanni (2002).

Test set prediction	X	Y	Time
Thesis results	67%	59%	76%
Scintu (2004) and Fanni (2002) results	100%	76%	54%

Table 3.5 : percentage of correct prediction of the neural models Test set for the thesis results and the previous works results (Scintu (2004) and Fanni (2002)).

As one can see in Table 3.5, the results of this work are determined by a good improvement of the duration of the activity approximation. In terms of percentage, worse results for the source position approximation are obtained, therefore it is important to underline that in Scintu (2004) and Fanni (2002) works curves of contaminants concentration in monitoring wells were used while in this work only one single value of contaminant concentration for each well is used.

The results achieved in this work shows that in most of the cases, the identification error is less than one cell size and the maximum error is less than two cells size (Table 3.4).

Figure 3.13 represents the error made by the artificial neural networks for time steps approximation.

Regarding the identification of contamination time, in most cases the network is able to detect 100% of the duration of the pollution activity. This is probably due to the different dynamics of the pollutant processes depending on the distance of the source from the boundaries and from the pumping well.

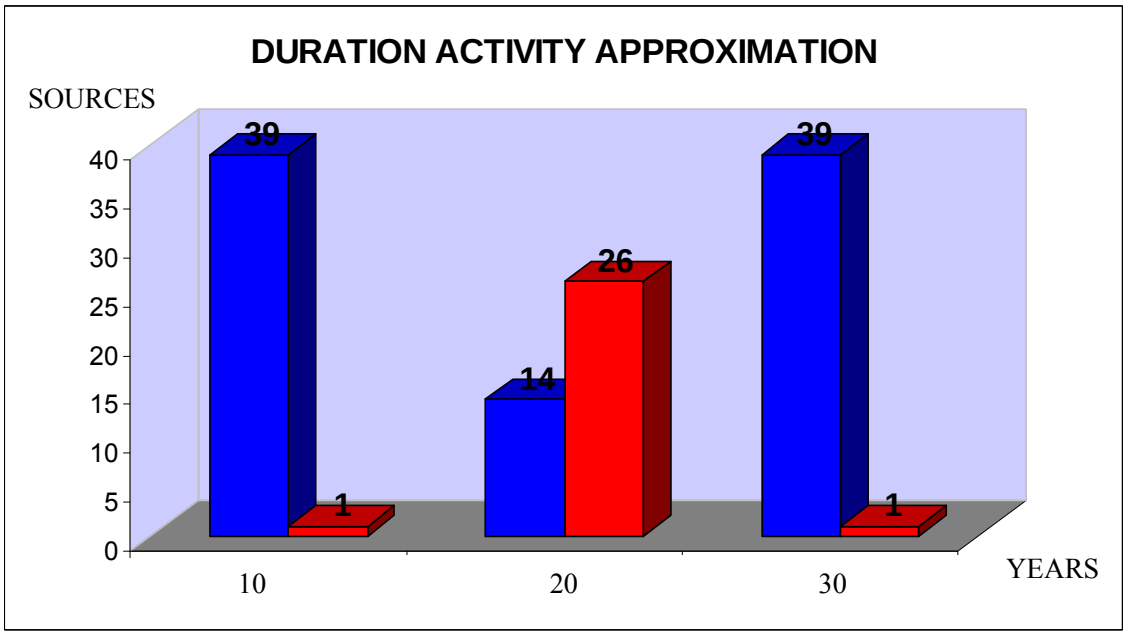


Figure 3.13: duration activity approximation of the artificial neural networks.

Figure 3.13 shows the duration activity approximation of the artificial neural networks. As one can see for the sources duration activity of 10 and 30 years, only one case time approximation is wrong.

For the 20 years sources duration activity, the wrong cases have been higher than the correct cases with 26 wrong cases out of a total of 40 cases. Nevertheless the mean error was of 2.66 years. The minimum and maximum errors were respectively of 6 months and 5 years and 3 months. Various trials performed to improve these results have shown that these results are strongly influenced by the instability of the ANNs.

Figure 3.14 show the position of the pollutant sauces where the ANN calculate duration activity for the 10, 20 and 30 years is wrong.

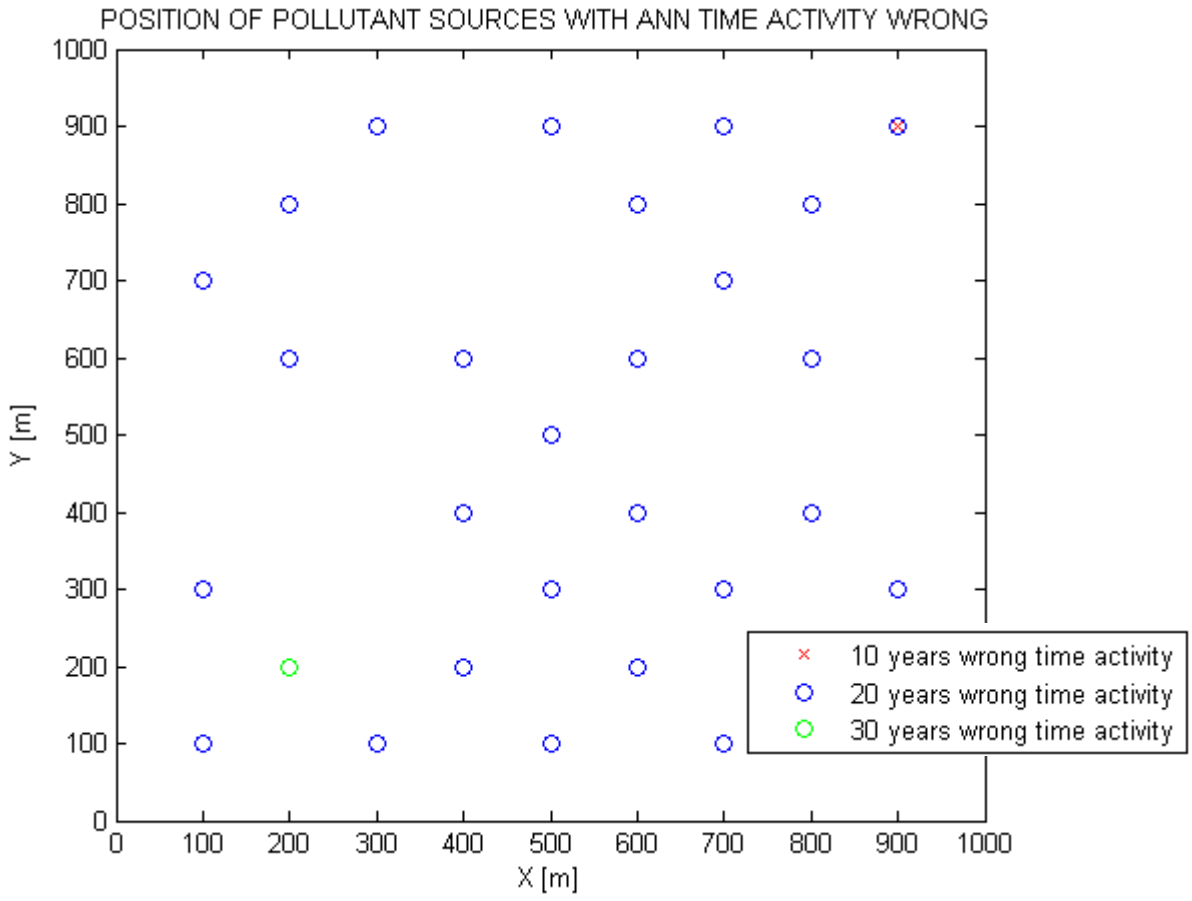


Figure 3.14: position of pollutant sources with ANN activity wrong

3.6 Summary

This chapter describes in details the estimation of the source behaviour in terms of spatial location (X,Y) and the duration of the activity (T). Various source scenarios have

been constructed in order to generate the examples patterns used for training and testing the ANNs able to solve the inverse problem. These scenarios have been performed by varying the pollutant source position and the duration of the source activity in the domain. An inverse method based on ANN technology has been used to identify unknown pollution sources. In particular, the inverse problem has been solved using measurements of contaminant concentration acquired in the monitoring wells at a certain time t.

4 ANN FOR ESTIMATING ALSATIAN AQUIFER POLLUTION SOURCE

This chapter proposes a new methodology that aims at solving the inverse problem in order to reconstruct the behaviour in time and space of the carbon tetrachloride unknown pollution source of the Alsatian aquifer (France).

The chapter provides a brief description of the Alsatian aquifer, the carbon tetrachloride tanker accident occurred in 1970 as well as the physical characteristics of this dangerous chemical. The numerical model of the Alsatian aquifer is described in order to explain the pattern construction and the ANN development.

The Alsatian aquifer flux and transport model used in this thesis is based on the work carried out by Taef Aswed in 2008 for his PhD thesis at the Institut de Mécanique des Fluides et Solides (IMFS) of Strasbourg.

4.1 Introduction

The purpose of this part of the research is to study the spreading of a dangerous chemical - carbon tetrachloride (CCl_4) - that contaminated, after a tanker accident in 1970, part of the largest aquifer in Western Europe: the Alsatian aquifer located in France.

The exact quantity of the chemical infiltrated is unknown and this constitutes the main issue for its location and remediation. Therefore, the objective is to reconstruct the temporal evolution of the Alsatian aquifer unknown contaminant source, determining whether or not ANNs are able to reconstruct the behaviour of the unknown contaminant source in the hydrological domain based on measurement of the contamination concentration curves in monitoring wells.

This task has been carried out on the basis of a previous study [Aswed, 2008] that focused on solving the inverse problem for the carbon tetrachloride Alsatian aquifer contamination. Aswed's work aims at modeling and simulating the transfer in the aquifer of contaminant and the determination of the source terms at the accident location. The pollution source was reconstructed based on the measurements of contaminants concentrations in the monitoring wells of the domain. The created model was calibrated using measured data of carbon tetrachloride concentration that were

collected during 12 years (from 1992 to 2004). Simulations were performed for a period of 54 years from 1970 to 2024.

4.1.1 Operative steps

The operational steps of this work can be described as follows:

- study of the flow and transport model of the Alsatian aquifer,
- creation of the examples patterns necessary for ANN development based on the Alsatian aquifer model,
- implementation of an ANN suitable to address the Alsatian aquifer pollution inverse problem.

The numerical model used to build the patterns for training, validating and testing the ANN are coming from Aswed's work (2008). The examples of patterns were created by developing different hydrological scenarios of the Alsatian aquifer pollution source.

4.2 Description of the Alsatian aquifer characteristics

4.2.1 General description of the Alsatian aquifer and the Upper Rhine Valley

The Alsatian aquifer is located North-East France in the southern part of the Upper Rhine Valley, part of Alsace region. A map of the aquifer area is available in Figure 4.1.

Figure 4.1: Alsace region and Alsatian aquifer in France.

The Upper Rhine Valley is a segment of the European Cenozoic rift system that develops in the north-western forelands of the Alps. It extends over 300 km, from Basel (Switzerland) in the south to Frankfurt (Germany) in the north with an average width of approximately 40 km. Upper Rhine Valley is flanked by the Vosges and Black Forest (Schwarzwald) mountains, respectively, to the west and the east [Bertrand et al, 2006]. Figure 4.2 provides a map of the Upper Rhine Valley.

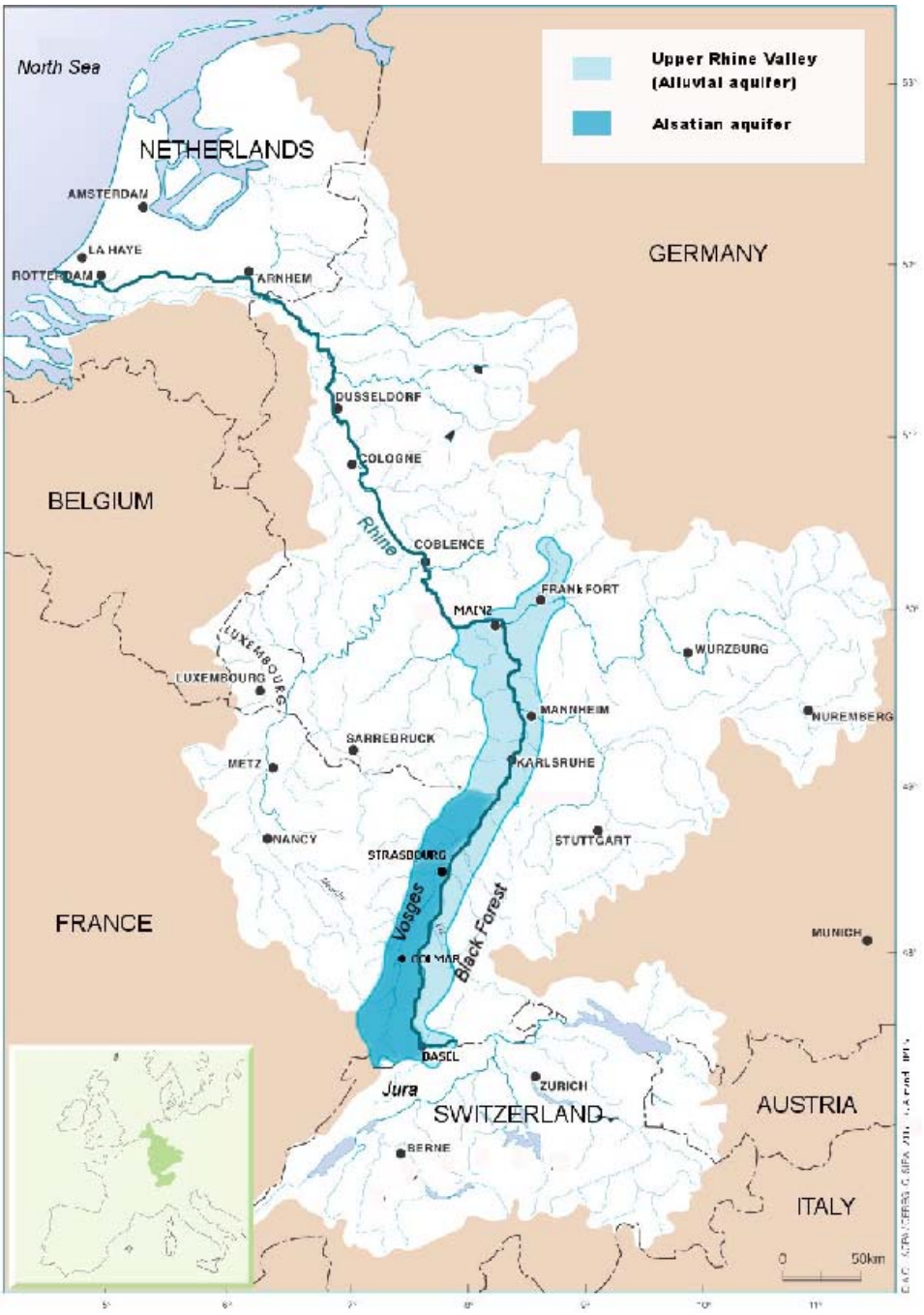


Figure 4.2: The Rhine valley.

The groundwater of the Rhine valley between Basel and Mainz is an essential compartment of the Upper Rhine hydrosystem that contains a volume of alluvial about 250 billion m^3, which makes this large alluvial aquifer the largest fresh water reserves in Europe [Guilley F., 2004].

The Alsatian aquifer forms an integral part of this immense hydrogeological system. It extends to the border between France and Germany. It is surrounded by the Jura Mountains in the south, the Vosges Mountains in the west, the Rhine in the east, and the Haguenau-Pechelbronn basin in the north-west [Hamond, 1995].

The aquifer surface is over 3000 km2. It has a length of 160 km and a maximum width of 20 km. This permeable alluvial has a thickness of a few meters at the Vosgean edge, and 150 m to 200 m in the centre of the Rhine plain. In Strasbourg region, it has an average thickness of around 80m [Hamond, 1995].

The groundwater reservoir contains about 50 billion m^3 of water, with an annual renewal of 1.3 billion m^3. The exploitation of the aquifer for collectives, industry and agriculture is almost the third of renewal volume, which is about 0.5 billion m^3. This large aquifer has a vital importance since it supplies to 75% of the drinking water requirements, 50% of the industrial water needs and 90% of the irrigation water needs in Alsace.

The numbers reported in this paragraph are extracted from the note of CIENPPA (1984), which highlights the importance of the phreatic aquifer and its economic role in the Alsace region.

4.2.2 Hydrodynamic and hydrographic system of the Alsace plain

The Alsatian aquifer is a phreatic aquifer filled by tertiary and quaternary sediments, mainly fluvial gravels and sands deposited from various origins and drained mainly by rivers and human activity. It is defined as an extensive alluvial aquifer with a layered structure composed by a random superposition of different alluviums (clay, sand fine to rough, gravels, coarse).

The groundwater in the aquifer flows mostly from south to north and with some local variations mainly due to heterogeneity in the permeable formation. The hydraulic gradient, however, is not uniform over the aquifer. It is about 0.7% to 0.9% in the centre of the plain and is higher at the edge of the aquifer, where the sediments are less thick [Hamond, 1995].

The Rhine valley is dominated by two main rivers: the Rhine and the Ill River. The channelled Rhine has a flow rate typically between 700 and 1500 m^3/s. The Ill River starts in Jura and runs northward through the Alsace. It has a discharge flow rate ranging between 5 and 10 m^3/s at Strasbourg. The flow direction of the Ill River is almost parallel to the Rhine. Before joining the Rhine aquifer at the north of Strasbourg, all the smaller rivers carrying the discharge water from the Vosges mountains flow directly into it [Eikenberg et al., 2001]. Figure 4.7 contains a map of the studied area where these rivers are represented.

These rivers have a classic hydraulic system. The precipitation of the basins varied with time: high water level in winter and spring, and low water level at the end of summer. The groundwater reservoir is part of a complex hydrographical system, which includes frequent exchanges between the rivers and the aquifer which vary with the seasons due to the proximity between the surface and the groundwater. The Alsatian aquifer is highly exposed to contamination from the neighbouring rivers ad their tributaries. The river-aquifer interactions are governed by the fluctuating water level of the rivers and it may be quantified in two ways [Aswed, 2008]:

- discharge of groundwater to surface water when the groundwater level is higher than the river stage,
- recharge of groundwater by surface water when the elevation of river stage is higher than groundwater level.

The interaction between the rivers and the aquifer at the Alsace region has been studied by the SEMA/DIREN. The exchange coefficient of the Rhine has been estimated to about 10^{-6} m/s by taking into account the discharge of the Rhine contra-cannel. The canalization of the Rhine in Alsace was established to maintain a constant water level of the river and for agriculture purposes.

Infiltration of rainwater is the major source of recharge for the Alsatian aquifer. A study of the mean precipitation in the modelled area that was carried out by the CEREG under the PIREN-Eau/Alsace program of the CNRS, showed that approximately 620 ML/yr of water recharges the groundwater at the centre of the Alsatian aquifer and that 680-740 ML/yr recharges the north-eastern margin of the aquifer. They generated a map of mean rainfall using five rainfall gauging stations. The groundwater recharge can be estimated to 5-10% of the precipitation.

4.3 History of the pollution by CCl_4 in the aquifer

The information contained in this paragraph are based on the report of Hamond (1995).

4.3.1 The accident of 1970 and pollution discovery

On December 11th 1970, a tanker truck - property of a Dutch company - containing carbon tetrachloride (CCl_4) capsized in the north of Benfeld, a small town located about 35 km south of Strasbourg (Figure 4.1). In spite of the efforts of the

firemen to control the spilling of the chemical, an important quantity of it could not be recovered (CCl_4 is an extremely volatile product).

The SGAL, however, shortly after the accident, hypothesized that the spill could reach the phreatic aquifer and described the probable migration mechanisms of CCl_4. According to a note of SGAL of December 21th1971, about 4000 litres of CCl_4 were spread in the area of the accident, infiltrating into the ground or disappearing by evaporation.

At first, the pollutant infiltration was contemplated into non-saturated medium, taking into consideration the optimistic hypothesis that the loess layers, present in the accident area, could act as a barrier between non-saturated and saturated aquifer. In a second moment, because of its high density, the amount of CCl_4 could reach the water table and quickly migrate into the aquifer. In the end, it was hypothesized the formation of a pollution plume of the aquifer of about 21 000 m^3 due to the phenomenon of convection, diffusion, dispersion, and solubility of the product.

However, no dynamic analysis to assess the propagation speed of such pollution was established. The migration of the pollutant downstream seemed not predictable due to the uncertainties about its behaviour. The idea, at the time, was that the pollution would be relatively limited, and that over time the pollutant may disperse, dilute and degrade before reaching the downstream Erstein's drinking water supplies. Erstein is a small town located downstream of Benfeld (Figure 4.1).

However, SGAL gave some recommendations concerning the installation of piezometers for monitoring the water quality and the plume spread in the accident area. These suggestions were transmitted to Mine Services and to the Direction department of agriculture (DDA) but without getting much attention, since they assumed that the chemical would be removed before being able to damage the supplies of water downstream.

After 20 years, in 1991, was carried out the first analysis in the drinking water wells by the BRGM. The analysis showed quantities of CCl_4 (about 15.6 µg/l) in the supplies of drinking water located in Erstein [Beyou, 1999]. Regular analyses of the drinking water wells showed that the level of CCl_4 has always been exceeding the safe limits recommended by the World Health Organization (WHO) (2 µg/l). In 1992, the pollution was confirmed with CCl_4 levels between 62.4 µg/l and 56.2 µg/l. This high level of CCl_4 concentrations has caused serious problem in the region by contaminating

an important drinking water source in the area. On July 10th 1992, the Ministry of Health orders to the public not to consume groundwater for drinking usage [Beyou, 1999].

4.3.2 Physical and chemical properties of CCl_4

CCl_4 is a man made volatile organic chemical (VOC) that can be classified as a dense non-aqueous phase liquid (DNAPL). CCl_4 has a density higher than water, implying that it has a tendency to penetrate the water table and move deeper into the aquifer. Plumes developing from these sources often travel large distances to eventually impact water supplies.

Decomposition of CCl_4 may produce phosgene, carbon dioxide, hydrochloric acid, methane tetrachloride, perchloromethane, tetrachloroethane, and benziform [HSG 108, 1998].

Some of the physical and chemical properties of CCl_4 are reported in Table 4.1.

Property	Value or Information	References
Molecular weight	153.84 g/mol	CPHF, 1998
Color	Colorless	NIOSH, 1994
Phase state	Liquid	NIOSH, 1994
Odor	Sweet, ether-like odor	NIOSH, 1994
Odor threshold (in water)	0.52 mg/l	U.S.EPA, 1998
Boiling point	76.7°C	CPHF, 1998
Melting point	-23°C	U.S.EPA, 1998
Solubility at 25°C	1160 mg/l	CPHF, 1998
Solubility at 20°C	800 mg/l	CPHF, 1998
Density	1.59 g/ml at 20°C	NIOSH, 1994
Log K_{ow}	2.64	CPHF, 1998
Soil Sorption Coefficient K_{oc}	K_{oc} = 71 (moves readily through soil)	U.S.EPA, 1998
Bioconcentration Factor	Log BCF = 1.24-1.48, not significant	U.S.EPA, 1998

Vapor Pressure	91.3 mm Hg at 20°C	CPHF, 1998
Henry's Law Constant	3.04 x 10^{-2} atm-m^3/mol at 24.8°C	U.S.EPA, 1998
Henry's Law Constant (dimensionless)	1.25 at 24.8	U.S.EPA, 1998

Table 4.1: Physical and chemical properties of carbon tetrachloride [Aswed (2008)].

Usage

Carbon tetrachloride is a synthetic chemical compound that has been widely used in different activities during the 20th century. Moreover, the product is often discharged without further precautions.

In 1910, it was used to extinguish fires [U.S. Patent 1,010,870]. The liquid vaporized extinguished the flames by inhibiting the chemical chain reaction of the combustion process. The carbon tetrachloride fire extinguishers were commonly used until the mid-20th century [U.S. Patent 1,105,263].

It is widely used in the production of refrigeration fluid (trichlorofluoromethane and dichlorodifluoromethane), in propellants for aerosol cans, in fabricating nylon, grain fumigant, to made petrol additives and semi-conductors. It is used as a solvent for fats, oils, and greases, for dry cleaning and for degreasing metals [U.S.EPA, 1998].

Carbon tetrachloride properties has made it readily usable in industrial chemical processes, since it is not dangerous in terms of handling accident (no risk of explosion or fire). Smaller amounts of this solvent can also be used in laboratories or for domestic activities DIY (Do-It-Yourself) and as a spot remover for clothing.

It was also used in agriculture through the mid-1980s as a fumigant to kill insects in grain. Its use as a pesticide was stopped in 1986.

All these uses are now banned and its usage is limited to some industrial applications because of its toxicity and its effect on the ozone layer [ATSDR, 1995]. The Montreal Protocol on substances that deplete the Ozone layer (1987) and its amendments (1990 and 1992) established a timetable for the phase out of the production and consumption of carbon tetrachloride. The manufacture of CCl_4 has, therefore, dropped and will continue to drop [UNEP, 1996; IPCS, 1999].

Regulation and recommendation

For about 15 to 20 years, Carbon Tetrachloride as well as Trichloroethylene and Tetrachloroethylene was a chlorinated solvent widely used and its use was not regulated at all. When the accident occurred in Benfeld, the French law did not contain any specific regulations on the matter.

The Public Health Code revised in 1989, provides the quality standard for the waters on 1 µg/l; if CCl_4 exceeds the limit, water monitoring should be strengthened. The World Health Organization, in 1999, recommended the level of 2 µg/l in the water and in air the CCl4 have to be less than 0.11 part per million (ppm). Humans cannot smell CCl4 if its level is less than 10 ppm.

The Environmental Protection Agency (EPA) has set a Maximum Contaminant Level Goals (MCLG) for carbon tetrachloride at zero parts par billion (ppb) of drinking water. Based on this MCLG, EPA has set an enforceable standard called a Maximum Contaminant Level (MCL). The MCL has been set at 5 ppb for general water usage.

Toxicity

The EPA has determined that carbon tetrachloride is a probable human carcinogen. CCl4 and some of its degradation products are considered carcinogens or suspected carcinogens. Exposure to high concentrations of carbon tetrachloride may cause liver, kidney, and central nervous system damage [ATSDR, 1995]. The biotransformation of CCl4 in the body is initiated by enzymatic reactions that transform it into trichloromethy radicals (CCl_3). CCl_3 is the active product responsible for liver cell damage [Macdonald, 1982].

Indeed, the CCl_4 works at the cellular metabolism level, causing an excessive production of water. This results in an increase in whole body fluid volume and an increase in blood pressure. These reactions cause oedema of the lung and brain that can be fatal. Moreover, chronic damage are carried in the blood (anaemia), skin, peripheral nerves (paralysis of the legs). Neuropsychological consequences, such as emotional and behaviour are also possible. The injection of 5 ml of this product provokes an acute intoxication that causes human death. Toxic effects of CCl_4 can occur after ingestion or breathing, and possibly from exposure to the skin. [Aswed (2008), Beyou, 1999].

Environmental impact

The high Henry's constant of CCl_4 indicates that it is extremely volatile. Because carbon tetrachloride evaporates easily, most of the compound released to the environment during its production and use reaches the air. It can remain in air for several years before it is broken down to other chemicals (from 30 to 100 years). In the air, it may react with other chemicals that have the potential to destroy upper atmosphere ozone layer. Small amounts of carbon tetrachloride are found in surface water. Evaporation from water is a significant removal process. Based upon field monitoring data, the estimated half-life in rivers is 3-30 days and in lakes and groundwater is 3-300 days.

If carbon tetrachloride spills onto the ground much of it will evaporate to the air. Some of it may be trapped into groundwater, where it can remain for months before it is broken down to other chemicals. Only a small amount sticks to soil particles; the rest evaporates or moves into the groundwater.

4.3.3 CCl_4 migration in the aquifer

In the Alsatian aquifer, CCl_4 and its toxic constituents behaviour are strongly influenced by aquifer heterogeneity that produce high uncertainty in the CCl_4 migration.

CCl_4 has low solubility in water, once it reaches the aquifer it migrates in two different phases: dissolved in water (aqueous) and liquid (non-aqueous phase). In the Alsatian aquifer carbon tetrachloride represents a long term source of contamination. It penetrates in the water table and continues to move vertically downward until gravitational movement is retained. Some of the contaminant is booked in the porous media by capillary forces. The trapped CCl_4 release miscible quantities of contaminant due to the contact with the groundwater flow.

The study site has complex heterogeneous and anisotropic hydrogeological conditions. In the groundwater system, carbon tetrachloride and its toxic constituents have actions of convection, dispersion, and diffusion. Volatilization of the CCl_4 at the site saturated zone is insignificant. The sorption can be neglected due to the low organic matter content in the ground. The chemical properties such as the solubility in water, diffusion, volatilization, and degradation coefficients are uncertain.

The exact amount of the chemical infiltrated and the source behaviour and how the pollutant feeds the contamination are unknown.

4.3.4 Cleanup approach

The information contained in this paragraph are from the report of Beyou (1999).

Since 1992, after the discovery of abnormal quantities of CCl_4, public authorities have worked to educate the community and recommended to stop drinking water from wells. To meet the urgency of the problem, a water supply decontamination was considered. A remediation system plant was installed in one residential well (Negerdorf) in Erstein town. To remove CCl_4 from groundwater, the remediation treatments include: air stripping and granular activated carbon absorption.

In a first moment, ground water is pumped through two packed towers for air-stripping. The towers are a forced draft system with air blown at up to 1700 m3/h. This blown air is used to separate CCl_4 from water by evaporating. The removal efficiency of CCl_4 at this stage of treatment is about 91%. Stripping towers also have the effect of removing 85% of the CO2 content in water drawn from the aquifer.

In a second moment, the stripped ground water is filtered by granular activated carbon (GAC). GAC is installed on a bed where the contaminated water is pumped through. During this process contaminant gets absorbed by GAC and 100% of CCl4 is removed from ground water.

The final processing step, before distributing water to consumers, is the disinfection with Chlorine.

Figure 4.3 shows a schematic diagram of the treatment plant installed on Negerdorf well in Erstein. This scheme is taken from the report of Aswed (2008).

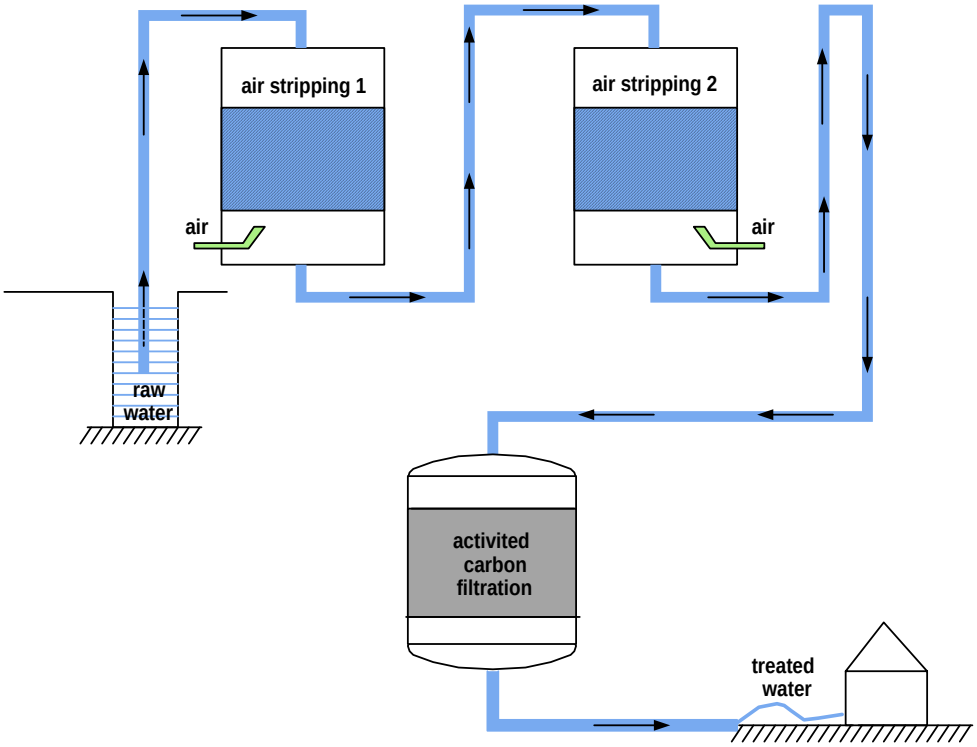


Figure 4.3: Schematic diagram of the treatment plant installed on Negerdorf well in Erstein.

4.4 The model of the CCl_4 pollution in the Alsatian aquifer

The CCl_4 Alsatian aquifer pollution has been the subject of several studies that address the hydrodynamic state of the aquifer and the pollution migration. Vigouroux et al. (1983) studied some cases of pollution in the aquifer of the Rhine Graben by using a 2D numerical model taking into account convection, dispersion and retention. Hamond (1995) presented a numerical model that approximated the hydraulic head and transport problems. His study was based on data measured between 1970 and 1993. Beyou (1999) completed the study by adding measured data during the period between 1993 and 1999.

None of the mentioned works showed satisfactory matching of the measured data.

Aswed (2008) developed a new 3 dimension (3D) numerical model of the CCl_4 pollution to simulate the concentration in the accident location. In particular the travel time between the source and measurement-wells was calculated by the method of moments. This numerical model was made on the basis of data measured in monitoring wells between 1999 and 2004. Aswed (2008) numerical models were constructed using the non commercial software “Transport of RadioACtiver Elements in the Subsurface” (TRACES) developed by Hoteit et al. (2004) at the IMFS. It was calibrated using

measured data of carbon tetrachloride concentration that were collected during 12 years (from 1992 to 2004). Simulations were performed from 1970 to 2024.

In this work, we try to solve the inverse problem for the Alsatian aquifer pollution using ANN technologies. In order to model the CCl_4 pollution behaviour with ANN, a model of groundwater flow and contaminant transport is necessary. To this end, Aswed (2008) model has been used.

4.4.1 Difficulties for solving the inverse problem for the CCl_4 pollution

Based on Aswed's work (2008), the main difficulties for solving the inverse problem for the pollution of the Alsatian aquifer are described below:

- The exact amount of the chemical infiltrated in the underground is unknown since some of the tanker volume was recovered and another part could have disappeared by vaporization.
- Carbon tetrachloride has low solubility in water. Some of the chemical could be trapped underneath of the accident site because of several physical and chemical mechanisms such as gravity, capillarity, and adsorption. The trapped CCl_4 may continue to release miscible quantities due to the rain or the contact with the underground water flow. Therefore, the initially trapped CCl_4 may act as a continuous source of contamination for the underground water.
- The source contaminant behaviour that continued to feed the contamination and, which is space and time dependent, is unknown.
- High uncertainty in the aquifer formation properties such as porosity and permeability. The aquifer is heterogeneous and has different permeable layers.
- Uncertainty in the chemical properties such as the solubility in water, diffusion, volatilization, and degradation coefficients.
- Complexity in defining a proper coupling between the water flow and pollutant transfer in the permeable medium and the corresponding boundary conditions.

4.4.2 Numerical solution of the flow and transport problems

Several types of numerical methods can be used to solve the groundwater flow and solute transport equations. Numerical simulators that are based on conventional finite difference (FD) of finite volume (FV) methods may not be ideal to solve the

Alsatian aquifer pollution. The aquifer geometry is complex and cannot be properly modelled by conventional finite difference method on structured (Cartesian) gridding. Conventional methods that intend to approximate the velocity by deriving the pressure head in a post-processing step may not be accurate in heterogeneous media. The accuracy of the predicted velocity field is thus crucial. The transport problem is dominated by convection. It is also known that first order approximation methods have poor convergence near chocks or sharp fronts of convection dominated problems. Fine gridding are thus required to reduce the numerical diffusion. To model the CCl_4 pollution, Aswed (2008) used the numerical model TRACES because it combines the mixed hybrid finite element (MHFE) and discontinuous Galerkin (DG) methods to solve the hydrodynamic state and mass transfer problems.

4.4.3 Characteristics of the flux and transport model of the Alsatian aquifer

The domain is highly heterogeneous due to the sedimentation effect, which provides anisotropic flow properties. The behaviour of the pollutant plume is strongly influenced by the contrast of permeability within the alluvial aquifer.

Several studies were preformed at IMFS to numerically model the subsurface water flow and contaminant migration in the Alsatian aquifer [Hamond (1995); Beyou, (1999)].

The aquifer structure between Benfeld and Erstein could be clearly visualized due to geological profiles carried out during the installation of 34 wells and piezometers. These data have been obtained from the BRGM database.

Hamond (1995) estimated the hydraulic head and water-flow velocity in the area between Kogenheim and Strasbourg by using a 2D steady-state model. A sketch of the domain is shown in Figure 4.4 (green polygon).

The accuracy of his model was validated using the average values of measured head data during the period between 1970 and 1994. He also analyzed the flow trajectories and the travel times of water particles between Benfeld and Strasbourg. The 2D model was calibrated with 34 points of measurement head, where the maximum difference between the measured and predicted piezometers was less than 10 cm for 31 points and 16 cm for the 3 other measurements.

The measurements in the aquifer showed that the contaminated zone is located approximately in the groundwater between Benfeld and Erstein, South-West/North-

East. The contaminated zone is confined within a rectangle domain of 6 km width and 20 km length. The contaminated domain is located between upstream of Benfeld and downstream of Erstein (see Figure 4.4).

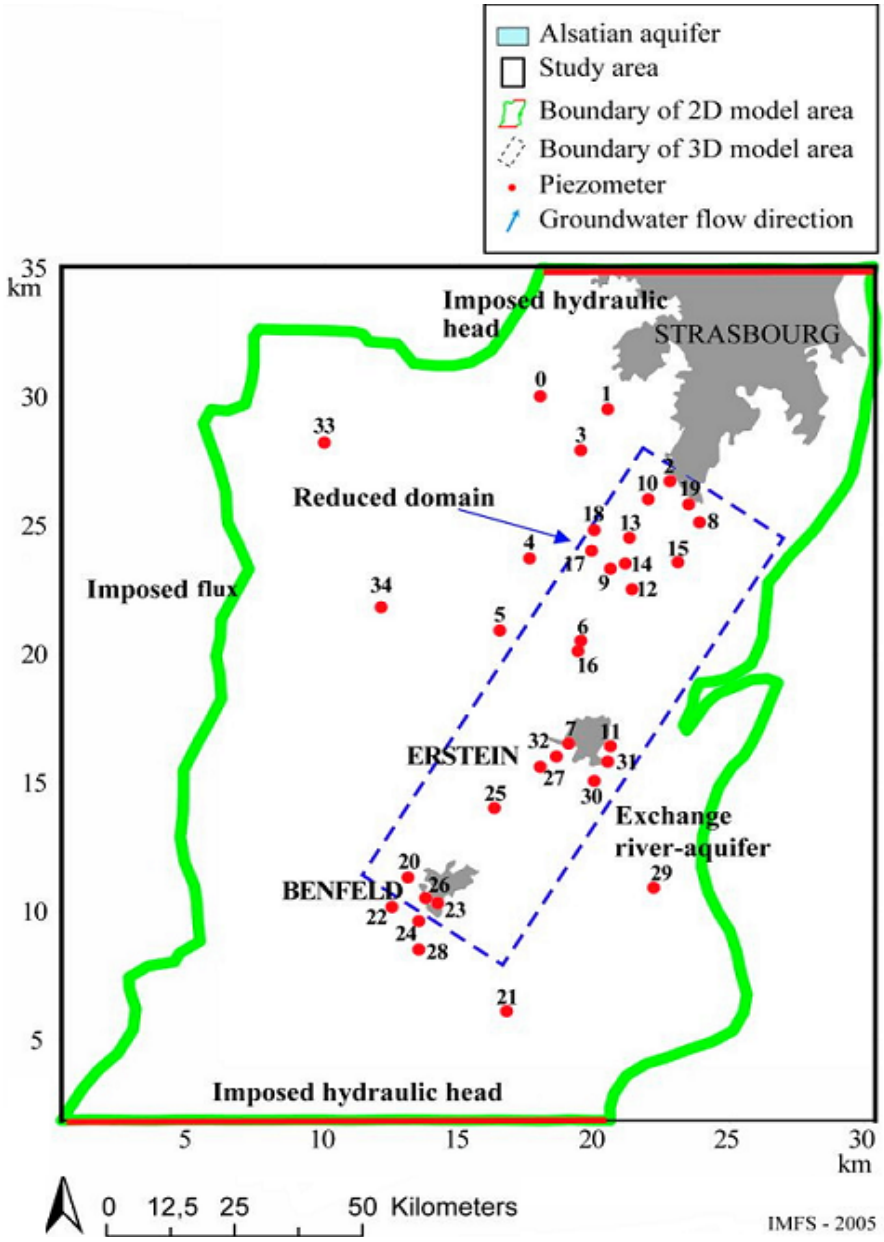


Figure 4.4: 2D and 3D domain with the related boundary conditions [Hamond, 1995].

In the work of Aswed (2008), the flow and contaminant transport problem are solved in the contaminated aquifer described by a three dimensional (3D) computational domain. In the 3D model, the aquifer is represented by layers of variable thickness and zones division of different hydrodynamic properties such as hydraulic conductivity and porosity. To be able to solve the water flow problem, the boundary conditions that involve the hydraulic head and water flow rates at the vertical boundaries of the 3D domain are based on the above mentioned 2D model. In the planer cross-section in

Figure 4.4, the contaminated domain (dashed rectangular zone) of the Alsatian aquifer is represented.

The planar area of the simulation domain is 20x6 km^2 with a depth of about 110 m. The simulation domain is discretized into a non uniform mesh with 25388 nodes and 45460 irregular prismatic elements (see Figure 4.5). The domain is divided into 10 successive layers according to the estimated geometry of the cross sections (the landfill site was divided into 8 zones by soil type). The layers have different depths (numbered from bottom to top) between 5 and 15 m.

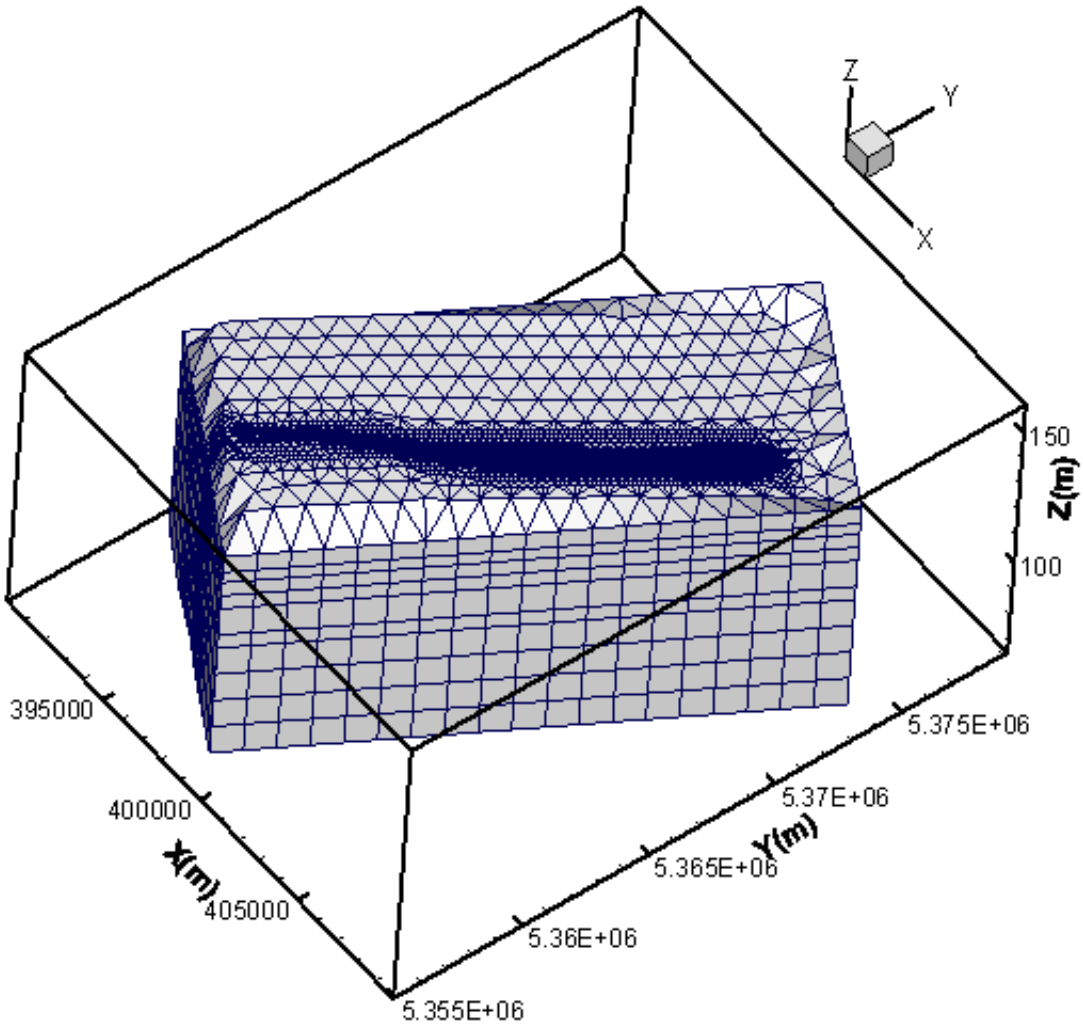


Figure 4.5: computational mesh of the 3D domain.

Figure 4.6 represents the hydraulic head in the top of the 3D domain. In the 3D domain, the upstream hydraulic head is 155 m and the downstream hydraulic head is 139 m. Hydraulic heads is constant along the depth of the aquifer at the vertical boundaries because water flow is essentially horizontal.

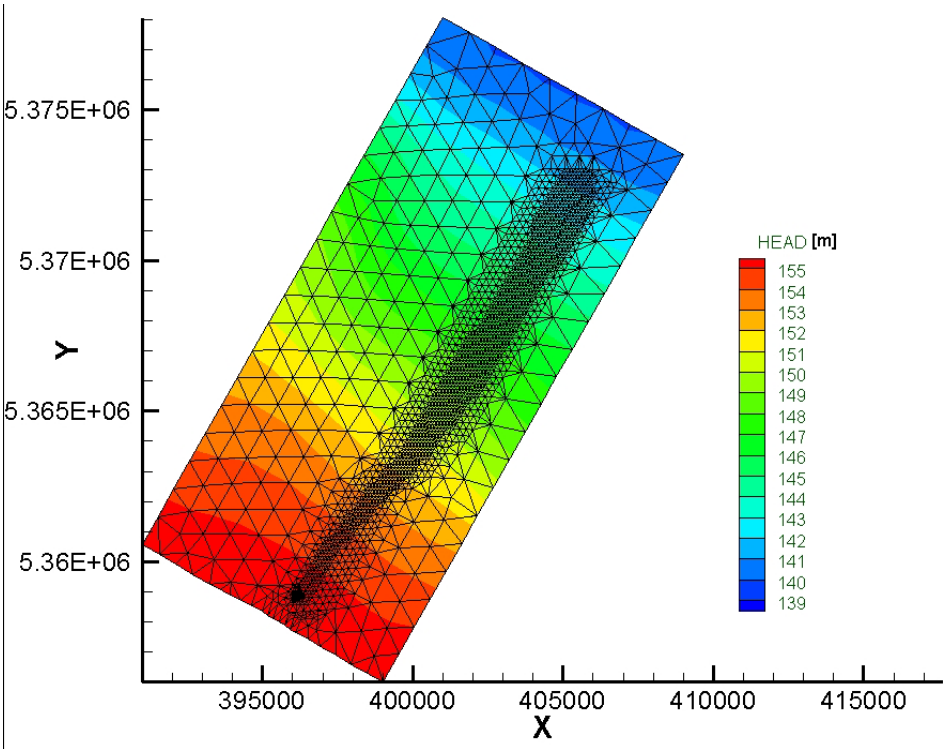


Figure 4.6: hydraulic head in the top of the domain.

Aswed (2008) interpreted the geological profiles of the domain. In total seven lithological units were identified that depend on the percentage of sands, the proportion of pebbles and gravels or their argillaceous characteristic. The entire thickness of the aquifer is about 80 m. The categories are classified below:

- loess, loam and clay
- compacted clay,
- sandy clay or clayey sand,
- fine sand to very fine (sand content > 70%),
- sandy alluvial (sand content between 50% and 70%, with 20% to 40% of pebbles or gravels),
- medium alluvial (sand content between 30% and 50% (coarse sand), medium gravels and pebbles),
- coarse alluvial (sand content < 20%, with medium to coarse gravels and pebbles (80% to 100%)).

The discrimination of the seven alluvial classes agrees fairly well with the assumed hydraulic conductivities in the different layers. From lithological profiles,

thirty cross-sections have been set according to this classification. The model area is divided into a number of zones such that a maximum and a minimum hydraulic conductivity coefficients are assigned to each zone.

Table 4.2 represents the categories of the hydraulic conductivity and their compatibility with the lithology of the aquifer formation between Benfeld and Erstein.

Lithology	**Hydraulic conductivity (m/s)**
Marly substratum: clay, sandy clay, clayey sand	10^{-8} to $2.5*10^{-6}$
Loess	$2.5*10^{-6}$ to $2.5*10^{-5}$
Fine sand to very fine (sand content > 70%)	$2.5*10^{-5}$ to $1*10^{-4}$
Sandy alluvial (50%< sand content>70%)	$1*10^{-4}$ to $5.5*10^{-4}$
Medium alluvial containing clay lens or high content of sand	$5.5*10^{-4}$ to $1.5*10^{-3}$
Medium alluvial (30% < sand content>50%)	$1.5*10^{-3}$ to $3.5*10^{-3}$
Coarse alluvial containing sandy lens or clayey	$3.5*10^{-3}$ to $1*10^{-2}$
Coarse alluvial (sand content <=20%)	$1*10^{-2}$ to $2*10^{-2}$

Table 4.2: categories of permeabilities considered in the model

In Aswed's work (2008), Monte-Carlo method is used to estimate the hydraulic conductivity for each grid block in the domain.

As it can be evinced from Table 4.2, the Alsatian aquifer consists essentially of sands and gravels. The estimated porosity varies slightly between 10% and 20% (Hamond, 1995; Beyou, 1999). The values of the porosity in the 3D model were estimated by Aswed (2008) through trial and error within the above provided range.

Aswed's work (2008) also consider initial longitudinal and transverse dispersivities based on prior information from similar geological formations (Gelhar et al., 1992). The dispersivity coefficient is estimated by trial and error technique using the software TRACES. The longitudinal dispersivity is between 10 and 20 m and the transverse horizontal and vertical dispersivity is between 0.5 and 3 m.

Field wells and source location

A groundwater monitoring network was installed in the domain to monitor water quality. Figure 4.7 shows a map of the studied domain with the piezometric wells where the red stars are monitoring wells and the blue triangle are pumping wells.

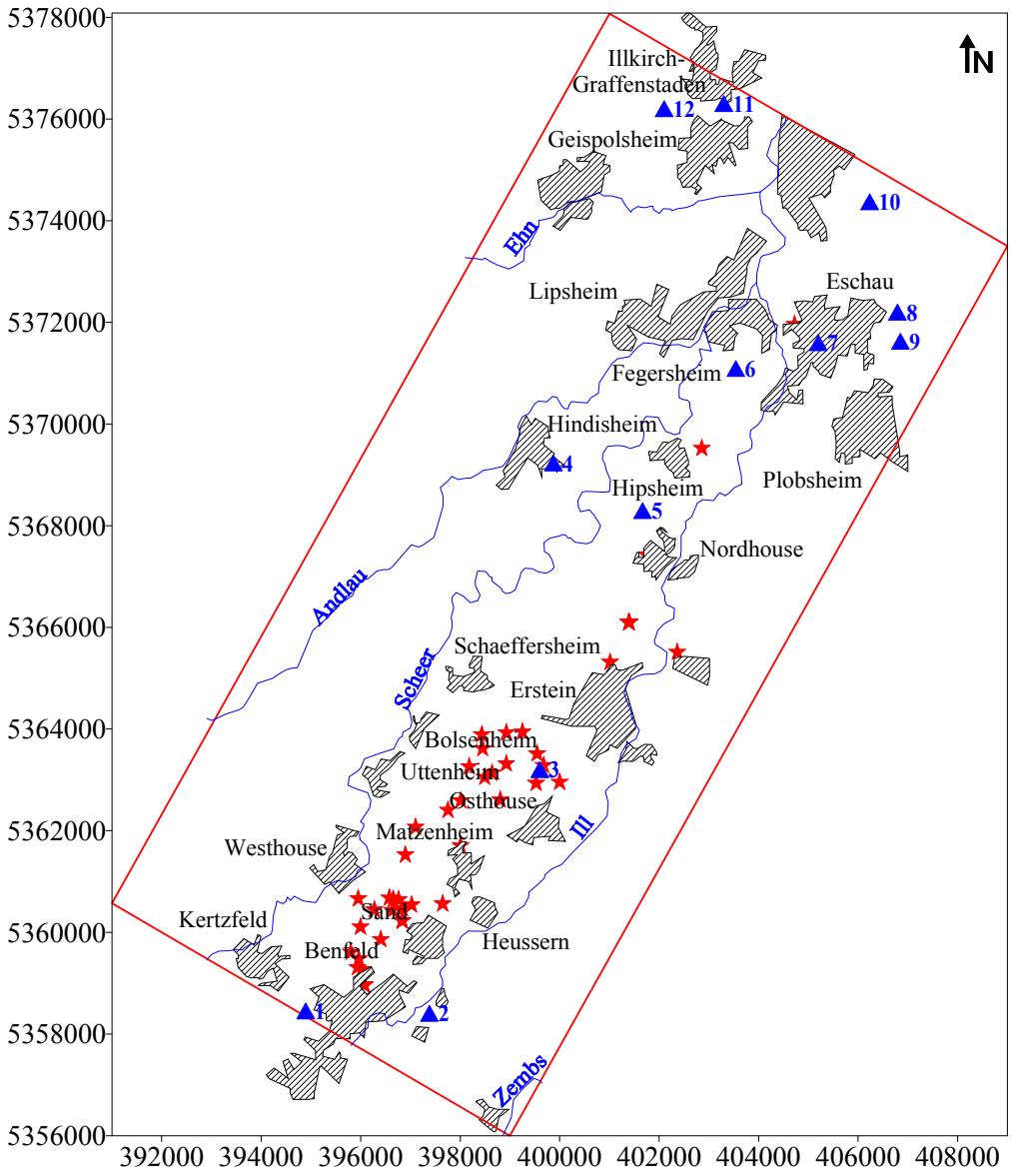


Figure 4.7: rivers and location of observation (stars) and pumping wells (triangles)

Three wells were constructed to supply the city of Erstein in drinking water: the Château d'eau, the Postal, and the Negerdorf wells. These wells are located in the south of Erstein. In particular:

- The old well Château d'eau, which was drilled in 1922, collected water at a depth between 12 and 26 m. This well was only used in necessary situations.

- The Postal well was used to supply water to the city. It was constructed in 1972 and collected water at a depth between 26 and 59 m.
- The Negerdorf well was constructed in 1991 to replace the Postal well. The Negerdorf well is located downstream of the pollution source and collects water at a depth between 49 and 79 m. It consists of three pumps with a pumping rate of 300 m^3/h each (BRGM, 1992). The analysis of the samples collected from the Negerdorf well is very important since it is representative of what is happening for all the piezometers and pumping wells.

The groundwater monitoring network is also composed of 12 pumping wells for home and industrial usage.

In 1996, three piezometers were installed and equipped with multi-level samplers that take samples from many small discrete zones in the aquifer in order to provide accurate vertical contaminant concentration profiles. Depth profiles were considered a necessity for groundwater-quality monitoring because contaminant concentrations can vary significantly in the vertical direction and, in some location, the entire zone of contamination may occupy only a small part of the total aquifer thickness.

Figure 4.9 shows the above mentioned multi-level piezometers located in Benfeld, Sand, between Erstein and Nordhouse. These multi-level piezometers are defined as follows, [Report of Alsace Region, 1997]:

- The piezometer PZ1 is located in Benfeld. The multi-level well reaches about 85 m in depth with 8 multi-level samplings, which are: 1.76-6.76 m, 7.76-12.76 m, 16.26-21.26 m, 28.26-33.26 m, 40.26-45.26 m, 52.26-57.26 m, 64.26-69.26 m, and 79.76-84.76 m below the surface.
- The piezometer PZ2 is located in Sand. The multi-level well reaches about 80 m in depth with 7 multi-level samplings, which are: 2.91-7.91 m, 14.91-19.91 m, 26.91-31.91 m, 38.91-43.91 m, 50.91-55.91 m, 62.91-67.91 m, and 74.91-79.91 m below the surface.
- The piezometer PZ3 is located between Erstein and Nordhouse. The multi-level well reaches about 78 m in depth with 8 multi-level samplings, which are: 1.71-6.71 m, 10.21-15.21 m, 22.21-27.21 m, 34.21-39.21 m, 42.71-47.71 m, 52.21-57.21 m, 64.21-69.21 m, and 72.71-77.71 m below the surface.

Besides an industrial piezometer the PZ4, which is 15 m deep, is located in Benfeld near piezometer PZ1. This is the nearest piezometer to the accident location that has been monitored for several years.

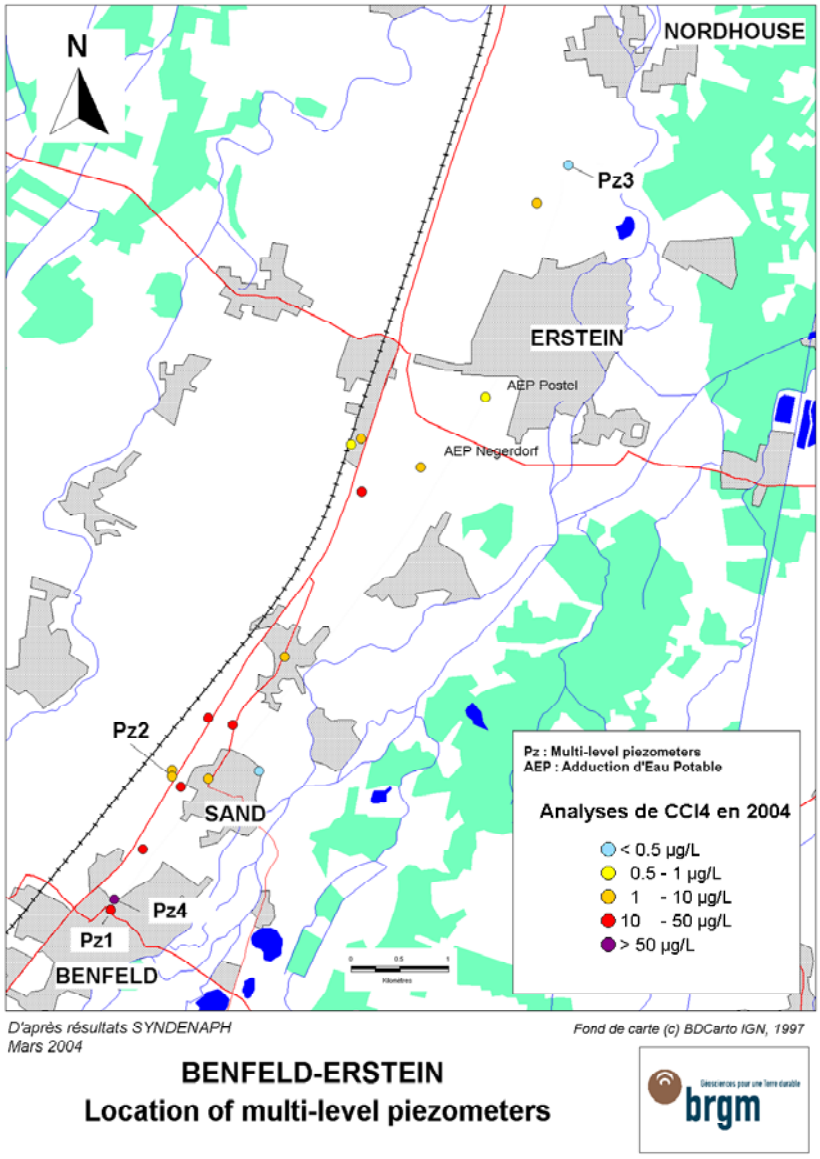


Figure 4.8: location of multi-level piezometers and water supply wells.

At the source zone (Benfeld), an additional well and surface water samplings were installed in 2004 to monitor the concentration of CCl_4 at the source area.

From the groundwater monitoring network concentrations data of carbon tetrachloride have been spasmodically collected between 1992 and 2004. However, only 16 piezometers were used for calibrating the developed model (BRGM, 1993).

From the new and old concentration of CCl_4 near the accident area (Benfeld) VILLIGER-Sytemtechnik (2004) expected two source zones, as sketched in Figure 4.9.

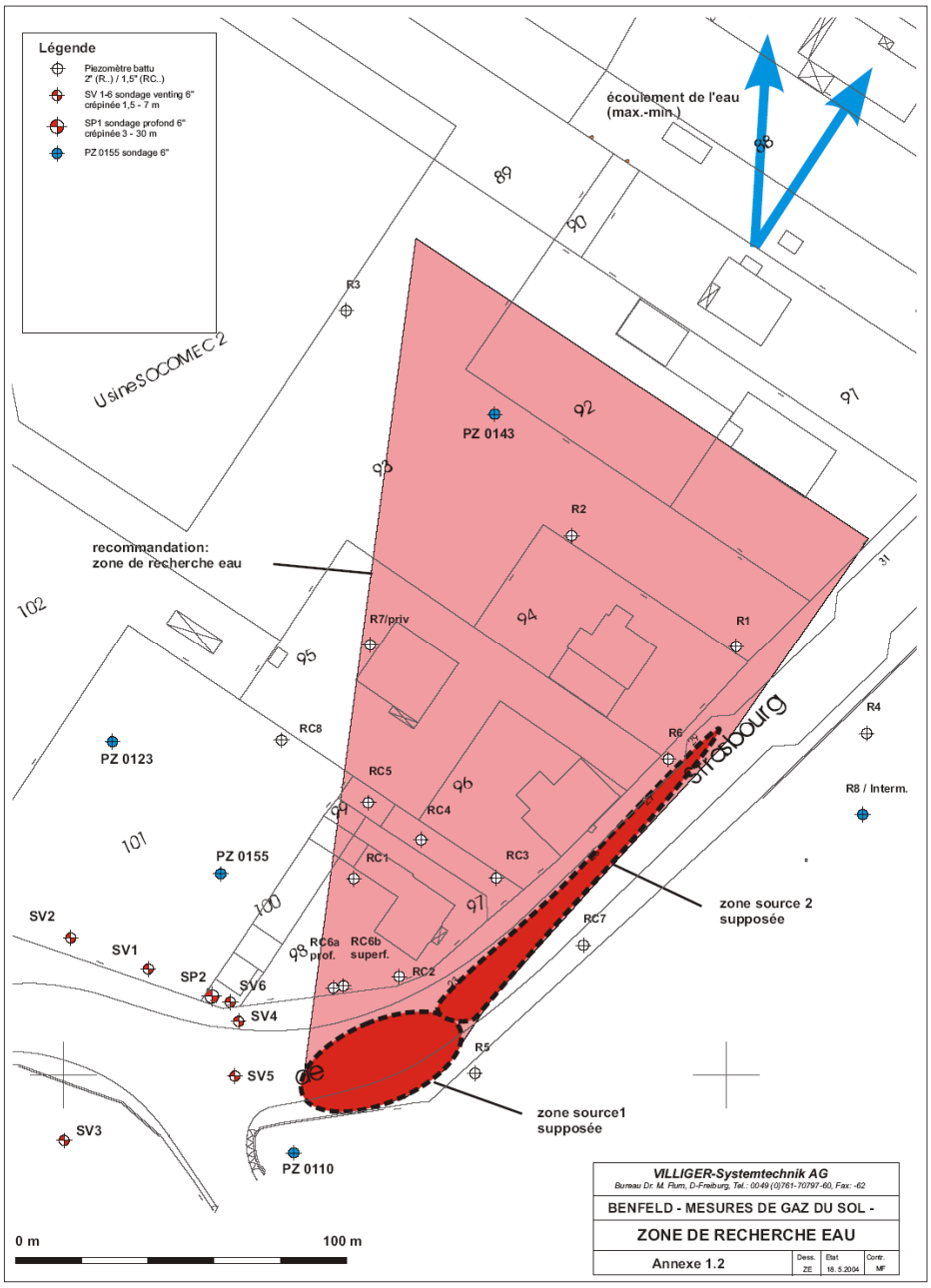


Figure 4.9: location of the source zones (VILLGER-Systemtechnik report, 2004).

The location and depth of the pollutant source are not known or uncertain. Numerous water samples within the area near the location of the accident were taken and analyzed by VILLIGER-Sytemtechnik (2004) as shown in Figure 4.10. The pollution source location and depth has been estimated by Aswed (2008) using the measured concentration of carbon tetrachloride collected in 2004 by VILLIGER-Sytemtechnik.

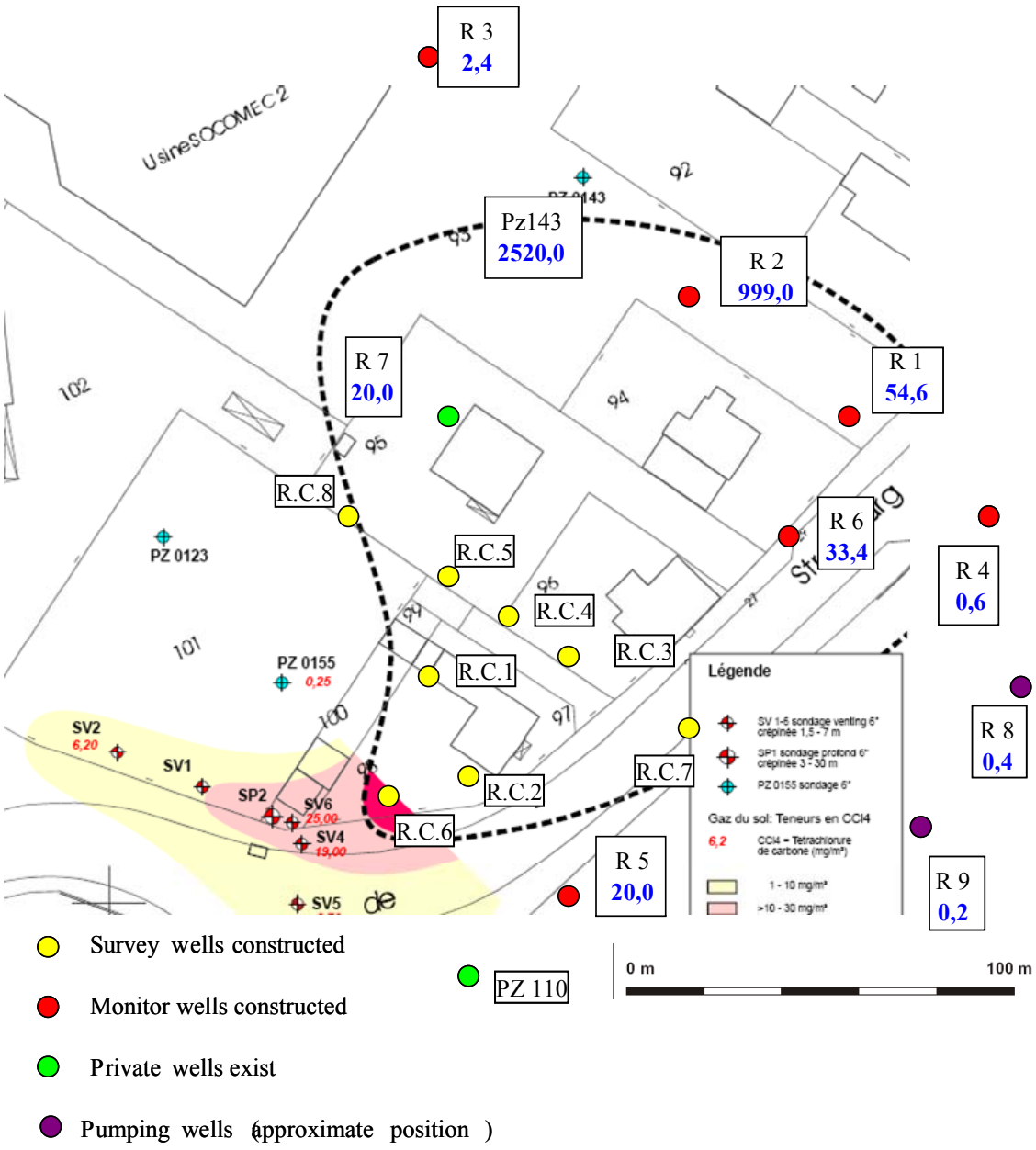


Figure 4.10: observed concentrations of CCl_4 [µg/l] collected on 18/05/2004 (VILLGER-Systemtechnik report, 2004).

In porous media for DNAPL, the pollution depth and area of the infiltration was approximated by Aswed (2008) using the following equation:

$$V_{HC} = S_r \theta_T V \qquad (4.1)$$

where,

V_{HC} : the volume of the pollutant [M^3];

S_r : the residual saturation [-];

θ_T : the porosity of the medium [-];

V : the volume of the contaminated aquifer, [M^3].

Based on a note of the SGAL in 1971, the volume of the infiltrated CCl_4 (V_{HC}) is about 4 m^3. The parameters S_r and θ_T are approximately 2-5% and 15-35%, respectively. Therefore, the volume of the contaminated aquifer is about 230 m^3 to 1300 m^3. To define the depth of the contaminant source, Aswed (2008) based the analysis on the results taken in 1997 from the multi-level piezometer PZ1 at Benfeld (EAT, 1997). These results showed that the concentration of CCl_4 is very low under 35 m depth. Consequently the source is described by the first four layers of the 3D numerical model with two mesh elements per layer; the thicknesses of the layers are respectively, 16, 4, 5, and 5 m from top to the bottom. The surface of infiltration is 7 to 37 m^2 assuming that the pollution depth is 35 m.

The three dimensional model is used by Aswed (2008) in order to estimate the behaviour of the contaminant source at different depths. At first, the concentration at the four layers of the source was fixed. Then, the code TRACES was used to calculate the concentration at each piezometer in the domain and for a given time. In order to find the concentration at the source, the calculated concentrations were matched with measured concentrations using the formula:

$$C_s(t-t_c) = C_{init}(t-t_c)\frac{C_{mes}(t)}{C_{cal}(t)}$$

Where:

C_s is the concentration at the cells representing the source,

t is the time of measurements,

t_c is travel time of the contaminant,

C_{init} is the constant value (100 μg/l) at the source,

C_{mes} is the concentration measured at the piezometers,

C_{cal} is the concentration calculated with TRACES at the piezometers.

The travel time of the contaminant t_c between the source and the measurement-wells is estimated by the temporal moments method. The time at the source t_s is given by:

$$t_s = t - t_c$$

The travel time of the contaminant between the source and the measurement-wells is calculated by method of moment as follows:

$$t_c = \frac{\int_0^\infty C_{cal}(t) \cdot t \cdot dt}{\int_0^\infty C_{cal}(t) \cdot dt}$$

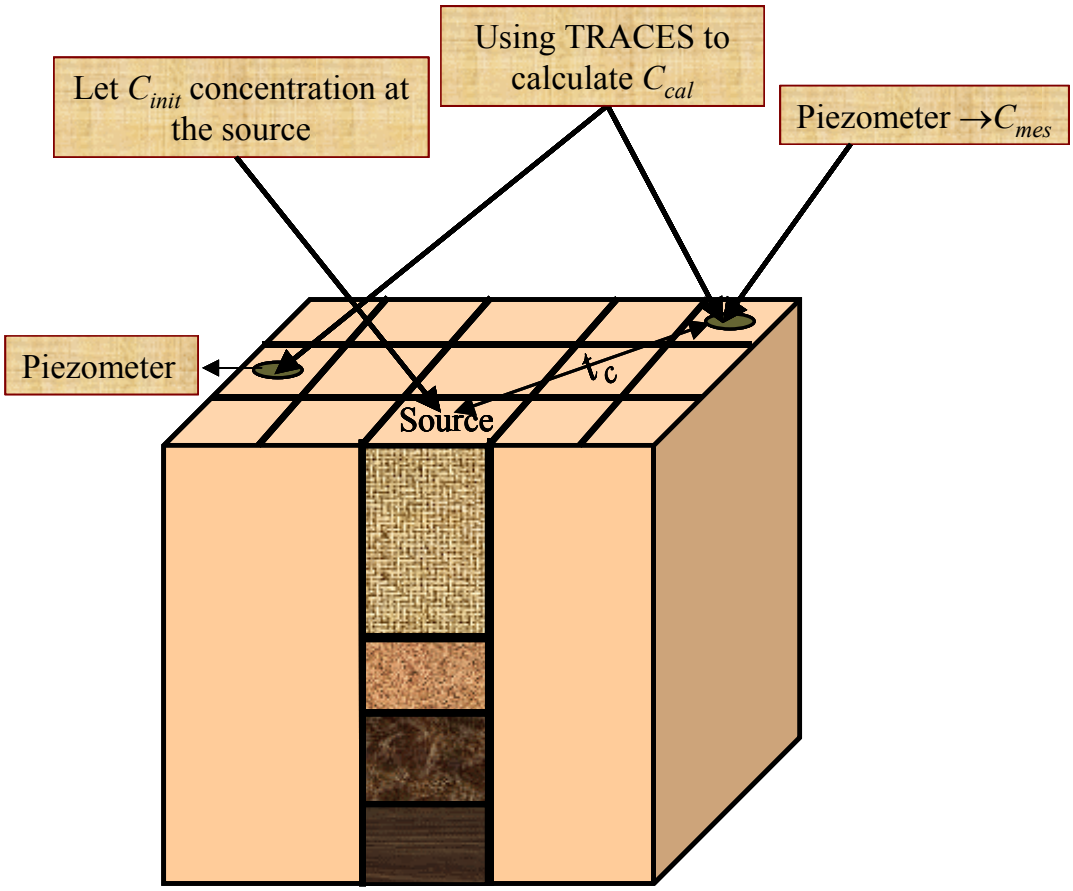


Figure 4.11: technique used to estimate the source term.

The approximated concentrations at the source term was very oscillatory. In order to have a smooth behavior of the source function in each layer, the predicted concentrations at the source are approached by using: the mean value interpolation and

a fitted exponential interpolation. The source functions in the four layers are depicted together in Figure 4.12 below:

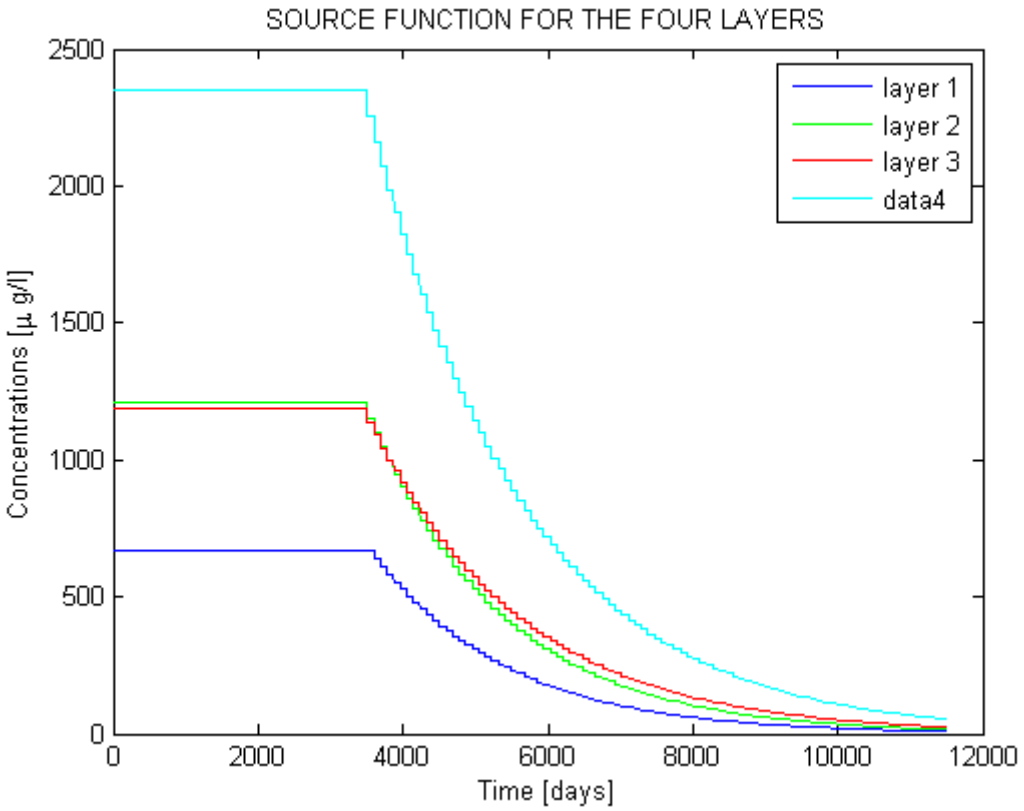


Figure 4.12: source function in the four layers.

4.4.4 Flux and transport model of the Alsatian aquifer

The following figures show the simulated CCl_4 plume, at 1825, 3650, 8010, 10200 and 20000 days after the accident occurred in 1970, performed by Aswed (2008) on the bases of the above mentioned estimated source.

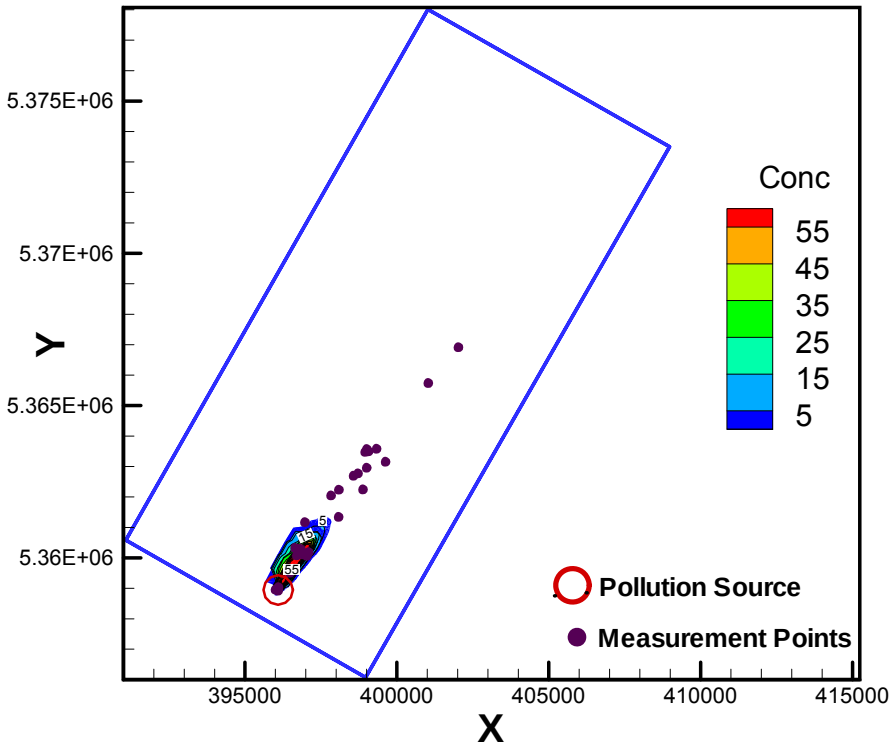


Figure 4.13: distribution of CCl_4 concentration [µg/l] after 1825 days of the accident.

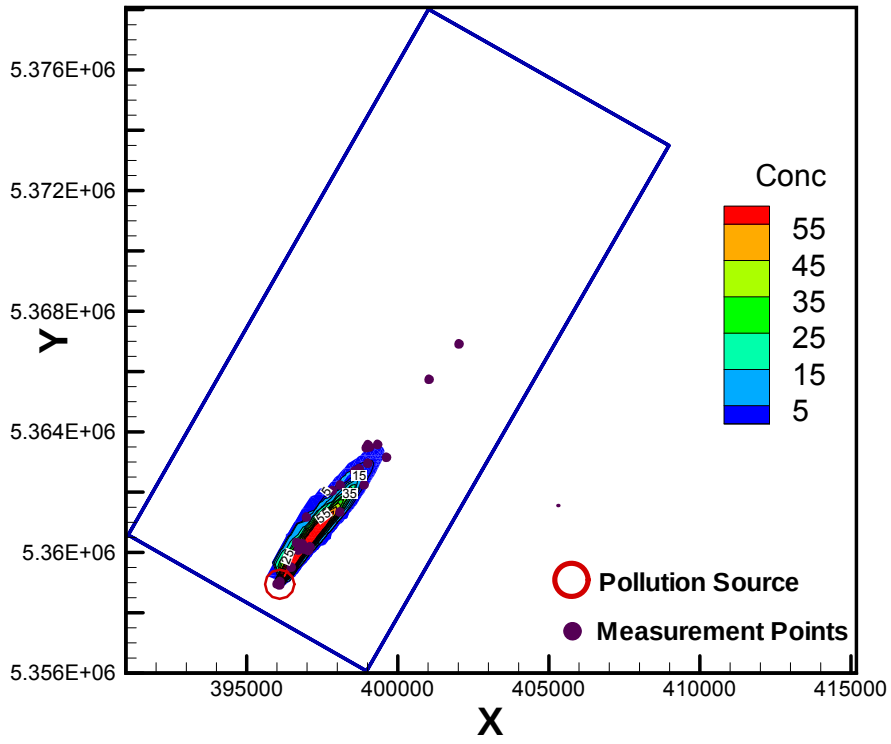


Figure 4.14: distribution of CCl_4 concentration[µg/l] after 3650 days of the accident.

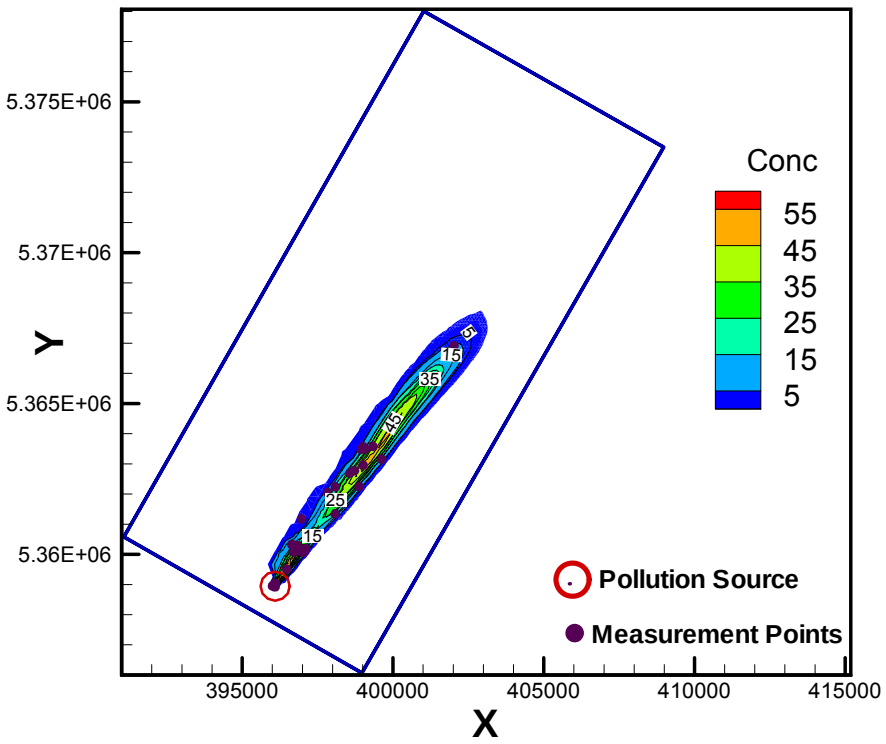


Figure 4.15: distribution of CCl_4 concentration [µg/l] after 8010 days of the accident.

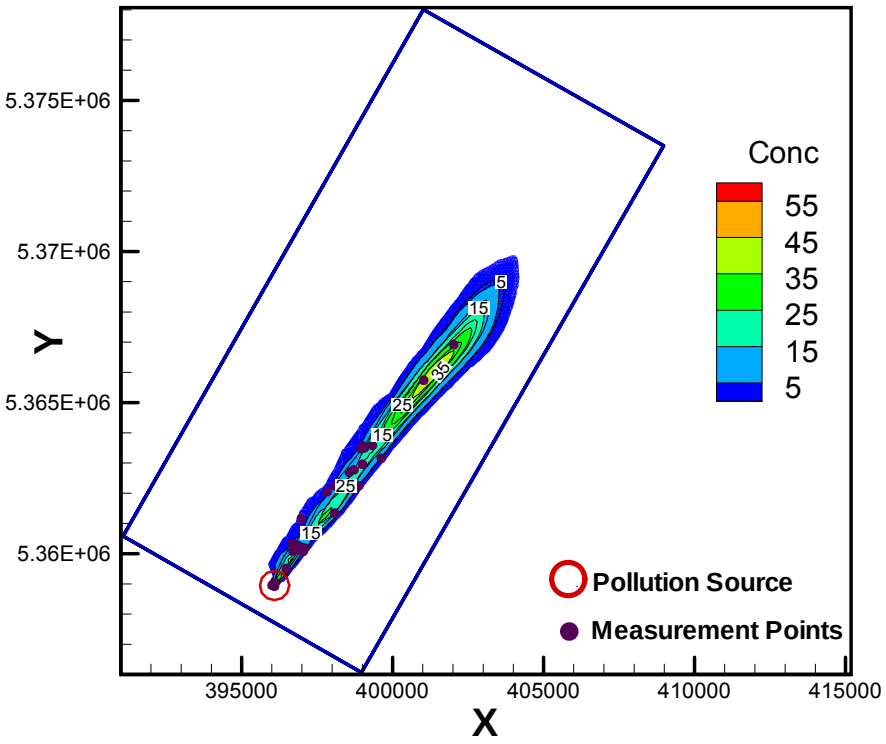


Figure 4.16: distribution of CCl_4 concentration [µg/l] after 10200 days of the accident.

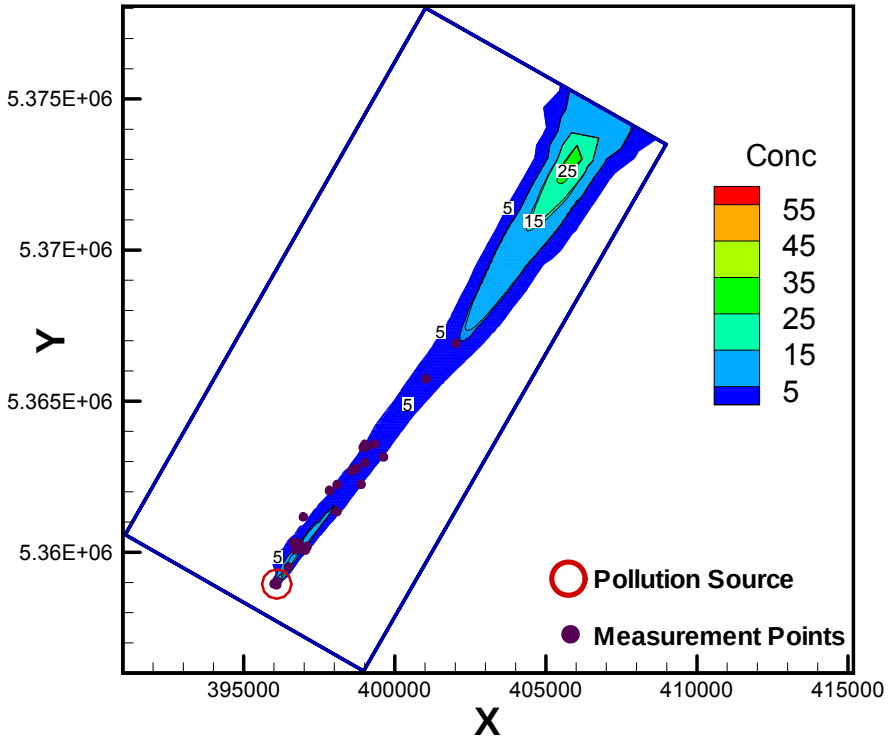


Figure 4.17: distribution of CCl_4 concentration [µg/l] after 20000 days of the accident.

4.5 ANN methodology to study the Alsatian aquifer pollution

As mentioned above, the pollution source behaviour at the accident location is unknown. It is characterized by time varying fluxs in the vertical position. The objective of this part of the research is to identify this unknown pollution source behaviour. To this end, concentration measurement data from monitoring wells may be utilized.

An artificial neural network based methodology has been developed to solve the inverse problem for the Alsatian aquifer contamination. Data for training the ANN are simulated using the groundwater flow and contaminant transport numerical model developed by Aswed (2008).

In a first step, the artificial neural network was trained to solve the direct problem. In this part of the application, the network was trained by means of examples, to associate the pollution sources features with the corresponding output contaminant concentration at monitoring wells. The input patterns were made of the pollution source features in terms of the injection rates in the vertical. The output patterns were contaminant concentration observation data at 45 monitoring wells. Sources and monitoring wells are related by a biunivocal relationship. It means that to each

contaminant concentration behaviour in monitoring wells corresponds only one source contaminant behaviour.

In a second time, the trained network was inverted in order to solve the inverse problem. On the basis of the known contaminant concentration data in monitoring wells, the pollution sources injection rates in the cross section have been identified.

The data set of patterns for training, validation and test has been constructed through a coherent number of hydrogeological scenarios, based on the Aswed's 3D model of the domain. Data pre-processing based on feature extraction techniques have been applied in order to reduce the size of the ANN patterns.

The ANN approximation implies different issues in input and in output matrices. In fact, the aim of the proposed approach is to reconstruct the profiles of the pollutant source. Therefore a great precision is needed for the input, whereas the output has to be calculated on the basis of measurements, so that we only need outputs corresponding to different cases are distinguishable. On the other hand, while the input has only four time-varying concentrations, the output corresponds to 45 wells, therefore we can expect that a greater number of components are necessary to describe the output rather than the input.

Figure 4.6 reported a schema of the applied methodology.

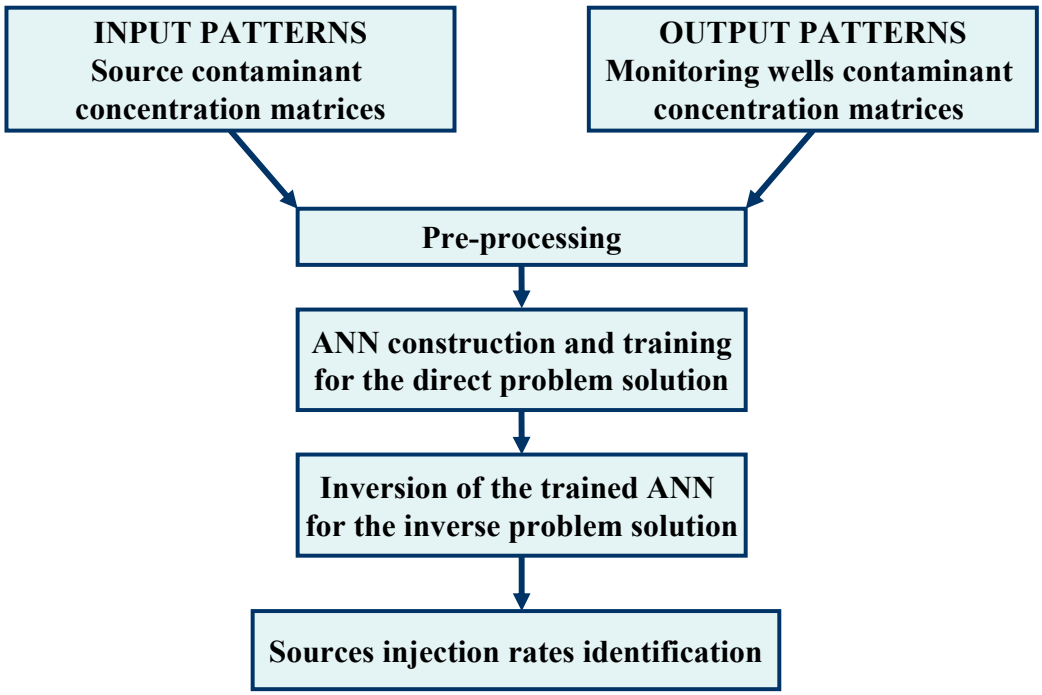


Figure 4.18: schema of the applied procedure.

4.6 ANN pattern construction: elaboration and reduction

4.6.1 ANN input and output data elaboration

Different scenarios of the contamination source behaviour have been constructed using the Alsatian aquifer 3D model developed by Aswed (2008). In other words, this model was the base for generating the ANN patterns. Various states of pollution sources have been constructed with Excel, adjusting the source characteristics in terms of injection rates over the vertical section. 104 examples were constructed.

All the 104 sources have the same duration of activity of about 31 years (11520 days) and were located in the same positions in the domain (accident site in Benfeld). So for each of the 104 sources, a different evolution of the contaminant concentration in the time for the 4 layer in the numerical domain has been considered. The total time of simulation is about 54 years (20000 days).

TRACES simulation samples have been saved as an ASCII matrix file before being processed with Matlab 7.1.

The sources scenarios provided to TRACES for the simulations were characterized by different injection rates and time intervals. In order to have uniform time intervals, it was considered appropriate to spread sources contaminant concentration evolution in the total time of the source activity duration (11520 days). From this process, we have one value a day of the CCl_4 concentration in the source location. This method has allowed us to extract a maximum of information from the input patterns and to make coherent comparison between the results.

The examples patterns obtained with TRACES consist of 208 matrices of contaminant concentration:

- 104 matrices correspondent to the pollution sources features. These had dimension of [11520x4]. 11520 represents the time (days) and 4 represents the layers in the source location,
- 104 matrices correspondent contaminant concentration in monitoring wells. These had dimension of [4000x45]. 4000 represents the time (one each five days) and 45 represents the monitoring wells in the domain. In this case it was taken only one value of contaminant concentration in monitoring wells each

five days, for computational needs: this time step could not be increased, due to numerical criterion.

Input and Output matrices were too large to be processed through the ANN, so a data pre-processing has been performed in order to reduce their dimension. The feature extraction techniques, applied in this case, have been choice on the bases of preliminary neural models trials. The following paragraph describes the feature extraction techniques used in this case.

4.6.2 ANN input and output data reduction

Each ANN example pattern was composed by two matrices: one for the input and one for the output. In order to make this data useful for the training of the ANN, it is necessary to reorganize the 104 input matrices and the 104 output matrices to form two unique big matrices: one for the input and one for the output. Several data pre-processing may be used in such cases. The chosen is strictly linked to the ANN approach selected.

The two groups of input and output matrices were considered separately. The pre-processing procedure used for each group is divided as follows:

- the application of the two-dimensional discrete fast Fourier transform (2D-FFT) to each matrices to the end of make the two input and output big matrices,
- the reduction of the two big matrices by a fixed threshold,
- the normalization and the application of the principal component analysis (PCA) to the two reduced matrices in order to further reduce their size.

A scheme of the pre-processing for input and output data is presented in the Figure 4.19.

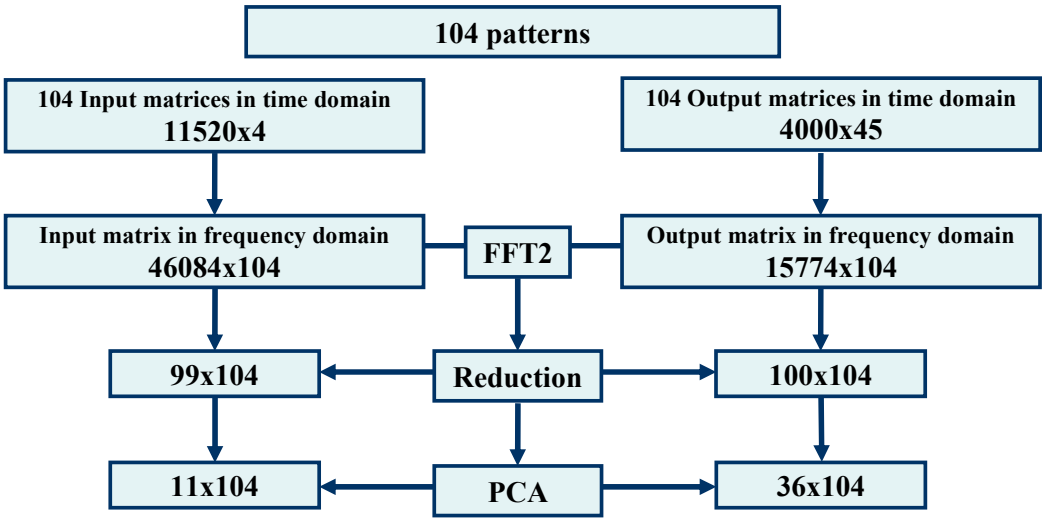


Figure 4.19: matrices reduction schema.

For the first step, the 2D-FFT has been calculated for each matrix, using Matlab. In this way, the matrix is transformed from time domain into frequency domain representation. The matrix in frequency domain are characterized by replics of the same information, so only one replica for each matrix was considered. For the four columns have been identified the elements that are not repeated and were stacked to form a column vector. In a second step, the amplitudes and the phases of the column vector have been calculated and stacked to form a big column vector where the first half of the column is represented by the components relating to the amplitudes and the second half of the components is related to the phases. For each pattern we have two vectors: one for the input that has dimension of [46084] elements and one for the output that has dimension of [15774] elements.

Vectors have been joined to form a big matrix where each column corresponds to one example. Consequently, the two groups of input and output matrices were reorganized to form two big matrices: one for the input patterns concerning of the sources features in frequency domain and one for the output patterns regarding the wells features in frequency domain. The big input matrix have dimension of [46084x104]. The big output matrix have dimension of [15774x104].

For the second step, the matrices frequency components of each matrix have been compared on the basis of their amplitude. For input and output matrices, a threshold has been set in order to select only the most significant 2D-FFT amplitude components. The remaining components and the corresponding phases components have been set to zero. The value of the threshold is determined by searching a crossover between the

approximation of the acquired data and the dimension of the input and output matrices. During this reduction it was necessary to push the reduction in order to have a number of the matrix row for the two matrices less or equal to the number of the examples. This dimension is the condition that we have to respect in order to apply the principal component analysis. As a result of this further reduction, the big input matrix has dimension of [99x104] and the big output matrix has dimension of [100x104].

For the third step, the last feature extraction based on normalization and principal component analysis was applied. Before applying the PCA, the matrix has been normalized so that they have means of zero and standard deviations of 1.

Principal component analysis is a powerful technique used as a mathematical tool for analysing, classifying and reducing numerical datasets. By means of PCA, systematic information initially dispersed over a large matrix of variable input (often intercorrelated) is extracted and condensed in a few abstract variables. Typically, PCA decomposes the primary data matrix by projecting the multi-dimensional data set onto a new coordinates base formed by the orthogonal directions with data maximum variance. The eigenvector of the data matrix are called principal components and they are uncorrelated among them. The magnitude of each eigenvector is expressed by its own eigenvalue, which gives a measure of the variance related to that principal component. As a result of the coordinate change, a data dimensionally reduction to the most significant components and the elimination of the less important ones are possible to achieve without any considerable information loss. The advantage of the PCA application is that it allows the elimination of the principal components that contribute less than a default value λ (expressed as a percentage of total variation in the datasets), so it is possible to define a priori the order of approximation due to loss of information. After various trials, to define the minimum number of principal components necessary to allow a good ANN performance, the minimum fraction variance component to keep our case was: for the input 0.002 and for the output 0.0000001.

The PCA application allowed an additional reduction with less loss of information. The big input matrix has a dimension of [11x104] and the big output matrix has a dimension of [36x104].

4.7 MLP network development and inverse problem solution

4.7.1 MLP networks development

The ANN was trained with the supervised learning by presenting the couples of input (pollution sources scenarios) and target patterns (contaminant concentration for the 45 monitoring wells).

The network architecture was a Multi Layer Perceptron (MLP) consisting in 3 layers: input, hidden and output layer. The input layer is composed by 11 neurons corresponding to the principal components of the pollution sources contaminant concentration for the 4 layers in the domain. The output layer is composed by 36 neurons corresponding to the principal components of the contaminant concentrations measured in the 45 monitoring wells. After the definition of the input and output layer dimension the number of hidden neurons have been defined. Several trainings are performed assuming a growing number of neurons in the hidden layer in order to identify the optimal number of hidden neurons. In our case, we have taken the same number of neurons of the input layer in the hidden layer.

The activation functions define the functional dependence between input and output. These determine the type of response of each neuron. In this case, an hyperbolic tangent activation function was used for the hidden layer and linear activation function for the output layer.

The learning algorithm selected is the Error Back Propagation (EBP) algorithm optimized by the Levenberg-Marquardt (LM) algorithm.

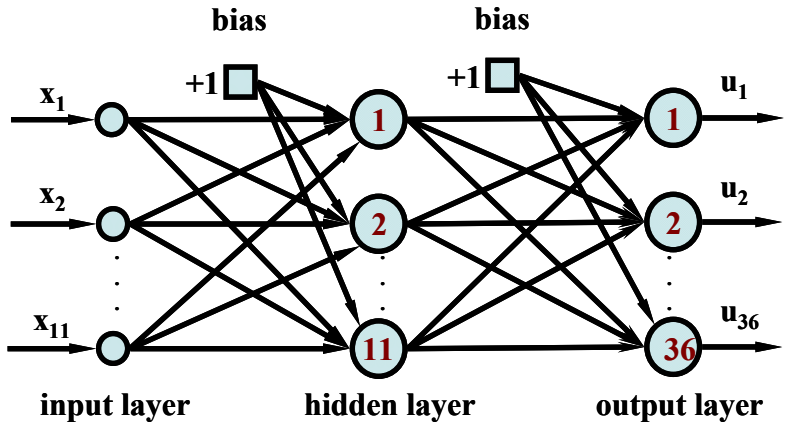


Figure 4.20: artificial neural network for the study of the Alsatian aquifer pollution source.

Figure 4.20 show the topology of the ANN generate for the Alsatian aquifer pollution source.

The learning rule used in this case is the “stopped training”. Based on the above mentioned learning rule, the example patterns are divided in three sets:

- The training set used during the training phase for adjusting the connection weights. It is composed of 74 examples.
- The validation set used for stopping the training phase. It is composed of 19 examples.
- The test set used for testing the trained ANN. It is composed of 11 examples.

Once the aforementioned key features are selected, the artificial neural network may be trained. The training of the ANN is the critical part of the proposed process. In fact, a particular attention has to be dedicated to train the ANN in such a way that it is able to generalize the information contained in the training set. The training is carried out with the presentation of a whole set of patterns: input and output pairs. The presentation of a whole set of patterns is called the epoch. The training is based on the iterative presentation of the epochs according to a random sequence of the patterns.

During the training phase the connection weights are modified in order to minimize the error between the target and the network output of the training set examples. In the same time the error is calculated in the validation set. When the error starts increasing, the training is interrupted.

Through the EBP algorithm, the error between the calculated output and the desired output, for a particular input state, is propagated backwards through the weights of the hidden layers, until it reaches the input layer (from output to input). The goal is to isolate the influence of each connection weight in the error between calculated and target. In a first step, on the basis of trainings trials, the number of epochs is assumed equal to 100. An additional number of epochs does not improve the ANN because after several trials, it has been observed that with the “stopped training” rule the number of 100 epochs was never totally reached.

Figure 4.21 shows the training (blue), validation (green) and test (red) error evolution of the ANN during the training phase after 19 epochs.

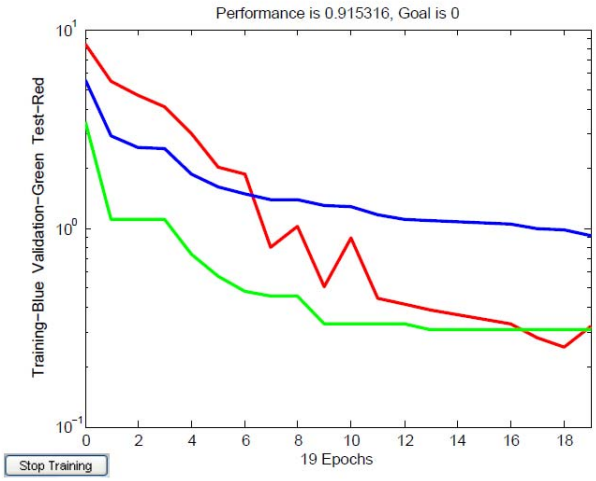


Figure 4.21: error representation during the training of the ANN.

4.8 Solution of the direct and inverse problem with the ANN

As in the case, presented in the chapter 3 of this work, the network was firstly trained to solve the direct problem. During the training phase all the connection weights have been determined and frozen. The trained network was inverted to solve the inverse problem and on the basis of the contaminant concentrations measured in the 45 monitoring wells the corresponding contaminant concentrations of the pollution source have been determined. On the basis of the three equations described in (3.2), starting from the output u, the vectors y and x have been determined. During the inversion process, the difference between the calculated input and the desired input was considered in order to verify if the trained ANN allowed to generalize the information contained in the training set.

In our case, the number of hidden neurons is lower than the number of output neurons therefore the system is overdetermined (see Figure 4.20). Provided that the matrix $\underline{\underline{W}}_2$ is full rank, the solution corresponding to the minimum sum squared error can be found as the equation (3.3). The second equation in (3.2), states a biunivocal relation between y and h, therefore the vector y can be calculated as in the equation (3.4). Finally, since we made the choice of having an equal number of neurons in the input and hidden layers, $\underline{\underline{W}}_1$ is a square matrix, and if it is full rank, the input pattern x can be calculated as in the equation (3.5).

The desired source injection rates for the 4 layers have been obtained by applying backward the pre-processing of the input data obtained from the inversion of the ANN.

In the Figure 4.22, Figure 4.23, Figure 4.24 and Figure 4.25 are showed the simulated (blue) and the inverted (green) ANN input patterns for the 4 layers. The simulated input pattern corresponds to the pollution source behaviour reconstructed during the inverse problem solution performed by Aswed (2008). As one can see in the following figures, the ANN approximated concentrations at the source term are very oscillatory due to the feature extraction.

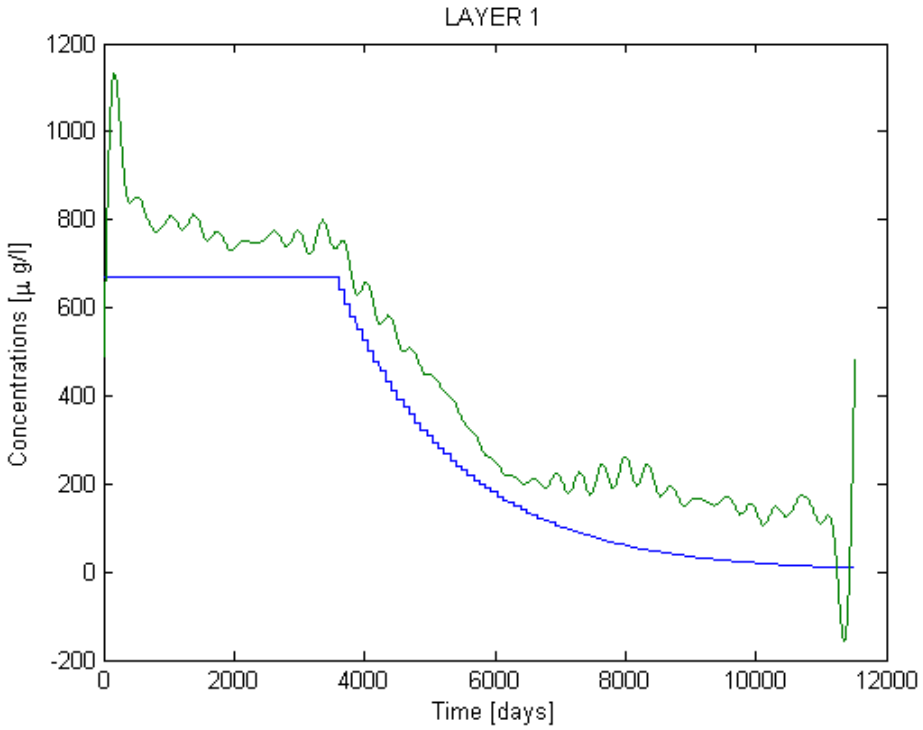


Figure 4.22: simulated (blue) and the inverted (green) pollution source behaviour of the first layer.

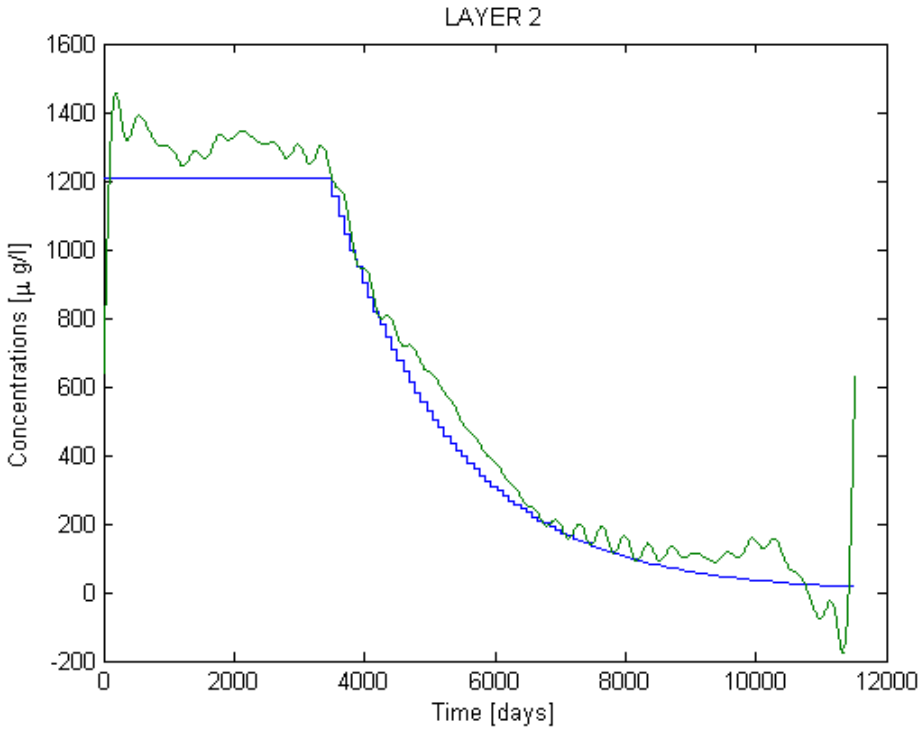


Figure 4.23: simulated (blue) and the inverted (green) pollution source behaviour of the second layer.

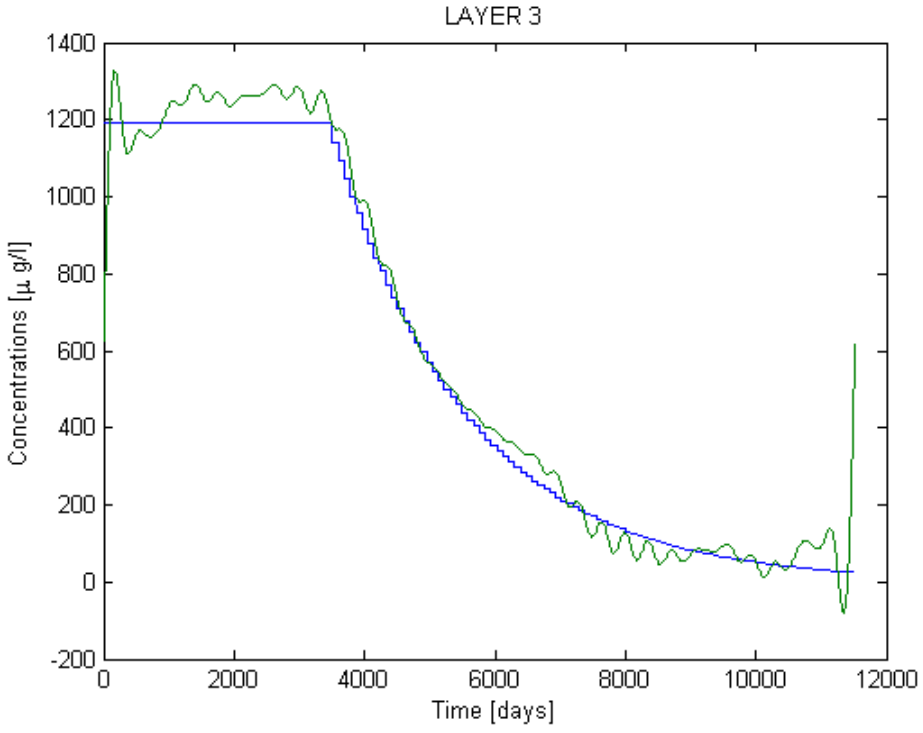


Figure 4.24: simulated (blue) and the inverted (green) pollution source behaviour of the third layer.

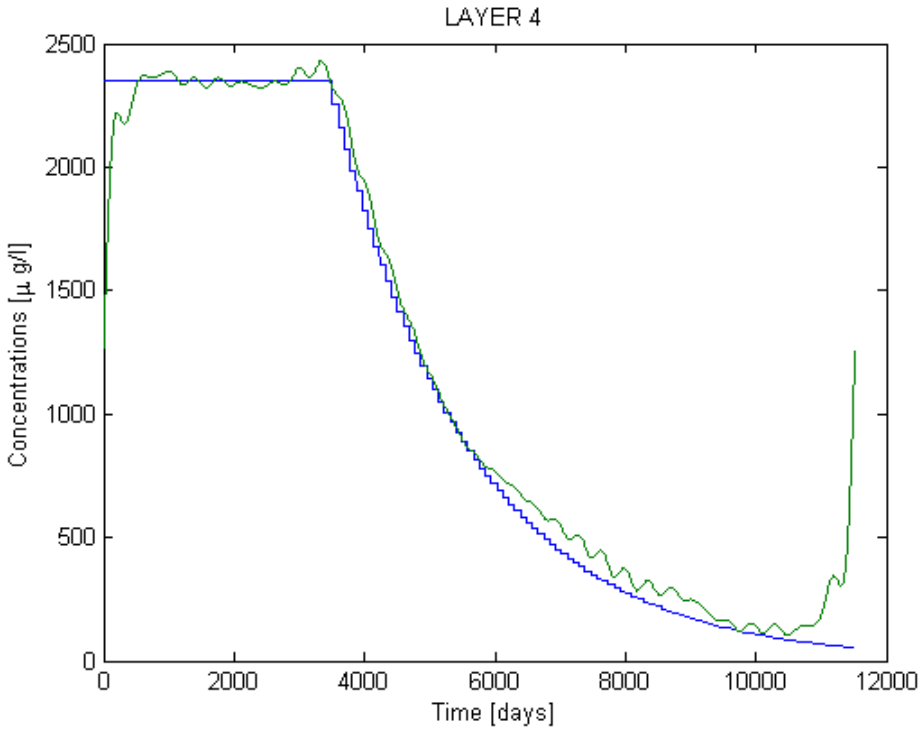


Figure 4.25: simulated (blue) and the inverted (green) pollution source behaviour of the fourth layer.

4.9 Summary

This chapter proposes a methodology that aims at solving the inverse problem for the Alsatian aquifer pollution. Based on ANN technology, the inverse problem is solved through the measurements of contaminant concentrations acquired in monitoring wells during 12 years (from 1992 to 2004). The behaviour at the four layers of the carbon tetrachloride unknown pollution source is reconstructed by inverting the trained ANN.

CONCLUSIONS

This work investigates the feasibility of a new approach for solving the inverse problem in aquifers contamination by using artificial neural networks. These computational tools have the ability to learn complex input-output relationships and have a very advantageous property of generalization.

In the field of hydrogeological studies, the literature, starting from the late '80s, but especially in the mid '90s, contains many examples of implementation of different ANN applications related to various issues. In literature, several studies are dedicated to the development of different models for solving the inverse problem, however practical works using the ANN approach are less popular.

In this thesis, data for training the ANNs have been simulated using the non-commercial software for modeling flow and contaminant transport TRACES. Neural networks have been implemented using the Neural Network Toolbox of Matlab 7.1.

The methodology developed in this work is applied on two different cases of continuous point contamination sources: a theoretical case and a real case(unknown pollution source of the Alsatian aquifer in France).

For the two studied cases, the operational steps are described as follows:

- study of the flow and contaminant transport models used for developing the suitable hydrogeological scenarios;
- creation of the ANN examples patterns based on the different hydrogeological scenarios;
- realization of the data pre-processing aimed at reducing patterns' size in order to make the patterns useful for ANNs;
- implementation of the ANNs suitable to address the aquifers pollution inverse problems;
- inverse problem solution.

The first part of the research shows how artificial neural network can be used as an efficient tool to locate in time and space an unknown pollutant source for a simple theoretical case.

120 different hydrogeological scenarios have been designed by considering 40 sources were located in different positions of the domain with a timing of activity of 10, 20 and 30 years. The samples obtained from the simulation model were the contaminant concentration curves acquired in the 50 cells distributed uniformly in the domain.

For output data reduction, the restricted hypothesis of the total absence of historical data concerning the aquifer pollution has been taken into consideration, in particular, the case of a contamination detection for the first time in a generic domain. Only the last value of contaminant concentration in time, obtained through simulations, was considered for the 50 cells. These 50 cells correspond to 50 hypothetical monitoring wells. An ANN based approach was applied in order to reduce the number and choose the best location of the monitoring wells in the domain: only 8 cells have been kept into consideration.

Based on the *"Leave one Out Cross Validation" (LOO)* training learning rule, 120 *3-8-8 MLP networks* have been trained through supervised learning in order to solve the direct problem. The learning algorithm chosen is the *Error Back Propagation algorithm* optimized by the *Levenberg-Marquardt algorithm*.

Each network was trained with 120 examples and tested with one example (kept out from the original training set). The test example allows us to estimate the network generalization capacity.

In our case, the *LOO* procedure is not used to train the network that will be applied to a particular case, but only to estimate the generalization ability of the 120 trained networks. If one wants to consider a new source not included in the 120 patterns, all the patterns will be used for the training set and the new case will be used for the test set. The methodology developed for this theoretical case allows us to reach the reasonable presumption that the error for the new case will not be greater than the errors experienced in the 120 networks already trained.

The 120 networks were firstly trained to solve the direct problem. Once the training phase is completed, each the trained networks were inverted using the test example to solve the inverse problem. On the basis of the measurements of the contaminant concentration acquired in the monitoring wells, the desired source characteristics have been found.

In general, the results show very good performances in locating the pollutant source. The results show that most of the time the identification error is less than the size of one cell, in fact the cell size is equal to 20*20 m^2. At the same time, the maximum error, which represents the worst case, is less than the size of two cells. Less satisfying results have been obtained concerning time step prediction with the 76% of correct duration activity approximation. Concerning the sources duration activity of 10 and 30 years, only one case time approximation was wrong. For the 20 years sources duration activity, the wrong cases were higher than the correct cases with 26 wrong cases out of a total of 40 cases.

The methodology applied for the theoretical case, however, may be useful not only to identify the location and activity of unknown pollution sources, but also to delimitate the study area and optimize the investigation costs by determining the best monitoring wells location.

In the real case considered in this study, the behaviour of the unknown pollution source, which following an accident in 1970 has polluted with carbon tetrachloride (CCl_4) the Alsatian aquifer (France), has been defined. 104 different hydrological scenarios have been build in order to generate the patterns for training, validating and testing an *11-11-36 MLP network*. The network has been firstly trained to solve the direct problem and, in a second moment, it has been inverted to solve the inverse problem.

On basis of the measures of the contaminant concentration curves acquired in the monitoring wells from Aswed's 3Dnumerical model (2008), the behaviour of the unknown pollution source for the 4 layer at the accident location was found. After the training procedure, the ANN has allowed to generalize the information contained in the training set. The research, however, is open to further developments, such as an improvement in the monitoring and implementation of training patterns scenarios based on few values of the real contaminant concentration in the monitoring wells in order to reconstruct the relationship between the source and the contaminant plume.

It is clear that ANNs represent an emerging new technology due to these main properties represented by the ability to be universal approximators. It appear obvious that ANN full potential for solving non-linear hydrogeological problems taking into account the effects of uncertainty in parameters and multi phases flow must be further explored.

Groundwater has become an extremely fragile resource: limited availability contrasts with the increasing demand due to population growth, especially in developing countries. Access to safe water is part the United Nations Millennium Development Goals as a key driver towards poverty eradication and mortality reduction, eventually increasing the total population with sustainable access to an improved water source.

These ambitious political goals set by the international community should be supported by technologies which offer concrete solutions and practical models to protect groundwater resources against pollution. In this regard, we believe that artificial neural network, even though a relatively new approach, should be further developed to timely and efficiently address pollution cases, reducing potential damages on groundwater resources. Together with an investment regarding the technical development of artificial neural network approach, attention should be paid to the public authorities and relevant water-sector stakeholders, making them aware of the potential applications of this technology in terms of detection of the source of groundwater pollution, contributing to the efficient mitigation of contaminated aquifers.

As a final recommendation, provided that sustainable water management including protection of groundwater resources is a major challenge for ensuring long-term social and economic development, artificial neural networks can offer a valuable contribution to the pool of existing solutions in the field of groundwater pollution.

REFERENCES

Anderson M. P. and Woessner W. M, “Applied Groundwater Modeling”, p. 381, Academic, San Diego, 1992.

ASCE - Task Committee on Application of Artificial neural Networks in Hydrology, “Artificial Neural Networks in Hydrology: II: Hydrologic Applications”, Journal of Hydrologic Engineering, Vol. 5, n° 2, pp. 124-137, April 2000;

Aswed T., PhD thesis “Modelisation de la pollution de la nappe d’alsace par solvants chlores”, 2008.

ATSDR, Agency for Toxic Substances and Disease Registry, “Center for Disease Control”. http://www.atsdr.cdc.gov/tfacts30.html, 1995.

Bashi-Azghadi S. N., Kerachian R., Bazargan-Lari M.R., Solouki K., “Characterizing an unknown pollution source in groundwater resources systems using PSVM and PNN”, Expert Systems with Applications, 2010.

Bear, J., “Hydraulic of groundwater”, Mc Grew-Hill Series in Water Resources and Environmental Engineering, New-York, 1979.

Bear, J., “Dynamics of Fluids in Porous Media”, Dover, New York, 1988.

Bertrand, G., Elsass, Ph., Wirsing, G., and Luz, A., “Quaternary faulting in the Upper Rhine Graben revealed by high-resolution multi-channel reflection seismic” C. R. Geoscience 338: 574–580, 2006.

Beyou, L., “Modélisation de l’impact d’une pollution accidentelle de la nappe phréatique d’Alsace. Cas du déversement de CCl_4 à Benfeld”, projet fin d’études pour un diplôme d’ingénieur à l’ENSAIS, 1999.

Bockelmann A., Ptak T., Teutsch G, “An analytical quantification of mass fluxes and natural attenuation rate constants at a former gaswork site”, ELSEVIER, Journal of Contaminant Hydrology, n° 53, pp 429-453, 2001.

Carcangiu S., Di Barba P., Fanni A., Mognaschi M.E., Montisci A., “Comparison of multi objective optimisation approaches for inverse magnetostatic problems”, COMPEL: Int. J. for Computation and Maths in Electrical and Electronic Eng., Vol. 26, N. 2, pp. 293-305, 2007.

Chavent, G., Cockburn, B., Cohen, G. and Jaffré, J., "Une méthode d'élémnets finis pour la simulation dans un réservoir de déplacements bidimensionnel d'huile par de l'eau", rapport INRIA, No 353, 1985.

Clement, T. P., Gautam, T. R., Lee, K. K. and Truex, M. J., "Modeling coupled NAPL-dissolution and rate-limited sorption reactions in biologically active porous media", Bioremediation Journal 8, no. 1–2: 1–18, 2004.

Cybenko G., "Approximation by Superposition of a Sigmoid Function ", Mathematics for Control, Signals and Systems,2, 4, pp. 303-314, 1989.

CPHF, "Health risk assessment of carbon tetrachloride (CTC) in California drinking water", State of California Contract No. 84-84571 to the California public health foundation 1998,.

De Marsily, G., "Hydrologie Quantitative", Masson, Paris, 215 p., 1981.

Domenico, P. A. and Schwartz, F. W., "Physical and chemical hydrogeology", New York, John Wiley and Sons, p. 824, 1990.

Eikenberg, J., Tricca, A., Vezzu, G., Stille, P., Bajo, S. and Ruethi, M., $^{228}Ra/^{226}Ra/^{224}Ra$ and $^{87}Sr/^{86}Sr$ isotope relationships for determining interactions between ground and river water in the upper Rhine valley, Journal of Environmental Radioactivity, 54, 133-163, 2001.

Fanni A., Uras G., Usai M., Zedda M.K., "Neural Network for monitoring. Groundwater". Fifth International Conference on Hidroinformatics, Cardiff, UK, pp. 687-692, 2002.

Fanni A., Montisci A. "A Neural Inverse Problem Approach for Optimal Design", IEEE Transaction on Magnetics, Vol. 39, No 3 pp. 1305 1308, 2003.

Govindaraju R.S. and Ramachandra Rao A., "Artificial neural networks in hydrology", Water Science and Technology Library, 2000.

Guilley, F., "L'emploi des huiles biodegradables dans les ecosystemes forestiers en remplacement des huiles minerales polluantes", projet fin d'études pour un diplôme d'ingénieur de l'ENSTIB, 2004.

Gümrah F., Öz B., Güler B. And Evin S., "The application of artificial neural networks for the prediction of water quality of polluted aquifer", Water, Air, and Soil Pollution 119: 275–294, 1990.

Hamond, M. C., “Modélisation de l’extension de la pollution de la nappe phréatique d’Alsace par le tétrachlorure de carbone au droit et à l’aval de Benfeld”, mémoire pour le DESS sciences de l’environnement à l’IMFS, 1995.

Holder TH., Teutsch G., Ptak T. and Schwarz R., “A new approach for source zone characterization: the Neckar Valley study”, Groundwater Quality: Remediation and Protection (Proceedings of the GQ'98 Conference held at Tubingen, Germany, September 1998). 1AHS Publ. no. 250, 1998.

Hoteit H. and Acherer Ph., TRACES user’s guide V 1.00, Institut mécanique des fluides et des solides de Strasbourg, 2004.

HSG 108, “IPCS International programme on chemical safety health and safety”, Guide No. 108 carbon tetrachloride health and safety guide,1998.

Ingrassia S. and Davino C., “Reti Neuronali e Metodi Statistici”, Franco Angeli, Milano, 2002.

Kinzelbach, W., “Groundwater Modelling, Developments in Water Science 25, Elsevier Science Publishers”, Amsterdam, 1986.

Macdonald, T. L., “Chemical mechanisms of halocarbon metabolism”, Crit. Rev. Toxicol. 11, pp. 85-120, 1982.

Mahar P. S. and Datta B., “Identification of Pollution Sources in Transient Groundwater Systems”, Water Resources Management, n° 14, pp. 209-227, Kuwer Academic Publishers, 2000;

Maier. H.R. and G.C.Dandy, “Neural networks for the prediction and forecasting of water resources variables: a review of modelling issues and applications”, Environmental Modelling & Software, 15: 101-124, 2000.

Mazzetti A., “Reti Neurali Artificiali”, Apogeo.

Mc-Culloch W. and Pitt W., “Brain Theory” Vol 1, Shaw and Palm, World Scintific, 1988.

Morshed. J. and J.J.Kaluarachchi, “Application of artificial neural network and generic algorithm in flow and transport simulations”, Advances in Water Resources, 22(2):145-158, 1998.

NIOSH, “Pocket Guide to Chemical Hazards”, U. S. Dept. of Health and Human Services, National Institute for Occupational Safety and Health, NIOSH publication, 94-116, 1994.

Principe J., N. Euliano, C. Lefebvre, “Innovating Adaptive and Neural Systems Instruction with Interactive Electronic Books”, Proceedings of the IEEE, special issue on engineering education, January 2000.

Rajanayaka C., Samarasinghe S. & Kulasiri D., “Solving the Inverse Problem in Stochastic Groundwater Modelling with Artificial Neural Networks”, available on line at http://www.iemss.org/iemss2002/proceedings/vol2.html, 2002.

Report of Alsace Region, 1997;

Rizzo, D.M. and D.E. Dougherty. "Characterization of aquifer properties using artificial neural networks: Neural Kriging." Accepted Water Resources Research. 30 (2), pp. 483-497, 1994.

Rojas R., “Neural Networks a systematic introduction”, Springer, 1996.

Sciuntu C., PhD thesis “Reti neurali artificiali: una applicazione nello studio di acquiferi contaminati”, 2004.

Schwarz R., Ptak T., Holder Th., Teutsch G., “Groundwater risk assessment at contaminated sites: a new investigation approach”, in Herbert M., Kovar K., Groundwater Quality: Remediation and Protection (Proceedings of the GQ ’98 Conference held at Tübingen, Germany, September 1998), IAHS Publ. n°250, pp. 68-71, 1998;

Singh R. M. and·Datta B., “Artificial neural network modeling for identification of unknown pollution sources in groundwater with partially missing concentration observation data”, Water Resources Management 21:557–572, 2006.

The mathworks, “Neural Networks Toolbox for use Matlab, v.7.1”, 2005

U.S. EPA, “Technical fact sheet on Carbon Tetrachloride”, National Primary Drinking Water Regulations, www://epa.gov/OGWDW/dwh/t-voc/carbonte.html, 1998.

U.S. Patent 1,010,870, “Process of extinguish fire” U.S Patent Office, USA,1911.

U.S. Patent 1,105,263, “Portable fire extinguisher” U.S Patent Office, USA,1914.

Wiedemeier, T. H., Rifai, H. S., Newell, C. J. and Wilson, J. T., "Natural attenuation of fuels and chlorinated solvents in the subsurface", John Wiley & Sons, Inc., New York, 1999.

Wolfe, W. J., Haugh, C. J., Webbers, A., and Diehl, T. H., "Preliminary conceptual models of the occurrence, fate and transport of chlorinated solvents in Karst regions of Tennessee", U.S. Geological Survey Water Resources Investigations, Report 97-4097, 1997.

Zhiqiang L., Rizzo D. and Haydenc N., "Utilizing Artificial Neural Networks to Backtrack Source Location", 2006.

Zio E., "Approaching the inverse problem of parameter estimation in groundwater models by means of artificial neural networks", Progress in Nuclear Energy, Elsevier Science Ltd, Vol 31, n° 3 pp. 303-315, 1997;