



UNIVERSITA' POLITECNICA DELLE MARCHE
FACOLTA' DI INGEGNERIA

Corso di Laurea Magistrale in Ingegneria Meccanica

**ALGORITMI DI MACHINE LEARNING NON SUPERVISIONATI
APPLICATI ALLA MANUTENZIONE PREDITTIVA NEL SETTORE
OIL & GAS**

**UNSUPERVISED MACHINE LEARNING ALGORITHMS
FOR PREDICTIVE MAINTENANCE IN OIL & GAS SECTOR**

Relatore:
Prof. Giovanni Mazzuto

Tesi di Laurea di:
Lorenzo Concetti

A.A. 2021 / 2022

Sommario

ELENCO DELLE FIGURE.....	IV
ELENCO DELLE TABELLE.....	VI
INTRODUZIONE.....	1
1. INTRODUZIONE AL CASO DI STUDIO.....	2
1.1. L'INDUSTRIA 4.0 E IL DIGITAL TWIN.....	2
1.2. LA MANUTENZIONE.....	3
1.3. LA MANUTENZIONE PREDITTIVA.....	4
1.4. CENNI SUI SISTEMI BIFASE GAS-LIQUIDO.....	4
2. DESCRIZIONE DELL'IMPIANTO.....	5
2.1. SCHEMA DELL'IMPIANTO E FUNZIONAMENTO.....	6
2.2. COMPONENTI IDRAULICI.....	8
2.2.1. <i>La pompa</i>	8
2.2.2. <i>L'eiettore</i>	9
2.2.3. <i>Il serbatoio</i>	10
2.2.4. <i>Le valvole d'intercettazione</i>	11
2.2.5. <i>Le tubazioni</i>	13
2.2.6. <i>Le valvole di regolazione</i>	13
2.3. COMPONENTI ELETTRONICI.....	14
2.3.1. <i>Revolution Pi</i>	14
2.3.2. <i>I sensori</i>	16
2.3.2.1. Endress+ Hauser Cerabar M PMP51.....	17
2.3.2.2. Endress+ Hauser Promag W.....	18
2.3.2.3. Setra 280E.....	19
2.3.2.4. Foxboro 841GM CI1.....	20
2.3.2.5. Foxboro Vortez DN 50.....	21
2.3.2.6. Foxboro IDP-10.....	22
2.3.2.7. Endress+ Hauser Prowirl 200.....	23
3. LA CALIBRAZIONE DEI SENSORI.....	24
3.1. CENNI TEORICI.....	24
3.1.1. <i>Cos'è un sensore</i>	24
3.1.2. <i>La taratura statica</i>	24

3.1.3.	<i>Il metodo dei minimi quadrati</i>	25
3.1.4.	<i>Il coefficiente di correlazione R^2</i>	26
3.2.	GLI STRUMENTI PER LA CALIBRAZIONE	27
3.2.1.	<i>Le misure di pressione</i>	27
3.2.2.	<i>Le misure di flusso</i>	28
3.2.3.	<i>Le misure di livello</i>	28
3.2.4.	<i>La misura dell'output</i>	28
3.3.	LA TARATURA DEI SENSORI.....	28
3.3.1.	<i>S1 - Endress+ Hauser Cerabar M PMP51</i>	29
3.3.2.	<i>S2 - Endress+ Hauser Promag W</i>	30
3.3.3.	<i>S3 - Setra 280E</i>	31
3.3.4.	<i>S4 - Foxboro 841GM CII</i>	32
3.3.5.	<i>S5 - Foxboro 841GM CII</i>	33
3.3.6.	<i>S6 - Foxboro Vortez DN 50</i>	34
3.3.7.	<i>S7 - Foxboro IDP-10</i>	35
3.3.8.	<i>S8 - Endress+ Hauser Prowirl 200</i>	36
4.	L'INTELLIGENZA ARTIFICIALE	37
4.1.	INTRODUZIONE	37
4.2.	MACHINE LEARNING	38
4.2.1.	<i>Apprendimento supervisionato</i>	39
4.2.2.	<i>Apprendimento non supervisionato</i>	39
4.2.2.1.	<i>Apprendimento semi-supervisionato</i>	40
4.2.3.	<i>Apprendimento per rinforzo</i>	40
4.3.	LE RETI NEURALI	40
4.3.1.	<i>Le reti neurali a singolo livello feedforward</i>	44
4.3.2.	<i>Le reti neurali multilivello feedforward</i>	45
4.3.3.	<i>Le reti neurali ricorrenti</i>	45
4.4.	LE SELF-ORGANIZING MAPS.....	46
4.4.1.	<i>La struttura</i>	48
4.4.2.	<i>L'algoritmo</i>	49
4.4.2.1.	<i>I pesi iniziali</i>	49
4.4.2.2.	<i>Similary Matching</i>	50
4.4.2.3.	<i>Aggiornamento dei pesi</i>	50
4.4.3.	<i>Qualità della SOM</i>	53

4.4.3.1.	Quantization Error	53
4.4.3.2.	Topographic Error.....	54
5.	CASO DI STUDIO	55
5.1.	INTRODUZIONE	55
5.2.	PARAMETRI DI ESERCIZIO.....	56
5.3.	RACCOLTA DATI	57
5.3.1.	<i>Tipologie di dati raccolti</i>	57
5.3.2.	<i>Dati a regime</i>	57
5.3.3.	<i>Dati anomali</i>	59
5.3.3.1.	Anomalia 3.....	60
5.3.3.2.	Anomalia 5.....	62
5.3.3.3.	Anomalia 6.....	64
5.3.3.4.	Anomalia 7.....	66
5.3.3.5.	Anomalia 8.....	68
5.3.3.6.	Anomalia 9.....	70
5.3.3.7.	Anomalia 10.....	72
5.3.4.	<i>Acquisizione dei dati</i>	74
5.3.5.	<i>Standardizzazione dei dati</i>	74
5.4.	LA RETE SOM.....	75
5.4.1.	<i>I parametri della rete</i>	76
5.4.1.1.	L'iper-parametrizzazione.....	76
5.4.2.	<i>L'addestramento e la validazione della rete</i>	77
5.4.3.	<i>I pesi</i>	79
5.4.4.	<i>La mappa di output</i>	80
5.4.4.1.	Area di regime	83
5.4.4.2.	Macroarea del transitorio aria e acqua	84
5.4.4.3.	Macroarea anomalia acqua	85
5.4.4.4.	Macroarea anomalia aria.....	86
5.4.4.5.	Area anomalia fluido bifase aria-acqua	87
5.4.4.6.	Area anomalia serbatoio	88
5.4.5.	<i>Mappa di output riadattata ai livelli 1 e 2</i>	89
5.4.5.1.	Anomalia 3 livello 1 e 2.....	90
5.4.5.2.	Anomalia 5 livello 1 e 2.....	91
5.4.5.3.	Anomalia 6 livello 1 e 2.....	92
5.4.5.4.	Anomalia 7 livello 1 e 2.....	93
5.4.5.5.	Anomalia 8 livello 1 e 2.....	94

5.4.5.6. Anomalia 9 livello 1 e 2.....	95
5.4.5.1. Anomalia 10 livello 1 e 2.....	96
6. CONCLUSIONI E SVILUPPI FUTURI.....	97
APPENDICE	98
BIBLIOGRAFIA.....	102

Elenco delle figure

FIGURA 1.1: IL DIGITAL TWIN.....	2
FIGURA 1.2: ESTRAZIONE DEL PETROLIO MEDIANTE EIETTORE	5
FIGURA 2.1: IMPIANTO SPERIMENTALE	5
FIGURA 2.2: SCHEMA IMPIANTO SPERIMENTALE	6
FIGURA 2.3: POMPA DI VERGANI 32-201.....	8
FIGURA 2.4: EIETTORE IN LABORATORIO.....	9
FIGURA 2.5: EIETTORE IN SEZIONE	9
FIGURA 2.6: SERBATOIO.....	11
FIGURA 2.7: VALVOLA D'INTERCETTAZIONE MANUALE	12
FIGURA 2.8: PERDITE DI CARICO VALVOLA D'INTERCETTAZIONE DN 50 PN 16.....	12
FIGURA 2.9: ELETTROVALVOLA COMANDATA	13
FIGURA 2.10: REVOLUTION PI.....	14
FIGURA 2.11: ENDRESS+ HAUSER CERABAR M PMP51.....	17
FIGURA 2.12: ENDRESS+ HAUSER PROMAG W.....	18
FIGURA 2.13: SETRA 280E	19
FIGURA 2.14: FOXBORO 841GM CII	20
FIGURA 2.15: FOXBORO VORTEZ DN 50	21
FIGURA 2.16: FOXBORO IDP-10.....	22
FIGURA 2.17: ENDRESS+ HAUSER PROWIRL 200.....	23
FIGURA 3.1: METODO DEI MINIMI QUADRATI	26
FIGURA 3.2: MACCHINA IDRAULICA PER TARATURA	27
FIGURA 3.3: CONTATORE PER LA TARATURA	28
FIGURA 3.4: TARATURA S1 - ENDRESS+ HAUSER CERABAR M PMP51	29
FIGURA 3.5: TARATURA S2 - ENDRESS+ HAUSER PROMAG W	30
FIGURA 3.6: TARATURA S3 - SETRA 280E.....	31
FIGURA 3.7: TARATURA S4 - FOXBORO 841GM CII.....	32
FIGURA 3.8: TARATURA S5 - FOXBORO 841GM CII.....	33

FIGURA 3.9: TARATURA S6 - FOXBORO VORTEZ DN 50	34
FIGURA 3.10: TARATURA S7 - FOXBORO IDP-10	35
FIGURA 3.11: TARATURA S8 - ENDRESS+ HAUSER PROWIRL 200	36
FIGURA 4.1: TIPI DI MACHINE LEARNING	38
FIGURA 4.2: IL NEURONE BIOLOGICO	40
FIGURA 4.3: SPAZIO SINAPTICO	41
FIGURA 4.4: MODELLO DI NEURONE DI MCCULLOCH-PITTS	42
FIGURA 4.5: MODELLO DI NEURONE DI ROSENBLATT (PERCEPTRON)	42
FIGURA 4.6: CLASSIFICAZIONE BINARIA (METODO PERCEPTRON)	43
FIGURA 4.7: RETE NEURALE SINGOLO LIVELLO FEEDFORWARD	44
FIGURA 4.8: RETE NEURALE MULTILIVELLO FEEDFORWARD	45
FIGURA 4.9: RETE NEURALE RICORRENTE	45
FIGURA 4.10: SELF-ORGANIZING MAPS	46
FIGURA 4.11: STRUTTURA DELLA SELF-ORGANIZING MAP	48
FIGURA 4.12: RETE SOM A TOPOLOGIA QUADRATA ED ESAGONALE	49
FIGURA 4.13: TIPI DI VICINATO IN FUNZIONE DELLA TOPOLOGIA DI GRIGLIA	52
FIGURA 5.1: IMPIANTO SPERIMENTALE PRESENTE IN LABORATORIO	55
FIGURA 5.2: VALVOLE D'INTERCETTAZIONE NELL'IMPIANTO SPERIMENTALE	59
FIGURA 5.3: ANOMALIA 3 – PORTATA ACQUA IN INGRESSO E PERCENTUALE DI CHIUSURA V1	61
FIGURA 5.4: ANOMALIA 5 – ACQUA NEL SERBATOIO E PERCENTUALE DI CHIUSURA V3	63
FIGURA 5.5: ANOMALIA 6 – ACQUA NEL SERBATOIO E PERCENTUALE DI CHIUSURA V3	65
FIGURA 5.6: ANOMALIA 7 – PORTATA ARIA IN INGRESSO E ANDAMENTO DELLA PRESSIONE NELL'EIETTORE... ..	67
FIGURA 5.7: ANOMALIA 8 – PRESSIONE DEL SERBATOIO E PERCENTUALE DI CHIUSURA V2	69
FIGURA 5.8: ANOMALIA 9 – PRESSIONE DEL SERBATOIO E PERCENTUALE DI CHIUSURA V2	71
FIGURA 5.9: ANOMALIA 10 – ACQUA NEL SERBATOIO E PERCENTUALE DI CHIUSURA V3	73
FIGURA 5.10: QUANTIZATION ERROR DELLA SELF-ORGANIZING MAP	78
FIGURA 5.11: TOPOGRAPHIC ERROR DELLA SELF-ORGANIZING MAP	78
FIGURA 5.12: RAPPRESENTAZIONE DELLA U-MATRIX TRAMITE MAPPA DI COLORE	79
FIGURA 5.13: LA MAPPA DI OUTPUT DELLA PROVA A REGIME E DELLE ANOMALIE A LIVELLO 3	80
FIGURA 5.14: ATTIVAZIONE DEI NEURONI	81
FIGURA 5.15: STUDIO DEL CONFINE DELL'AREA ANOMALIA 3 LIVELLO 3	82
FIGURA 5.16: AREA DI REGIME	83
FIGURA 5.17: MACROAREA DEL TRANSITORIO	84
FIGURA 5.18: MACROAREA ANOMALIA ACQUA	85
FIGURA 5.19: MACROAREA ANOMALIA ARIA	86
FIGURA 5.20: AREA ANOMALIA FLUIDO BIFASE ARIA-ACQUA	87
FIGURA 5.21: AREA ANOMALIA SERBATOIO	88
FIGURA 5.22: MAPPA DI OUTPUT DELLE ANOMALIE LIVELLO 1 E 2	89

FIGURA 5.23: PROIEZIONE SULLA MAPPA DI OUTPUT DELL'ANOMALIA 3 LIVELLO 1 E 2.....	90
FIGURA 5.24: PROIEZIONE SULLA MAPPA DI OUTPUT DELL'ANOMALIA 5 LIVELLO 1 E 2.....	91
FIGURA 5.25: PROIEZIONE SULLA MAPPA DI OUTPUT DELL'ANOMALIA 6 LIVELLO 1 E 2.....	92
FIGURA 5.26: PROIEZIONE SULLA MAPPA DI OUTPUT DELL'ANOMALIA 7 LIVELLO 1 E 2.....	93
FIGURA 5.27: PROIEZIONE SULLA MAPPA DI OUTPUT DELL'ANOMALIA 8 LIVELLO 1 E 2.....	94
FIGURA 5.28: PROIEZIONE SULLA MAPPA DI OUTPUT DELL'ANOMALIA 9 LIVELLO 1 E 2.....	95
FIGURA 5.29: PROIEZIONE SULLA MAPPA DI OUTPUT DELL'ANOMALIA 10 LIVELLO 1 E 2.....	96

Elenco delle tabelle

TABELLA 2.1: COMPONENTI IMPIANTO.....	6
TABELLA 2.2: CARATTERISTICHE DELLA POMPA DI VERGANI 32-201	8
TABELLA 2.3: CARATTERISTICHE SERBATOIO DI ACCUMULO	11
TABELLA 2.4: CARATTERISTICHE VALVOLA D'INTERCETTAZIONE	12
TABELLA 2.5: INTERFACCE REVPI CORE 3 +.....	15
TABELLA 2.6: SPECIFICHE REVPI CORE 3 +	15
TABELLA 2.7: SPECIFICHE REVPI AIO	15
TABELLA 2.8: SENSORI DELL'IMPIANTO	16
TABELLA 2.9: CARATTERISTICHE CERABAR M PMP51	17
TABELLA 2.10: ENDRESS+ HAUSER PROMAG W.....	18
TABELLA 2.11: SETRA 280E.....	19
TABELLA 2.12: FOXBORO 841GM CI1	20
TABELLA 2.13: FOXBORO VORTEZ DN 50	21
TABELLA 2.14: FOXBORO IDP-10	22
TABELLA 2.15: ENDRESS+ HAUSER PROWIRL 200	23
TABELLA 3.1: SPECIFICHE S1 - ENDRESS+ HAUSER CERABAR M PMP51	29
TABELLA 3.2: DATI DI TARATURA S1 - ENDRESS+ HAUSER CERABAR M PMP51.....	29
TABELLA 3.3: SPECIFICHE S2 - ENDRESS+ HAUSER PROMAG W.....	30
TABELLA 3.4: DATI DI TARATURA S2 - ENDRESS+ HAUSER PROMAG W.....	30
TABELLA 3.5: SPECIFICHE S3 - SETRA 280E	31
TABELLA 3.6: DATI DI TARATURA S3 - SETRA 280E	31
TABELLA 3.7: SPECIFICHE S4 - FOXBORO 841GM CI1.....	32
TABELLA 3.8: DATI DI TARATURA S4 - FOXBORO 841GM CI1	32
TABELLA 3.9: SPECIFICHE S5 - FOXBORO 841GM CI1.....	33
TABELLA 3.10: DATI DI TARATURA S5 - FOXBORO 841GM CI1	33
TABELLA 3.11: SPECIFICHE S6 - FOXBORO VORTEZ DN 50	34
TABELLA 3.12: DATI DI TARATURA S6 - FOXBORO VORTEZ DN 50	34

TABELLA 3.13: SPECIFICHE S7 - FOXBORO IDP-10.....	35
TABELLA 3.14: DATI DI TARATURA S7 - FOXBORO IDP-10.....	35
TABELLA 3.15: SPECIFICHE S8 - ENDRESS+ HAUSER PROWIRL 200.....	36
TABELLA 3.16: DATI DI TARATURA S8 - ENDRESS+ HAUSER PROWIRL 200.....	36
TABELLA 5.1: PARAMETRI DI ESERCIZIO DEL CASO DI STUDIO	56
TABELLA 5.2: VARIABILI CAMPIONATE.....	57
TABELLA 5.3: LETTURE A REGIME.....	58
TABELLA 5.4: DATABASE DI PARTENZA	74
TABELLA 5.5: DATABASE STANDARDIZZATO	75

Introduzione

L'obiettivo del seguente studio è quello di sviluppare una rete neurale in grado di monitorare sia lo stato di salute di un impianto industriale che la presenza di eventuali anomalie. L'impianto bifase oggetto di studio è di tipo sperimentale ed è posto all'interno del dipartimento di Ingegneria Industriale e Scienze Matematiche (DIISM) dell'Università Politecnica delle Marche (Ancona, Italia). La sua realizzazione è datata intorno ai primi anni '90. Lo scopo dell'impianto sperimentale è quello di essere un'alternativa ai sistemi di pompaggio tradizionali per l'estrazione del petrolio nei giacimenti senza pressione. Il processo di estrazione degli idrocarburi dal giacimento esaurito viene effettuato sfruttando la pressione di un altro pozzo di petrolio al culmine della sua vita utile. La pressione di quest'ultimo è superiore alla pressione di trasporto ed è quindi in grado di aspirare petrolio dal giacimento esaurito che, al contrario, non ha una pressione sufficiente per il trasporto sulla linea. Per motivi di sicurezza, all'interno del dipartimento, il sistema è stato riadattato ad un impianto bifase aria-acqua.

Dopo aver spiegato il funzionamento, viene trattato il retrofit dell'impianto. Tramite l'ausilio di sensori collegati ad un hardware centrale per la gestione dei dati, l'impianto diventa "connesso" ed è possibile visualizzare il suo stato tramite il gemello digitale sullo schermo della workstation.

La prima fase di studio è dedicata alla trattazione delle specifiche e alla taratura dei sensori. Questi, infatti, sono stati nuovamente tarati in laboratorio per avere una lettura più precisa nelle acquisizioni registrate.

La seconda parte dell'elaborato è dedicata allo studio delle reti neurali artificiali. Dopo aver classificato e caratterizzato, in maniera generale, le diverse tipologie di reti neurali, lo studio si incentra sulla realizzazione di un algoritmo particolare di intelligenza artificiale per la gestione dell'impianto. Viene sviluppata una rete neurale non supervisionata dal nome *Self-Organizing Map* in grado di proiettare i dati d'ingresso ad alta dimensione in una mappa di output bidimensionale. Per registrare i dati necessari per la fase di addestramento della rete, sono state disposte all'interno dell'impianto diverse valvole d'intercettazione manuale con lo scopo di simulare eventuali perdite o occlusioni nelle condotte del sistema. Dopo aver ricreato artificialmente le diverse situazioni anomale, i dati vengono raccolti in un unico database per poi essere standardizzati tramite il metodo *z-score*.

In base alla somiglianza delle letture fornite in ingresso, l'algoritmo riesce a riordinare le informazioni nella mappa di output svolgendo una sorta di classificazione.

1. Introduzione al caso di studio

1.1. L'industria 4.0 e il Digital Twin

Un impianto industriale è un complesso di capitali, macchine, mezzi e addetti atti a sfruttare delle risorse materiali ed energetiche con lo scopo di trasformarle in prodotti finiti con maggior valore aggiunto tramite delle trasformazioni chimico-fisiche o dei processi di fabbricazione e montaggio. Il termine “industria 4.0” (in inglese, *Industry 4.0*) indica una tendenza dell'automazione industriale ad integrare alcune nuove tecnologie produttive con lo scopo di migliorare le condizioni di lavoro e aumentare la produttività e la qualità produttiva degli impianti.

Questa evoluzione industriale si basa sul concetto di *Smart Factory* (in italiano *fabbrica intelligente*) che si compone di tre parti:

1. **Smart production:** sono le nuove tecnologie produttive che mettono in comunicazione tutti i soggetti presenti all'interno dell'impianto produttivo come gli operatori, le macchine e gli strumenti;
2. **Smart services:** rappresentano le infrastrutture informatiche che permettono di far comunicare i vari elementi presenti nei sistemi produttivi;
3. **Smart energy:** l'organizzazione produttiva viene realizzata con un occhio attento ai consumi energetici, cercando di creare dei sistemi sempre più performanti sia a livello produttivo sia a livello energetico.

Alla base della *Smart Factory* c'è il *Digital Twin* (in italiano *gemello digitale*), una rappresentazione digitale di un elemento dell'impianto. Tramite la sincronizzazione dell'elemento fisico con il suo gemello digitale è possibile tenere traccia del lavoro dell'elemento dell'impianto. L'impiego del *Digital Twin* aiuta a rilevare le possibili anomalie nei processi di produzione per raggiungere obiettivi funzionali come il controllo in tempo reale, la manutenzione predittiva, l'adattamento in-process, l'analisi dei Big Data e l'apprendimento automatico.

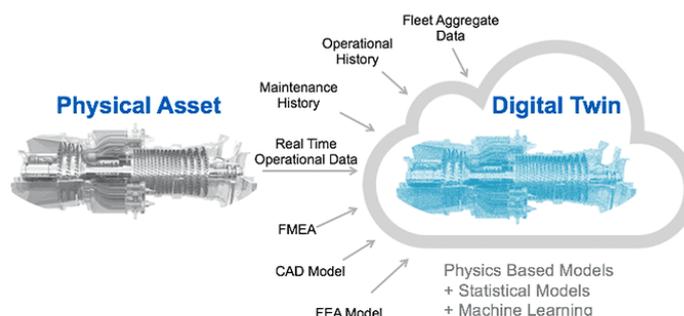


Figura 1.1: Il Digital Twin

1.2. La manutenzione

Il termine manutenzione fa riferimento alla scienza della conservazione che prende il nome di terotecnologia (dal greco *teros*, *logos* = conservare, prendersi cura di). Nel 1970 l'ente normatore inglese British Standard Institution associò alla scienza della conservazione questa definizione:

“La terotecnologia è una combinazione di direzione, finanza, ingegneria e altre discipline, applicate ai beni fisici per perseguire un economico costo del ciclo di vita ad esse relativo. Tale obiettivo è ottenuto con il progetto e l'applicazione della disponibilità e della manutenibilità agli impianti, alle macchine, alle attrezzature, ai fabbricati e alle strutture in genere, considerando la loro progettazione, installazione, manutenzione, miglioramento, rimpiazzo con tutti i conseguenti ritorni di informazioni sulla progettazione, le prestazioni e i costi”

Nel 2003 con la norma **UNI EN 13306** la manutenzione viene descritta come *“combinazione di tutte le azioni tecniche, amministrative e gestionali, previste durante il ciclo di vita di un'entità, destinate a mantenerla o riportarla in uno stato in cui possa eseguire la funzione richiesta”*. (1)

Quindi, lo scopo della manutenzione è quello di conservare il sistema in ottima salute al fine di avere continuità nel processo produttivo e di rendere massima l'esperienza all'operatore. All'interno delle diverse realtà industriali, la manutenzione può avere un impatto significativo in termini di costi e affidabilità, influenzando anche l'abilità dell'azienda di essere competitiva sul mercato.

Oggi esistono diverse tipologie di manutenzione come:

- **manutenzione autonoma**: attività che hanno il compito di giudicare e verificare il funzionamento di un determinato asset;
- **manutenzione preventiva**: attività che hanno il compito di riconoscere il funzionamento attuale del sistema per adottare eventualmente una modifica tempestiva della componentistica e rallentare il normale degrado del funzionamento;
- **manutenzione programmata**: viene disposto un periodo di fermo della macchina per svolgere gli opportuni interventi di manutenzione dopo che questa abbia superato determinate ore di esercizio, chilometri percorsi o numero di pezzi prodotti;
- **manutenzione a guasto**: attività che richiedono l'intervento da parte di un operatore al fine di sostituire determinati componenti danneggiati per ristabilire il corretto funzionamento della macchina e il ripristino delle usure;
- **manutenzione migliorativa “retrofit”**: attività che puntano a migliorare tecnicamente le prestazioni previste in sede di progetto dei diversi asset.

1.3. La manutenzione predittiva

Il concetto di manutenzione predittiva rimanda a quello di *Prognostic and Health Management (PHM)*, disciplina che esegue previsioni di eventi futuri sia in termini di guasti che in termini di qualità e anomalie presenti nel sistema industriale. Infatti, in un contesto industriale, se si riuscisse a prevedere l'eventuale danneggiamento di una determinata macchina si potrebbe intervenire tempestivamente con delle operazioni di ripristino durante un suo periodo fermo, senza incorrere nell'interruzione della produzione, risparmiando sia in termini di tempo che in termini di costi.

Questi sistemi sono in grado di fornire previsioni (prognostica) grazie allo studio di errori storici (diagnostica). Dopo aver disposto i diversi sensori sull'impianto e averli collegati ad un'unità centrale, questi forniranno una serie di dati che costituiranno il nostro database. L'approccio più usato è quello basato sui dati. Infatti, dopo aver strutturato e filtrato il database di partenza si può ricorrere o ad un approccio statistico che basa la sua valutazione sulle variabili come la media, la varianza e lo studio delle distribuzioni normali oppure si può ricorrere al machine learning.

1.4. Cenni sui sistemi bifase gas-liquido

In un contesto industriale petrolifero, quando all'interno di un pozzo il petrolio è in esaurimento, la pressione delle rocce circostanti immette all'interno della sacca anche l'acqua proveniente dalle falde acquifere limitrofe. In queste circostanze per non intaccare il rendimento dell'impianto, una possibile soluzione potrebbe essere quella di installare, sia in superficie che alla base del pozzo, delle apposite pompe che accelerino il processo di estrazione. Il costo delle pompe risulta essere molto elevato rispetto ai volumi di idrocarburi prodotti. Per questo motivo nella maggior parte delle volte, si preferisce abbandonare il pozzo.

Una soluzione indubbiamente più efficiente, sia a livello energetico che a livello economico, è quella di installare un sistema di pompaggio per miscele bifase gas-liquido. Il processo di estrazione del petrolio dal giacimento esaurito viene effettuato sfruttando la pressione di un giacimento al culmine della sua vita utile. La pressione di quest'ultimo è superiore alla pressione di trasporto ed è quindi in grado di aspirare petrolio dal giacimento esaurito che, al contrario, non ha una pressione sufficiente per il trasporto sulla linea.

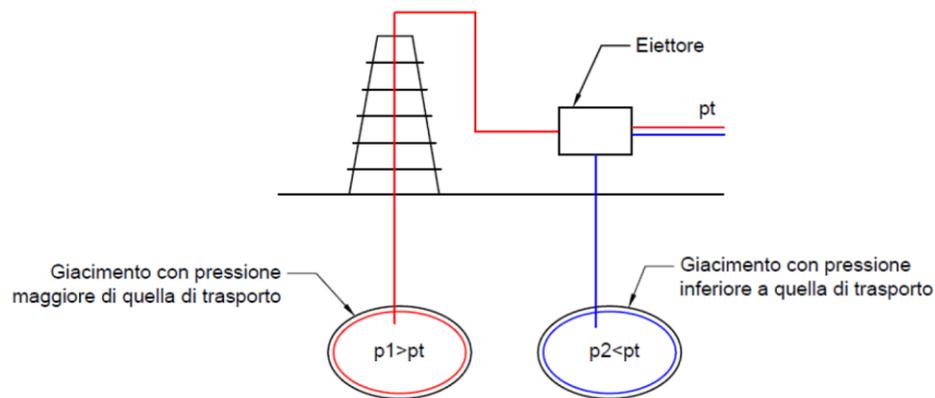


Figura 1.2: Estrazione del petrolio mediante eiettore

L'elemento principale dell'impianto mostrato in Figura 1.2 è l'eiettore gas-liquido. Un dispositivo in grado di miscelare due flussi a diversa pressione e di imprimere l'energia necessaria al trasporto. La mancanza di parti in movimento e la conseguente estrema semplicità dell'apparato, i bassi costi di manutenzione e di installazione, l'assenza di problemi di tenute, fanno dell'eiettore gas-liquido una macchina che può offrire notevoli possibilità di applicazioni nel trasporto sia di miscele di idrocarburi che di gas corrosivi, tossici o radioattivi.

2. Descrizione dell'impianto

L'impianto esaminato in questo elaborato di tesi è di tipo sperimentale ed è posto all'interno del dipartimento di Ingegneria Industriale e Scienze Matematiche (DIISM) dell'Università Politecnica delle Marche (Ancona, Italia).



Figura 2.1: Impianto sperimentale

È stato realizzato intorno ai primi anni '90 con lo scopo di riprodurre artificialmente l'estrazione del petrolio e del gas da un pozzo esaurito tramite la pressione di uno attivo. Per motivi di sicurezza all'interno dell'impianto sperimentale vengono utilizzati acqua e aria a temperatura ambiente, e non petrolio greggio e gas naturale. L'acqua presente nella vasca e la pompa volumetrica simulano il comportamento del pozzo pressurizzato, mentre l'aria il gas naturale del giacimento esaurito.

2.1. Schema dell'impianto e funzionamento

Dalla Figura 2.2 è possibile visualizzare lo schema dell'impianto sperimentale presente in laboratorio.

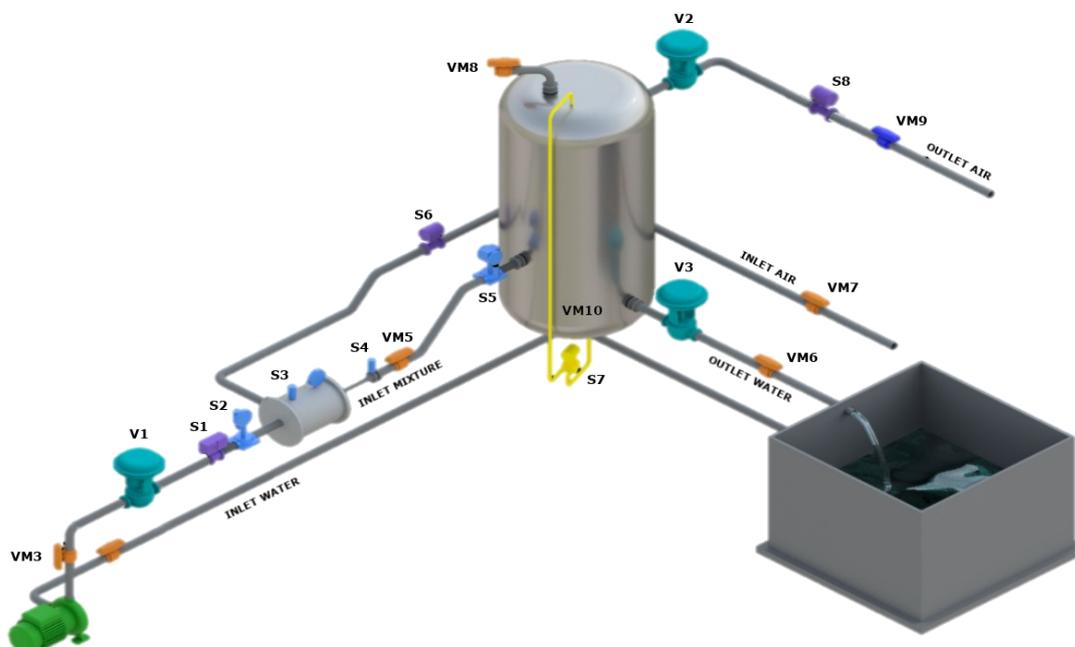


Figura 2.2: Schema impianto sperimentale

Tabella 2.1: Componenti impianto

ID	Descrizione	UM
S1	Sensore di pressione assoluta Endress+ Hauser Cerabar M PMP51	[bar]
S2	Sensore di portata Endress+ Hauser Promag W	[m ³ /h]
S3	Sensore di pressione assoluta Setra 280E	[bar]
S4	Sensore di pressione assoluta Foxboro 841GM-CI1	[bar]

S5	Sensore di pressione relativa Foxboro 841GM-CI1	[bar]
S6	Sensore di portata volumetrica Foxboro Vortez DN 50	[m ³ /h]
S7	Sensore di pressione differenziale Foxboro IDP-10	[mm]
S8	Sensore di portata volumetrica Endress+ Hauser Prowirl 200	[m ³ /h]
V1	Elettrovalvola pneumatica Spirax Sarco 9126E Pneumatic Valve	[%]
V2	Elettrovalvola pneumatica ECKARDT MB6713	[%]
V3	Elettrovalvola pneumatica ECKARDT MB6713	[%]
S1_rif	Pressione di riferimento in acqua in ingresso	[bar]
S5_rif	Pressione di riferimento miscela nel serbatoio	[bar]
S7_rif	Livello di riferimento acqua nel serbatoio	[mm]
PID_S1	Controllore di pressione dell'acqua in ingresso	on/off
PID_S5	Controllore del livello dell'acqua all'interno del serbatoio	on/off
PID_S7	Controllore di pressione dell'aria all'interno del serbatoio	on/off
VM	Valvole d'intercettazione manuali	on/off

La pompa, dopo aver prelevato l'acqua dalla vasca, manda in pressione "INLET WATER" nell'eiettore, all'interno del quale si viene a creare un vuoto che consente l'ingresso di una certa quantità d'aria ambientale chiamata "INLET AIR". La miscela, ora bifasica, prende il nome di "INLET MIXTURE". Quest'ultima, quindi, viene convogliata in un serbatoio verticale che ha il compito di separare la miscela tra la fase liquida ("OUTLET WATER") e fase gassosa ("OUTLET AIR").

Sempre dallo schema è possibile notare che l'impianto è dotato di tre elettrovalvole:

- V1 controlla la pressione dell'acqua in ingresso al sistema;
- V2 controlla la pressione all'interno del serbatoio;
- V3 controlla il livello dell'acqua all'interno del serbatoio.

Le valvole contrassegnate con "VM" sono di tipo manuale e sono state impiegate per simulare le possibili anomalie dell'impianto.

2.2. Componenti idraulici

Di seguito verranno descritti in maniera approfondita tutti i componenti idraulici presenti nell'impianto.

2.2.1. La pompa

La pompa presente nell'impianto è di tipo centrifugo monoblocco flangiata direttamente al motore.

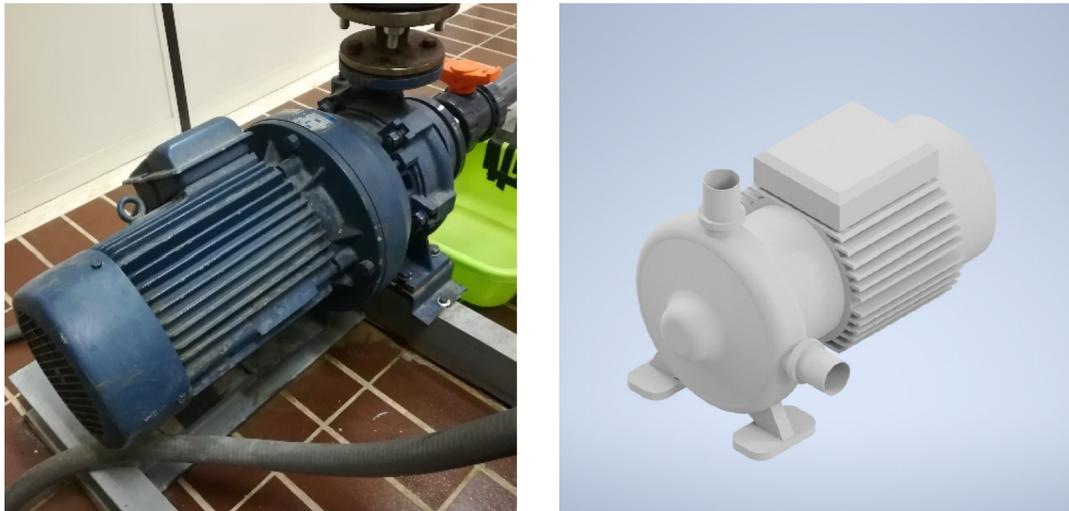


Figura 2.3: Pompa di Vergani 32-201

Il corpo è diviso verticalmente, la bocca aspirante è assiale mentre quella premente radiale è rivolta verso l'alto; l'albero motore prolungato è supportato da cuscinetti a sfere lubrificati a grasso, che non necessitano di lubrificazione periodica. La girante è calettata direttamente sull'albero motore.

Tabella 2.2: Caratteristiche della pompa di Vergani 32-201

Caratteristiche Pompa	
Modello	Vergani 32-201
Giri al minuto	2900 rpm
Capacità	233 l/min
Portata	14 m ³ /h
Potenza	7,5 Kw
Prevalenza	54,25 m

2.2.2. L'eiettore

L'eiettore è un dispositivo meccanico in grado di innalzare la pressione del fluido che passa al suo interno. Il principio di funzionamento dell'eiettore si basa sul fenomeno idrodinamico detto *effetto Venturi*, per cui la pressione di una corrente di un fluido aumenta con il diminuire della velocità. L'eiettore quindi, lavorando come una pompa, prende anche il nome di “Pompa Venturi”.



Figura 2.4: Eiettore in laboratorio

Come illustrato dalla Figura 2.5, le parti principali dell'eiettore sono:

1. **il bocchello di aspirazione:** consente l'ingresso del fluido d'aspirazione;
2. **l'ugello convergente:** permette l'espansione e l'accelerazione del fluido motore;
3. **il bocchello del fluido motore:** consente l'ingresso del fluido motore;
4. **la camera di miscelazione:** consente il miscelamento, lo scambio di energia e quantità di moto tra fluido primario e secondario;
5. **il diffusore divergente:** innalza la pressione della corrente fluida.

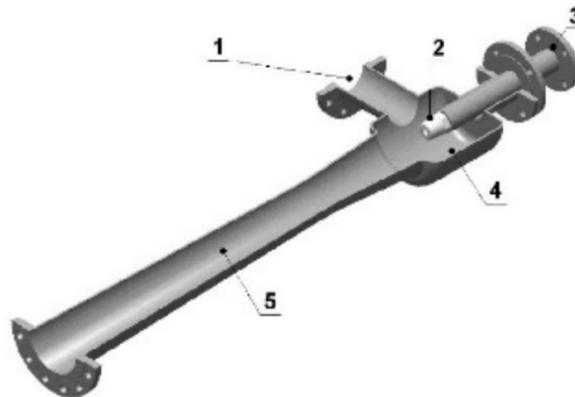


Figura 2.5: Eiettore in sezione

All'interno della pompa di Venturi sono presenti due fluidi:

- **il fluido motore:** "INLET WATER";
- **il fluido d'alimentazione** (fluido secondario): "INLET AIR".

Si considerino le seguenti proprietà:

- p è la pressione del fluido;
- ρ è la densità del fluido;
- u è la velocità di scorrimento;
- g è la costante gravitazionale;
- h è la quota media della sezione.

Il fluido motore entrato dal relativo bocchello (3) incontra l'ugello convergente (2) per poi arrivare alla camera di miscelazione (4). Muovendosi all'interno dell'ugello, il fluido subirà una riduzione di sezione di passaggio che per il Teorema di Bernoulli si trasformerà in un aumento di velocità e una riduzione di pressione.

$$p + \rho \frac{u^2}{2} + \rho gh = cost$$

Il fluido fuoriuscito dall'ugello è caratterizzato da un'alta velocità che secondo l'effetto Venturi, permette la nascita di una depressione nella camera di miscelazione che promuove l'ingresso del fluido di aspirazione ("INLET AIR"). Nella miscela che si è appena formata l'elevata quantità di moto del fluido motore viene trasmessa a quella del fluido di aspirazione. Mediante il cono divergente, la miscela rallenta e l'energia cinetica viene riconvertita in energia statica di pressione. Dopo che l'"INLET MIXED" attraverserà il diffusore, presenterà in uscita una pressione maggiore rispetto a quella d'ingresso.

2.2.3. Il serbatoio

Nella Figura 2.6 è rappresentato il serbatoio presente in laboratorio. Oltre al compito di accumulo di acqua in pressione, svolge anche il ruolo di separatore verticale della miscela liquido gassosa in ingresso dal punto 1. L'uscita per la componente liquida è contrassegnata con 2, mentre per la componente gassosa con 3. Il punto 4 rappresenta una valvola manuale d'intercettazione che permette la simulazione di un eventuale foratura della cisterna.



Figura 2.6: Serbatoio

Al serbatoio sono collegati sia un manometro meccanico a tubo di Bourdon al fine di monitorare la pressione al suo interno sia una valvola di sicurezza a molla tarata ad una pressione del 10% minore del valore della pressione massima di prova come da norma.

Tabella 2.3: Caratteristiche serbatoio di accumulo

Caratteristiche Serbatoio	
Modello	Elbi 780-I
Capacità	780 l
Pressione massima	9,8 bar
Temperatura esercizio	-10/+50 _C
Materiale	Acciaio Zincato

2.2.4. Le valvole d'intercettazione

Dalla Figura 2.2 e dalla Tabella 2.1 nell'impianto si possono notare sei valvole d'intercettazione manuale contrassegnate con la sigla VM. Sono state posizionate in maniera da simulare potenziali occlusioni che si possono verificare durante la vita utile dell'impianto. Essendo delle comuni valvole a sfera, le ostruzioni sono state registrate tramite la rotazione dell'otturatore dotato di una cavità cilindrica coassiale al flusso.



Figura 2.7: Valvola d'intercettazione manuale

Per simulare un'occlusione completa di un tubo presente all'interno dell'impianto occorre ruotare la testa della valvola fino ad un quarto di giro. La sua corsa ridotta permette da una parte un azionamento immediato, ma dall'altra impedisce la regolazione del flusso in maniera precisa creando quindi dei livelli di anomalia approssimativi. Per il calcolo delle perdite di carico è possibile fare riferimento alla Figura 2.8.

Tabella 2.4: Caratteristiche Valvola d'intercettazione

Caratteristiche Valvole Intercettazione	
Modello	FIP DN 50 PN 16
Capacità	780 l
Pressione massima	16 bar
Temperatura esercizio	0/+60 C
Materiale	PVC-U

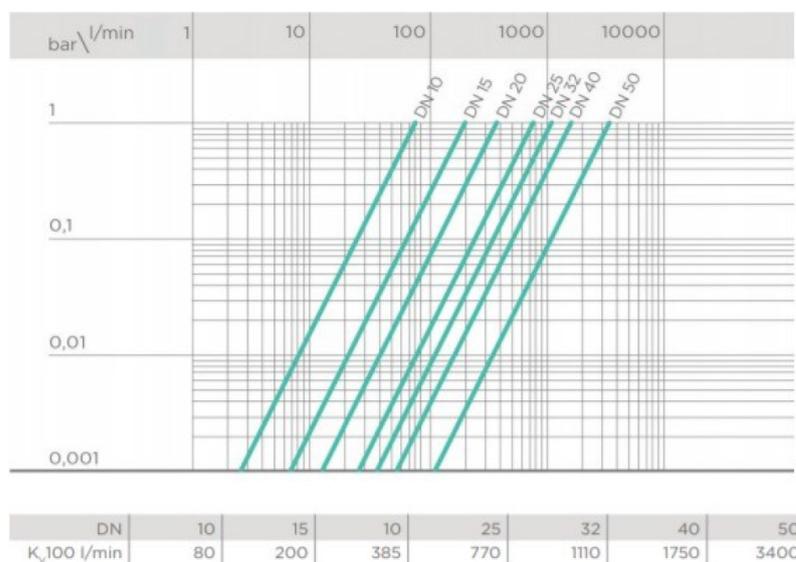


Figura 2.8: Perdite di carico valvola d'intercettazione DN 50 PN 16

2.2.5. Le tubazioni

Le tubazioni installate sull'impianto sono in materiale PVC-U che, oltre a garantire un'ottima resistenza chimica e alla corrosione, permettono di contenere il flusso fino ad una pressione di 16 bar a temperatura ambiente. La pressione di esercizio registrata sull'impianto è di cinque bar, quindi al di sotto del valore limite. Nelle fasi di simulazione delle anomalie, per non danneggiare l'impianto, si è stabilito di non superare i 2 bar di pressione assoluta all'uscita dell'eiettore dato che non si conoscano i limiti operativi del cono divergente in perspex, realizzato su commissione.

2.2.6. Le valvole di regolazione

Lo schema in Figura 2.2 testimonia la presenza di tre elettrovalvole contrassegnate con V1, V2, V3 che hanno lo scopo di gestire l'acqua in ingresso all'impianto e il flusso in uscita sia del liquido che dell'aria della miscela bifase all'interno del serbatoio di contenimento.



Figura 2.9: Elettrovalvola comandata

Dalla Figura 2.9 si può notare che l'apertura della valvola è gestita da un attuatore pneumatico indicato con (1) a sua volta controllato da un posizionatore (2). La regolazione della posizione dello stantuffo che permette il passaggio del fluido è gestita da un segnale elettrico che varia tra 0 mA quando la valvola è completamente aperta e 20 mA quando è chiusa.

Per la gestione dei diversi livelli di acqua e aria si è ricorso all'utilizzo di tre controllori PID che, ricevendo il segnale dalla scheda di controllo RevPi, permettono all'utente la completa amministrazione dell'impianto mediante l'apposita interfaccia grafica.

2.3. Componenti Elettronici

Di seguito vengono descritti in maniera approfondita tutti i componenti elettronici presenti nell'impianto.

2.3.1. Revolution Pi

Revolution Pi è un controllore industriale open source basato su Raspberry Pi Compute Module. L'idea di RevPi nasce nel 2016 con la missione di sviluppare un controller a basso costo, scalabile e adatto all'industria.



Figura 2.10: Revolution Pi

La Figura 2.10 raffigura il dispositivo presente in laboratorio composto da 4 schede:

- **RevPi Core 3:** scheda principale che collegata al computer, nel caso in esame mediante connessione Ethernet, permette la trasmissione dei segnali dei diversi asset;
- **RevPi MIO:** scheda di acquisizione dotata di 8 ingressi analogici, 8 uscite analogiche e 4 canali digitali al sistema;
- **RevPi AIO:** scheda di acquisizione dotata di 4 ingressi a 24 bit e 2 uscite analogiche con una risoluzione di 16 bit.

Nelle seguenti Tabelle sono riportate tutte le caratteristiche tecniche dell'Hardware in questione.

Tabella 2.5: Interfacce RevPi Core 3 +

Interfacce RevPi Core 3 ⁺	N° prese
Ethernet RJ45	1
USB 2.0	2
micro-HDMI	1
Micro USB 2.0	1
PiBridge	2

Tabella 2.6: Specifiche RevPi Core 3 +

Specifiche Tecniche RevPi Core 3⁺	
Processore	Broadcom BCM2837B0 con ARM Cortex-A53 quad-core
Frequenza di clock	1,2 GHz
RAM	1 GB DDR2 (RevPi Core 3+/Core 3)
Archiviazione	8 GB – 32 GB
Alimentazione elettrica	12 – 24 V
Dimensioni (L x S x A)	111 x 22,5 x 96 mm
Temperatura di esercizio	-25°C...+55°C
Temperatura di conservazione	-40°C...+85°C
Classe di protezione	IP20
Protezione ESD	4 kV/8 kV

Tabella 2.7: Specifiche RevPi AIO

Input Analogici		Output analogici	
Voltage range	Current range:	Voltage range	Current range:
0 – 5 V	4 – 20 mA	0 – 5 V	4 – 20 mA
0 – 10 V	0 – 20 mA	0 – 10 V	0 – 20 mA
+/- 5 V	0 – 24 mA	+/- 5 V	0 – 24 mA
+/- 10 V	+/- 25 mA	+/- 10 V	

2.3.2. I sensori

Per completare il processo di *smartizzazione*, oltre all'acquisto dell'unità centrale RevPi, sono stati disposti diversi sensori in grado controllare in maniera continua lo stato dell'impianto. Una volta collegati all'hardware tramite gli appositi pin della scheda analogica, la gestione dei livelli di acqua e aria è rimandata all'operatore tramite l'apposita interfaccia web. I sensori installati sull'impianto sono riportati nella Tabella 2.8 e possono essere divisi in:

- sensori con uscita in tensione;
- sensori con uscita in corrente.

Tabella 2.8: Sensori dell'impianto

ID	Descrizione	UM	Uscita	Tag
S1	Pressione dell'acqua in ingresso	[bar]	Corrente	Endress+ Hauser Cerabar M PMP51
S2	Portata dell'acqua in ingresso	[m3/h]	Corrente	Endress+ Hauser Promag W
S3	Pressione dell'eiettore	[bar]	Tensione	Setra 280E
S4	Pressione della miscela dopo il diffusore	[bar]	Corrente	Foxboro 841GM CII
S5	Pressione nel serbatoio	[bar]	Corrente	Foxboro 841GM-CII
S6	Portata dell'aria in ingresso	[m3/h]	Corrente	Foxboro Vortez DN 50
S7	Livello di acqua nel serbatoio	[mm]	Corrente	Foxboro IDP-10
S8	Portata di aria in uscita	[m3/h]	Corrente	Endress+ Hauser Prowirl 200

Di seguito vengono descritti in maniera più approfondita tutte le tipologie di sensori.

2.3.2.1. Endress+ Hauser Cerabar M PMP51

Il trasmettitore di pressione digitale Cerabar PMP51 viene tipicamente utilizzato nelle applicazioni industriali in cui viene richiesta una misura di pressione, livello, volume o massa dei liquidi o gas. La misura della pressione è affidata ad un sensore piezo-resistivo che seguendo la deformazione di una membrana a cui è fissato varia la resistenza elettrica e restituisce il segnale in funzione della di quest'ultima.



Figura 2.11: Endress+ Hauser Cerabar M PMP51

Tabella 2.9: Caratteristiche Cerabar M PMP51

Caratteristiche Cerabar M PMP51	
Alimentazione	220 V
Output	4 - 20 mA
Campo di misura	1 bar - 10 bar

2.3.2.2. Endress+ Hauser Promag W

Il flussimetro Endress+ Hauser Promag W è un dispositivo in grado di registrare la portata di un liquido senza che ci sia perdita di pressione. La misura della portata è affidata ad un trasmettitore e ad un sensore interno al dispositivo che misurando la tensione indotta generata dall'acqua che scorre nella tubazione la trasforma in una misura di portata.



Figura 2.12: Endress+ Hauser Promag W

Tabella 2.10: Endress+ Hauser Promag W

Caratteristiche Promag W	
Alimentazione	220 V
Diametro	50 mm
Output	4 - 20 mA
Campo di misura	0,076 – 10 m ³ /h

2.3.2.3. Setra 280E

Il sensore di pressione Setra 280E nell'impianto è disposto in modo tale da monitorare la pressione all'interno dell'eiettore. Formato da acciaio inox, la misura della pressione è affidata alla presenza di un elettrodo isolato che forma un condensatore variabile. All'aumentare della pressione all'interno della camera di miscelazione la capacità diminuisce. Questa variazione di capacità viene rilevata e convertita in un segnale continuo lineare in uscita da parte del sensore.



Figura 2.13: Setra 280E

Tabella 2.11: Setra 280E

Caratteristiche Setra 280E	
Alimentazione	15-32 V DC
Output	0-5V DC
Campo di misura	0-100 psi

2.3.2.4. Foxboro 841GM CII

I trasduttori di pressione Foxboro 841GM CII sono presenti in due parti diverse dell'impianto: uno all'uscita del tubo divergente, l'altro è collegato al serbatoio. Similmente a quello che avviene per il sensore Endress+ Hauser Cerabar M PMP51, la conversione della pressione in segnale elettrico viene ricavata dalla deformazione fisica degli estensimetri connessi alla membrana del trasduttore. La pressione flette il diaframma che a sua volta deforma gli estensimetri. Questa deformazione varierà la resistenza elettrica dando un segnale in uscita da 4 a 20 mA in DC.

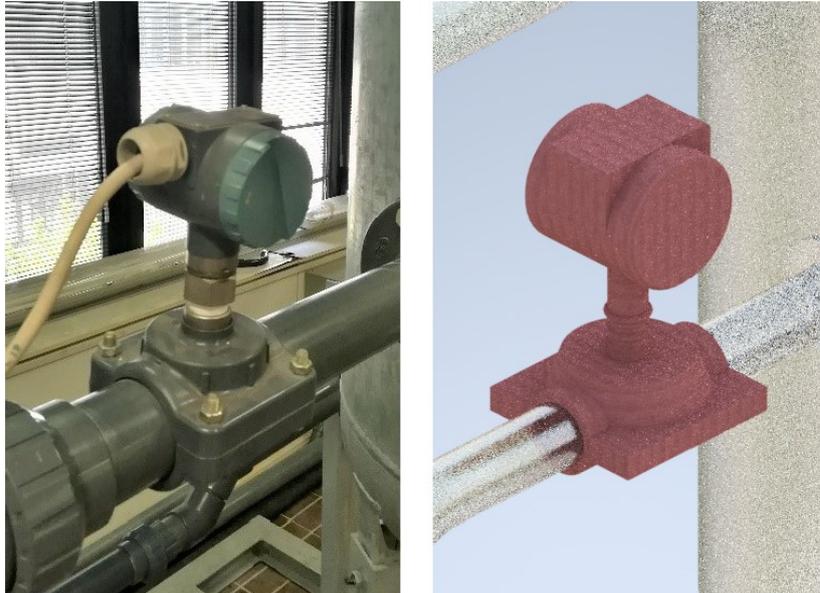


Figura 2.14: Foxboro 841GM CII

Tabella 2.12: Foxboro 841GM CII

Caratteristiche Foxboro 841GM CII	
Alimentazione	12,5 - 36 V DC
Diametro	50 mm
Output	4-20 mA
Campo di misura	0-300 psi

2.3.2.5. Foxboro Vortez DN 50

Questo tipo di sensore misura la portata dell'aria in ingresso all'impianto. Il suo funzionamento si basa sulla scia di Karman che si forma dal distacco dei vortici creati ai lati del corpo deflettore. La frequenza di distacco è direttamente proporzionale alla velocità del flusso e inversamente proporzionale alla larghezza del corpo deflettore. Il rilevamento delle variazioni di pressione e sua conversione in segnale elettrico compreso tra i 4 e i 20 mA, è affidato ad un sensore piezoelettrico.

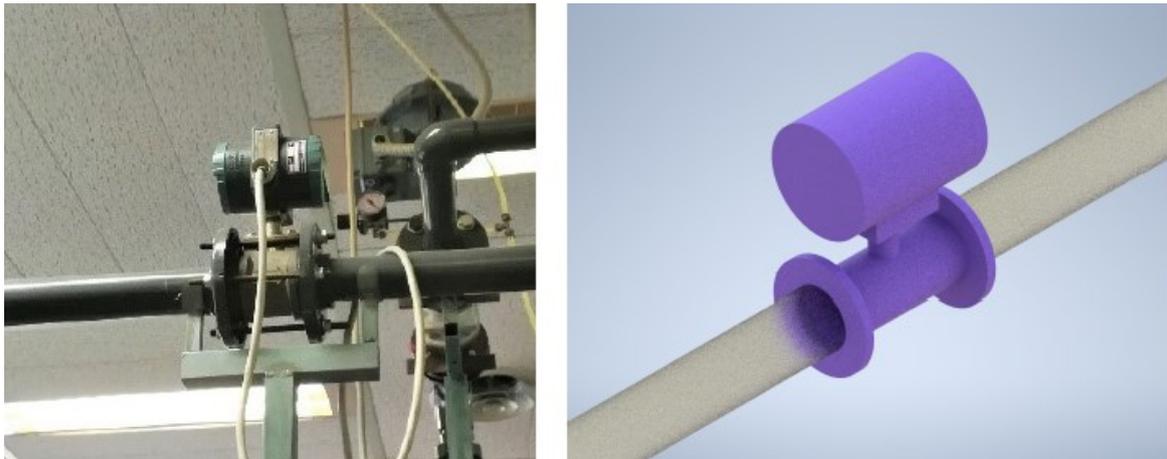


Figura 2.15: Foxboro Vortez DN 50

Tabella 2.13: Foxboro Vortez DN 50

Caratteristiche Foxboro Vortez DN 50	
Alimentazione	24 V DC
Diametro	50 mm
Output	4-20 mA
Campo di misura	0-200 m ³ /h
Temperatura max	200 °C

2.3.2.6. Foxboro IDP-10

Il sensore di pressione differenziale Foxboro IDP-10 viene utilizzato per misurare il livello della miscela presente all'interno del serbatoio. Mediante un estensimetro al silicio presente nel dispositivo, viene rilevata la differenza di pressione tra il punto 1 e il punto 2 segnati in Figura 2.16.



Figura 2.16: Foxboro IDP-10

La variazione di pressione registrata dall'estensimetro viene convertita in un cambiamento di resistenza che si trasforma in una variazione di corrente nel range 4-20 mA. La differenza di pressione registrata è misurata tra la testa e il fondo del serbatoio. Durante la calibrazione del sensore, il segnale prodotto in uscita invece di essere espresso in bar, viene convertito in mm così da conoscere l'altezza del livello dell'acqua nel serbatoio.

Tabella 2.14: Foxboro IDP-10

Caratteristiche Foxboro IDP-10	
Alimentazione	12,5-42 V DC
Pressione max	250 bar
Output	4-20 mA
Campo di misura	0-300 psi

2.3.2.7. Endress+ Hauser Prowirl 200

Il flussimetro Endress+ Hauser Prowirl 200 è un dispositivo multi-variabile in grado di rilevare la portata volumetrica dell'aria in ingresso al serbatoio di contenimento. Il principio di funzionamento è analogo a quanto descritto nel capitolo 2.3.2.5 con il sensore Foxboro Vortez DN 50.



Figura 2.17: Endress+ Hauser Prowirl 200

Tabella 2.15: Endress+ Hauser Prowirl 200

Caratteristiche Endress+ Hauser Prowirl 200	
Alimentazione	12-35 V DC
Diametro	50 mm
Output	4-20 mA
Campo di misura	0-200 m ³ /h
Temperatura max	180°C

3. La calibrazione dei sensori

3.1. Cenni teorici

3.1.1. Cos'è un sensore

Un sensore è un dispositivo capace di misurare una grandezza fisica interagendo direttamente con essa. L'energia immagazzinata dall'apparecchio modifica il suo stato iniziale cambiandone proprietà come lunghezza, resistenza elettrica e così via. Spesso al termine sensore viene associato quello di trasduttore, perché oltre al rilevamento di una certa misura, il dispositivo è in grado di restituire in uscita un segnale di natura diversa rispetto a quello d'ingresso, convertendolo tramite una determinata legge. È possibile quindi risalire alla grandezza d'ingresso conoscendo la funzione di trasferimento e la grandezza in uscita e viceversa.

Oggi, essendoci sempre di più la necessità di monitorare grandezze fisiche, chimiche e biologiche nell'ambito industriale, civile e della diagnostica, la tecnologia installata nei sensori risulta essere sempre più all'avanguardia.

Nell'impianto oggetto di studio, l'utilizzo dei sensori ha permesso di aumentare la sicurezza e la capacità di prevenzione ai guasti controllando costantemente la sua condizione di salute. Questi dispositivi vengono classificati in base alle grandezze (meccaniche, elettriche o termiche) che sono in grado di rilevare e al principio di funzionamento. Oggi i sensori più utilizzati sono di tipo elettrico, le cui uscite possono essere di tipo analogico o digitale.

Le caratteristiche principali di un sensore sono:

- campo di misura;
- sensibilità;
- precisione;
- finezza;
- rapidità nella risposta.

All'ingegnere spetta cercare, nella fase di progettazione, un giusto compromesso tra tutte queste caratteristiche, in vista dei compiti specifici che il sensore dovrà svolgere.

3.1.2. La taratura statica

La relazione che persiste tra ingresso e uscita in uno strumento di misura è descritta dalla taratura. Lo scopo del processo di taratura è quello di studiare le prestazioni del sistema curando in particolar

modo la qualità della misurazione. In base ai tipi di grandezze da misurare, le tipologie di taratura possono essere divise in due categorie:

- Taratura statica: misura entità costanti o che variano lentamente nel tempo;
- Taratura dinamica: misura entità che variano velocemente nel tempo.

Mediante la taratura statica si determinano diverse proprietà dello strumento come la sensibilità statica o coefficiente di taratura, la linearità, l'accuratezza e la sensibilità ai disturbi.

Lo strumento con cui viene effettuata la taratura deve essere almeno di una classe superiore allo strumento da calibrare. Dopo aver posto lo strumento di taratura in un ambiente controllato, viene variato l'ingresso in maniera quasi statica e vengono registrati i valori di uscita al sensore.

Una volta eseguite le misurazioni sulle uscite, supponendo che la funzione di trasferimento ingresso-uscita sia di tipo lineare, la relazione risulta essere del tipo:

$$q_o = mq_i + b$$

Dove:

- q_o = segnale in uscita
- q_i = segnale in ingresso
- m = coefficiente angolare della retta
- b = punto d'intersezione tra retta e asse verticale

3.1.3. Il metodo dei minimi quadrati

Il metodo dei minimi quadrati è una tecnica di ottimizzazione che ha lo scopo di trovare la funzione ottima (nel nostro caso in esame, una retta) che approssimi il più possibile la distribuzione spaziale dei dati registrati. La funzione trovata deve essere quella che minimizza la somma dei quadrati delle distanze tra i dati osservati e quelli della curva che rappresenta la funzione stessa. Praticamente il metodo minimizza la somma delle distanze al quadrato tra i punti sperimentali y e i punti sulla curva stessa $\hat{y}_i(2)$.

$$\min \sum_{i=1}^{i=n} (y_i - \hat{y}_i)^2$$

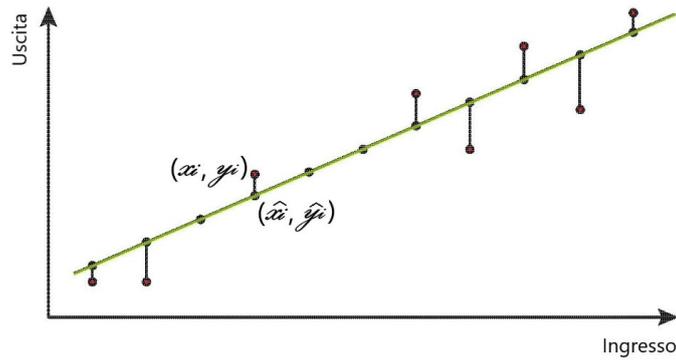


Figura 3.1: Metodo dei minimi quadrati

Il caso di studio tratta di regressione lineare, quindi, il calcolo del coefficiente angolare e della costante b è dato da:

$$m = \frac{N \sum_i q_i q_o - (\sum_i q_i)(\sum_i q_o)}{N \sum_i q_i^2 - (\sum_i q_i)^2}$$

$$b = \frac{(\sum_i q_o)(\sum_i q_i^2) - (\sum_i q_i q_o) \sum_i q_i}{N \sum_i q_i^2 - (\sum_i q_i)^2}$$

Una volta aver calcolato i valori di m (sensibilità) e b , è opportuno calcolare le rispettive deviazioni standard trovate tramite le formule:

$$s_m^2 = \frac{N s_{q_o}^2}{N \sum_i q_i^2 - (\sum_i q_i)^2}$$

$$s_b^2 = \frac{s_{q_o}^2 \sum_i q_i^2}{N \sum_i q_i^2 - (\sum_i q_i)^2}$$

I termini s_m e s_b rappresentano l'incertezza di linearità della sensibilità e dello zero. È possibile infine calcolare anche l'incertezza della misura s_{q_i} tramite:

$$s_{q_i} = \frac{s_{q_o}}{m}$$

Dove

$$s_{q_o} = \sqrt{\frac{1}{N} \sum_{i=1}^{i=n} (m q_i + b - q_o)^2}$$

3.1.4. Il coefficiente di correlazione R^2

Per analizzare la bontà del modello lineare costruito tramite i dati sperimentali si può ricorrere all'utilizzo del coefficiente di correlazione lineare R^2 . Nel campo della statistica, questo coefficiente

viene chiamato anche indice di correlazione di Pearson ed esprime la relazione di linearità che sussiste tra due variabili.

Può essere calcolato come:

$$R^2 = \frac{\sum_i [\hat{y}_i - y_m]^2}{\sum_i [y_i - y_m]^2}$$

Dove:

- y_m = valore medio delle letture effettuate

Il valore di R^2 può essere compreso tra 0 e 1. In particolare, quando R^2 è prossimo ad 1 allora i punti sperimentali giacciono sulla retta, mentre quando R^2 è prossimo a 0 i punti sono distribuiti sul piano in maniera casuale. In conclusione, più l'indice è alto maggiore sarà l'affidabilità della retta estrapolata tramite il metodo dei minimi quadrati mentre più l'indice è basso minore sarà la probabilità di ricadere in una relazione lineare che approssima il comportamento del sensore.

3.2. Gli strumenti per la calibrazione

Di seguito vengono riportati gli strumenti utilizzati per la taratura dei sensori presenti all'interno dell'impianto. In base alle tipologie di variabili fisiche da misurare, ci si è avvalsi di apposite macchine di taratura.

3.2.1. Le misure di pressione

Per la taratura dei sensori di pressione sia assoluti che relativi si è ricorso all'utilizzo di una macchina idraulica presente in Figura 3.2 in cui il solo azionamento manuale è in grado di generare delle pressioni nel range compreso tra 0 e 10 bar assoluti.



Figura 3.2: Macchina idraulica per taratura

3.2.2. Le misure di flusso

Per i sensori che hanno il compito di registrare la portata all'interno dell'impianto, la calibrazione è stata effettuata tramite l'utilizzo di un contatore e di un cronometro. Il contatore è stato inserito in serie al sensore e ha contato i metri cubi di fluido che lo hanno attraversato. Dopo aver portato il sistema a regime e aver generato un flusso costante, con il contatore si registra la quantità di fluido passata per il sensore, mentre con il cronometro il tempo impiegato.



Figura 3.3: Contatore per la taratura

3.2.3. Le misure di livello

Per la taratura del trasduttore di livello, dato che sul serbatoio è installata una presa di pressione che permette di monitorare a vista il livello del liquido, è stata utilizzata una semplice barra millimetrata.

3.2.4. La misura dell'output

Come già precedentemente descritto nella sezione 2.3.2, i sensori presenti nel laboratorio sono stati collegati ad un'unità centrale Revolution Pi in grado di registrare i loro segnali di uscita sia in tensione sia in corrente.

3.3. La taratura dei sensori

Di seguito vengono riportati per ogni sensore le tabelle che testimoniano:

- Dettagli costruttivi;
- Dati acquisiti per la taratura;
- Grafico della curva di taratura;
- Equazione curva di taratura lineare;
- Coefficiente di correlazione R^2 .

3.3.1. S1 - Endress+ Hauser Cerabar M PMP51

Tabella 3.1: Specifiche S1 - Endress+ Hauser Cerabar M PMP51

Codice	S1
Grandezza	Pressione assoluta
Marca	Endress+ Hauser
Modello	Cerabar M PMP51
Range	0-10 bar
Excit.	220 V
Output	4 - 20 mA
Data Taratura	28/03/2022

Tabella 3.2: Dati di taratura S1 - Endress+ Hauser Cerabar M PMP51

Pressione [bar]	Corrente [mA]
0	4000
1,005	5614
1,012	5624
6,153	13859
10	20000

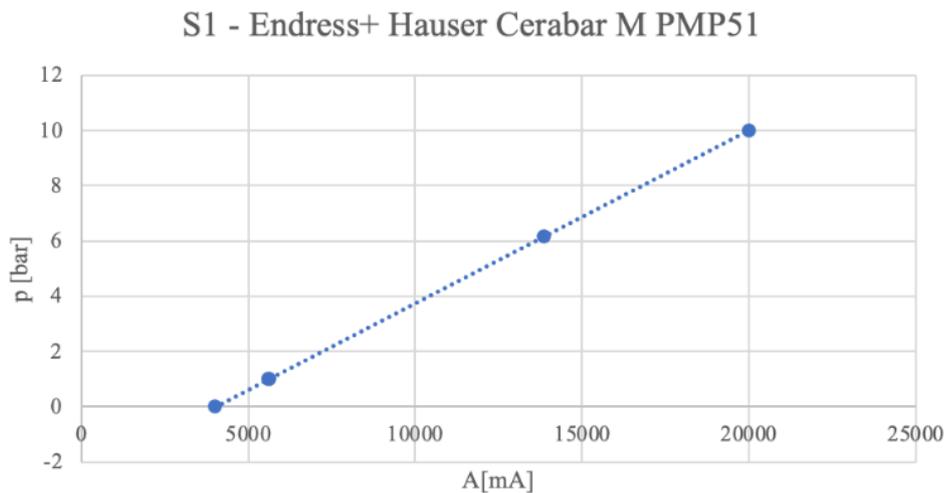


Figura 3.4: Taratura S1 - Endress+ Hauser Cerabar M PMP51

- **Funzione p(A):** $p = 0.0006A - 2.5027$
 - **Correlazione:** $R^2 = 1$

3.3.2. S2 - Endress+ Hauser Promag W

Tabella 3.3: Specifiche S2 - Endress+ Hauser Promag W

Codice	S2
Grandezza	Portata d'acqua
Marca	Endress+ Hauser
Modello	Promag W
Range	0-10 m ³ /h
Excit.	220 V
Output	4 - 20 mA
Data Taratura	28/03/2022

Tabella 3.4: Dati di taratura S2 - Endress+ Hauser Promag W

Portata [m ³ /h]	Corrente [mA]
0	4003
9,3144	8146
9,706	8322
9,735	8338
10,1595	8524
10,5845	8713
10,625	8730

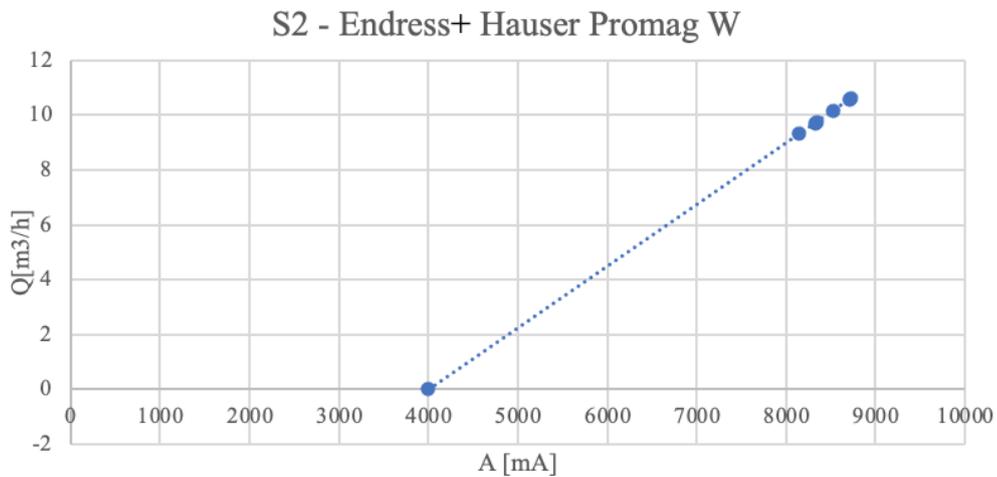


Figura 3.5: Taratura S2 - Endress+ Hauser Promag W

- **Funzione Q(A):** $Q = 0.0022A - 8.9957$
- **Correlazione:** $R^2 = 1$

3.3.3. S3 - Setra 280E

Tabella 3.5: Specifiche S3 - Setra 280E

Codice	S3
Grandezza	Pressione assoluta
Marca	Endress+ Hauser
Modello	280 E
Range	0-100 psi
Excit.	24V
Output	0-5 V
Data Taratura	28/03/2022

Tabella 3.6: Dati di taratura S3 - Setra 280E

Portata [m ³ /h]	Tensione [mV]
1,013	777
1,513	1128
2,013	1494
3,013	2215
4,013	2951
5,013	3690
6,013	4395
7,013	5050
7,513	5460

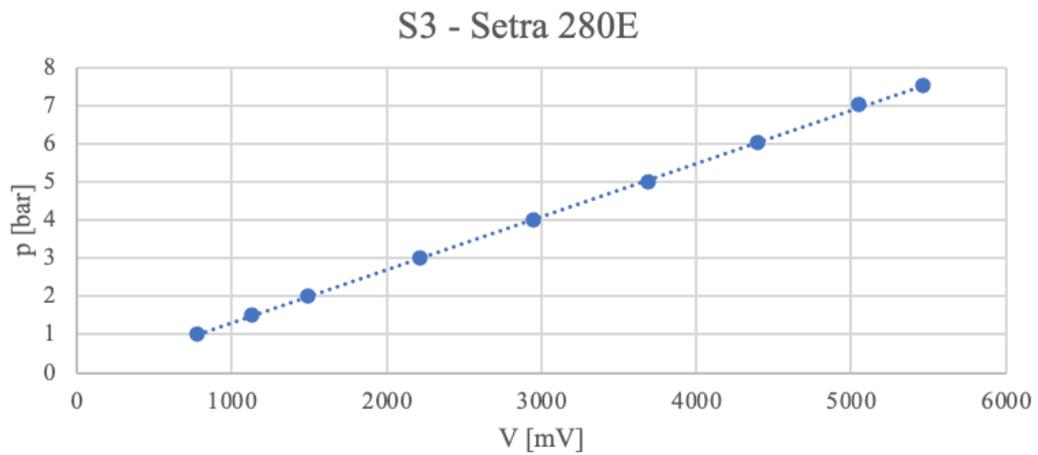


Figura 3.6: Taratura S3 - Setra 280E

- **Funzione p(V):** $p = 0.001388V - 0.065475$
 - **Correlazione:** $R^2 = 0,9998$

3.3.4. S4 - Foxboro 841GM CI1

Tabella 3.7: Specifiche S4 - Foxboro 841GM CI1

Codice	S4
Grandezza	Pressione relativa
Marca	Foxboro
Modello	841GM-CI1
Range	0-300 psi
Excit.	24V
Output	4-20 mA
Data Taratura	28/03/2022

Tabella 3.8: Dati di taratura S4 - Foxboro 841GM CI1

Pressione rel. [bar]	Pressione ass. [bar]	Corrente [mA]
0	1,013	3969
0,5	1,513	4749
1	2,013	5548
2	3,013	7189
3	4,013	8779
4	5,013	10434
5	6,013	11981
6	7,013	13567
8	9,013	16722
10	11,013	19924

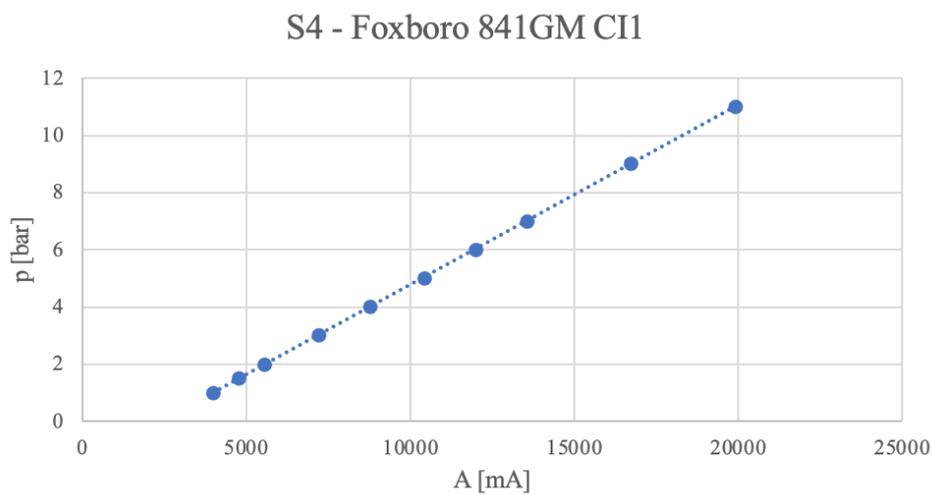


Figura 3.7: Taratura S4 - Foxboro 841GM CI1

- **Funzione p(A):** $p = 0.000627A - 1.47462$
- **Correlazione:** $R^2 = 1$

3.3.5. S5 - Foxboro 841GM CI1

Tabella 3.9: Specifiche S5 - Foxboro 841GM CI1

Codice	S5
Grandezza	Pressione relativa
Marca	Foxboro
Modello	841GM-CI1
Range	0-300 psi
Excit.	24V
Output	4-20 mA
Data Taratura	28/03/2022

Tabella 3.10: Dati di taratura S5 - Foxboro 841GM CI1

Pressione rel. [bar]	Pressione ass. [bar]	Corrente [mA]
0	1,013	3978
0,5	1,513	4764
1	2,013	5575
2	3,013	7162
3	4,013	8849
4	5,013	10400
5	6,013	11963
6	7,013	13496
8	9,013	16772
10	11,013	19950

S5 - Foxboro 841GM CI1

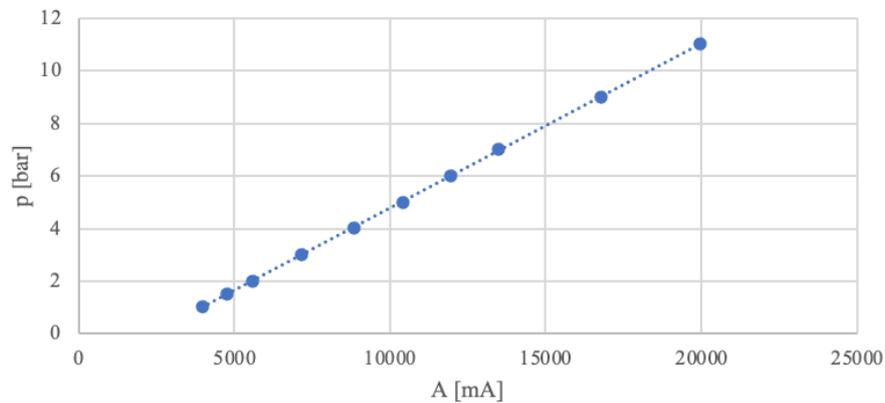


Figura 3.8: Taratura S5 - Foxboro 841GM CI1

- **Funzione p(A):** $p = 0.000626A - 1.47761$
- **Correlazione:** $R^2 = 1$

3.3.6. S6 - Foxboro Vortez DN 50

Tabella 3.11: Specifiche S6 - Foxboro Vortez DN 50

Codice	S6
Grandezza	Portata
Marca	Foxboro
Modello	Vortez DN50
Range	0-50 m ³ /h
Excit.	24V
Output	4-20 mA
Data Taratura	28/03/2022

Tabella 3.12: Dati di taratura S6 - Foxboro Vortez DN 50

Portata [m ³ /h]	Corrente [mA]
0	4000
50	20000

S6 - Foxboro Vortez DN 50

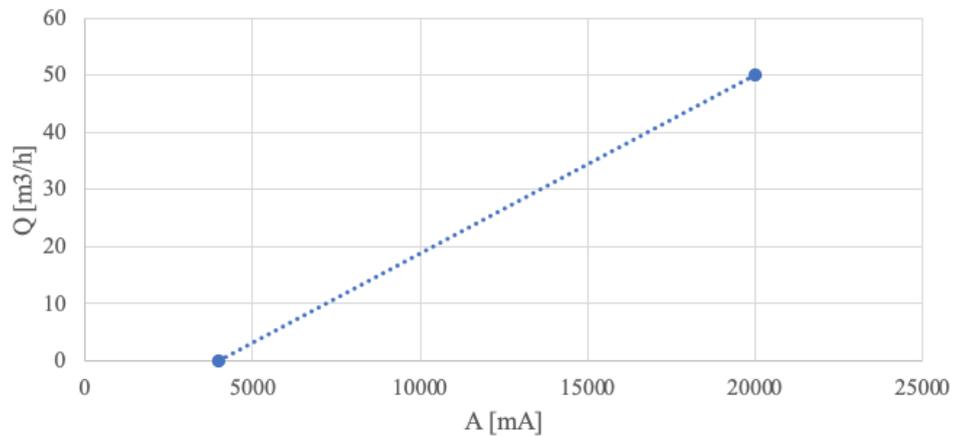


Figura 3.9: Taratura S6 - Foxboro Vortez DN 50

- **Funzione Q(A):** $Q = 0.003125A - 12.5$
- **Correlazione:** $R^2 = 1$

3.3.7. S7 - Foxboro IDP-10

Tabella 3.13: Specifiche S7 - Foxboro IDP-10

Codice	S7
Grandezza	Livello acqua
Marca	Foxboro
Modello	IDP-10
Range	0-10 bar
Excit.	24 V
Output	4 - 20 mA
Data Taratura	28/03/2022

Tabella 3.14: Dati di taratura S7 - Foxboro IDP-10

Livello [mm]	Corrente [mA]
0	3864
83	5793
105	6285
151	7279

S7 - Foxboro IDP-10

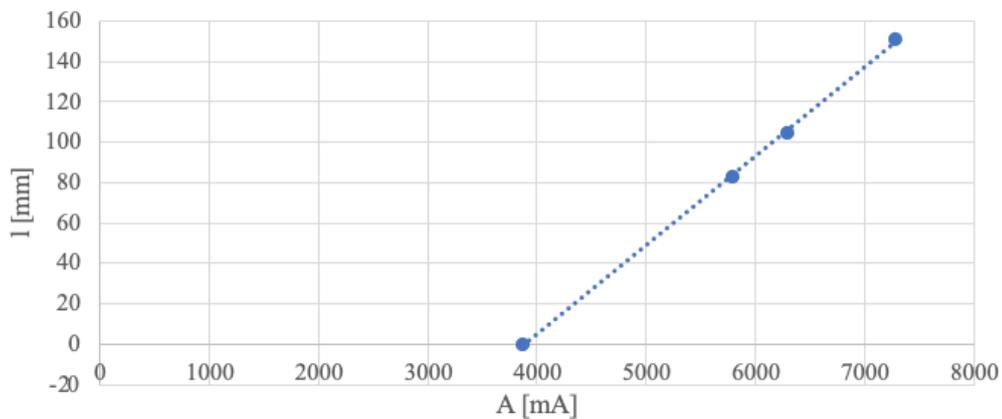


Figura 3.10: Taratura S7 - Foxboro IDP-10

- **Funzione I(A):** $I = 0.044217A - 170.8533$
 - **Correlazione:** $R^2 = 0,9996$

3.3.8. S8 - Endress+ Hauser Prowirl 200

Tabella 3.15: Specifiche S8 - Endress+ Hauser Prowirl 200

Codice	S8
Grandezza	Portata
Marca	Endress+ Hauser
Modello	Prowirl 200
Range	0-125 m ³ /h
Excit.	24 V
Output	4 - 20 mA
Data Taratura	28/03/2022

Tabella 3.16: Dati di taratura S8 - Endress+ Hauser Prowirl 200

Portata [m ³ /h]	Corrente [mA]
0	4000
125,2	20000

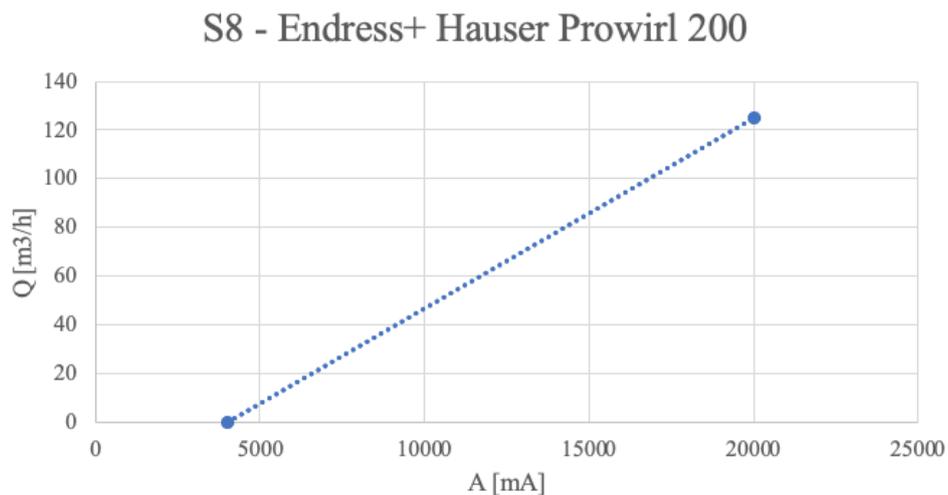


Figura 3.11: Taratura S8 - Endress+ Hauser Prowirl 200

- **Funzione Q(A):** $Q = 0.0078A - 31.3$
- **Correlazione:** $R^2 = 1$

4. L'intelligenza artificiale

4.1. Introduzione

L'intelligenza artificiale (in inglese, *Artificial Intelligence, AI*) è una disciplina che cerca di emulare il comportamento del cervello umano tramite un algoritmo appositamente designato e un'infrastruttura digitale ottimizzata. La parola *intelligenza* deriva dal latino *interlegere*, cioè saper scegliere tra diverse opzioni, selezionare, discernere l'informazione utile in mezzo a tante altre. Il professor Marvin Lee Minsky nel 1990 la definisce come “la scienza di far fare alle macchine delle cose che richiederebbero intelligenza se fossero fatte dagli uomini”. Se fosse comparata all'uomo, l'intelligenza artificiale dovrebbe compiere in modo razionale le seguenti azioni:

- agire umanamente rispetto a un essere umano;
- pensare umanamente attraverso funzioni cognitive simili a quello dell'uomo;
- pensare razionalmente ricorrendo alla logica;
- agire razionalmente attraverso delle tecniche che permettono all'intelligenza artificiale di estrarre il miglior risultato atteso in base ai dati messi a disposizione dall'utente.

Per fare ciò un sistema intelligente deve essere dotato di:

- **comprensione**: sistemi che simulano capacità cognitive di correlazione dati ed eventi per il riconoscimento di testi, immagini, tabelle, video, voce da cui estrapolare informazioni;
- **ragionamento**: sistemi che sfruttano il processo logico per considerare contemporaneamente molteplici informazioni raccolte mediante precisi algoritmi;
- **apprendimento**: sistemi di machine learning che definiscono come trattare gli input di dati e come restituirli, in grado di apprendere da esempi forniti dall'uomo e di migliorare le prestazioni con l'apprendimento continuo;
- **interazione uomo-macchina**: sistemi per interagire attraverso le modalità di percezione dell'uomo (questo è uno dei punti focali che anima il dibattito filosofico negli ultimi anni).

Negli anni nascono due scuole di pensiero che si distinguono in base al ruolo svolto dalla macchina intelligente:

- AI deboli: le macchine svolgono compiti complessi per l'uomo agendo come se fossero intelligenti e avessero una mente;
- AI forti: il sistema è dotato di una coscienza che si basa su una serie di programmi atti a riprodurre le prestazioni e conoscenze delle persone esperte in un determinato campo.

4.2. Machine learning

Un algoritmo tradizionale è un codice alfanumerico che passato ad un computer, svolge un determinato compito prefissato dal programmatore. Un algoritmo di intelligenza artificiale, invece, è in grado di percepire tramite dei sensori l'ambiente circostante e di interagirvi tramite degli attuatori prendendo autonomamente scelte ottimali. Il machine learning è un sottoinsieme dell'intelligenza artificiale che ha il compito di creare dei sistemi in grado di apprendere e ottimizzare le loro prestazioni in base ai dati forniti in ingresso. In sostanza, gli algoritmi di machine learning sfruttano dei metodi matematico-computazionali per ricavare le informazioni direttamente dai dati. In questo modo migliorano anche le loro prestazioni in modo "adattivo", man mano che gli "esempi" da cui apprendere aumentano.

Questa tipologia di algoritmi si divide in tre macrogruppi (3) :

- Apprendimento supervisionato;
- Apprendimento non supervisionato;
 - Apprendimento semi-supervisionato;
- Apprendimento per rinforzo

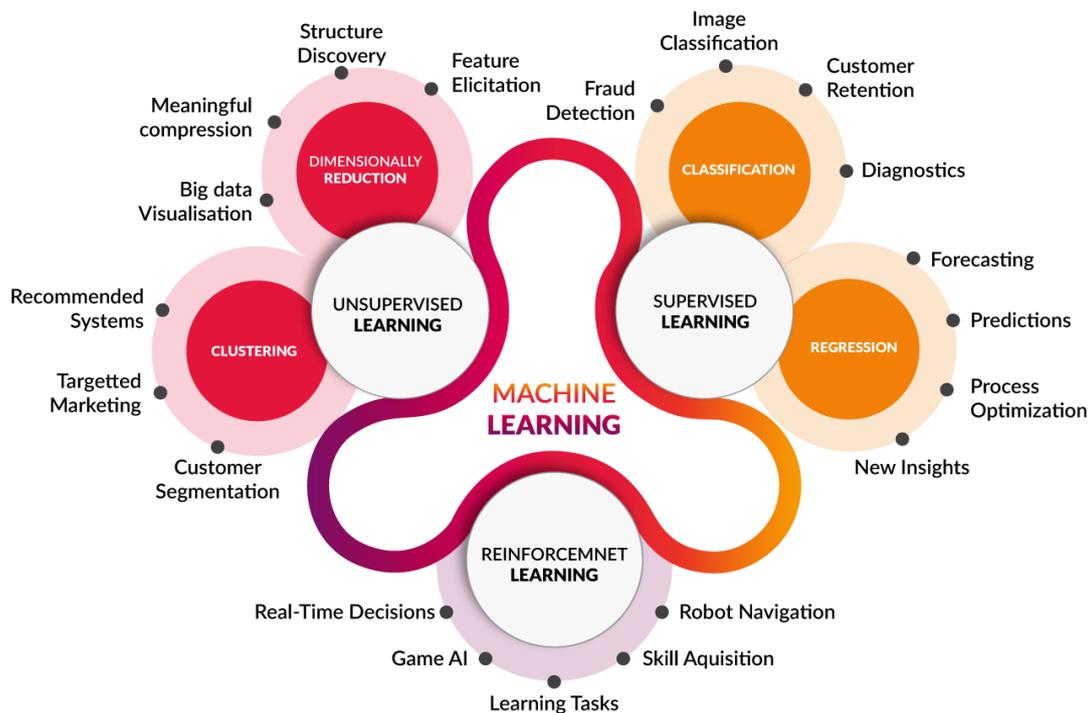


Figura 4.1: Tipi di Machine learning

4.2.1. Apprendimento supervisionato

Il machine learning supervisionato è un tipo di apprendimento automatico in cui l'algoritmo viene addestrato utilizzando dei dati strutturati ed etichettati che fungono da supervisori per la previsione degli output. Ciò significa che i dati di input sono già contrassegnati con gli output corretti. I dati di input vengono divisi in due gruppi: *training set* formato dall'80% dei dati con cui addestrare l'algoritmo, *test set* formato dalla restante parte per valutare le prestazioni. Dopo aver analizzato i dati d'ingresso, lo scopo dell'algoritmo è quello di trovare una funzione di mappatura in grado di correlare ingresso e uscita. Bisogna però prestare attenzione a come vengono divisi i dati nei due insiemi in quanto non devono esserci sbilanciamenti e tutte le casistiche devono essere ricoperte da entrambi gli insiemi.

Come suggerito dalla Figura 4.1, gli algoritmi di ML supervisionato possono essere divisi in due principali categorie di:

- algoritmi di classificazione: gli output vengono divisi in due o più classi. Il programma deve essere in grado di creare un modello capace di attribuire ai dati di input una o più classi tra quelle disponibili. Esempi di algoritmi di classificazione sono: K-NN, macchina a vettori di supporto, alberi delle decisioni, regressione logistica.
- algoritmi di regressione: simili a quelli per classificazione, ma con la differenza che l'output si presenta in un dominio continuo e non discreto. Esempi di algoritmi di regressione sono: regressione lineare, regressione polinomiale.

4.2.2. Apprendimento non supervisionato

Il machine learning non supervisionato è una tecnica di apprendimento in cui, a differenza del machine learning supervisionato, l'utente non funge più da supervisore e non fornisce all'algoritmo dei dati etichettati. Questo metodo consente al modello di lavorare autonomamente alla ricerca della divisione in gruppi ottimale per i dati d'ingresso. La capacità di scoprire somiglianze e differenze nelle informazioni lo rendono la soluzione ideale per l'analisi esplorativa dei dati, le strategie di cross-selling, la segmentazione dei clienti e il riconoscimento delle immagini. Un esempio di machine learning non supervisionato è rappresentato dalle mappe auto-organizzanti (in inglese, *Self-Organizing-Map*, SOM) sfruttate per la tutela della salute e l'identificazione delle anomalie dell'impianto oggetto d'esame.

4.2.2.1. Apprendimento semi-supervisionato

Il machine learning semi-supervisionato è la combinazione dei due approcci appena spiegati: l'algoritmo prevede una prima fase supervisionata su dei dati etichettati e una successiva fase non supervisionata.

4.2.3. Apprendimento per rinforzo

Questo tipo di machine learning basa il suo funzionamento sul concetto di premio. L'algoritmo interagendo costantemente con l'ambiente circostante riceve continui feedback in base all'azione intrapresa. L'algoritmo riceverà da parte dell'utente un *positive reward* se eseguirà un'azione corretta e un *negative reward* in caso contrario.

4.3. Le reti neurali

Le reti neurali artificiali (*Artificial Neural Network*, ANN) sono dei modelli matematici basati sul lavoro svolto dai neuroni artificiali. Quest'ultimi si ispirano al funzionamento dei neuroni biologici presenti nel cervello. Ogni cellula nervosa biologica riceve dai neuroni vicini dei continui stimoli. Una volta integrati gli stimoli, la cellula si eccita e trasmette il suo potenziale sotto forma di impulso elettrico ai neuroni circostanti.

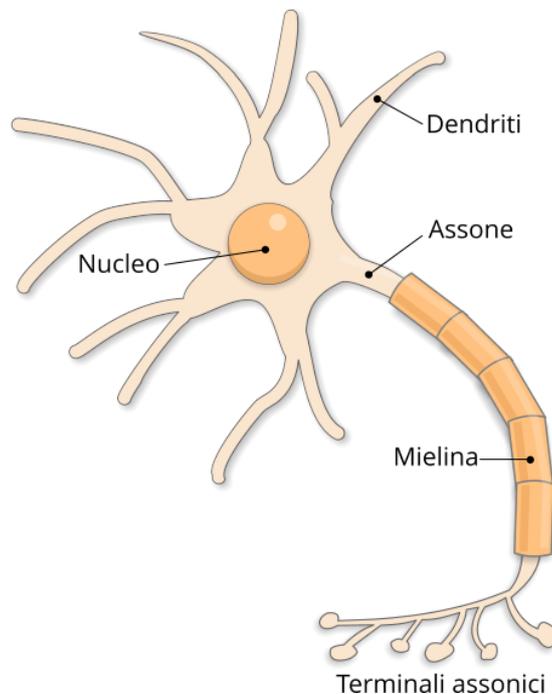


Figura 4.2: Il neurone biologico

Dalla Figura 4.2 è possibile visualizzare le parti che formano il neurone biologico:

- **dendriti:** si dipartono dal corpo della cellula e funzionano da recettori dell'impulso nervoso, ovvero catturano i segnali elettrici degli assoni di altri neuroni;
- **corpo cellulare:** in esso sono contenuti molti elementi cellulari, come i mitocondri, i ribosomi e il nucleo, che racchiude il DNA. Quest'ultimo contiene le istruzioni per sintetizzare le proteine che determinano la forma e la funzione della cellula;
- **assone:** è il prolungamento nervoso del neurone che trasmette l'impulso elettrico dal corpo cellulare. Spesso è rivestito da una guaina mielinica che isola la carica elettrica e accelera la trasmissione dell'impulso nervoso;
- **sinapsi:** è lo spazio in cui si produce la connessione tra due neuroni. Quando l'impulso nervoso giunge al bottone sinaptico, situato all'estremità dell'assone, piccole sacche piene di fluido, dette *vescicole sinaptiche*, liberano delle sostanze chimiche note come neurotrasmettitori. Tali molecole attivano le proteine situate nei dendriti del secondo neurone, che si trasformano in recettori del segnale elettrico. In questo modo l'impulso nervoso viaggia da neurone all'altro e in neuroni si connettono per formare le reti neurali;

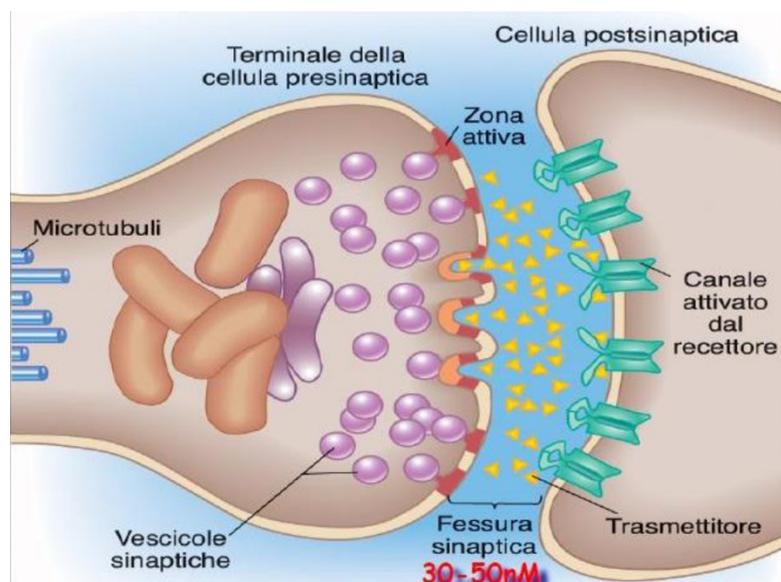


Figura 4.3: Spazio sinaptico

Cercando di comprendere il funzionamento del cervello biologico, Warren McCulloch e Walter Pitts pubblicano nel 1943 una prima rappresentazione schematica di una cellula semplificata del cervello, il *neurone di McCulloch - Pitts (MCP)* (4). Il neurone è rappresentato come un semplice gate logico con un'uscita binaria. Nella Figura 4.4 si vede come ai dendriti arrivano diversi segnali, che poi entrano nel corpo cellulare. Se il segnale accumulato supera una certa soglia, si produce un segnale di uscita.

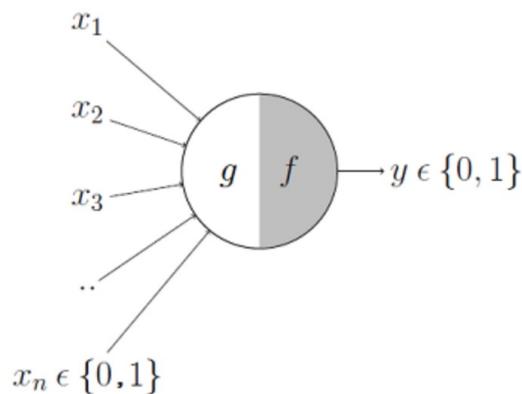


Figura 4.4: Modello di neurone di McCulloch-Pitts

Questo modello permetteva però di prevedere solo funzioni lineari con input e output di tipo binario, con una soglia di attivazione impostata manualmente e con dei pesi tutti uguali per gli input.

Dopo pochi anni, Frank Rosenblatt pubblica il primo concetto di *apprendimento Perceptron* basato sul modello del neurone MCP (5). Tramite l'utilizzo di questa regola, l'algoritmo riesce a ricercare autonomamente i coefficienti peso ottimali da moltiplicare ai nodi di input così da essere in grado di decidere sull'attivazione del neurone.

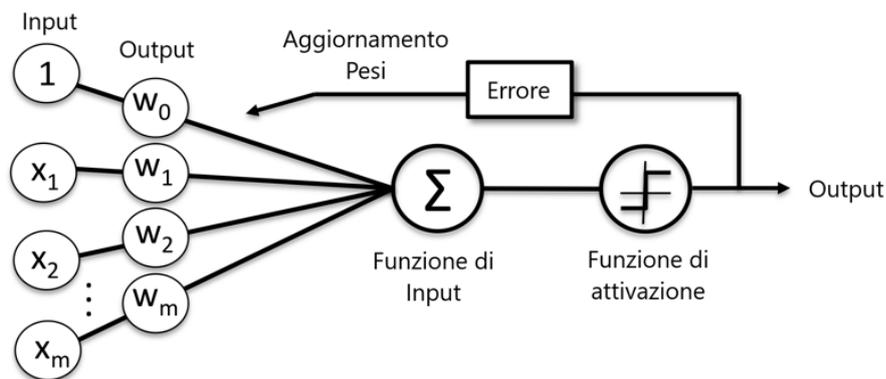


Figura 4.5: Modello di neurone di Rosenblatt (*Perceptron*)

Il metodo *Perceptron*, quindi, può essere utilizzato come algoritmo di classificazione binaria dove le due classi di appartenenza sono espresse tramite:

- 1 (classe positiva);
- -1 (classe negativa).

Viene definita una *funzione di attivazione* $\phi(z)$, dove z è l'input fornito alla rete che è formato da una combinazione lineare dei valori di ingresso x e i loro corrispondenti vettori peso w .

$$z = w_1x_1 + w_2x_2 + \dots + w_nx_n$$

Siano:

$$w = \begin{bmatrix} w_1 \\ \dots \\ w_n \end{bmatrix}, x = \begin{bmatrix} x_1 \\ \dots \\ x_n \end{bmatrix}$$

Se l'output di $\phi(z)$ è maggiore di una soglia stabilita θ , allora l'algoritmo sarà in grado di prevedere che quella determinata uscita rientra nella classe 1, altrimenti nell'altra classe.

La $\phi(\cdot)$ è una funzione definita a tratti:

$$\phi(z) = \begin{cases} 1 & \text{se } z \geq \theta \\ -1 & \text{altrimenti} \end{cases}$$

Allora è possibile riscrivere sia $\phi(z)$ che z in forma compatta:

$$z = w_0x_0 + w_1x_1 + w_2x_2 + \dots + w_nx_n = \sum_{j=0}^m x_jw_j = w^T x \text{ dove } w_0 = -\theta, x_0 = 1$$

$$\phi(z) = \begin{cases} 1 & \text{se } z \geq 0 \\ -1 & \text{altrimenti} \end{cases}$$

La lettera T messa all'apice di w sta per *trasposizione*, un'operazione che trasforma un vettore colonna in un vettore riga e viceversa.

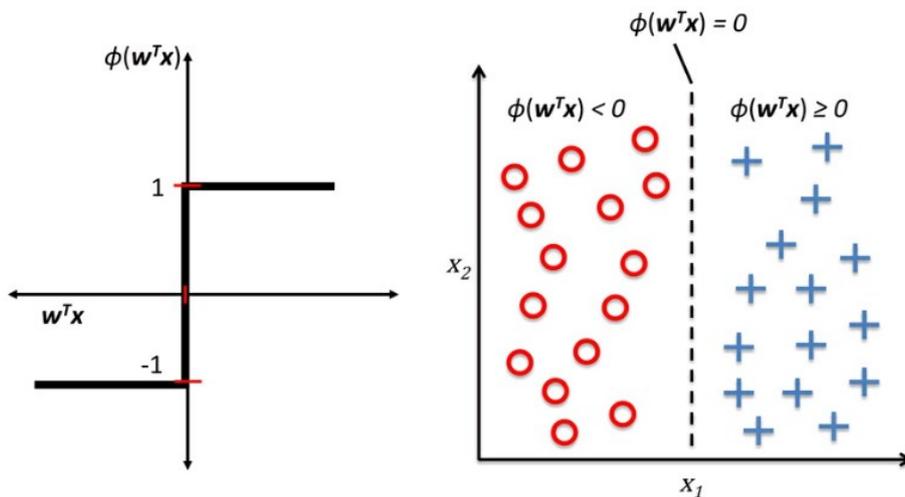


Figura 4.6: Classificazione binaria (metodo Perceptron)

La regola del Perceptron iniziale di Rosenblatt può essere riassunta dai seguenti passi:

1. inizializzare i pesi a 0 o a numeri casuali piccoli;
2. per ogni campione di addestramento svolgere seguenti passi.

1. calcolare il valore di output;
2. aggiornare i pesi.

Il valore di output è l'etichetta della classe prevista dalla funzione $\phi(z)$ e il simultaneo aggiornamento di ogni peso w_j nel vettore dei pesi w può essere scritto come:

$$w_j := w_j + \Delta w_j$$

Il valore di Δw_j viene utilizzato per aggiornare il peso w_j e si calcola come:

$$\Delta w_j = \eta(y^i - \hat{y}^i)x_j^i$$

Con η è indicato il *tasso di apprendimento* (un valore costante compreso fra 0.0 e 1.0), y^i è la vera etichetta della classe del campione di apprendimento i -esimo e \hat{y}^i è l'etichetta della classe prevista. La convergenza della rete è garantita solo se le due classi sono separabili linearmente e se il tasso di apprendimento è sufficientemente ridotto. Se le classi non possono essere separate da un confine decisionale lineare, si definisce un numero massimo di passi sul dataset di apprendimento (epoche) e/o una soglia per il numero di errate classificazioni tollerate. In caso contrario, l'algoritmo così strutturato non smetterebbe mai di correggere i pesi (6).

4.3.1. Le reti neurali a singolo livello feedforward

In una rete neurale i nodi di input vengono organizzati sotto forma di livelli (*layers*). Questa tipologia di rete prende il nome *feedforward* in quanto le informazioni si muovono dai neuroni di input a quelli di output e non viceversa.

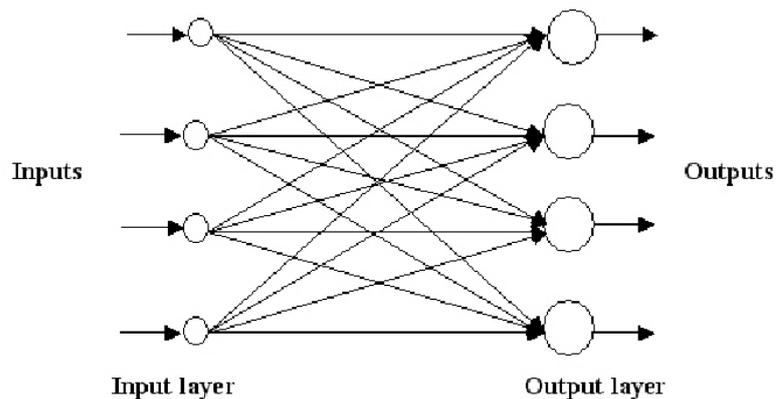


Figura 4.7: Rete neurale singolo livello feedforward

La Figura 4.7 rappresenta il caso più semplice, dove il livello di input è collegato direttamente al livello di output. Nonostante la presenza di due livelli, la rete viene comunque chiamata a singolo strato perché i neuroni di input non eseguono operazioni.

4.3.2. Le reti neurali multilivello feedforward

Le reti neurali multilivello feedforward, chiamate anche *multy layer perceptron (MLP)*, presentano uno o più livelli nascosti (*hidden layers*) tra il livello di input e quello di output. I nodi presenti nel layer di input vengono connessi a quelli del layer nascosto e quest'ultimi a quelli di output. Se la rete presenta più di un livello nascosto prende il nome di rete neurale artificiale profonda. A differenza delle reti a singolo livello, le *multy layer perceptron* sono in grado di trovare soluzioni a problemi non lineari.

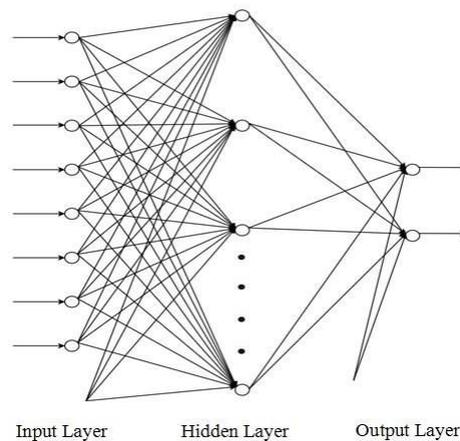


Figura 4.8: Rete neurale multilivello Feedforward

4.3.3. Le reti neurali ricorrenti

Le reti neurali ricorrenti (*Recurrent Neural Network, RNN*) sono una tipologia di reti neurali artificiali in cui i valori di uscita di un determinato strato successivo vengono utilizzati come ingresso ad uno strato di livello inferiore. L'interconnessione tra i diversi layer consente l'utilizzo di uno degli strati come memoria. Trovano largo impiego nell'analisi predittiva.

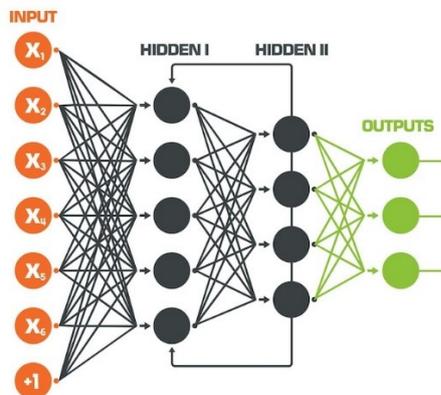


Figura 4.9: Rete neurale ricorrente

4.4. Le Self-Organizing Maps

Le mappe auto-organizzanti (in inglese *SOM*, *Self-Organizing Maps*) sono delle particolari reti neurali che sfruttano l'apprendimento non supervisionato. In base alla somiglianza delle informazioni fornite in ingresso, le mappe riescono a riordinare i dati svolgendo una sorta di classificazione in uscita.

Esistono due tipologie di reti SOM:

- **monodimensionali**: i neuroni vengono organizzati in una struttura monodimensionale;
- **bidimensionali**: i neuroni vengono organizzati in una griglia a due dimensioni.

Questa tipologia di rete è stata teorizzata per la prima volta nel 1982 dal professor dell'Accademia di Finlandia Tuevo Kohonen. Prende il nome di “*Self-Organizing*” perché non è richiesta alcuna supervisione, “*Maps*” perché la rete tenta di mappare i propri pesi per conformarsi ai dati forniti in ingresso. Il metodo è ampiamente utilizzato come strumento di clustering e di visualizzazione nell'analisi dei dati. Tra le aree di applicazione delle SOM ci sono la strumentazione industriale, il riconoscimento dei modelli, l'estrazione e l'elaborazione dei dati, il controllo dei processi, la robotica, le telecomunicazioni, le applicazioni mediche, l'ottimizzazione e la gestione dei prodotti. L'obiettivo principale della rete è trasformare uno spazio di input complesso ad alta dimensione in uno spazio di output più semplice e discreto a bassa dimensione preservando le relazioni dei dati d'ingresso.

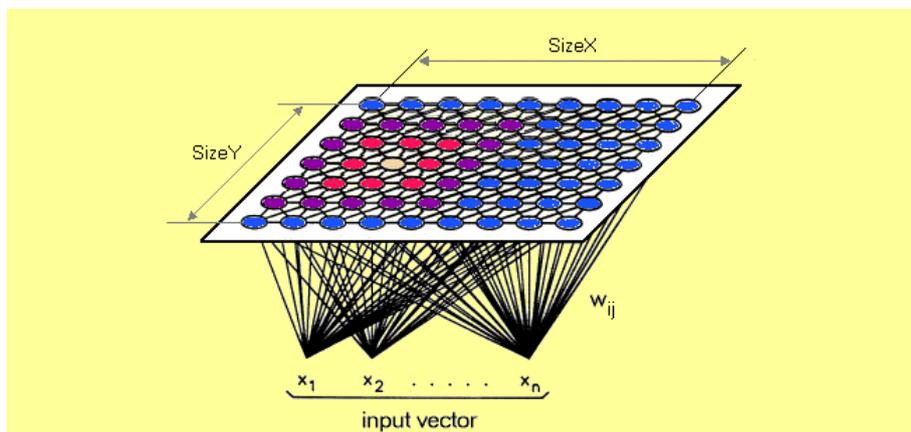


Figura 4.10: Self-Organizing Maps

Nello spazio di output, le coordinate dei nodi vengono assegnate in base alle caratteristiche intrinseche dei dati d'ingresso. Le mappe auto-organizzanti sono simili ad alcune tecniche di riduzione dimensionale come l'analisi delle componenti principali (*Principal Component Analysis, PCA*) (7) e il ridimensionamento multidimensionale (*Multidimensional Scaling, MDS*) (8). Tuttavia, le SOM presentano numerosi vantaggi rispetto a queste tecniche. Infatti, le mappe auto-organizzanti

sono facili da implementare, in grado di risolvere problemi molto complessi ed elaborano in modo più efficace i dati rumorosi.

Come spiegato nel capitolo 4.2.2 dedicato all'Apprendimento non supervisionato, la SOM non classifica gli output mediante degli esempi forniti in ingresso, ma sfrutta una forma di apprendimento non supervisionato, che prende il nome di "apprendimento competitivo" (*competitive learning*), in cui i nodi di output competono tra loro per rappresentare i diversi modelli presenti nello spazio di input. La natura feedforward della rete fa sì che il flusso d'informazioni sia unidirezionale. Dalla Figura 4.10 si può notare che ogni nodo dello strato di ingresso è collegato (tramite dei pesi w_{ij}) a ogni nodo dello strato di uscita.

La formazione di una SOM comporta tre processi caratteristici, che possono essere riassunti come segue:

- **competizione:** i nodi di uscita competono tra loro per rappresentare al meglio il campione in ingresso. L'efficacia della classificazione nella griglia di output viene misurata utilizzando una funzione discriminante, in cui un vettore di input viene confrontato con il vettore dei pesi di ciascun nodo di output. Il nodo con il peso più simile al campione in ingresso viene dichiarato vincitore. La funzione più utilizzata per determinare la migliore unità di corrispondenza (*Best Matching Unit, BMU*) è la distanza euclidea.
- **cooperazione:** analogamente a quello che avviene nei neuroni umani, che trattano informazioni simili attraverso l'impiego di neuroni vicini così da sfruttare connessioni sinaptiche brevi, l'organizzazione topografica della SOM prevede che nello spazio di uscita i neuroni vicini rappresentino input con proprietà simili. Il best match unit determina la posizione spaziale del quartiere di nodi cooperanti. Questi nodi, condividendo caratteristiche comuni, si attivano l'un l'altro per imparare qualcosa dallo stesso input.
- **adattamento:** i vettori peso del nodo vincitore e quelli del vicinato vengono adattati a favore dei valori più alti della funzione discriminante. Attraverso questo processo di apprendimento, i nodi vincitori diventano più simili al campione in ingresso. Pertanto, i nodi che hanno una forte risposta ad un particolare input avranno maggiori possibilità di rispondere a dati di input simili in futuro.

Questi processi vengono ripetuti per tutti i campioni di addestramento che vengono passati in sequenza alla SOM. In questo modo, le diverse posizioni della mappa vengono addestrate per produrre una risposta forte ai diversi tipi di dati in ingresso.

4.4.1. La struttura

La struttura della mappa auto-organizzante è tipicamente costituita da due strati, lo strato d'ingresso e quello di uscita. A differenza della maggior parte delle altre reti neurali, nella SOM i nodi d'ingresso sono collegati direttamente a quelli di uscita senza la presenza di un hidden layer. I nodi del livello d'ingresso includono gli "attributi" (*features*), cioè tutte le variabili contenute nei dati di input (9).

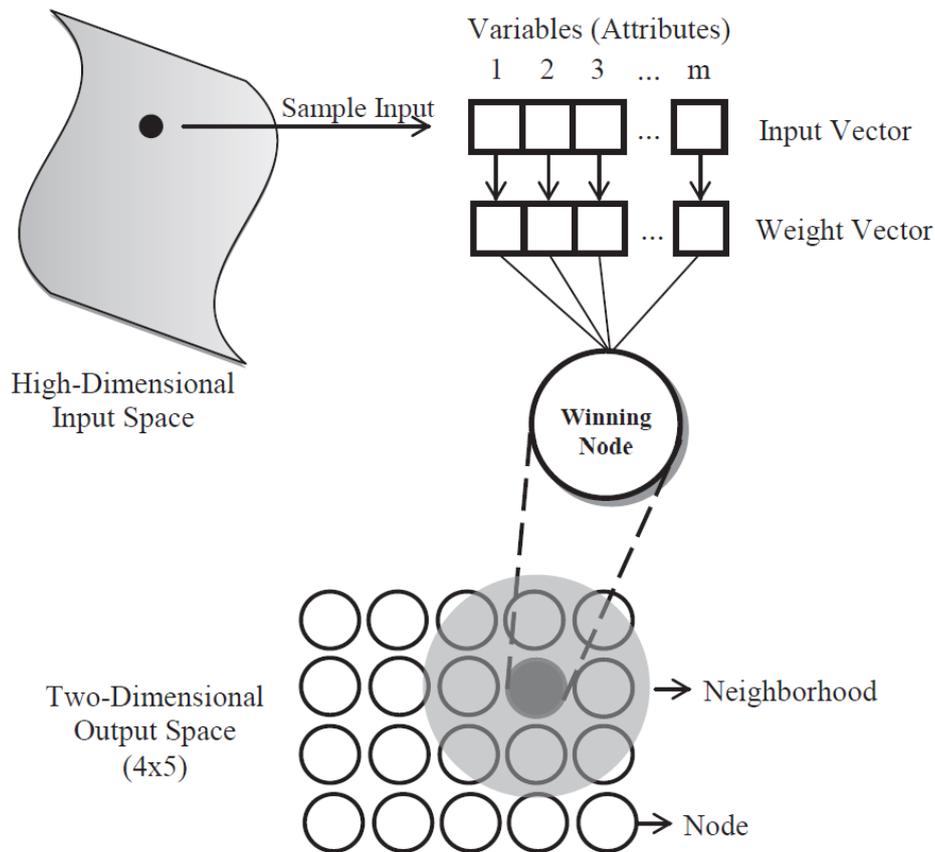


Figura 4.11: Struttura della Self-Organizing Map

Ogni dato d'ingresso è rappresentato da un vettore m -dimensionale, $x = (x_1, x_2, \dots, x_m)$ dove gli elementi indicano le caratteristiche dell'input. Se il database di partenza è vasto e ci sono grandi differenze tra i vari valori degli attributi, occorre una normalizzazione totale per evitare la dominanza di una particolare feature o sottoinsieme di features. In letteratura sono stati proposti diversi metodi per la normalizzazione del dataset, uno dei più comuni è la trasformazione *z-score* che converte i valori x_j di ogni dato in punteggi standard, con una media μ_{x_j} pari a zero e una deviazione standard unitaria σ_{x_j} , sottraendo la media e dividendo per la deviazione standard ogni attributo.

$$\frac{x_j - \mu_{xj}}{\sigma_{xj}}$$

Lo strato di uscita, noto anche come "*Kohonen layer*" o "*SOM layer*", è una rappresentazione a bassa dimensione dei dati di input. I nodi dello strato di uscita sono disposti secondo un'architettura solitamente bidimensionale organizzata come una griglia composta da righe e colonne. Il numero di nodi nello strato di uscita denota il numero massimo di cluster e influenza l'accuratezza della SOM. Secondo uno studio riportato in letteratura (10), sia N la dimensione del dataset d'ingresso, il numero dei nodi in uscita è pari a $5\sqrt{N}$. Se si rappresenta la mappa sotto forma di schema quadrato, la lunghezza delle righe e delle colonne risulta $\sqrt{5\sqrt{N}}$.

In genere, come mostrato nella Figura 4.12 la topologia della rete può essere quadrata o esagonale.

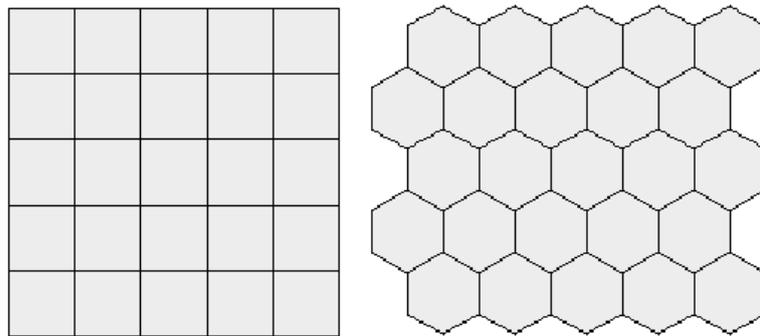


Figura 4.12: Rete SOM a topologia quadrata ed esagonale

Queste topologie hanno proprietà diverse. Nella topologia rettangolare ogni nodo ha quattro vicini, mentre in quella esagonale sei. Presentando un maggior numero di vicini, la struttura esagonale è solitamente la preferita.

4.4.2. L'algoritmo

4.4.2.1. I pesi iniziali

La formazione della SOM comincia con l'inizializzazione dei vettori peso $w_i = (w_{i1}, w_{i2}, \dots, w_{im})$ dove con $i = 1, 2, \dots, m$ si intende il numero di nodi di uscita della rete. I pesi collegano i nodi di ingresso a quelli di uscita aggiornandosi periodicamente attraverso il processo di apprendimento. I pesi iniziali, $w_i(t)$ per $t = 0$, possono essere determinati arbitrariamente o sistematicamente. Il metodo di inizializzazione più comune assegna ai vettori peso dei valori casuali, preferibilmente piccoli prossimi a 0. Essendo scelti in maniera randomica, alla rete non viene imposto nessun tipo di ordine durante la fase di inizializzazione. L'unica limitazione imposta da questo tipo di approccio è

di evitare la presenza di pesi simili nel dataset. Parametri come quelli del tasso di apprendimento e della dimensione del vicinato devono essere scelti ponderatamente con lo scopo di far convergere l'algoritmo in tempi ragionevoli attraverso un certo numero di iterazioni (almeno 1000).

4.4.2.2. Similary Matching

L'apprendimento non supervisionato della rete vede competere tra loro i nodi d'uscita per essere attivati. Sarà attivato e dichiarato vincitore solo il nodo il cui vettore peso è più simile al vettore di input. Per la ricerca del BMU, vengono calcolate le distanze tra un dato di ingresso x e tutti i vettori peso w_i , utilizzando diversi metodi di misurazione, come la distanza di Manhattan, la distanza di Chebyshev, la distanza euclidea. Secondo Kohonen, tra i vari metodi disponibili, quello ideale per la rappresentazione visiva è la distanza euclidea perché con essa si ottiene una visualizzazione più isotropa del dataset (11).

La distanza euclidea tra un campione x , scelto a caso dal set di dati di input, e tutti i vettori peso al tempo t si calcola con la seguente formula:

$$d_i(t) = \|x(t) - w_i(t)\| = \sqrt{\sum_{j=1}^m (x_{tj} - w_{tji})^2} \quad i = 1, 2, \dots, n$$

Dove con $\|\cdot\|$ è indicata la norma euclidea. Il vettore dei pesi w_i di ciascun nodo di uscita ha la stessa dimensione del vettore di input x . Una volta terminato l'addestramento, il nodo c che risulterà vincitore all'iterazione t , sarà determinato utilizzando il criterio della minima distanza euclidea:

$$c(t) = \arg \min\{\|x(t) - w_i(t)\|\}$$

4.4.2.3. Aggiornamento dei pesi

Una volta trovato il BMU, i vettori peso del vincitore e quelli dei nodi vicini vengono adattati con lo scopo di diventare più rappresentativi e fedeli dello spazio di ingresso.

Per svolgere questo processo di aggiornamento, la rete deve disporre di due parametri: il *learning rate* (tasso di apprendimento) e la *neighborhood size* (dimensione del vicinato). Nelle SOM, a differenza delle altre reti convenzionali, il tasso di apprendimento $\alpha(t)$ assume valori che variano tra 0 e 1. Il ruolo del learning rate è di controllare il tasso di variazione dei pesi e la dimensione del vicinato.

Il tasso di apprendimento diminuisce gradualmente in funzione delle iterazioni t che servono ad addestrare la rete. Ciò vuol dire che, mentre la rete è in fase di addestramento, il tasso di

apprendimento diminuisce gradualmente. Infatti, nella fase iniziale, dove ai pesi vengono assegnati valori casuali, il tasso di apprendimento ha un valore alto, mentre all'aumentare delle iterazioni si riduce progressivamente a valori più piccoli.

Utilizzando un formalismo a tempo discreto, sia $w_i(t)$ il vettore peso del neurone vincente al tempo t , il vettore peso $w_i(t + 1)$ al tempo $t + 1$ è definito da:

$$w_i(t + 1) = w_i(t) + \alpha(t)[x(t) - w_i(t)]$$

Le proprietà topologiche dello spazio di partenza si conservano in quello di arrivo grazie al neighborhood size. Infatti, al processo di apprendimento partecipano anche i vettori dei nodi vicini al BMU. Il tasso di adattamento dei pesi diminuisce allontanandosi dal nodo vincente, secondo una funzione di decadimento spazio-temporale chiamata funzione di vicinato $h_{ci}(t)$, dove i indica l'indice dell'unità vicina. In letteratura, si distinguono due tipi di funzioni di vicinato. La più semplice di esse, la funzione di vicinato discreto, definisce un insieme N_c di elementi intorno al nodo vincente, dove $h_{ci}(t) = b$ se $i \in N_c$ e $h_{ci}(t) = 0$ se $i \notin N_c$. Il valore di $b(t)$ rappresenta il grado di partecipazione all'aggiornamento dei pesi. Nella Figura 4.13 vengono rappresentati diversi esempi di vicinato per topologie di output rettangolari ed esagonali. Un altro tipo di funzione ampiamente utilizzata per l'aggiornamento del neighborhood size è la funzione di vicinato continuo, che porta a un output più omogeneo. La funzione scelta è di tipo gaussiano, che diminuisce sia nel dominio dello spazio che in quello del tempo.

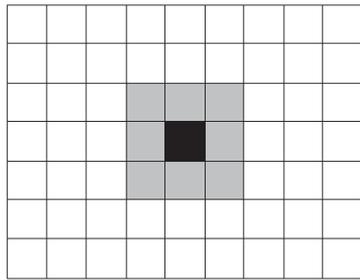
$$h_{ci}(t) = \exp\left(-\frac{d_{ci}^2}{2\sigma^2(t)}\right)$$

d_{ci}^2 rappresenta la distanza laterale tra il neurone vincente c e il neurone eccitato i , mentre $\sigma(t)$ è il raggio del vicinato all'iterazione t . La funzione gaussiana è simmetrica intorno al punto massimo definito da $d_{ci} = 0$ e diminuisce in modo lineare fino a zero all'aumentare della distanza laterale. Qui, $\sigma(t)$ è una funzione monotonamente decrescente e fa sì che il valore del neighborhood size diminuisca nel tempo. È stato osservato sperimentalmente che la scelta di un valore iniziale troppo piccolo per la dimensione del vicinato impedisce la convergenza ottimale dell'algoritmo disponendo in modo non appropriato i nodi all'uscita. Kohonen (11) suggerisce di iniziare con un valore di $\sigma(0)$ abbastanza grande.

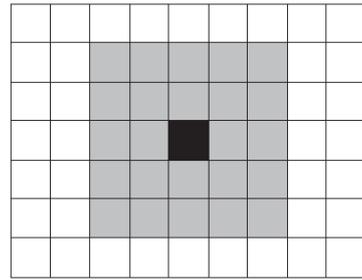
Infine, dopo aver definito l'insieme dei vicini o il kernel attorno al nodo vincente c , si sceglie una funzione di vicinato per aggiornare i vettori peso dei rispettivi nodi. La seguente formula esprime l'aggiornamento del vettore peso $w_i(t + 1)$ con tutti i parametri della rete.

$$w_i(t + 1) = w_i(t) + \alpha(t)h_{ci}(t)[x(t) - w_i(t)]$$

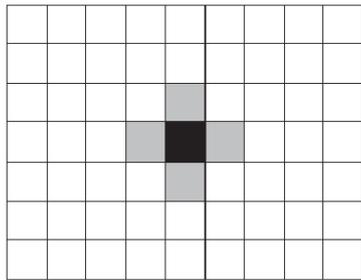
I tre processi di competizione, cooperazione e adattamento sono ripetuti per tutti i dati di addestramento finché i pesi non convergono e non si osservano dei cambiamenti evidenti nello strato di uscita a bassa dimensione.



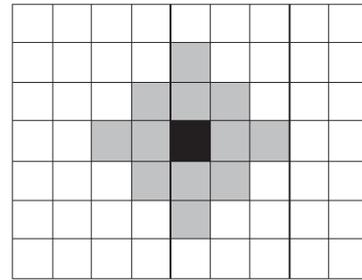
Rectangular SOM grid (size 1)



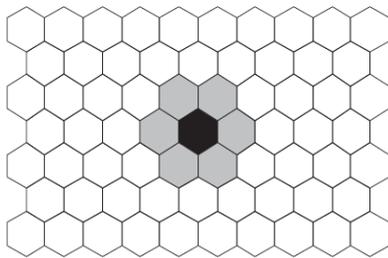
Rectangular SOM grid (size 2)



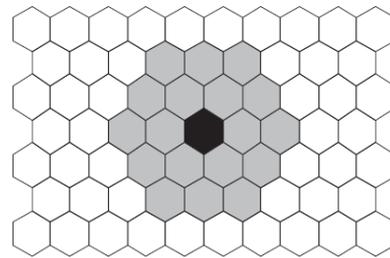
Rectangular SOM grid (size 1)



Rectangular SOM grid (size 2)



Hexagonal SOM grid (size 1)



Hexagonal SOM grid (size 2)

Figura 4.13: Tipi di vicinato in funzione della topologia di griglia

4.4.3. Qualità della SOM

La rete SOM è in grado di produrre in uscita una rappresentazione a bassa dimensione dei dati d'ingresso ad alta dimensione, senza perdita di informazione. I risultati della SOM variano in base ai dati iniziali dei parametri come il numero di nodi di uscita, il tasso di apprendimento e la dimensione del vicinato. Risulta quindi necessario esaminare la qualità dei risultati per determinare i valori dei parametri che forniscono la corrispondenza più appropriata dell'insieme di partenza.

Tuttavia, non esistono regole fisse per determinare i parametri a priori. In letteratura, per verificare la qualità della rete, sono stati sviluppati diversi metodi di misurazione. Principalmente della SOM vengono studiate due proprietà: la qualità dell'apprendimento e la qualità della proiezione. La prima è indicata tramite il *vector quantization* (quantizzazione vettoriale) e la seconda tramite il *topology preservation* (conservazione topologica). Esiste una relazione tra queste due proprietà: la qualità della conservazione della topologia diminuisce con l'aumentare dell'accuratezza della rappresentazione dei dati.

4.4.3.1. Quantization Error

L'errore di quantizzazione (*Quantization Error, QE*) è uno strumento che serve a valutare la qualità dell'apprendimento della rete e mostra quanto la mappa si adatti ai dati di partenza. Viene calcolato determinando la distanza media dei vettori campione verso il BMU (11). La formula che esprime l'errore medio di quantizzazione è la seguente:

$$QE = \frac{\sum_{t=1}^T \|x(t) - w_c(t)\|}{T}$$

dove $w_c(t)$ è il vettore di riferimento associato al BMU di $x(t)$.

Aumentando il numero di nodi di uscita, il QE può essere ridotto visto che i campioni vengono distribuiti in maniera più sparsa sulla mappa (12). In questo caso, la misura dell'errore di quantizzazione ignora completamente le relazioni topologiche e, a causa del compromesso tra quantizzazione vettoriale e qualità della proiezione, l'eccessiva riduzione della QE può portare a delle distorsioni topologiche della mappa. Inoltre, Sun (13) suggerisce di assegnare un valore piccolo per il neighborhood size con lo scopo di ottenere un piccolo errore di quantizzazione e una buona approssimazione della distribuzione degli input.

4.4.3.2. Topographic Error

La conservazione topologica (*Topology Preservation*), che rappresenta la qualità della proiezione, è una proprietà più complessa da misurare rispetto alla valutazione dell'apprendimento della rete. Viene sfruttato l'errore topografico (*Topographic Error, TE*) che definisce la percentuale di vettori per i quali la prima e la seconda BMU non sono unità adiacenti. L'errore topografico è calcolato come:

$$TE = \frac{1}{T} \sum_{t=1}^T u(x(t))$$

dove T indica il numero di campioni in ingresso e $u(x(t))$ è uguale ad 1 se il primo e il secondo BMU di $x(t)$ non sono adiacenti, altrimenti $u(x(t))$ è 0. Concludendo, minore è la violazione della relazione di vicinato sulla griglia, migliore è la conservazione topologica da parte della SOM.

5. Caso di studio

5.1. Introduzione

Il caso di studio prende in esame l'impianto descritto nel capitolo 2. L'obiettivo del seguente elaborato di tesi è quello di sviluppare un algoritmo in grado di identificare e mappare su una griglia di output i diversi dati presenti nel dataset di partenza. Per la raccolta dei dati, come descritto nella sezione 3, sono stati utilizzati dei sensori opportunamente collegati all'hub centrale Revolution Pi. Una volta unificate e standardizzate le diverse letture della prova a regime e delle prove anomale sono state date in pasto alla rete SOM che le ha mappate in una griglia di uscita bidimensionale. La rete ha classificato le anomalie in quattro macroaree sulla griglia: anomalia dell'acqua, anomalia dell'aria, anomalia del livello del serbatoio e anomalia del fluido bifase aria-acqua.

Le varie anomalie sono state simulate tramite l'utilizzo di alcune valvole d'intercezione manuale poste in diverse parti dell'impianto con l'intento di occludere il passaggio del fluido nelle tubazioni.

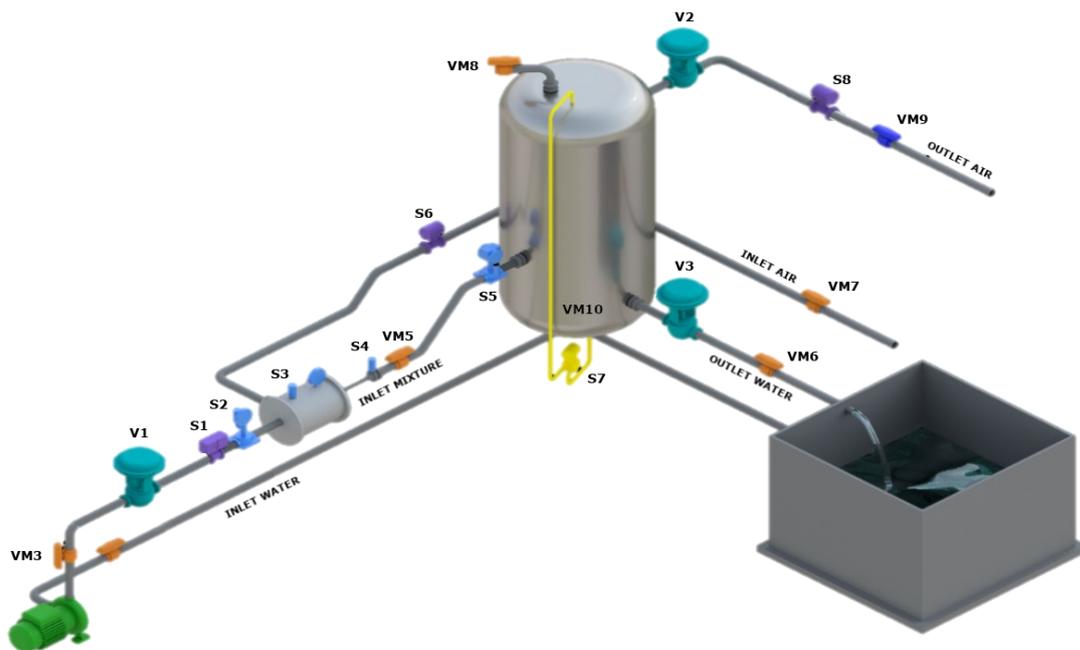


Figura 5.1: Impianto sperimentale presente in laboratorio

In Figura 5.1 è riportato lo schema dell'impianto. Le valvole utilizzate per la generazione delle anomalie sono: VM3, VM5, VM6, VM7, VM8, VM9, VM10 per un totale di 7 anomalie.

5.2. Parametri di esercizio

L'impianto oggetto d'esame attende che da parte dell'utente vengano impostati diversi parametri di esercizio come:

- **pi** : pressione di esercizio del fluido motore rilevata dal sensore di pressione S2;
- **ls** : livello del serbatoio rilevato dal sensore S7;
- **ps** : sovra-pressione all'interno del serbatoio S5;
- **Kpp, Kip, Kdp** : parametri per il PID di controllo della portata dell'acqua in ingresso;
- **Kpl, Kil, Kdl** : parametri per il PID di controllo del livello del serbatoio;
- **Kps, Kis, Kds** : parametri per il PID di controllo della pressione del serbatoio.

Tramite l'utilizzo dell'elettrovalvola V1 si mantiene costante la portata d'acqua in ingresso all'impianto, con l'elettrovalvola V2 si conserva la pressione all'interno del serbatoio, mentre con la V3 si controlla il livello ls della miscela bifase all'interno del serbatoio. I parametri che caratterizzano i vari PID sono la costante proporzionale, integrativa e derivativa. Sono state scelte in maniera tale da controllare accuratamente la chiusura e l'apertura delle elettrovalvole presenti nel laboratorio. Nella Tabella 5.1 sono espressi i parametri di esercizio per la raccolta dei dati a regime.

Tabella 5.1: Parametri di esercizio del caso di studio

Descrizione	Variabile	UM
Pressione fluido motore	pi	5.5 bar
Livello serbatoio	ls	300 mm
Pressione serbatoio	ps	1.3 bar
PID ingresso acqua	Kpp	0.8
PID ingresso acqua	Kip	0.4
PID ingresso acqua	Kdp	0
PID livello serbatoio	Kpl	1.7
PID livello serbatoio	Kil	0.7
PID livello serbatoio	Kdl	0.1
PID pressione serbatoio	Kps	1
PID pressione serbatoio	Kis	0.7
PID pressione serbatoio	Kds	0

5.3. Raccolta dati

5.3.1. Tipologie di dati raccolti

I dati presenti nel dataset fornito alla rete sono stati registrati sotto il controllo delle elettrovalvole V1, V2, V3 e tramite l'ausilio dei sensori descritti nel capitolo 2. Nella Tabella 5.2 vengono riportate le diverse variabili raccolte durante il processo di campionamento. Grazie alla tecnologia del Revolution Pi, è stato possibile raccogliere informazioni a 16 bit ogni 0,1 secondi di tempo.

Tabella 5.2: Variabili campionate

ID	Descrizione	Zona	UM
S1	Pressione Assoluta	Ingresso Acqua	[bar]
S2	Portata Volumetrica	Ingresso Acqua	[m ³ /h]
S3	Pressione Assoluta	Eiettore	[bar]
S4	Pressione Assoluta	Tubo Miscelatore	[bar]
S5	Pressione Relativa	Ingresso Serbatoio	[bar]
S6	Portata Volumetrica	Ingresso Aria	[m ³ /h]
S7	Livello Acqua	Serbatoio	[mm]
S8	Portata Volumetrica	Uscita Aria	[m ³ /h]
V1	Elettrovalvola Pneumatica	Out Pompa	[%]
V2	Elettrovalvola Pneumatica	Out Aria	[%]
V3	Elettrovalvola Pneumatica	Out Acqua	[%]

5.3.2. Dati a regime

Sperimentalmente si è osservato che dopo circa cento secondi dall'accensione dell'impianto, dai dati registrati dai sensori, il sistema si porta ad una condizione di regime stazionario. In queste condizioni, campionando a 0.1 s, sono stati registrati un totale di 2748 campioni.

Tabella 5.3: Letture a regime

Sensore	Media	Min	Max	Dev. Std σ
S1	5.4983	5.437	5.563	0.034429
S2	9.9575	9.894	10.024	0.035907
S3	0.9708	0.97	0.971	0.000364
S4	1.3425	1.323	1.363	0.007442
S5	1.2999	1.286	1.318	0.005792
S6	26.4049	25.556	27.106	0.311141
S7	299.9596	297.294	302.625	1.204324
S8	21.5479	20.713	22.442	0.416975

5.3.3. Dati anomali

Per registrare i dati riguardanti le prove anomale sono state sfruttate le valvole d'intercettazione manuale contrassegnate con VM3, VM5, VM6, VM7, VM8, VM9, VM10. Come mostrato nella Figura 5.2, ciascuna valvola ha il ruolo di portare il sistema in instabilità:

- **VM3** ostruisce il flusso dell'acqua in ingresso;
- **VM5** ostacola il flusso del fluido bifase aria-acqua dall'uscita dell'eiettore all'ingresso del serbatoio;
- **VM6** ostruisce il flusso d'acqua dal serbatoio alla vasca di contenimento;
- **VM7** blocca l'ingresso dell'aria nell'impianto;
- **VM8** simula una perdita d'aria all'interno del serbatoio;
- **VM9** ostacola la fuoriuscita dell'aria dal serbatoio all'atmosfera;
- **VM10** simula una perdita d'acqua del serbatoio.

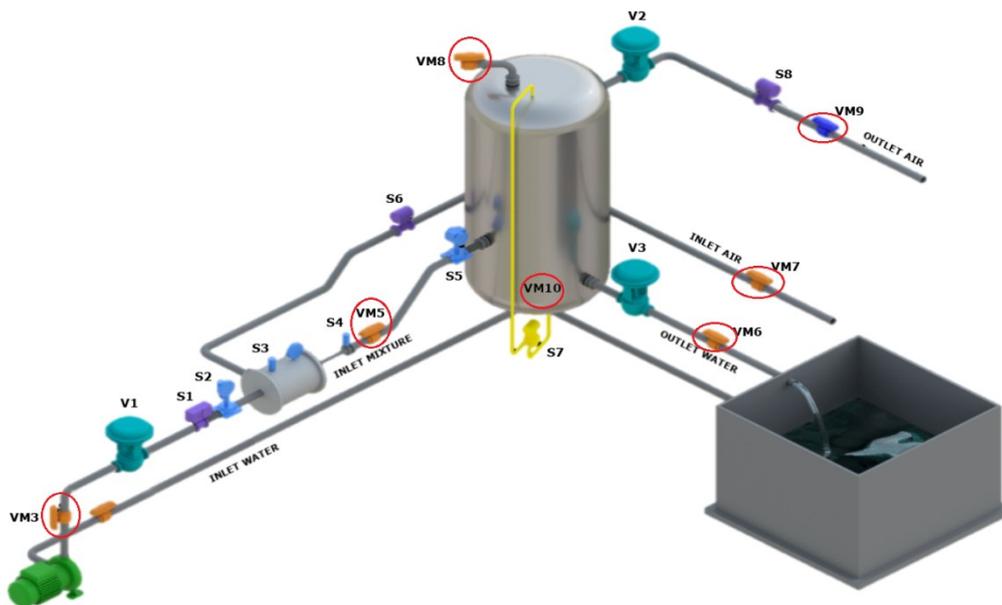
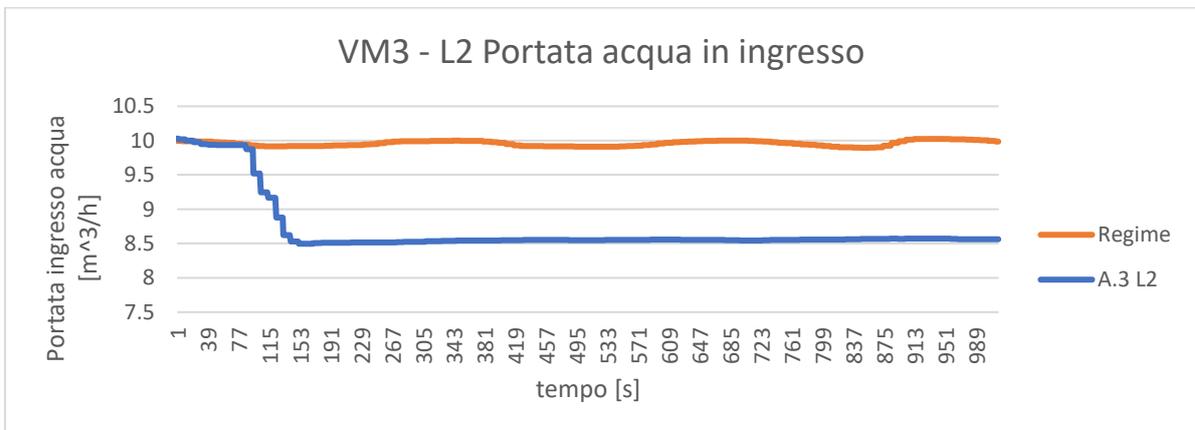
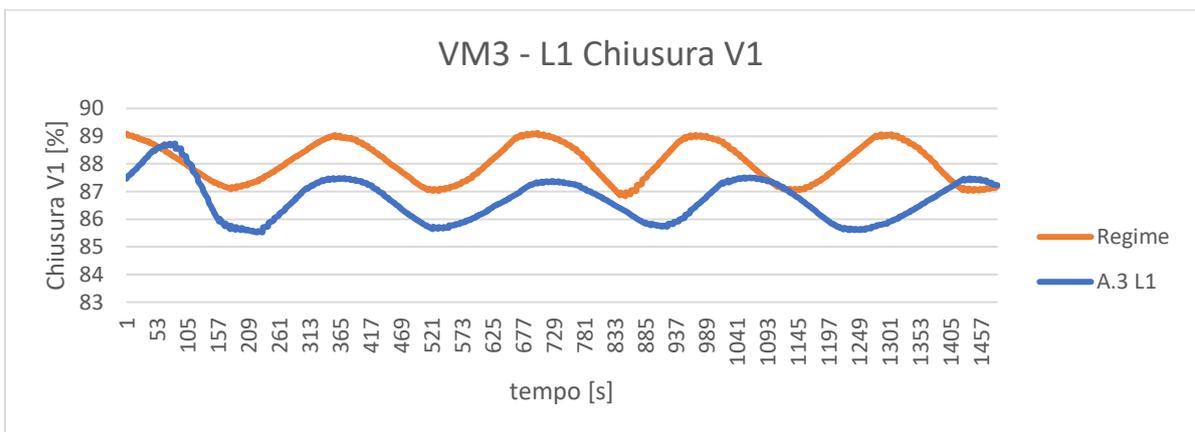
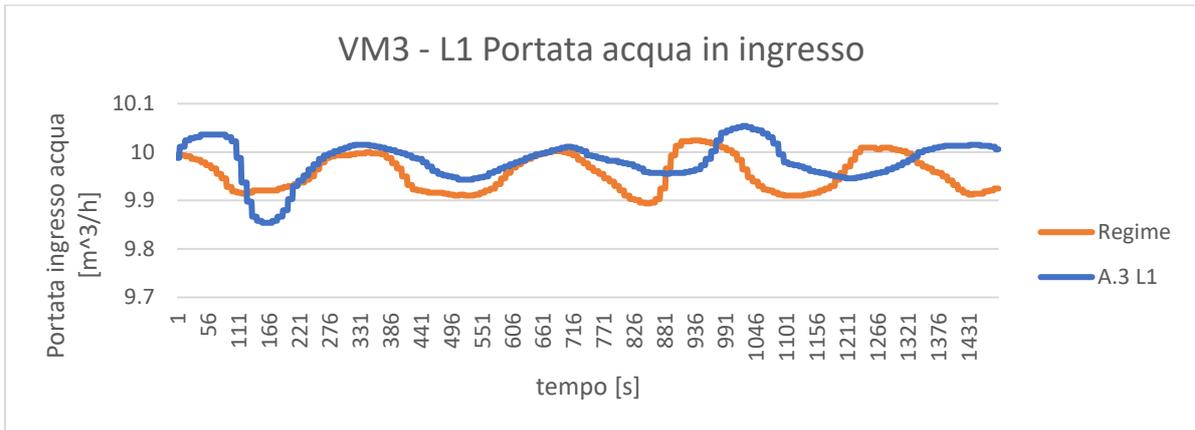


Figura 5.2: Valvole d'intercettazione nell'impianto sperimentale

Chiudendo sempre di più le teste delle varie valvole d'intercettazione, per ogni tipo di anomalia sono stati registrati tre diversi livelli di occlusione: piccola, media, grande. Una volta raggiunto lo stato di regime da parte dell'impianto, viene chiusa la valvola e viene fatta partire la registrazione dei dati dal terminale. La rete neurale SOM è stata addestrata utilizzando i livelli d'intensità maggiore in modo tale da mappare in maniera più efficiente le diverse anomalie sulla griglia di output.

5.3.3.1. Anomalia 3

L'acqua in ingresso all'impianto viene ostruita dalla valvola d'intercettazione VM3 che porta il sistema in una nuova condizione di equilibrio. Nella Figura 5.3 vengono raffigurati i diversi gradi di intercettazione del flusso in ingresso.



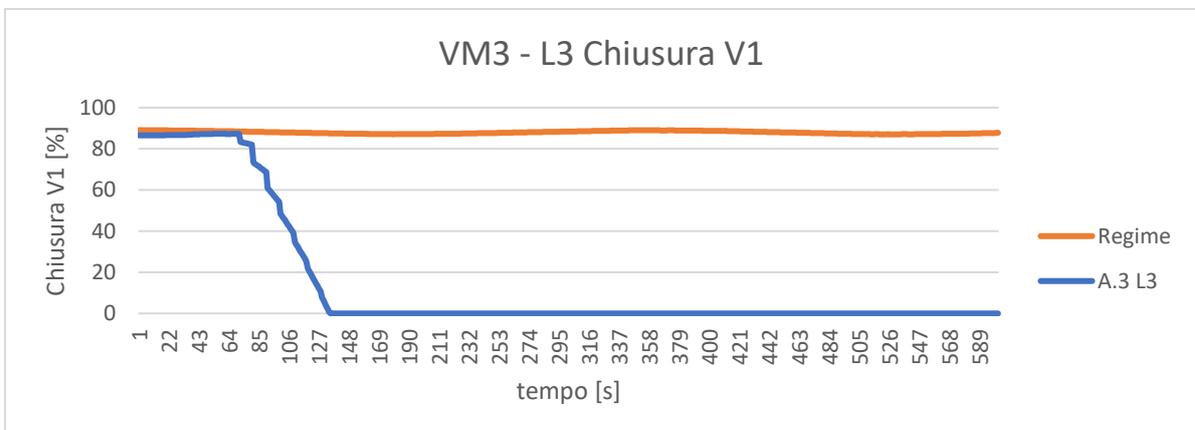
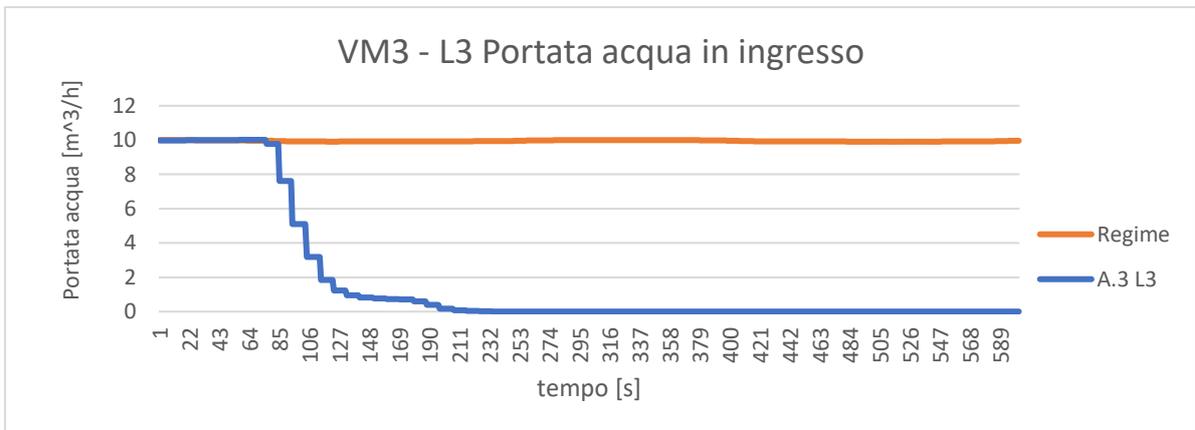
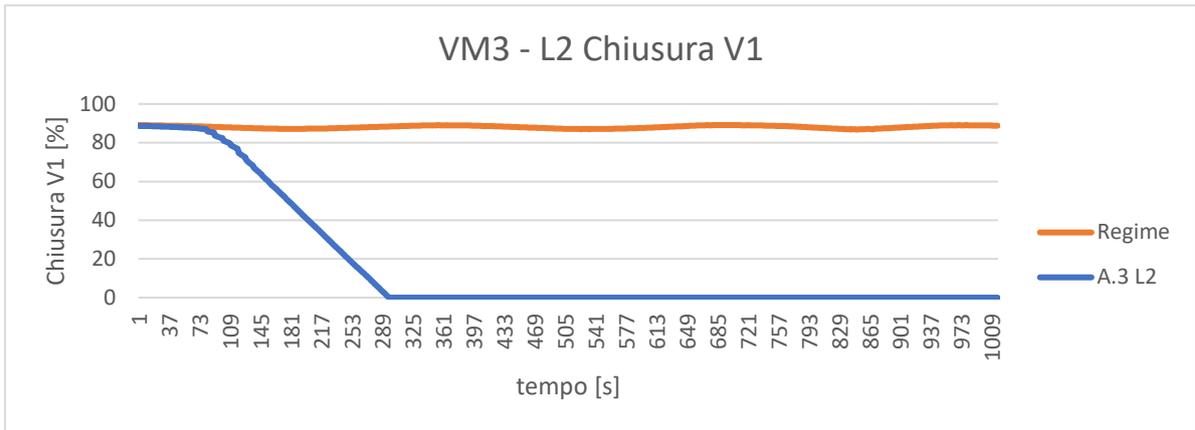
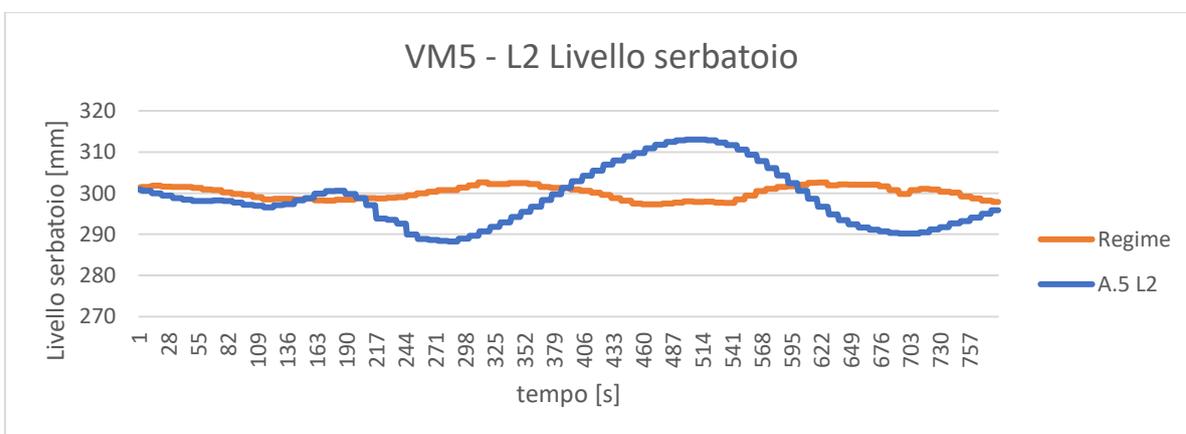
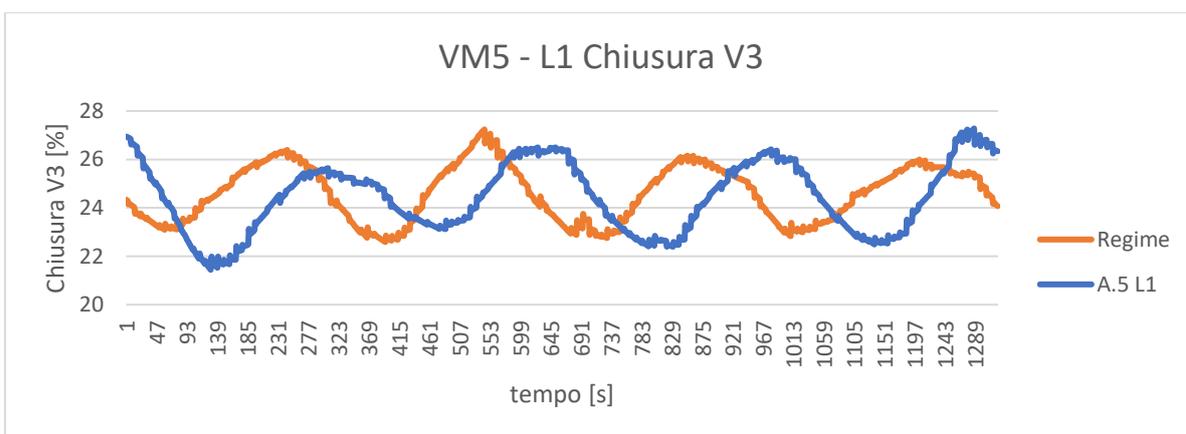
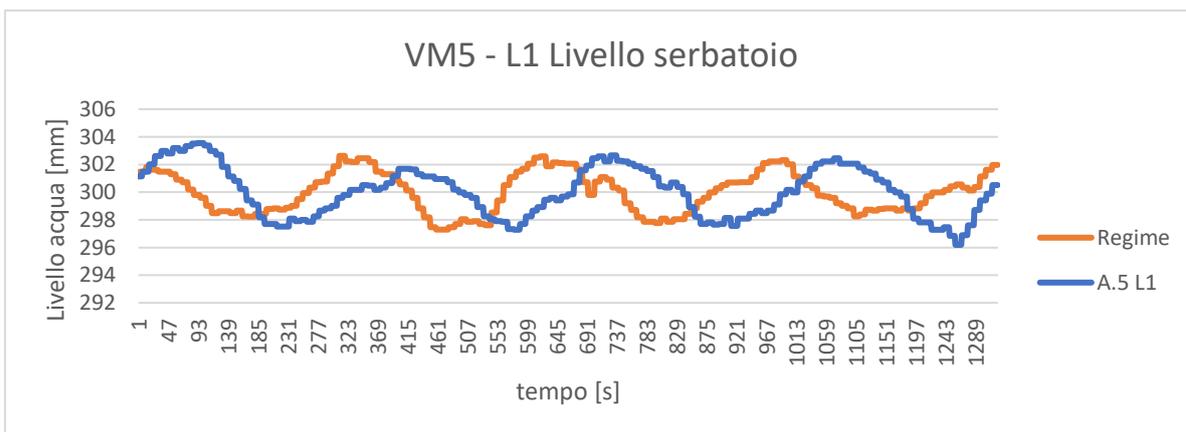


Figura 5.3: Anomalia 3 – Portata acqua in ingresso e percentuale di chiusura V1

Nei primi due casi, la valvola non ostruisce completamente il passaggio dell'acqua e tramite il lavoro svolto dal PID, il sistema riesce a riportarsi in condizioni simili a quelle di equilibrio stazionario. Quando invece la valvola dell'acqua viene chiusa completamente, come avviene nel caso del livello 3, l'elettrovalvola V3 si chiude cercando di far tornare il sistema nello stato di partenza, ma l'impianto va in avaria. Proprio per questa ragione, per non danneggiare il sistema è stato deciso di non andare oltre i 50 secondi di campionamento.

5.3.3.2. Anomalia 5

La valvola VM5 è interposta tra l'uscita del tubo miscelatore e l'ingresso del serbatoio di contenimento. È stata utilizzata per simulare un ingorgo all'ingresso del serbatoio. Di seguito sono riportati gli andamenti dei livelli del serbatoio con le rispettive percentuali di chiusura dell'elettrovalvola registrate per le diverse prove anomale.



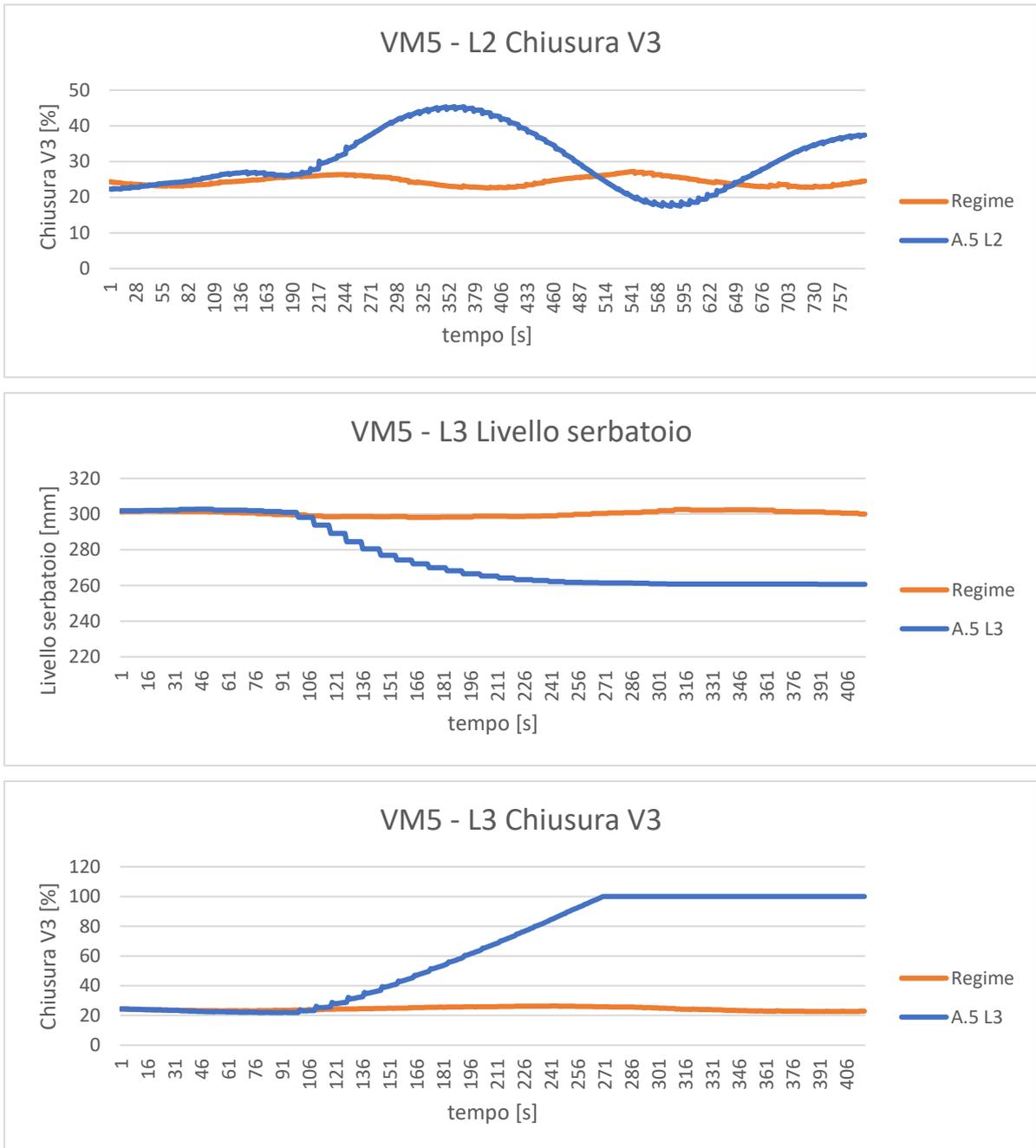
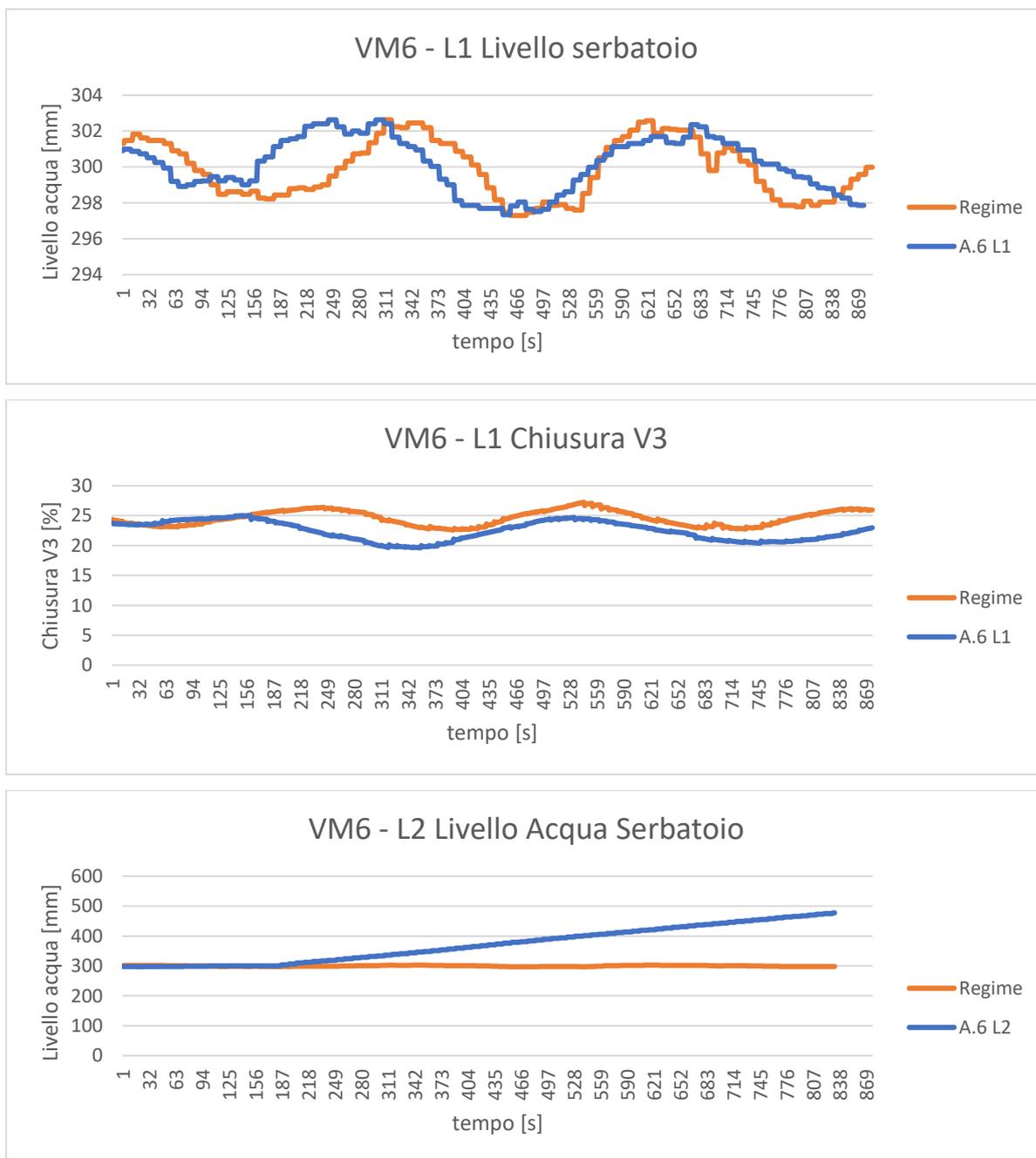


Figura 5.4: Anomalia 5 – Acqua nel serbatoio e percentuale di chiusura V3

Nei primi due casi, il sistema si riporta nelle condizioni di regime tramite l'azionamento dell'elettrovalvola che regola il livello dell'acqua all'interno del serbatoio. Nell'ultimo caso invece, essendo stato occluso quasi totalmente il passaggio del fluido, il PID che regola l'elettrovalvola V3 si chiude completamente così da rendere stabile il livello dell'acqua presente all'interno del serbatoio, mentre l'impianto va in avaria a causa del flusso retrogrado da parte dell'acqua che, ritornando nell'eiettore, allaga le condotte di aspirazione dell'aria in ingresso.

5.3.3.3. Anomalia 6

L'anomalia generata dalla valvola d'intercettazione VM6 consiste nell'ostruzione del flusso dell'acqua dall'uscita del serbatoio all'ingresso nella vasca. Nella Figura 5.5 vengono riportati gli andamenti del livello dell'acqua misurato dal sensore S7 e lo stato dell'elettrovalvola V3 durante le diverse prove.



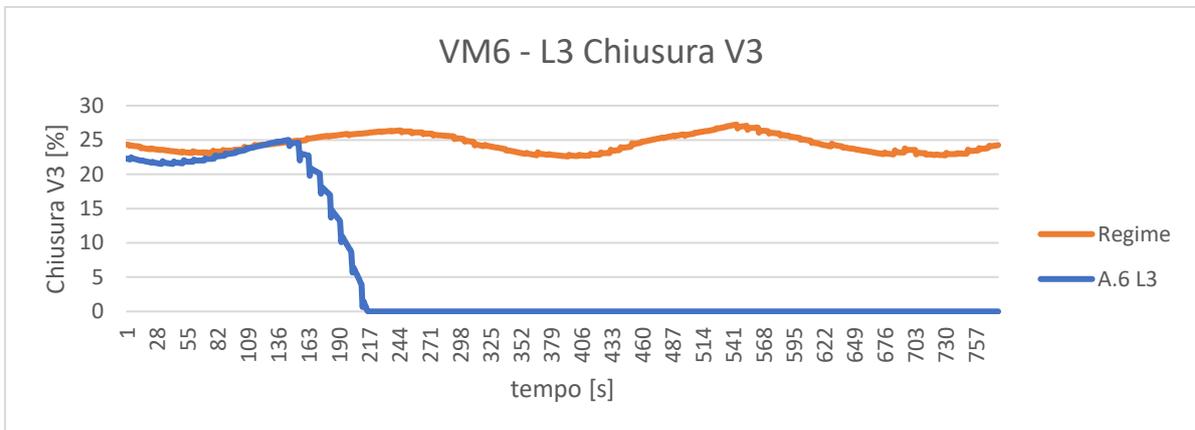
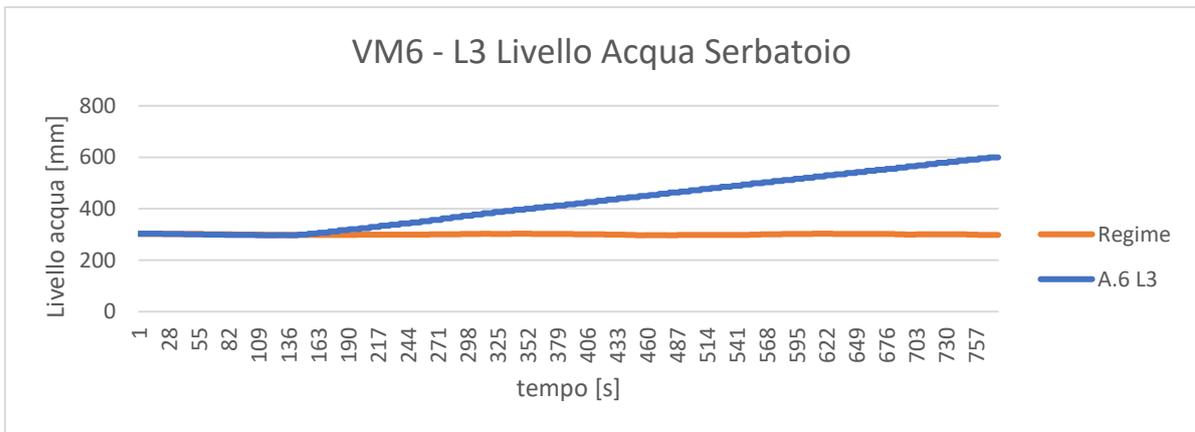
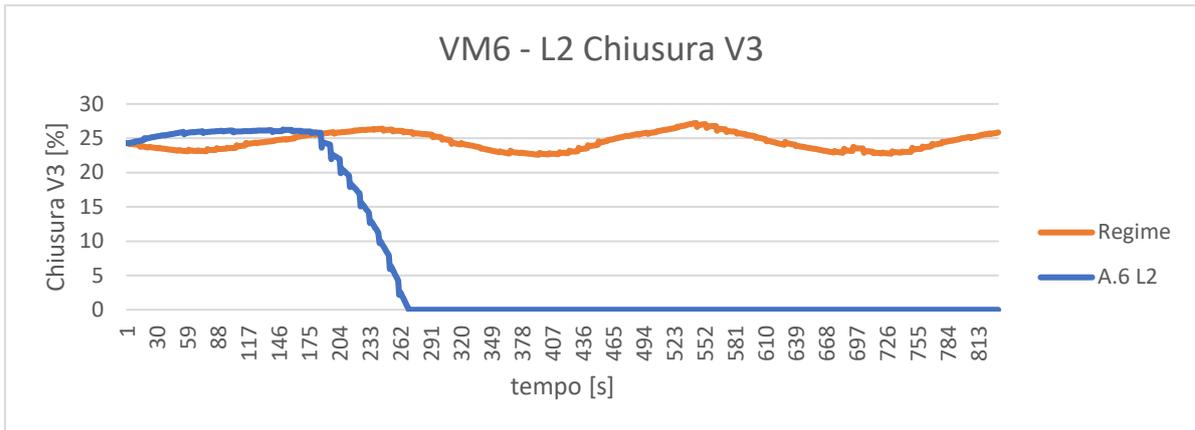
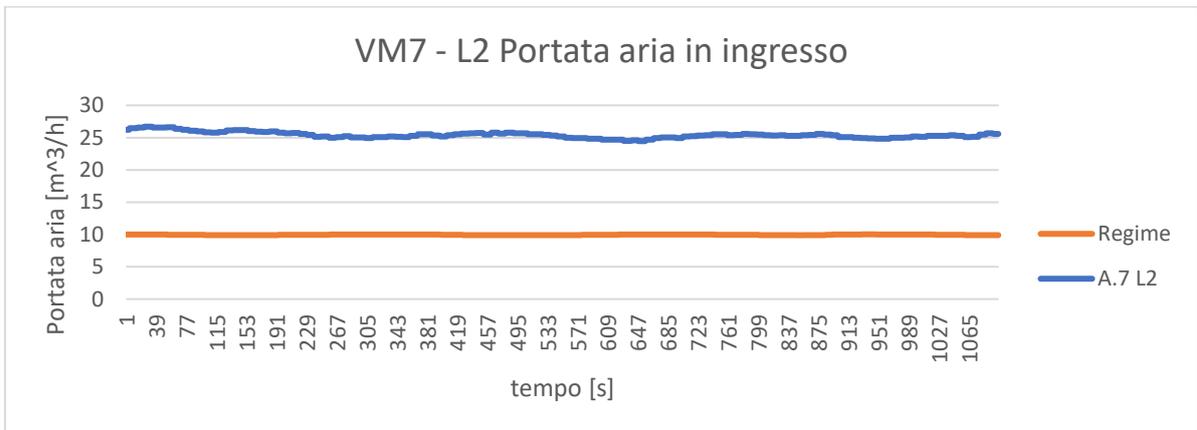
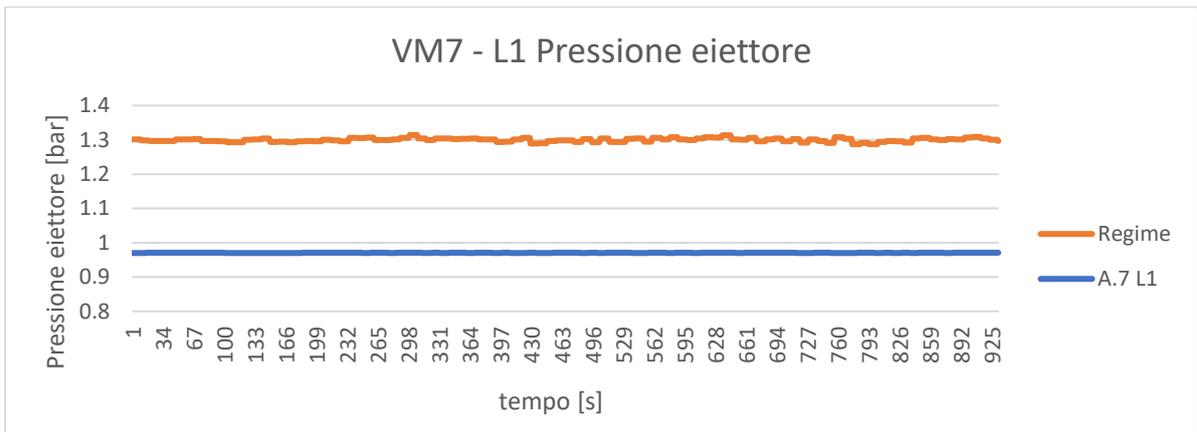
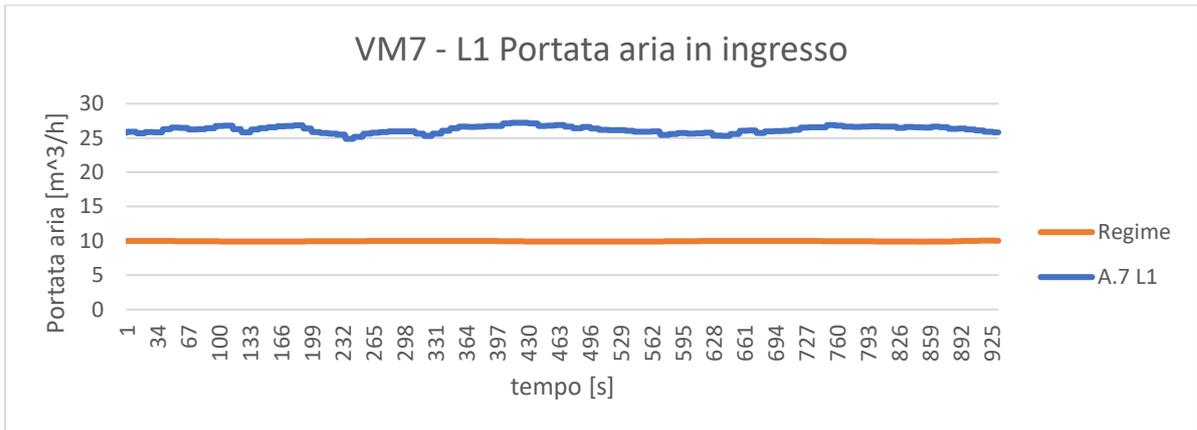


Figura 5.5: Anomalia 6 – Acqua nel serbatoio e percentuale di chiusura V3

Nel primo caso, grazie all'azione dell'elettrovalvola che controlla il livello dell'acqua nel serbatoio, il sistema non risente dell'ostruzione e si riporta nelle condizioni di regime. Nel secondo e terzo caso, l'ostruzione è più marcata e nonostante l'elettrovalvola sia aperta completamente, il livello dell'acqua nel serbatoio continua a crescere. Per non danneggiare l'impianto si è scelto di non superare i 600mm di acqua nel serbatoio.

5.3.3.4. Anomalia 7

La valvola d'intercettazione manuale VM7 presente sull'impianto ha lo scopo di ostruire il condotto di aspirazione dell'aria in ingresso. Di seguito vengono riportati gli andamenti della portata d'aria in ingresso e della pressione all'interno dell'eiettore per ogni livello di anomalia.



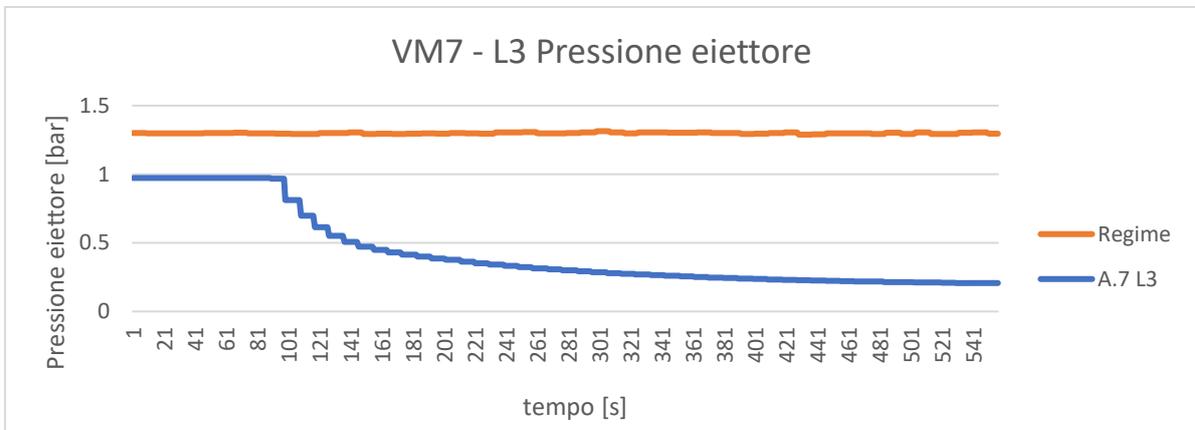
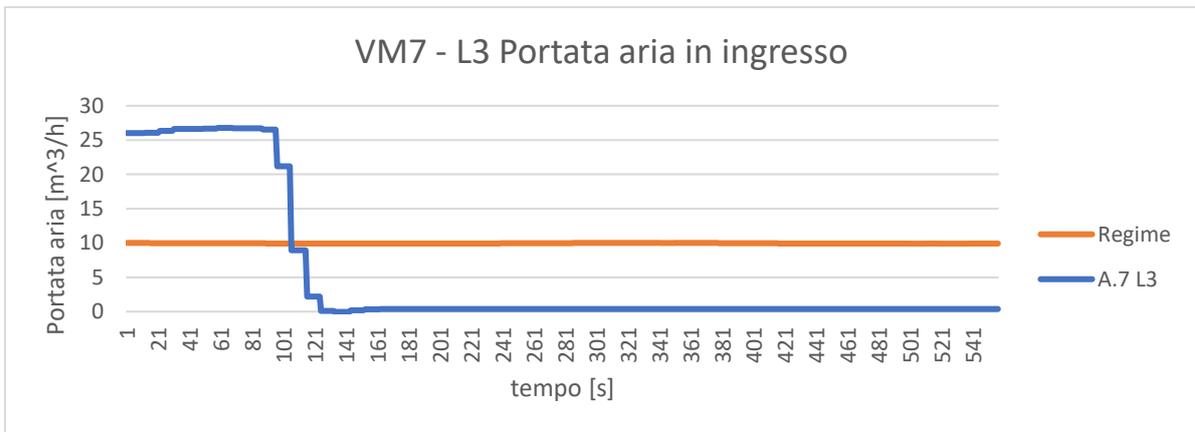
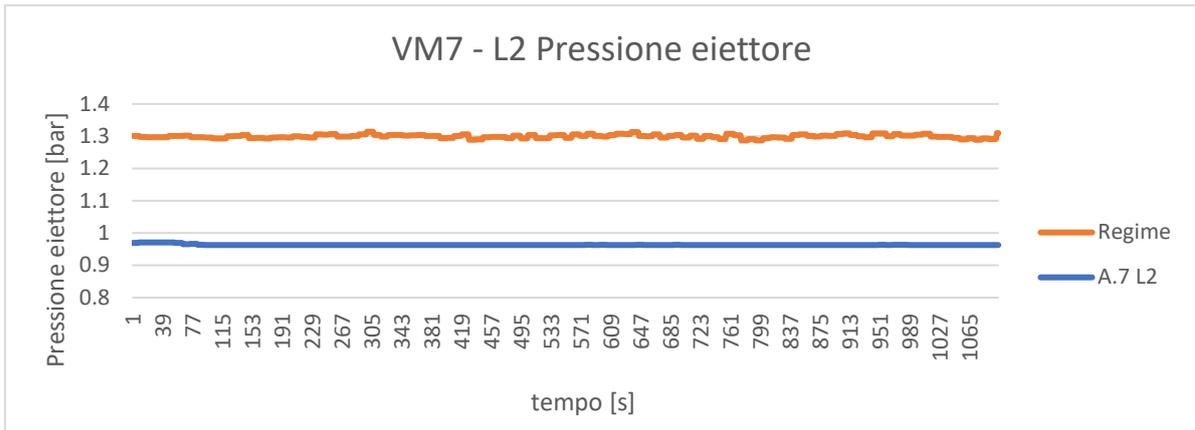
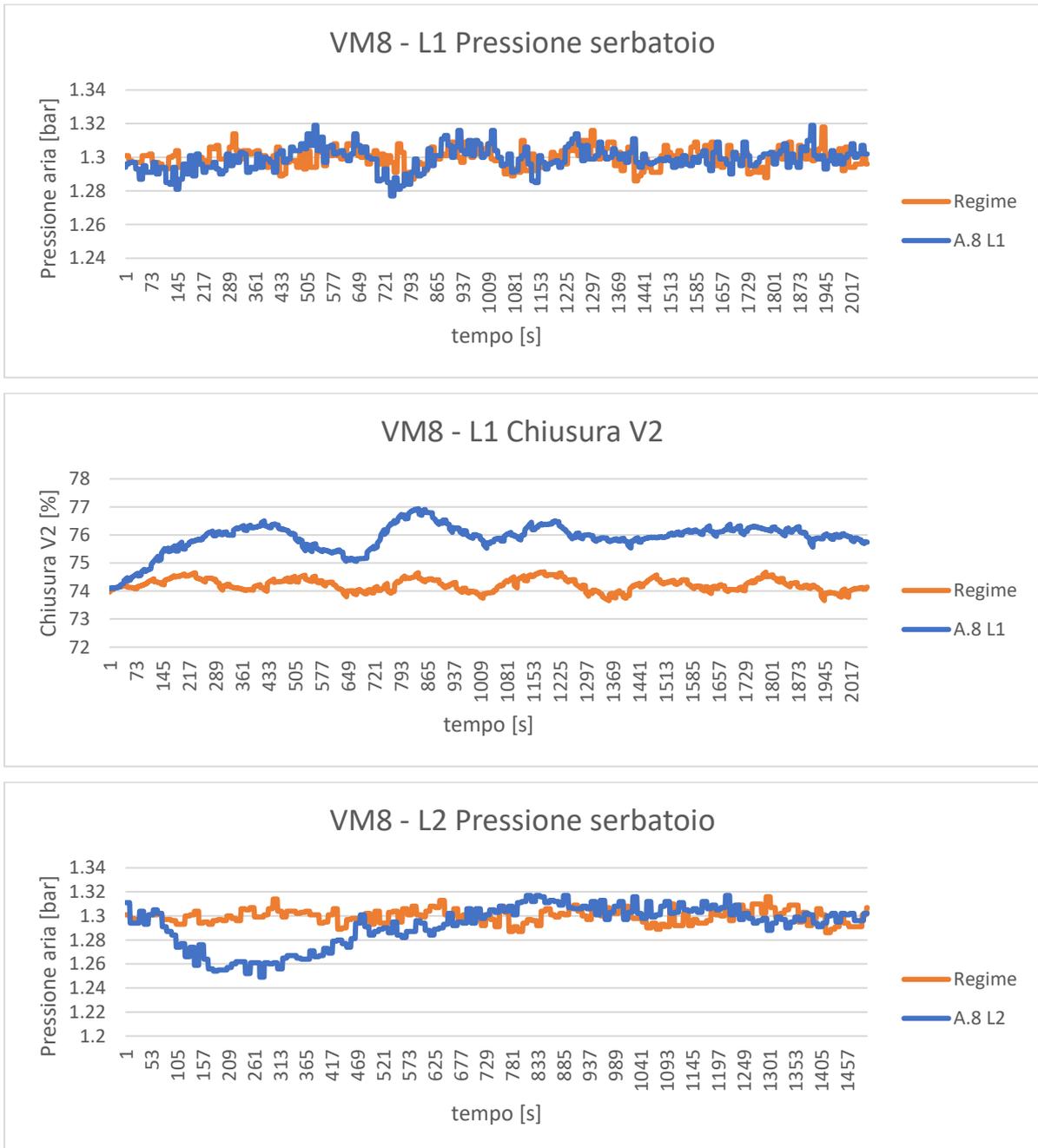


Figura 5.6: Anomalia 7 – Portata aria in ingresso e andamento della pressione nell'eiettore

Per i primi due livelli di occlusione, grazie al lavoro svolto dai PID presenti nel sistema, l'impianto si riporta nello stato di regime stazionario. Invece, quando la valvola VM7 è chiusa completamente, all'interno della camera di miscelazione dell'eiettore non si crea depressione e, non essendoci aria, in ingresso al serbatoio arriva solo acqua.

5.3.3.5. Anomalia 8

La valvola d'intercettazione manuale VM8 ha lo scopo di simulare una perdita d'aria nel serbatoio di contenimento. Nella Figura 5.7 vengono raffigurati gli andamenti della pressione all'interno del serbatoio e l'andamento dell'elettrovalvola dell'aria V2.



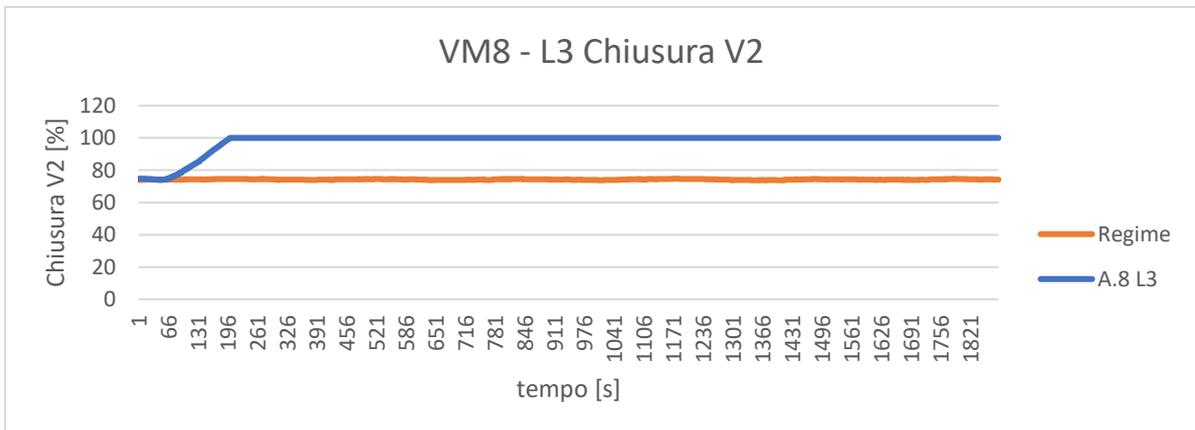
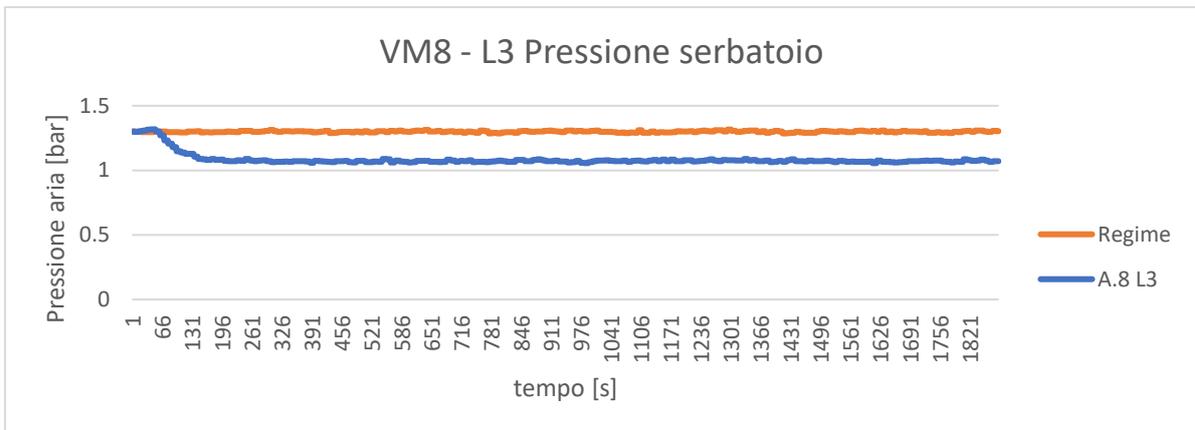
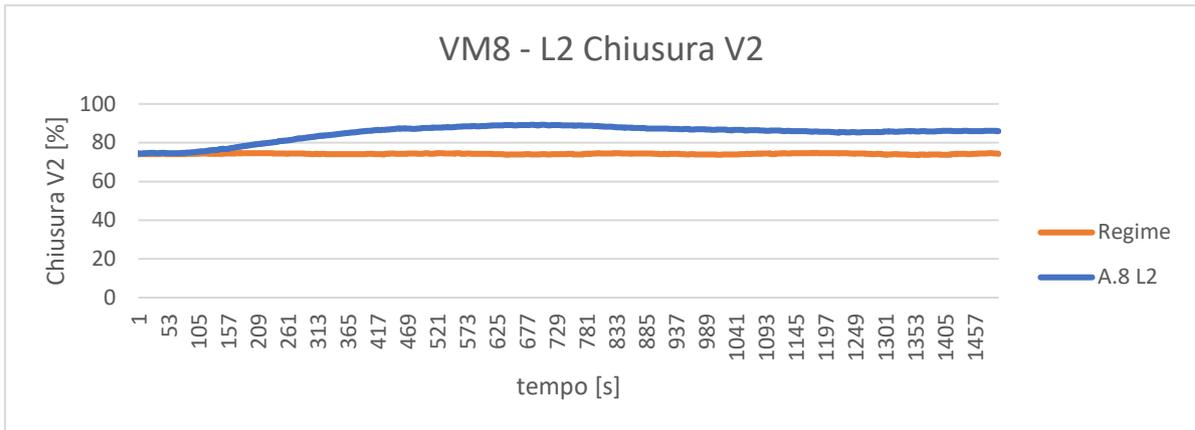
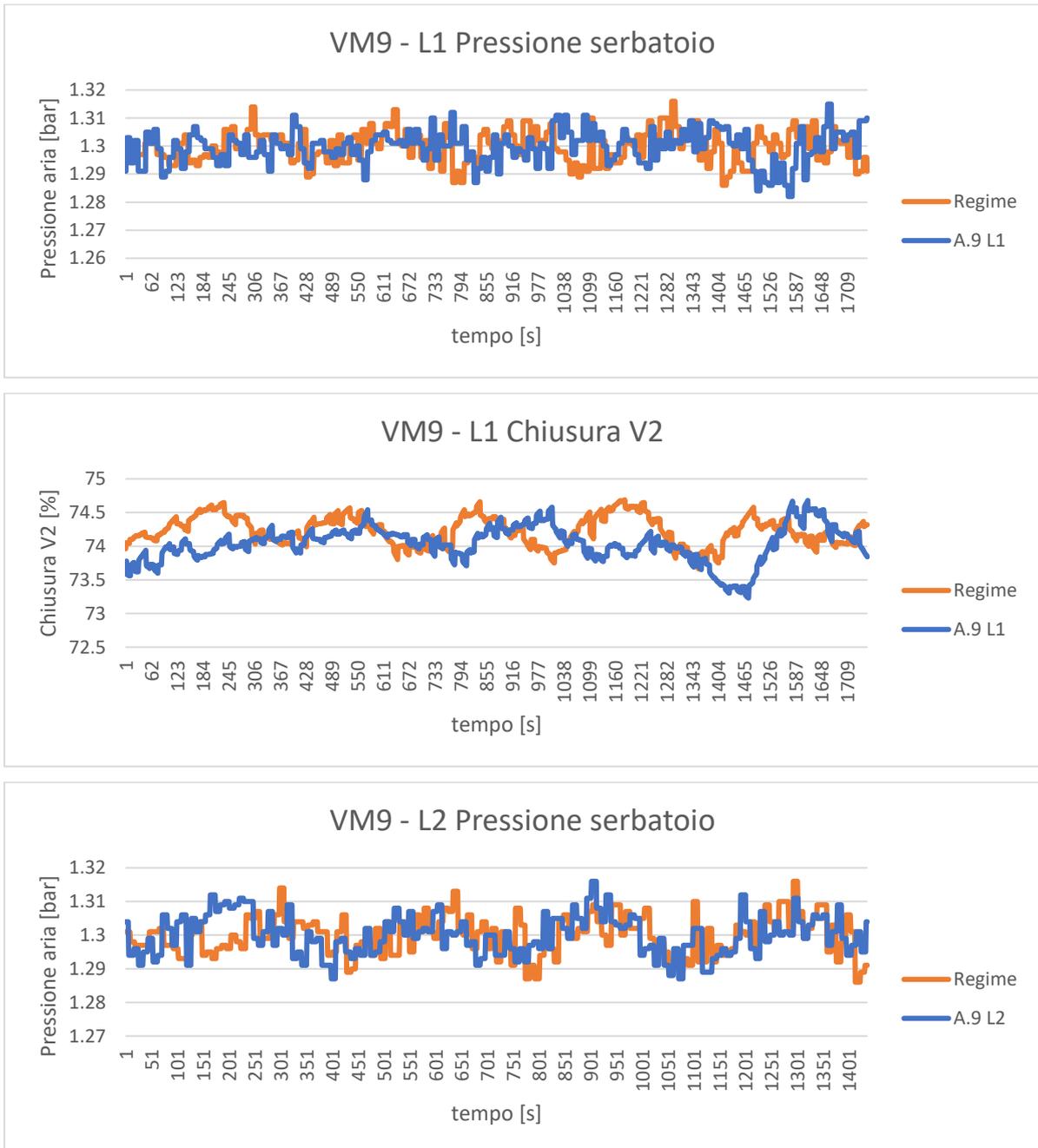


Figura 5.7: Anomalia 8 – Pressione del serbatoio e percentuale di chiusura V2

Per i primi due livelli di anomalia, grazie al lavoro svolto dal PID dell'elettrovalvola V2, l'impianto si riporta nello stato di regime stazionario con una pressione all'interno del serbatoio di 1,3 bar. Quando la valvola VM8 è aperta completamente, nonostante l'elettrovalvola V2 sia chiusa al 100% chiude e all'interno del serbatoio, il sistema si porta in una nuova condizione di equilibrio dove all'interno del serbatoio la pressione è di 1,1 bar.

5.3.3.6. Anomalia 9

L'anomalia generata dalla valvola d'intercettazione VM9 consiste nell'ostruzione del condotto di uscita dell'aria del serbatoio. Nella Figura 5.8 vengono riportati gli andamenti della pressione all'interno del serbatoio e dell'elettrovalvola V2 durante le diverse prove.



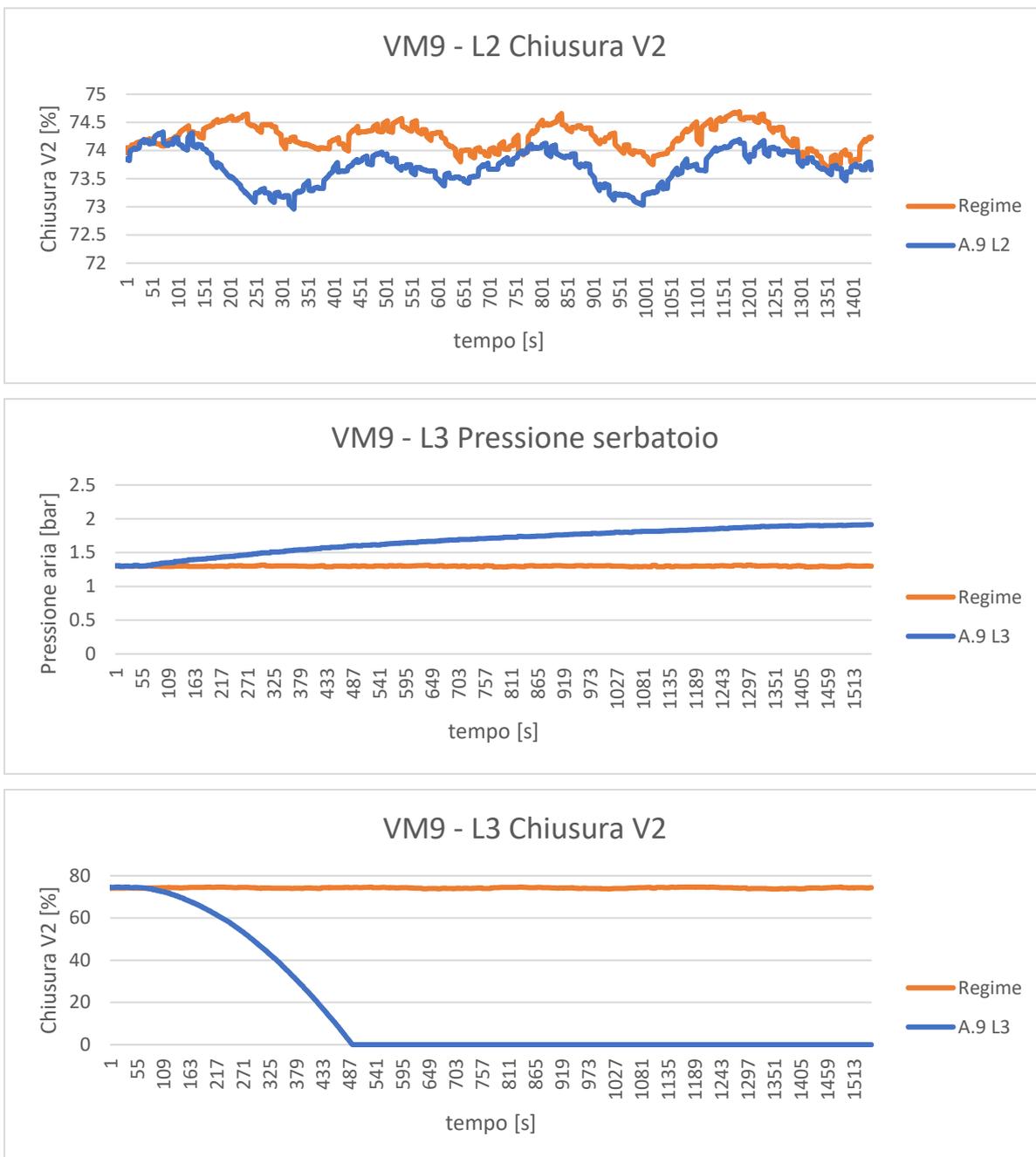
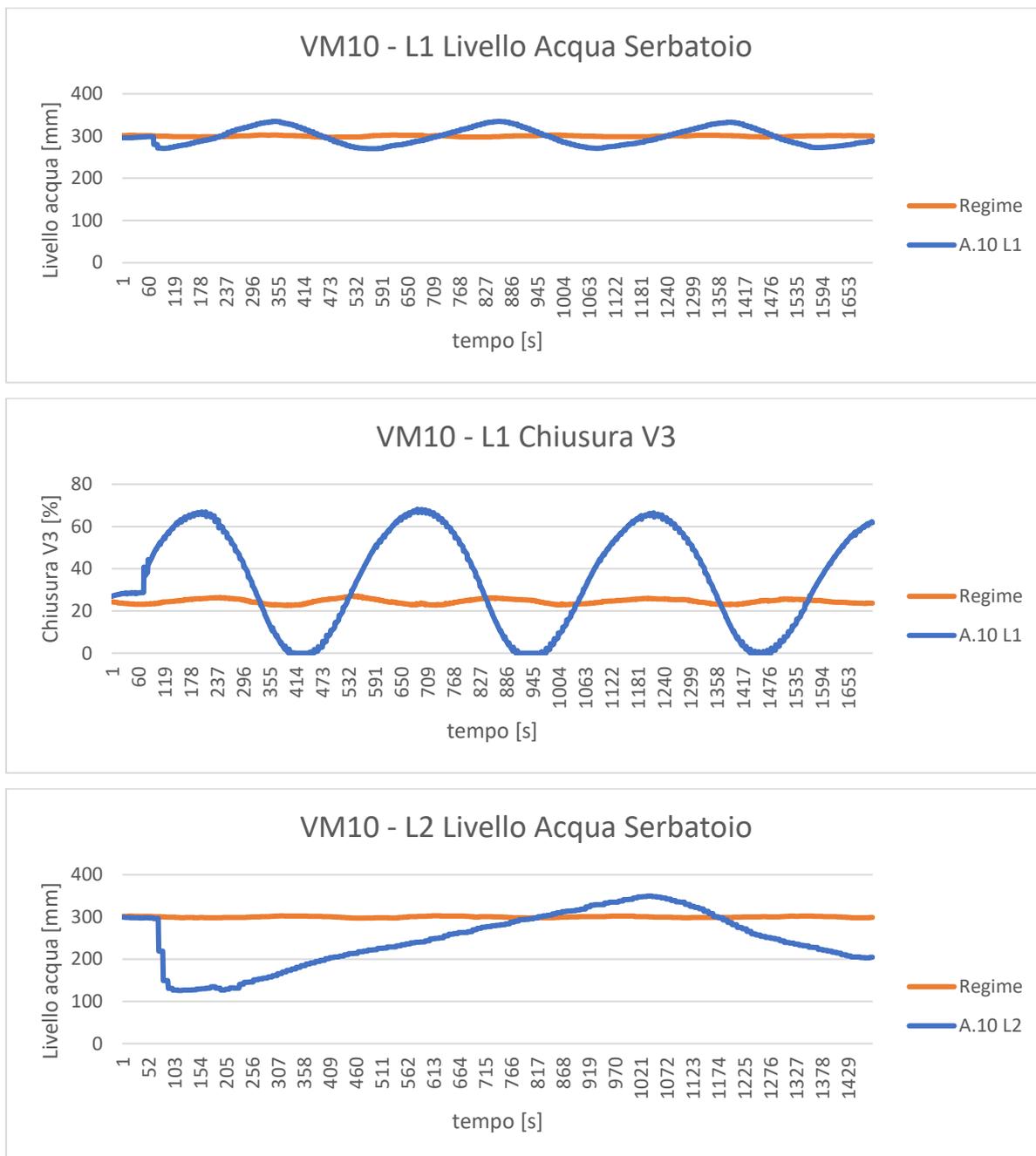


Figura 5.8: Anomalia 9 – Pressione del serbatoio e percentuale di chiusura V2

Nei primi due casi, il sistema si riporta nelle condizioni di regime tramite l'azionamento dell'elettrovalvola V2. Nel caso tre invece, essendo stato occlusa totalmente la via di uscita dell'aria, l'elettrovalvola si apre completamente, ma l'aria non avendo il passaggio libero, torna nel serbatoio e ne aumenta la pressione fino a 1.9 bar.

5.3.3.7. Anomalia 10

Il ruolo della valvola VM10 è di simulare un'eventuale perdita d'acqua all'interno del serbatoio. Nella Figura 5.9 vengono rappresentate le diverse perdite e i modi in cui la valvola V3 cerca di far tornare il sistema nelle condizioni di partenza.



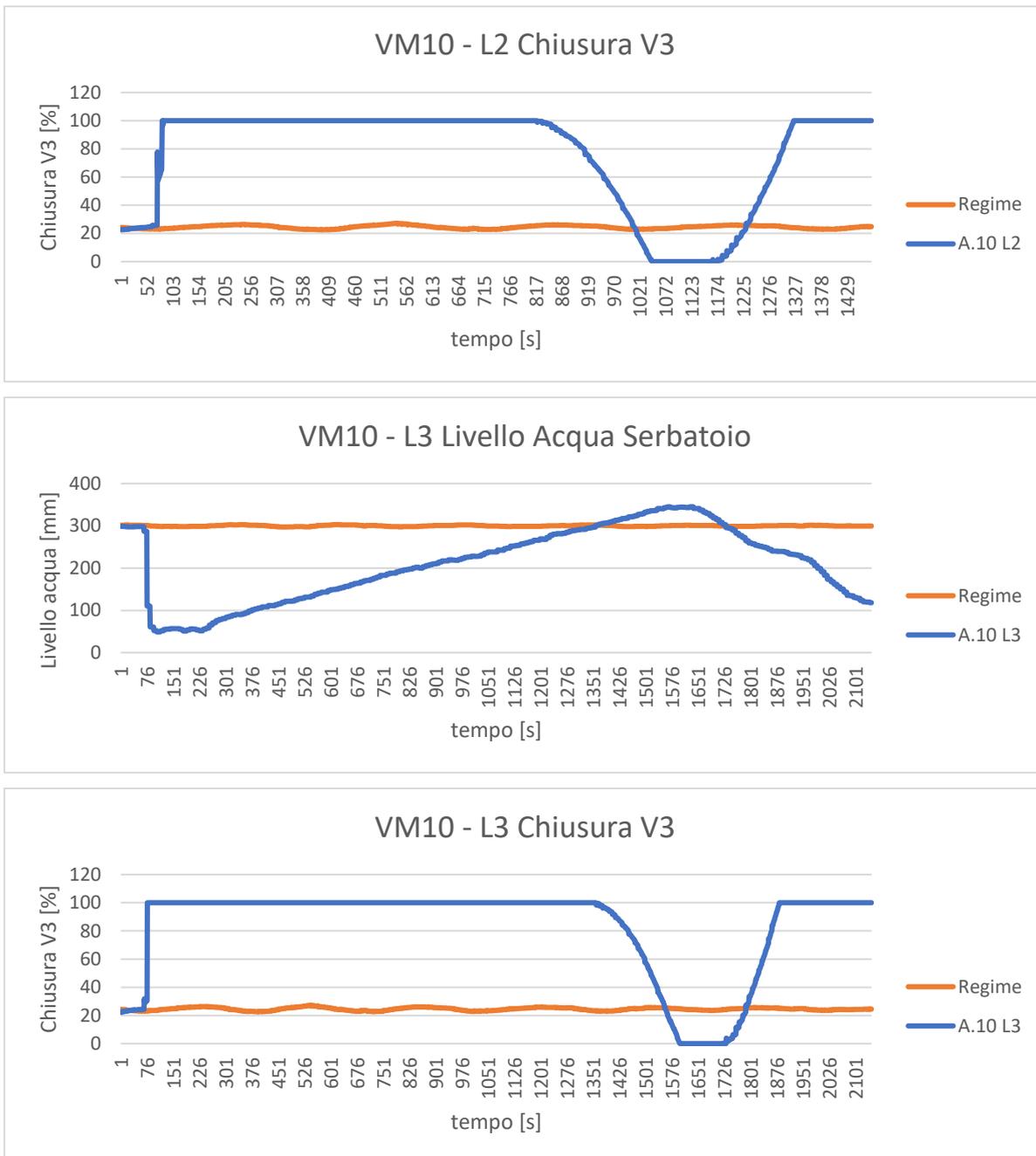


Figura 5.9: Anomalia 10 – Acqua nel serbatoio e percentuale di chiusura V3

Per simulare il Livello 1, la testa della valvola VM10 è stata aperta solo di qualche grado e grazie all'azione del PID dell'elettrovalvola V3, il sistema si riporta in condizioni simili a quelle iniziali. Per il Livello 2 e Livello 3, invece, l'apertura è più pronunciata e nonostante l'azione correttiva dell'elettrovalvola, il sistema non si stabilizza intorno ad un nuovo punto di equilibrio.

5.3.4. Acquisizione dei dati

Una volta accesi l'impianto, i sensori e l'hub centrale RevPi è stato possibile registrare i dati delle diverse prove. Gli elementi presenti nel database fanno riferimento a delle grandezze di diverso tipo, che hanno scale differenti tra loro. Infatti, grazie alla varietà di sensori installati nell'impianto, ad ogni istante di campionamento sono presenti misure di portata, pressione e livello che caratterizzano i fluidi circolanti nel sistema. Inoltre, nel database vengono riportate anche le percentuali di chiusura delle elettrovalvole che tramite il controllo dei PID cercano di riportare il sistema nelle condizioni iniziali quando su di esso viene simulata un'anomalia.

Di seguito, nella Tabella 5.4 vengono riportate le prime cinque letture della prova di regime.

Tabella 5.4: Database di partenza

S1	S2	S3	S4	S5	S6	S7	S8	%V1	%V2	%V3
5.508	9.995	0.971	1.349	1.301	25.931	301.304	21.933	89.08	73.96	24.33
5.5	9.995	0.971	1.349	1.301	25.931	301.304	21.933	89.08	74.05	24.31
5.5	9.995	0.971	1.349	1.301	25.931	301.48	21.933	89.02	74.05	24.29
5.5	9.995	0.971	1.349	1.301	25.931	301.48	21.933	89.01	74.05	24.15
5.5	9.995	0.971	1.349	1.301	26.006	301.48	21.933	89.01	74.05	24.18

Quando l'algoritmo d'intelligenza artificiale viene addestrato con dati di diversa natura, non converge ad una soluzione ottimale. Per occorrere a questo problema, le letture presenti nel database sono state standardizzate tramite la tecnica z-score.

5.3.5. Standardizzazione dei dati

Sia n il numero totale di campioni all'interno di una variabile, la media aritmetica μ si calcola sommando tutti i valori a disposizione e dividendo il risultato finale per il numero complessivo dei dati.

$$\mu = \frac{1}{n} \sum_{i=1}^n x_i$$

Si definisce σ scarto quadratico medio (o deviazione standard) quell'indice di dispersione statistico che esprime la differenza dei valori del dataset rispetto alla media della variabile. Il quadrato della deviazione standard prende il nome di varianza σ^2 .

$$\sigma = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \mu)^2}$$

Nel campo della statistica e dell'analisi dei dati, con il termine “standardizzazione” s'intende quel procedimento che trasforma una variabile aleatoria distribuita secondo una media μ e varianza σ^2 , ad una variabile aleatoria avente media 0 e varianza pari a 1. Il calcolo prevede di sottrarre alla variabile d'interesse X la sua media μ e dividere il tutto per la deviazione standard σ .

$$Z = \frac{X - \mu}{\sigma}$$

Questa formula prende il nome di standard-score o z-score.

Di seguito, vengono riportate le prime cinque letture dei dati standardizzati in ingresso alla rete.

Tabella 5.5: Database standardizzato

S1	S2	S3	S4	S5	S6	S7	S8	%V1	%V2	%V3
0.3394	0.2887	0.1552	-0.0728	0.0166	0.2176	0.1904	0.7245	0.5623	0.0718	-0.4091
0.3313	0.2887	0.1552	-0.0728	0.0166	0.2176	0.1929	0.7245	0.5604	0.0750	-0.4103
0.3313	0.2887	0.1552	-0.0728	0.0166	0.2232	0.1929	0.7245	0.5602	0.0750	-0.4135
0.3313	0.2878	0.1552	-0.0728	0.0166	0.2232	0.1929	0.7245	0.5602	0.0750	-0.4147
0.3313	0.2878	0.1552	-0.1001	0.0166	0.2232	0.1929	0.7264	0.5602	0.0750	-0.4158

5.4. La rete SOM

Per monitorare lo stato di salute dell'impianto presente in laboratorio, è stata sviluppata una rete neurale di tipo SOM. Come spiegato nel capitolo 4.4, dedicato alla struttura della mappa auto-organizzante, l'algoritmo è di tipo non supervisionato e sfrutta l'apprendimento competitivo.

Una volta acquisiti i dati in formato `.csv`, vengono standardizzati e successivamente unificati in un data frame finale. Il campionamento avviene ogni 0,1 secondi e le letture complessive di tutte e ventidue le prove sono 27447. La dimensione della griglia della mappa di output dipende strettamente dal numero dei campioni forniti alla rete. Per non incorrere in una mappa di dimensione troppo elevata, sono stati filtrati i dati e all'algoritmo sono state fornite soltanto 10259 letture, cioè quelle della prova a regime e dei livelli 3 delle anomalie con un tempo di campionamento pari a 0,3 secondi.

La Self-Organizing Map è stata scritta in codice Python. Per la realizzazione dell'algoritmo, oltre ai pacchetti di Numpy e Pandas necessari per la manipolazione dei dati, per la creazione della mappa è stata utilizzata la libreria MiniSom e per la rappresentazione grafica il pacchetto Bokeh.

5.4.1. I parametri della rete

Il pacchetto open source MiniSom è un'implementazione minimalista basata su Numpy delle Self-Organizing Maps. La funzione MiniSom si presenta come:

$$SOM = \text{MiniSom}(x, y, \text{input len}, \text{topology}, \text{sigma}, \text{learning rate}, \text{neighborhood function}, \text{seed})$$

I parametri che caratterizzano la rete sono:

- **x**: numero di righe della griglia di output;
- **y**: numero di colonne della griglia di output;
- **input len**: lunghezza del database fornito alla rete;
- **topology**: topologia della rete;
- **sigma**: raggio di vicinato;
- **learning rate**: tasso di apprendimento;
- **neighborhood function**: funzione di vicinato;
- **seed**: seme della rete.

Per il calcolo del numero di righe e di colonne che caratterizzano la mappa di output viene sfruttata la formula sperimentale presente nello studio riportato in letteratura (10). Siano 10259 le letture fornite alla rete, considerando una mappa quadrata, il numero delle righe e delle colonne è di:

$$\sqrt{5\sqrt{10259}} \cong 18.$$

Per la realizzazione di una mappa isotropa, ripetibile e in grado di evidenziare in maniera marcata le relazioni di vicinato, viene scelta la topologia esagonale, la funzione di vicinato gaussiano e il seme è impostato a 0. Di seguito, viene approfondito il processo di iper-parametrizzazione della rete oggetto d'esame.

5.4.1.1. L'iper-parametrizzazione

Per la scelta ottima dei parametri che caratterizzano il tasso di apprendimento e la dimensione del vicinato, l'algoritmo viene iper-parametrizzato tramite la funzione `hyperopt.fmin()`.

Il compito svolto da `hyperopt` è di trovare il miglior valore di una o più funzioni scalari tra un insieme di possibili argomenti. Mentre molti pacchetti di ottimizzazione presuppongono che gli input siano ricavati da uno spazio vettoriale, `hyperopt`, incoraggiando l'utente a descrivere in maniera accurata lo spazio della funzione obiettivo, è in grado di cercare i parametri in modo più efficiente.

Per usare `hyperopt` è necessario definire:

- **la funzione obiettivo da minimizzare;**
- **lo spazio su cui cercare;**
- **il database in cui archiviare tutte le valutazioni della ricerca;**
- **l'algoritmo di ricerca da utilizzare.**

Essendo il numero di campioni molto elevato e la dimensione della mappa grande, il range di valori accettabili definito per il raggio di vicinato e il tasso di apprendimento varia da 0,001 a 5. Il compito svolto dall'iper-parametrizzazione è di testare la rete SOM con diversi valori di learning rate e neighborhood size al fine di trovare la migliore combinazione in grado di minimizzare l'errore di quantizzazione e di restituire una mappa di output che preservi le relazioni topologiche dei dati di partenza. Quindi i parametri in ingresso alla rete sono:

- *x*: 18;
- *y*: 18;
- *input len*: 10259;
- *topology*: esagonale;
- *sigma*: 2.382866878925671;
- *learning rate*: 1.422871364551101;
- *neighborhood function*: funzione gaussiana;
- *seed*: 0.

5.4.2. L'addestramento e la validazione della rete

Dopo aver ricavato e impostato i vari parametri della funzione MiniSom, per la realizzazione della mappa occorre inizializzare i pesi e avviare l'addestramento della rete. I pesi associati ad ogni neurone vengono inizializzati in maniera randomica a valori piccoli prossimi allo zero e considerate 10000 iterazioni, l'algoritmo inizia la fase di addestramento.

Per valutare la bontà della mappa di output, è necessario fare uno studio approfondito su due parametri:

- **quantization error;**
- **topographic error.**

L'errore di quantizzazione permette di valutare la qualità dell'apprendimento della rete e mostra quanto la mappa si adatti ai dati di partenza. Come spiegato nella sezione 4.4.3 dedicata alla qualità della rete, il quantization error si calcola determinando la distanza media dei vettori campione verso

il BMU. Più l'errore è prossimo a zero, più la qualità della mappa è alta e i dati di output sono disposti in maniera corretta.

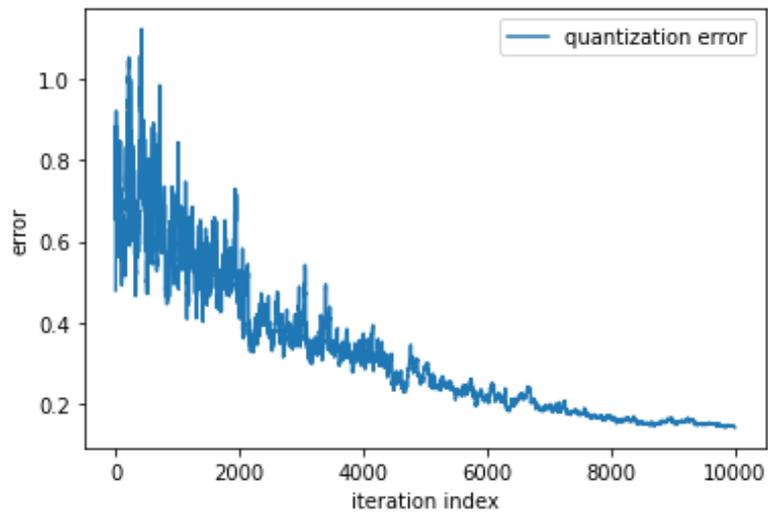


Figura 5.10: Quantization error della Self-Organizing Map

La Figura 5.11 rappresenta l'andamento dell'errore topografico in funzione delle iterazioni. L'errore topografico definisce la percentuale di vettori per i quali la prima e la seconda BMU non sono unità adiacenti. Un campione per il quale questi due nodi non sono adiacenti conta come un errore. Se l'errore topografico è 0, non si è verificato alcun errore. Se 1, la topologia non è stata conservata per nessuno dei campioni e nella mappa di output vengono perse le relazioni del dataset d'ingresso.

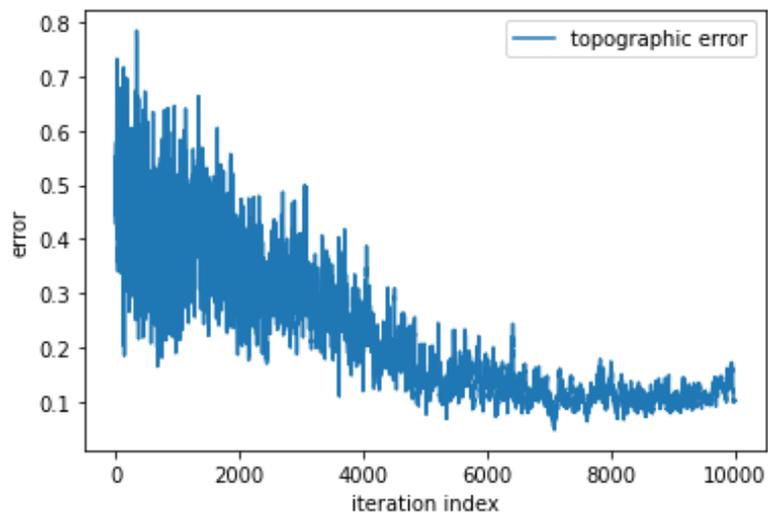


Figura 5.11: Topographic error della Self-Organizing Map

5.4.3. I pesi

Dopo aver validato la rete in maniera analitica, è necessario fare uno studio approfondito sulla griglia di output per verificare in che maniera le relazioni topologiche dei dati di input vengono conservate nella mappa. Compiuto da parte della rete il processo di addestramento basato sull'apprendimento competitivo, il peso $w_i(t)$ associato al campione $x(t)$ viene aggiornato al tempo $t + 1$ tramite la seguente formula:

$$w_i(t + 1) = w_i(t) + \alpha(t)h_{ci}(t)[x(t) - w_i(t)]$$

Dove $\alpha(t)$ rappresenta il valore del learning rate al tempo t e $h_{ci}(t)$ è la funzione di vicinato gaussiana. Le coordinate x e y dei dati proiettati sulla mappa di output si ricavano tramite la funzione *som. get_euclidean_coordinates ()*. Per l'analisi topologica è necessario studiare la matrice delle distanze dei pesi (U-Matrix) che si ricava tramite il metodo *som. distance_map ()*. La U-Matrix ha la stessa dimensione della griglia e rappresenta la distanza tra i vari cluster nella mappa di output.

Nella Figura 5.12, la matrice delle distanze viene rappresentata tramite una mappa di calore. Più il colore è chiaro e tendente al bianco maggiore è la somiglianza del cluster della casella con i dati delle caselle vicine, più risulta intenso e tendente al blu scuro maggiore è la diversità del cluster con i vicini.



Figura 5.12: Rappresentazione della U-Matrix tramite mappa di colore

5.4.4. La mappa di output

La Figura 5.13 rappresenta la mappa di output. Di seguito, viene riportata la legenda che permette di distinguere le diverse prove e consente la lettura della mappa.

- **regime**: contrassegnato con un asterisco arancione;
- **anomalia 3 L3**: contrassegnata con un pallino verde;
- **anomalia 5 L3**: contrassegnata con un pallino rosso;
- **anomalia 6 L3**: contrassegnata con un pallino viola;
- **anomalia 7 L3**: contrassegnata con un pallino marrone;
- **anomalia 8 L3**: contrassegnata con una croce rosa;
- **anomalia 9 L3**: contrassegnata con una linea grigia;
- **anomalia 10 L3**: contrassegnata con un rombo giallo.

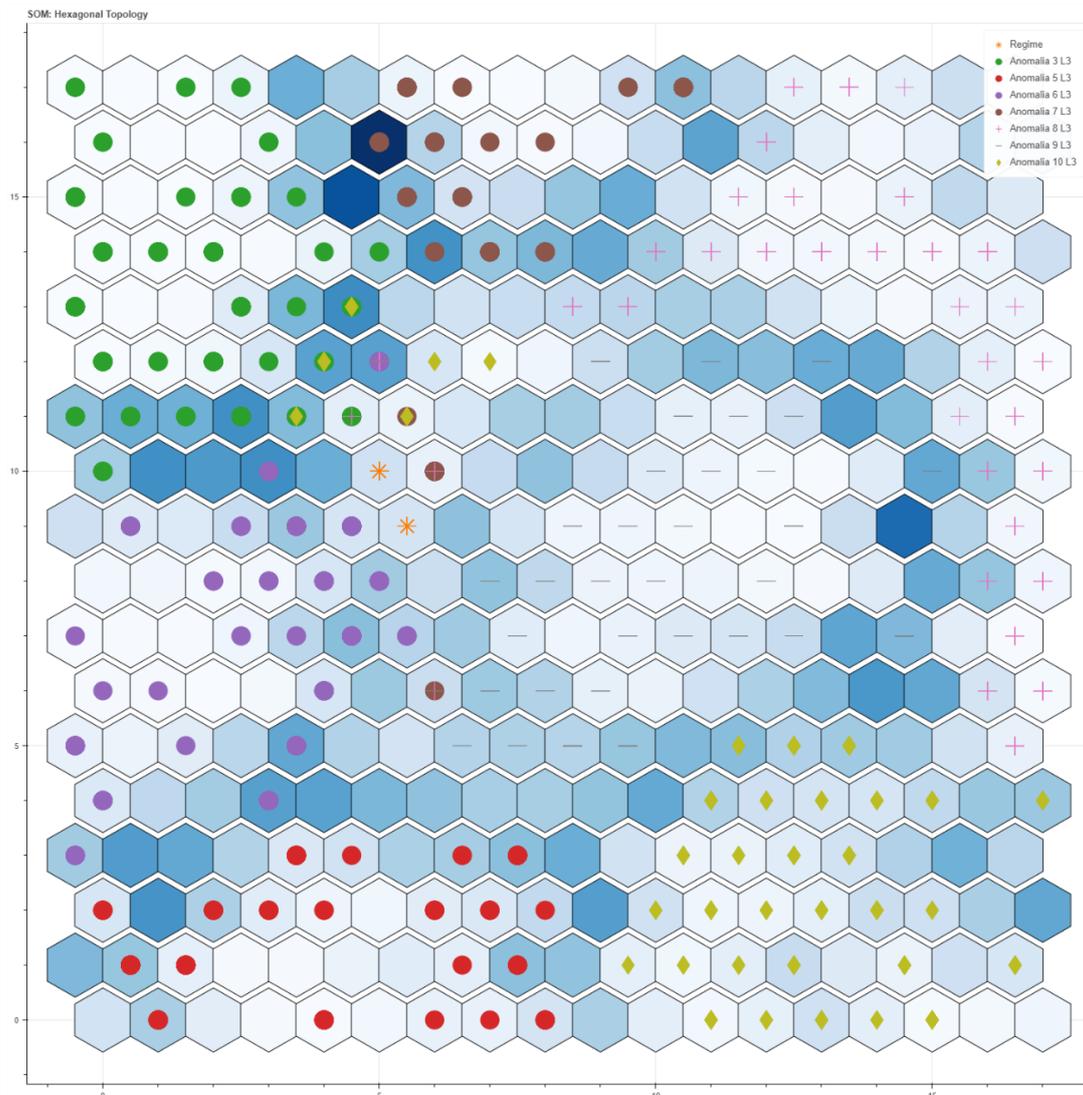


Figura 5.13: La mappa di output della prova a regime e delle anomalie a livello 3

Quando i dati d'ingresso vengono forniti alla rete, prima di essere proiettati sulla mappa di output, vengono confrontati con tutti i BMU presenti su di essa. Tramite il metodo *som.activation_response()* si ricava la matrice delle occorrenze delle risposte dei BMU. Dalla Figura 5.14 si evince che i nodi più interrogati della mappa sono quelli di coordinate (5,10) e (6,10). In queste due posizioni sono collocate tutte le letture dello stato di regime del sistema. Quindi, per il posizionamento dei dati delle letture successive, l'algoritmo fa riferimento principalmente alle letture di regime. In base alle differenze tra le letture in ingresso e quelle di regime, è in grado di disporre i dati anomali in zone diverse sulla mappa.

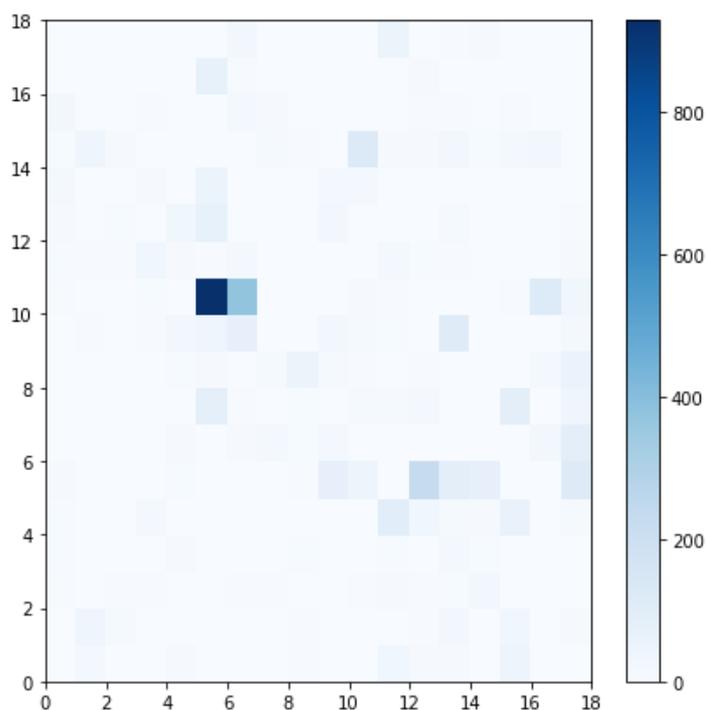


Figura 5.14: Attivazione dei neuroni

I dati proiettati sulla mappa di output formano dei cluster sulle singole caselle esagonali. A loro volta, l'unione di più caselle forma un'area d'interesse e l'unione di più aree formano una macroarea d'interesse. La Tabella A. 1 presente in Appendice riporta il tipo e il numero di letture d'ingresso presenti in ogni singola casella della mappa di output.

Nella Figura 5.13, oltre ai dati d'ingresso, viene rappresentata sullo sfondo la U-Matrix sotto forma di mappa di calore. Se una casella ha un colore tendente al bianco, allora le sei celle vicine appartengono alla stessa area, mentre se il colore s'intensifica e si avvicina al blu notte la casella si trova in prossimità del confine tra due o più aree. Dalla Figura 5.15 si evince che i bordi dell'area dell'anomalia 3 livello 3 sono di colore più scuro rispetto agli esagoni presenti all'interno, ciò testimonia che oltre la frontiera ci sono altre aree, quella dell'anomalia 7 e dell'anomalia 6.

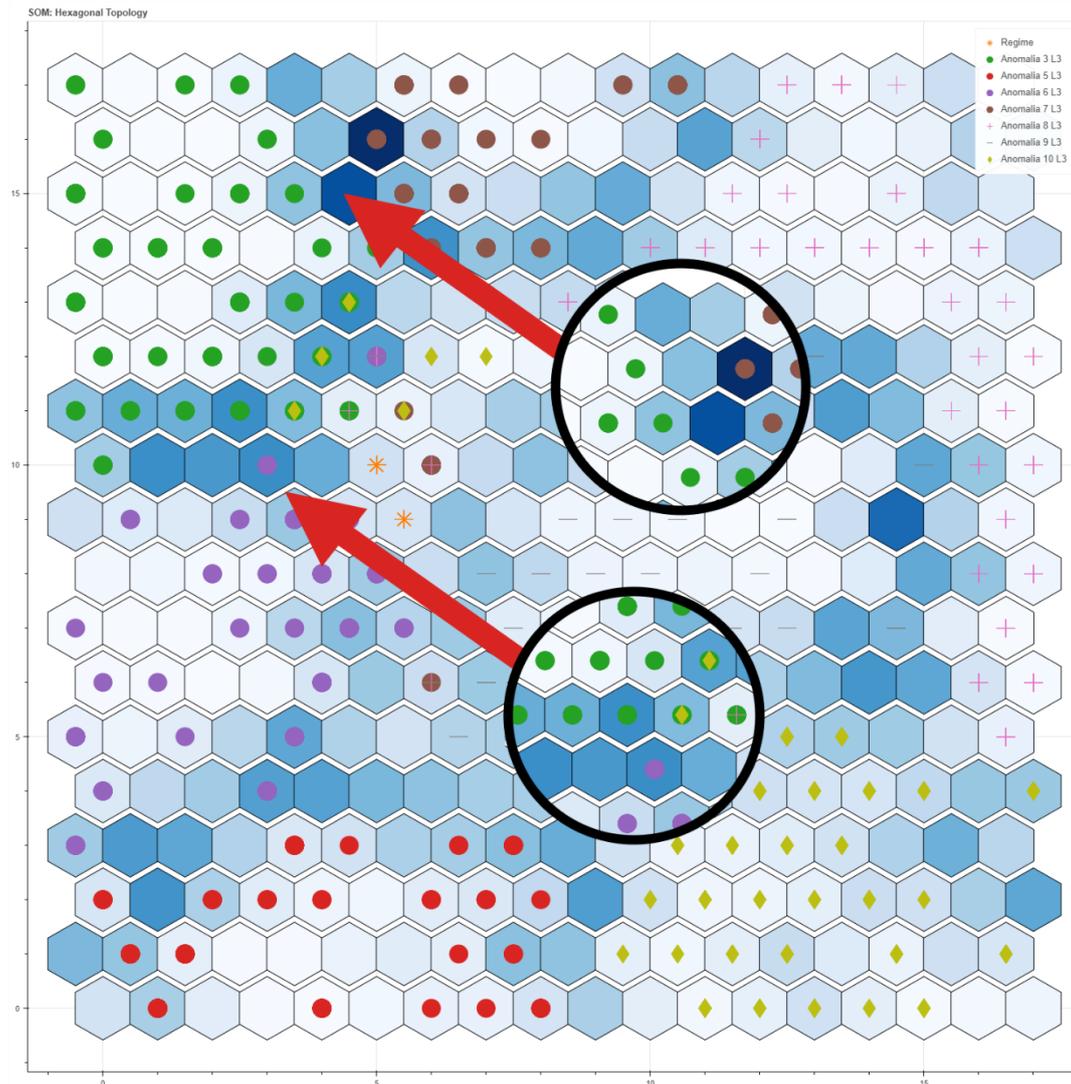


Figura 5.15: Studio del confine dell'area Anomalia 3 livello 3

Dallo studio dei risultati è emerso che la mappa auto-organizzante dispone i dati d'ingresso in sei macroaree della mappa, quali:

- area di regime;
- macroarea del transitorio aria e acqua;
- macroarea anomalia acqua;
- macroarea anomalia aria;
- area anomalia fluido bifase aria-acqua;
- area anomalia perdita del serbatoio;

5.4.4.1. Area di regime

Le letture dello stato di regime del sistema, come spiegato nella sezione precedente, vengono inserite all'interno delle caselle di coordinate (5,10) e (6,10). Per facilitare la lettura della mappa, nelle seguenti immagini i markers presenti nella Figura 5.13 vengono sostituiti con la colorazione dell'intera cella esagonale.

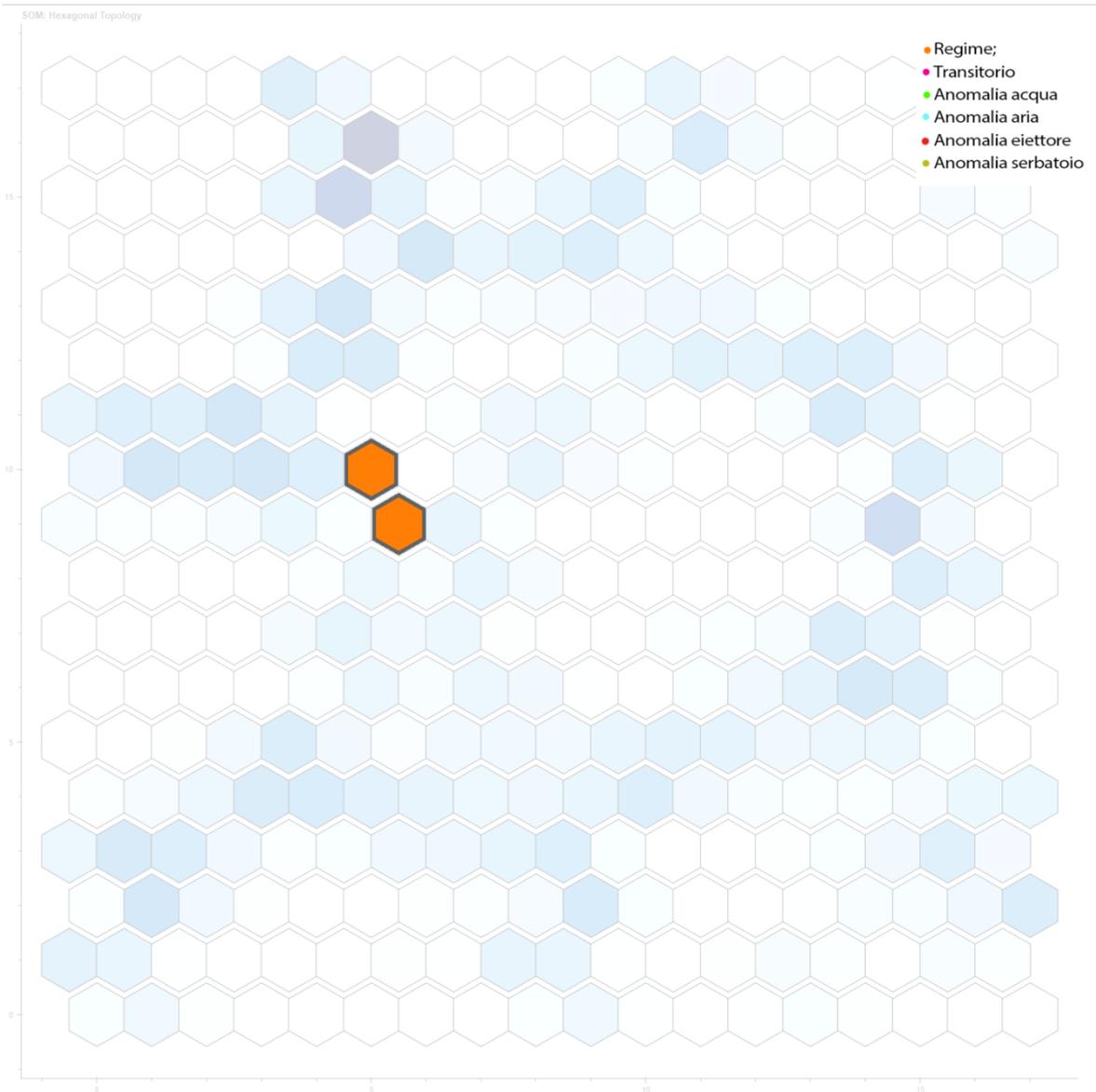


Figura 5.16: Area di regime

5.4.4.2. Macroarea del transitorio aria e acqua

Andando ad analizzare i dati proiettati sulla mappa di output, vengono disposte vicino alla zona di regime le prime letture delle prove anomale dell'aria e dell'acqua. Le valvole d'intercettazione vengono chiuse manualmente e il sistema passa dallo stato di regime ad un nuovo punto di funzionamento. Quindi, nella registrazione di ogni singola prova anomala sono presenti anche le letture che caratterizzano il transitorio. Nella seguente mappa, l'area del transitorio dell'aria viene contrassegnata con il colore viola.

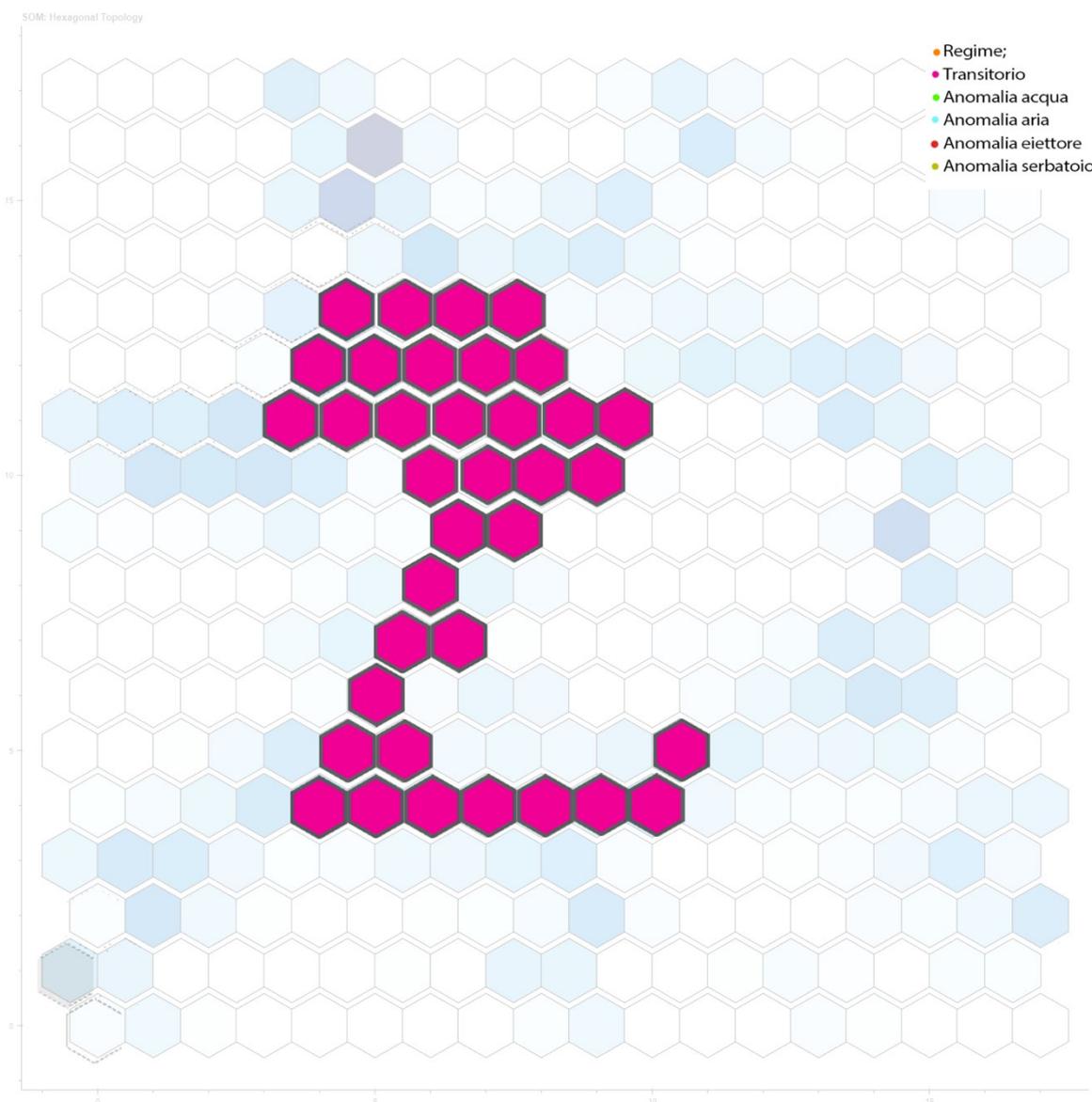


Figura 5.17: Macroarea del transitorio

5.4.4.3. Macroarea anomalia acqua

La rete dispone i dati delle anomalie generate dalla VM3 e dalla VM6 nella parte sinistra della mappa. Le due valvole di intercettazione occludono il passaggio dell'acqua in ingresso all'impianto e all'uscita del serbatoio. Dai dati forniti in ingresso, la mappa auto-organizzante è in grado di riconoscere che in entrambi i casi si registra nell'impianto un'anomalia dell'acqua. La zona della mappa contrassegnata in verde rappresenta la macroarea dell'anomalia dell'acqua.



Figura 5.18: Macroarea anomalia acqua

5.4.4.4. Macroarea anomalia aria

Le anomalie generate tramite la chiusura delle valvole d'intercettazione VM7, VM8 e VM9 occludono il passaggio dell'aria. La valvola VM7 impedisce il flusso d'aria in ingresso, la VM9 quello in uscita e la VM8 simula una perdita d'aria all'interno del serbatoio. L'algoritmo riconosce che tutte e tre le anomalie riguardano l'aria e dispone le letture nella parte destra della mappa. La macroarea dell'anomalia dell'aria è contrassegnata in celeste.

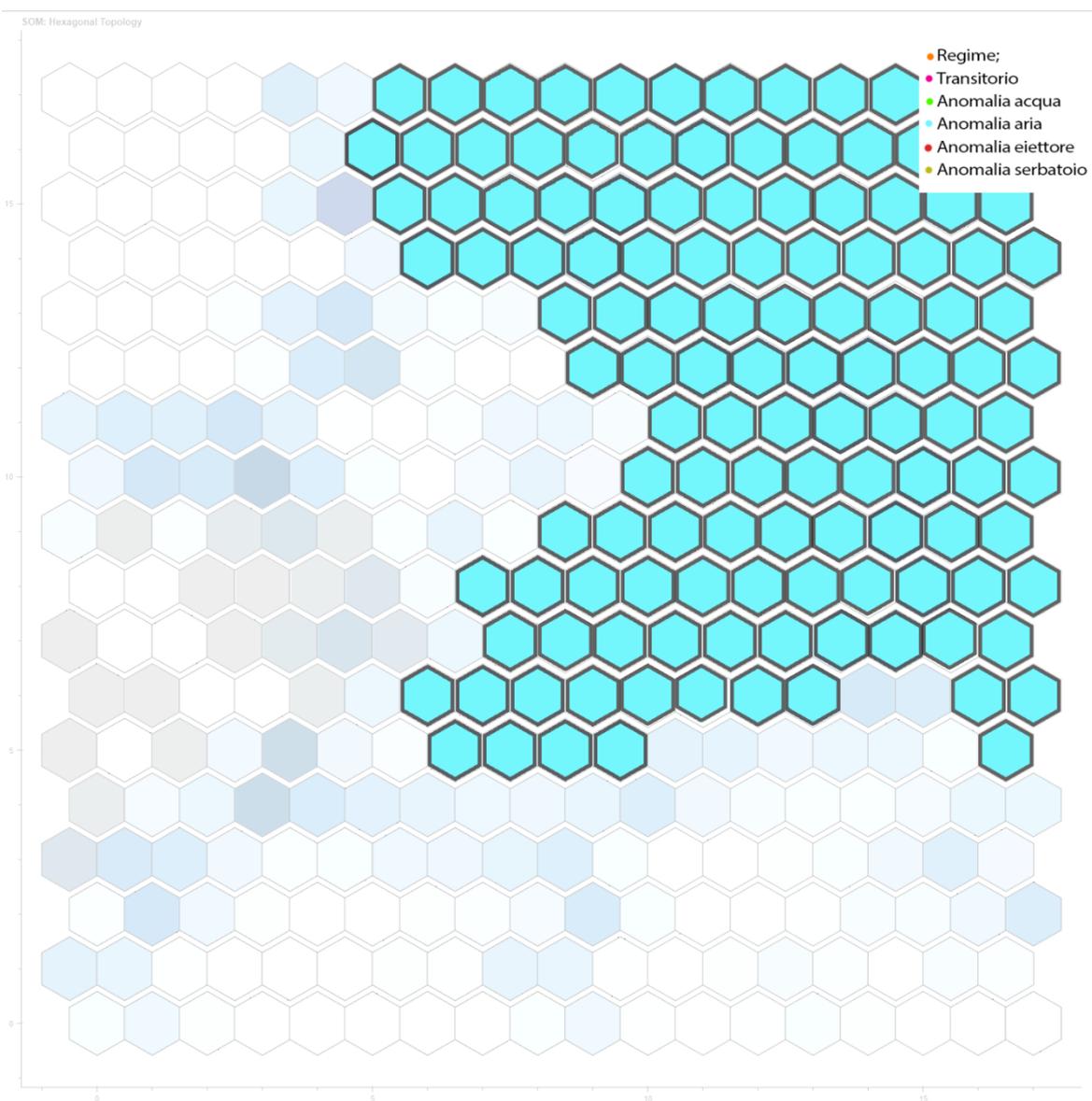


Figura 5.19: Macroarea anomalia aria

5.4.4.5. Area anomalia fluido bifase aria-acqua

La valvola d'intercettazione VM5 occlude il passaggio del fluido bifase aria-acqua dall'eiettore al serbatoio. L'area è contrassegnata sulla mappa in rosso ed è disposta in basso a sinistra, sotto la macroarea dell'anomalia dell'acqua e dell'aria.

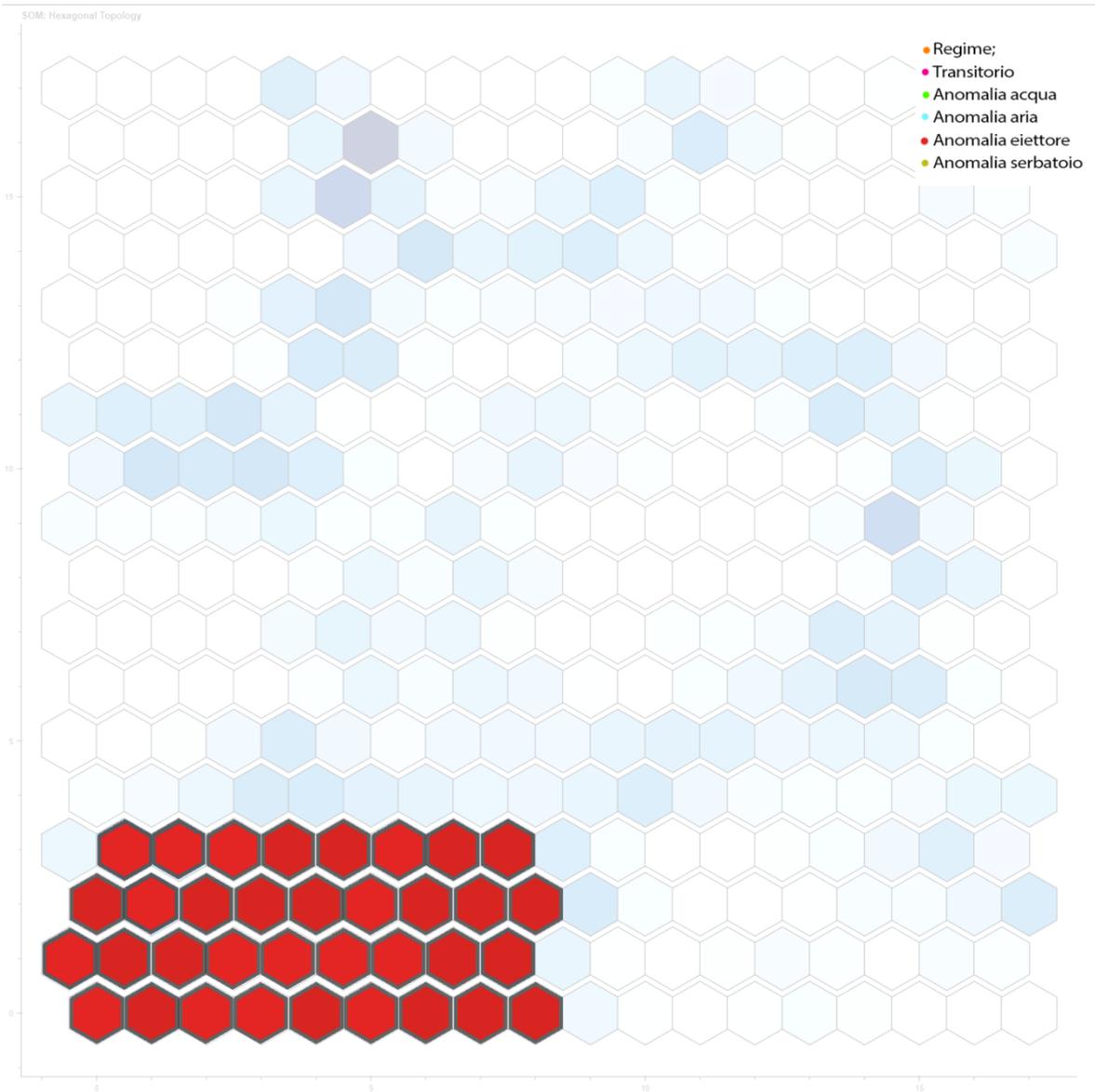


Figura 5.20: Area anomalia fluido bifase aria-acqua

5.4.4.6. Area anomalia serbatoio

La valvola VM10 simula una perdita d'acqua all'interno del serbatoio. La diversità di questo tipo di anomalia rispetto alle altre descritte in precedenza spinge la rete auto organizzante a disporre le letture in una nuova zona. L'area sulla mappa è colorata di giallo ed è disposta in basso a destra.

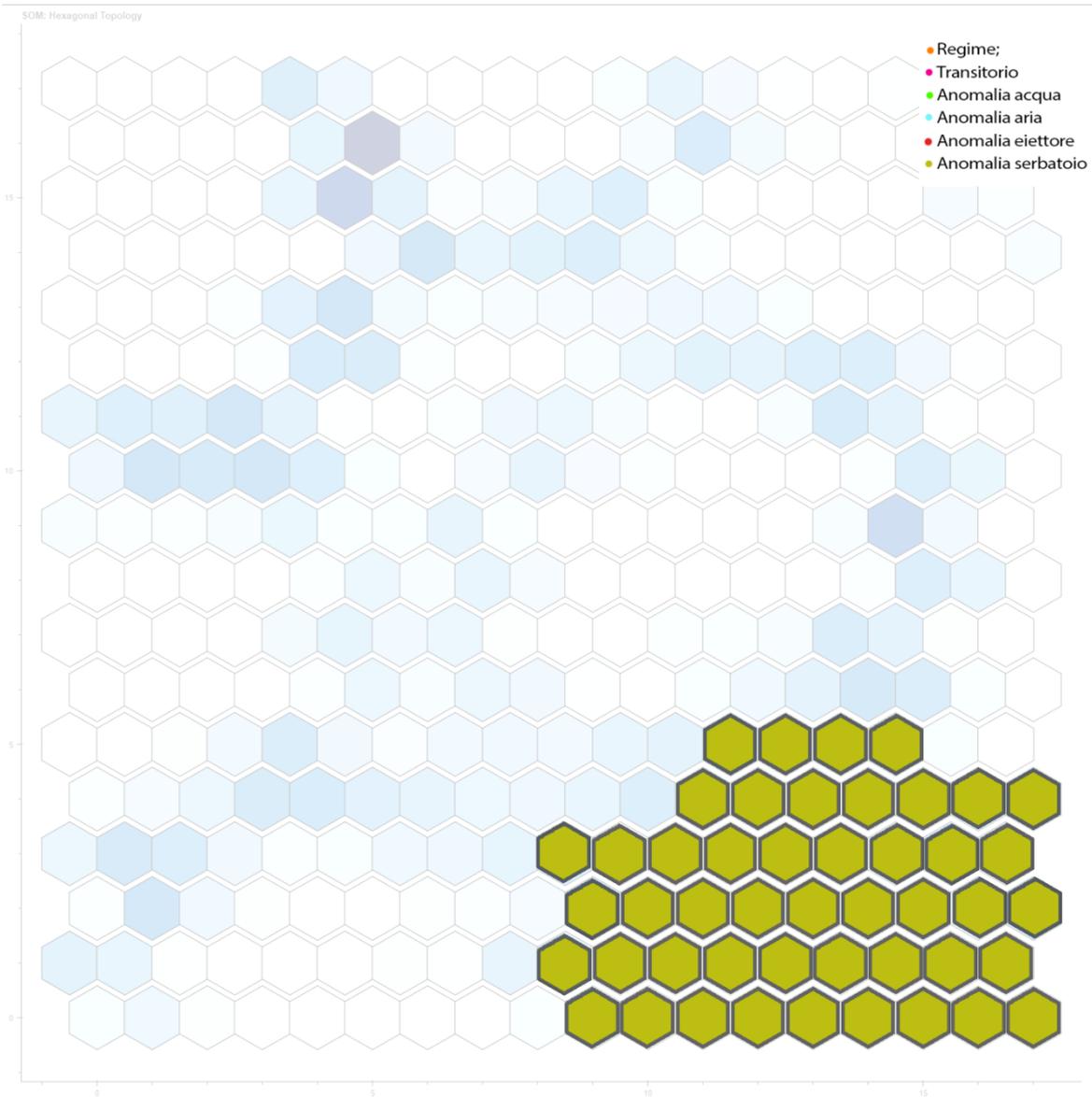


Figura 5.21: Area anomalia serbatoio

5.4.5. Mappa di output riadattata ai livelli 1 e 2

Dopo aver sviluppato l'algoritmo e aver studiato le macroaree che si formano nella mappa di output con i dati provenienti dalle registrazioni della prova a regime e delle anomalie di livello massimo, vengono inserite in ingresso le restanti prove anomale di livello 1 e livello 2.

Come avvenuto per le letture necessarie per l'addestramento, i dati delle prove anomale di livello 1 e livello 2 sono stati prima unificati e successivamente standardizzati con media nulla e varianza unitaria. Di seguito vengono riportate tutte le letture divise per le diverse aree.

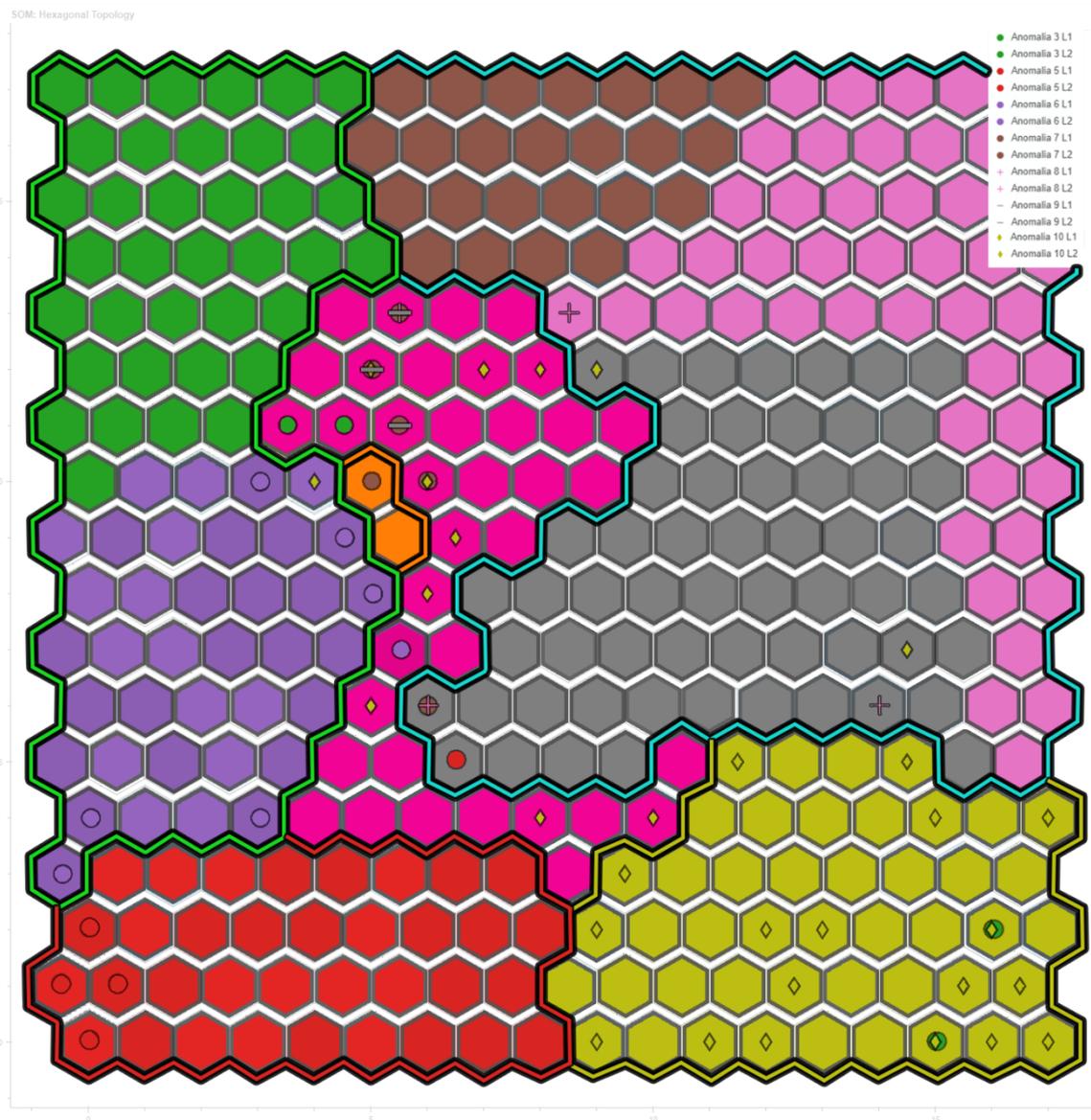


Figura 5.22: Mappa di output delle anomalie livello 1 e 2

5.4.5.1. Anomalia 3 livello 1 e 2

La Figura 5.23 presenta sullo sfondo le aree e le macroaree ricavate dalla rete SOM addestrata in base ai dati della prova a regime e delle prove anomale di livello 3. Tramite dei piccoli cerchi verdi, vengono raffigurate le letture delle prove anomale di livello 1 e livello 2 generate dalla valvola VM3. Ricordando che l'area contrassegnata in verde rappresenta la zona in cui l'algoritmo classifica le letture dell'anomalia 3 livello 3, i nuovi dati d'ingresso vengono disposti in quattro cluster distinti. I primi due cluster vengono posti in prossimità della frontiera dell'area verde e vicino a quella del regime colorata in arancio, poiché le letture dei primi due livelli della VM3 sono simili a quelle di questi due gruppi. Gli ultimi due cluster, contenendo le ultime letture del livello 2, vengono collocati all'interno dell'area dell'anomalia del serbatoio, poiché il sistema in entrambe le situazioni presenta la stessa portata d'acqua in ingresso.

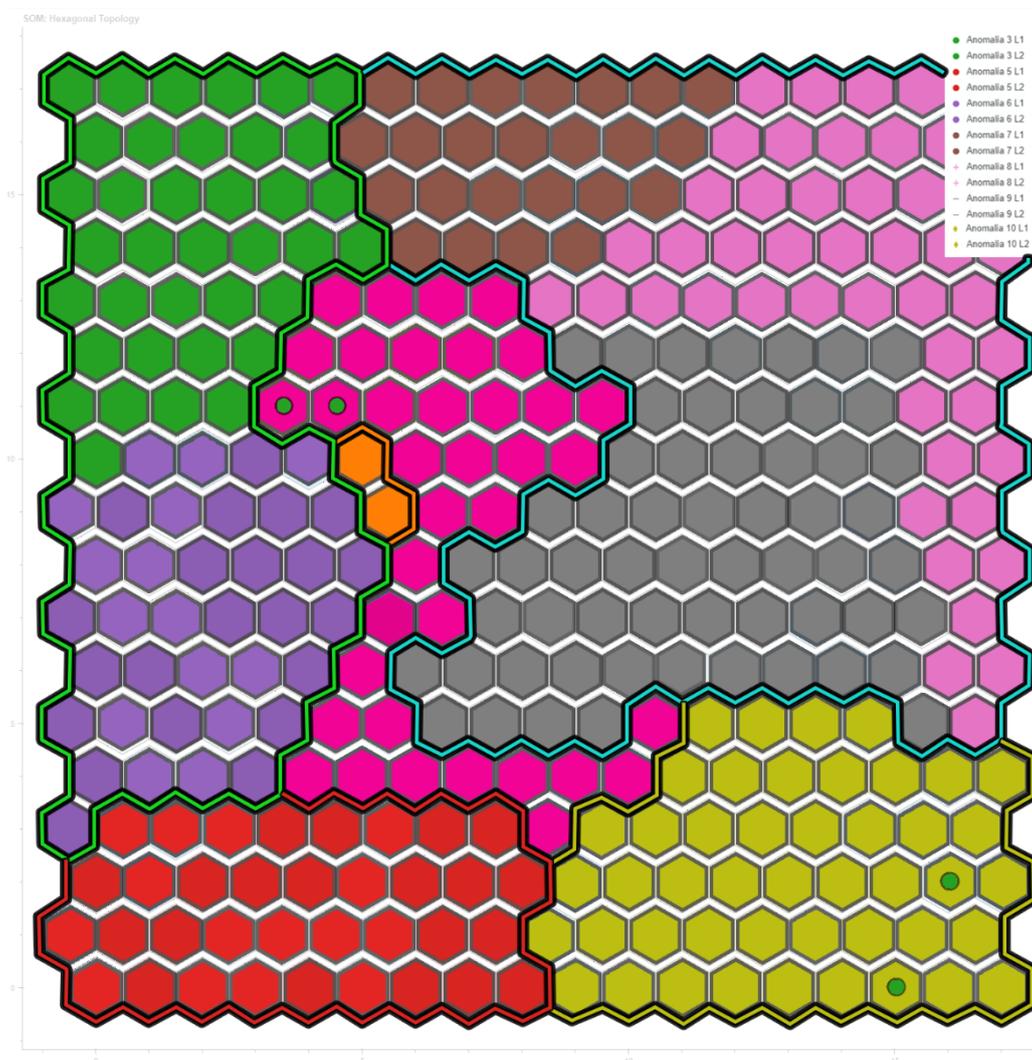


Figura 5.23: Proiezione sulla mappa di output dell'anomalia 3 livello 1 e 2

5.4.5.2. Anomalia 5 livello 1 e 2

L'area rossa presente sulla mappa rappresenta il cluster in cui la rete di partenza distribuisce le letture anomale dell'impianto causate dall'occlusione del condotto che collega l'uscita dell'eiettore all'ingresso del serbatoio di contenimento. Tramite i cerchi rossi, vengono raffigurati i dati delle prove anomale di livello 1 e livello 2 della valvola VM5. Questi nuovi dati in ingresso, registrando occlusioni inferiori rispetto a quelle con cui è stata addestrata la rete, si distribuiscono ai lati dell'area rossa. Le prime letture dell'anomalia 5 livello 1 si scostano dalla frontiera dell'area rossa e si collocano al confine della macroarea dell'anomalia dell'aria.

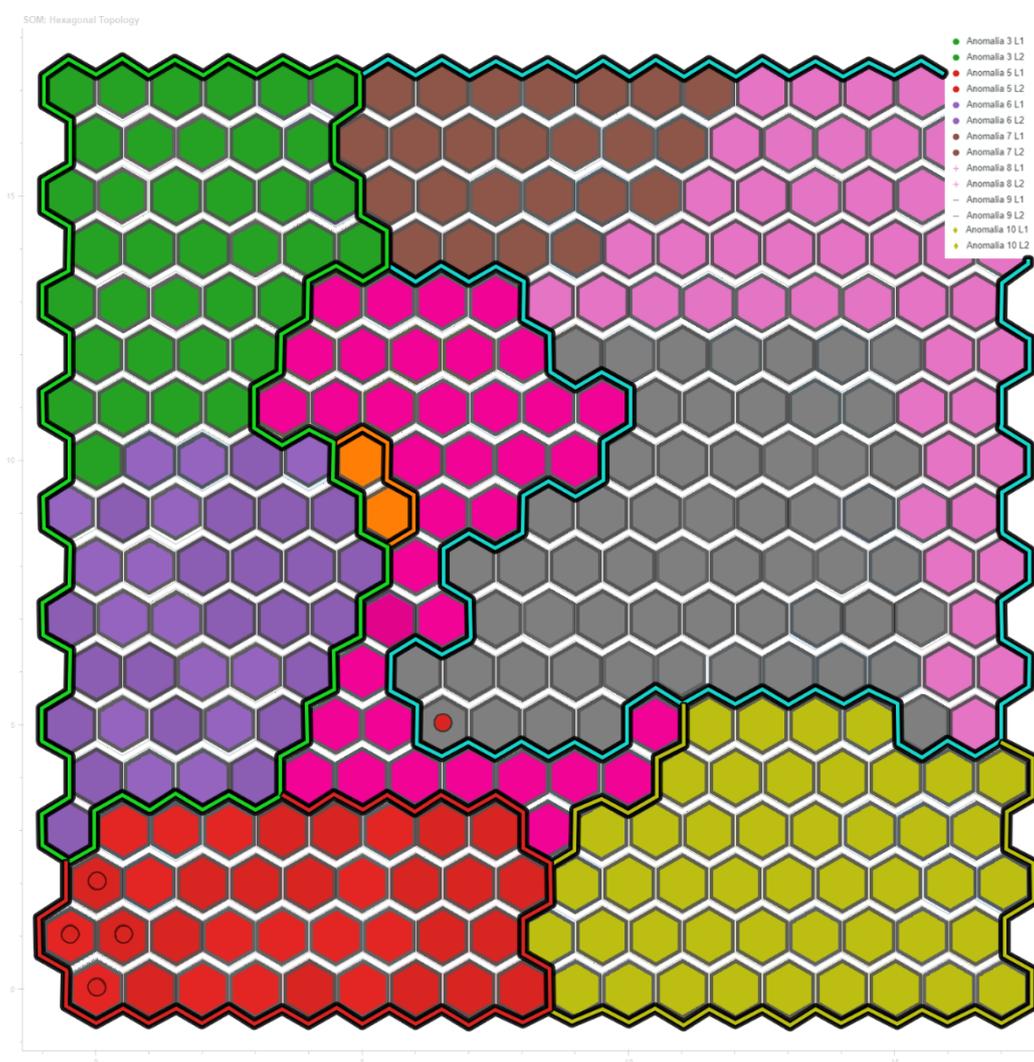


Figura 5.24: Proiezione sulla mappa di output dell'anomalia 5 livello 1 e 2

5.4.5.3. Anomalia 6 livello 1 e 2

L'area colorata in viola delimita il perimetro in cui la mappa auto-organizzante colloca le letture anomale causate dall'occlusione del condotto che collega il flusso dell'acqua dal serbatoio alla vasca. I cerchi viola rappresentano i dati delle prove anomale di livello 1 e livello 2 della valvola VM6. Essendo le occlusioni di grado inferiore rispetto a quelle di livello 3, i nuovi dati vengono proiettati sulla frontiera della rispettiva area anomala.

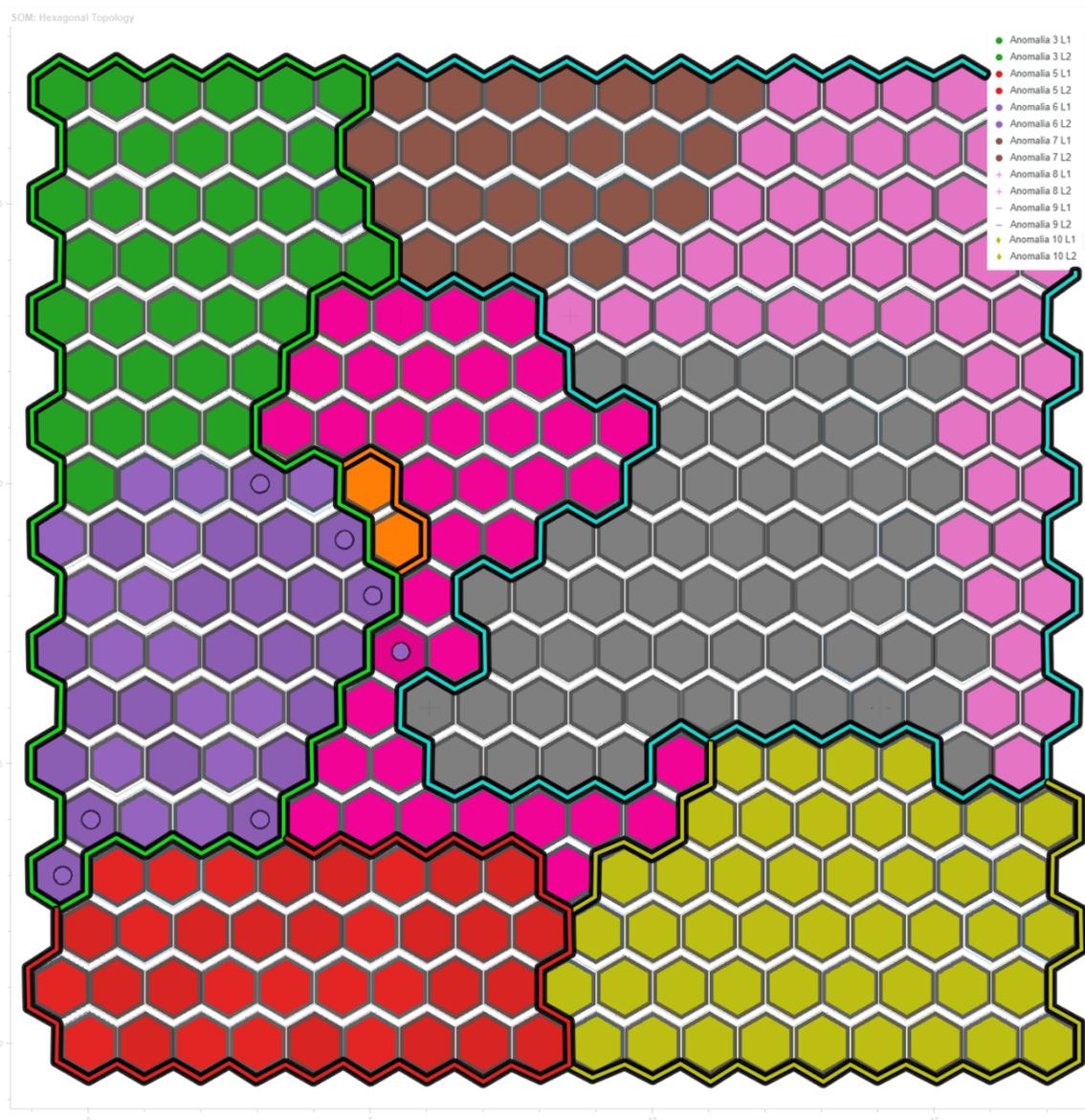


Figura 5.25: Proiezione sulla mappa di output dell'anomalia 6 livello 1 e 2

5.4.5.4. Anomalia 7 livello 1 e 2

Nella Figura 5.26 l'area colorata in marrone rappresenta il cluster in cui la mappa colloca le letture registrate sull'impianto dopo aver chiuso completamente il condotto d'aspirazione dell'aria in ingresso tramite la valvola d'intercettazione manuale VM7. I primi due gradi di occlusione vengono raffigurati sulla mappa di output tramite dei piccoli cerchi marroni. I dati registrati da queste due prove, presentano un'ostruzione ridotta del condotto di aspirazione dell'aria. Il sistema risulta pronto e si riporta subito nelle condizioni di regime; proprio per questo le letture del livello 1 e livello 2 vengono proiettate tra l'area marrone e quella arancione dei dati a regime.

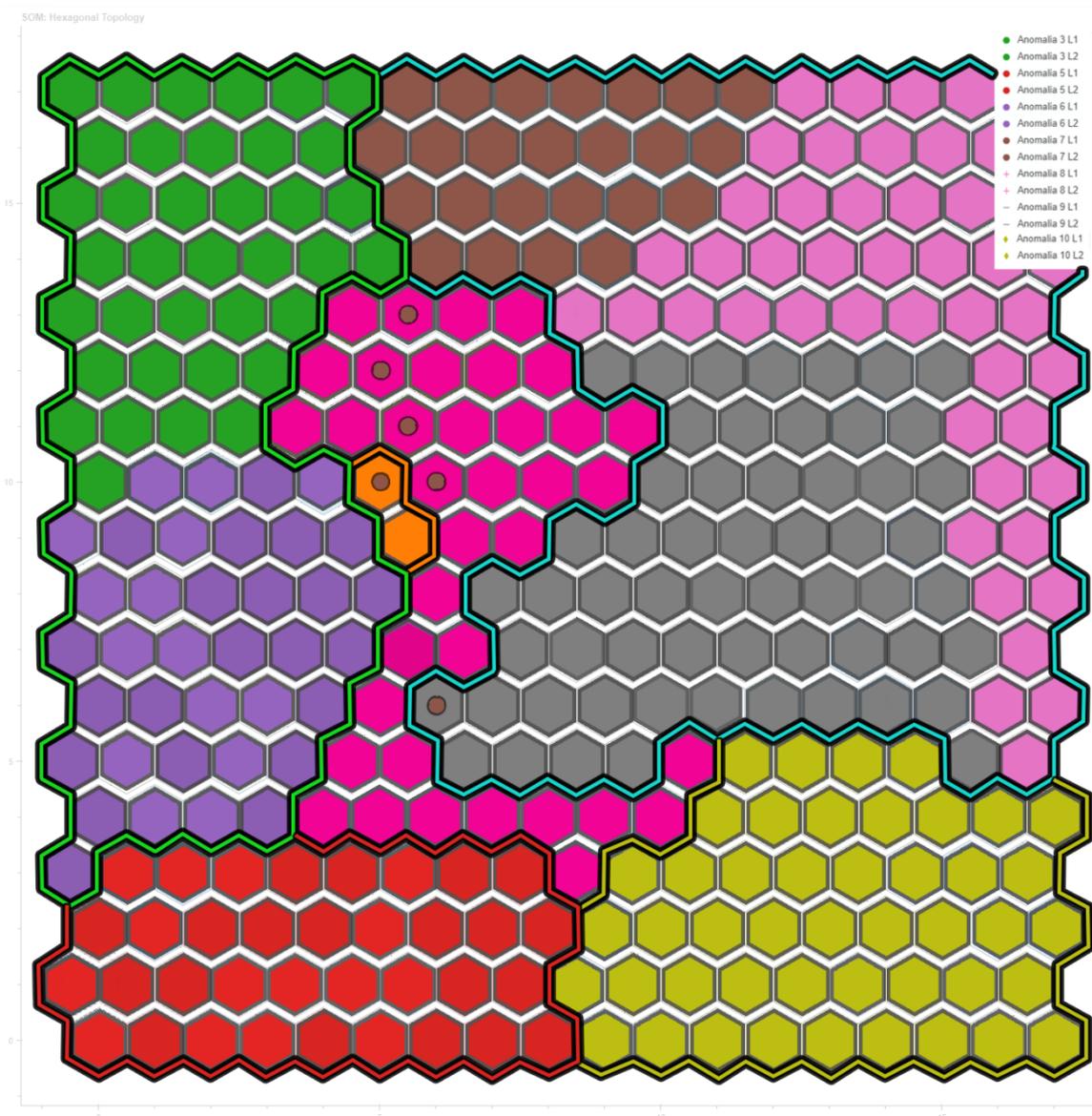


Figura 5.26: Proiezione sulla mappa di output dell'anomalia 7 livello 1 e 2

5.4.5.5. Anomalia 8 livello 1 e 2

Aperto completamente la valvola d'intercettazione manuale VM8 si simula il caso in cui il serbatoio perda aria, quindi, pressione. Nella mappa di output, la rete contrassegna in rosa queste letture anomale e le colloca nella parte destra. Somministrando all'algoritmo i dati dei livelli 1 e 2 dell'anomalia 8, la rete li divide in quattro cluster principali rappresentati tramite una croce rosa. Tutte le letture del livello 1, essendo d'intensità minore, vengono proiettate vicino alla zona in cui sono rappresentati i dati di regime. Le letture del livello 2, invece, presentando un'apertura maggiore della valvola si collocano vicino alla frontiera dell'anomalia 8 livello 3.

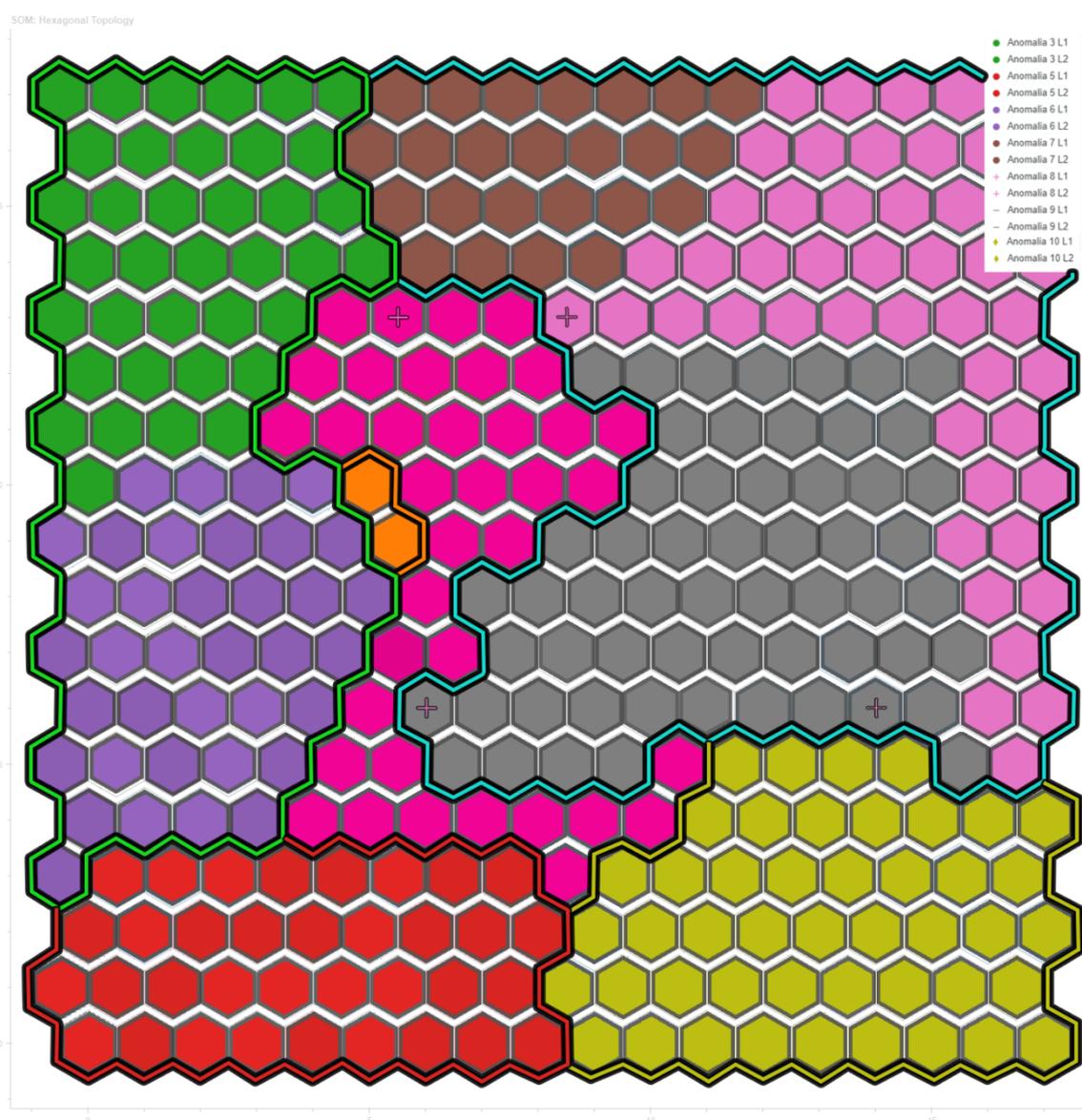


Figura 5.27: Proiezione sulla mappa di output dell'anomalia 8 livello 1 e 2

5.4.5.6. Anomalia 9 livello 1 e 2

La valvola d'intercettazione manuale VM9 ostruisce il passaggio dell'aria in uscita dal serbatoio. Le letture registrate chiudendo del tutto la valvola formano sulla mappa un'area colorata in grigio.

Fornendo all'algoritmo i dati dei livelli 1 e 2 dell'anomalia 9, tutte le letture vengono proiettate in 3 cluster contrassegnati con una linea grigia. Anche in questo caso, essendo le occlusioni di livello inferiore rispetto a quelle del livello 3, i dati vengono rappresentati in prossimità della zona di regime, più specificatamente nella zona del transitorio.

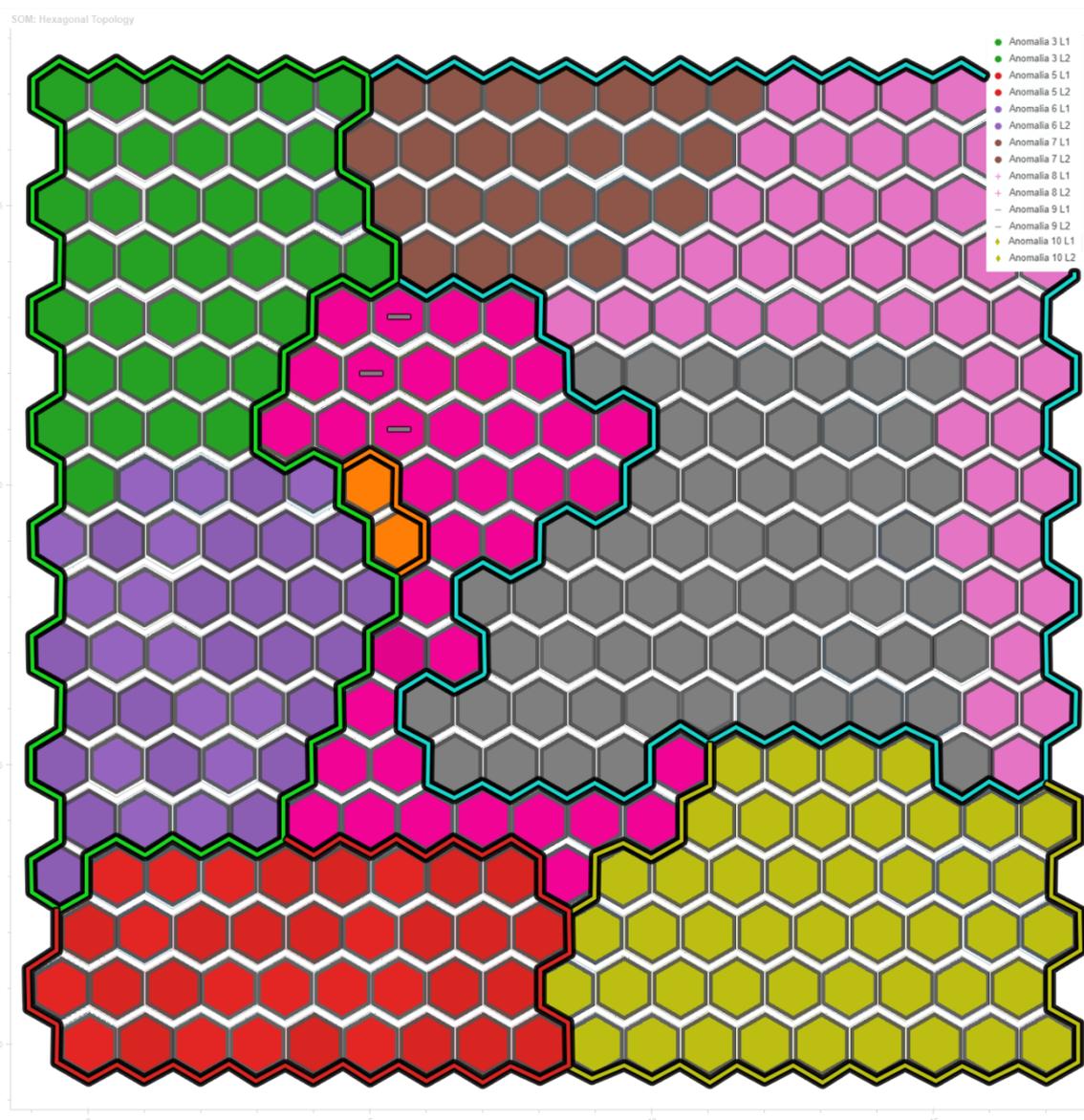


Figura 5.28: Proiezione sulla mappa di output dell'anomalia 9 livello 1 e 2

5.4.5.1. Anomalia 10 livello 1 e 2

Per mezzo della valvola d'intercettazione manuale VM10 si simula la perdita d'acqua del serbatoio. I dati del livello 3, forniti alla rete d'ingresso per il processo di addestramento, vengono proiettati nella mappa di output nella zona a destra in basso, formando un'area gialla. Fornendo alla rete le letture del livello 1 e livello 2, la mappa le raffigura con dei rombi gialli costruendo diversi cluster posizionati in varie parti della griglia. Aprendo, seppur di poco, la VM10, il sistema è instabile e non raggiunge un nuovo punto di equilibrio. Le letture del livello 1 vengono proiettate a ridosso della zona di regime, mentre quelle del livello 2, essendo d'intensità maggiori, ricadono pienamente all'interno dell'area gialla.

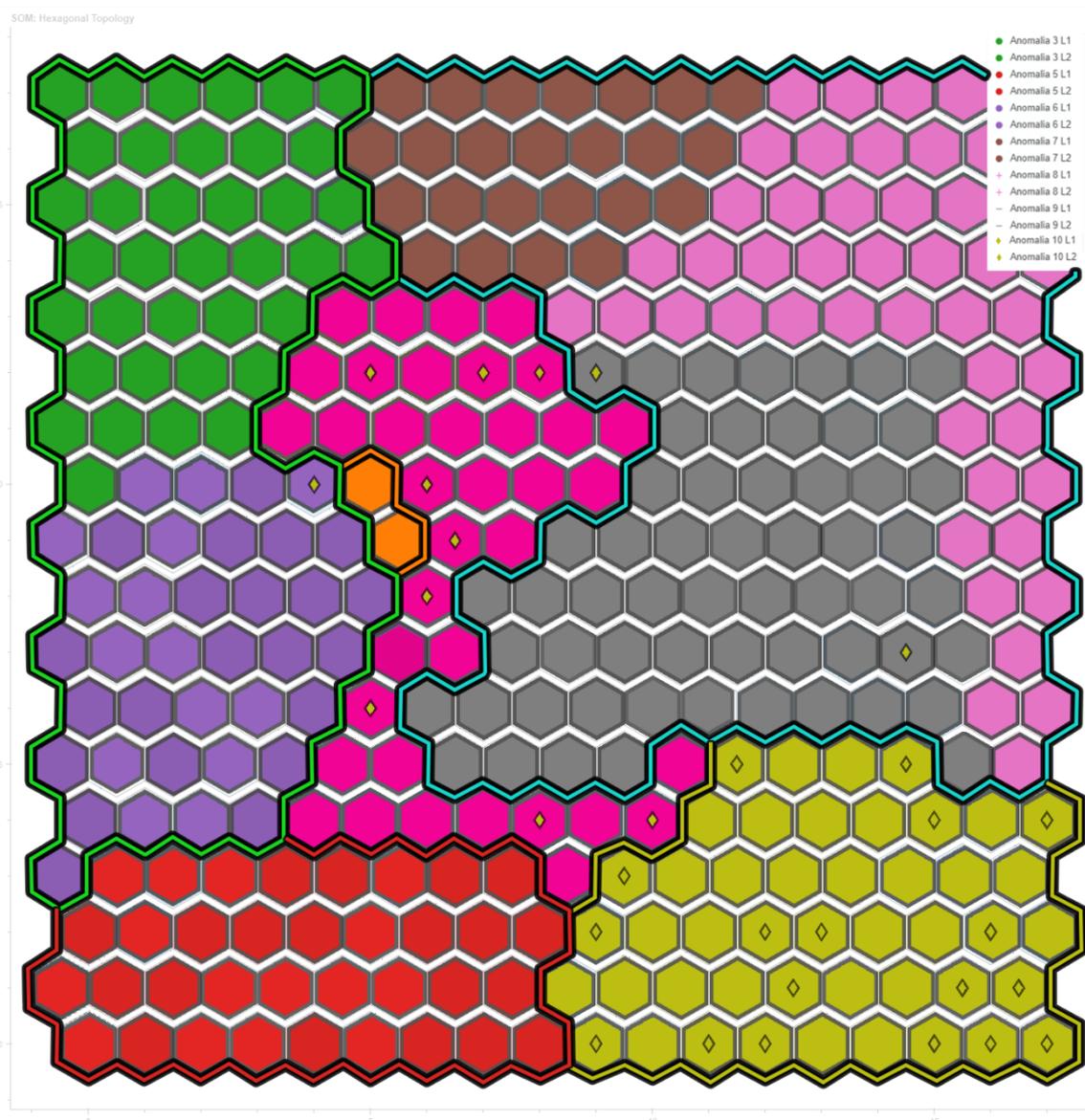


Figura 5.29: Proiezione sulla mappa di output dell'anomalia 10 livello 1 e 2

6. Conclusioni e sviluppi futuri

Lo scopo dell'elaborato è stato quello di realizzare una rete neurale artificiale di tipo SOM con il compito di monitorare lo stato di salute dell'impianto bifase aria-acqua situato all'interno del dipartimento di Ingegneria Industriale e Scienze Matematiche (DIISM) dell'Università Politecnica delle Marche (Ancona, Italia).

Le Self-Organizing Maps vengono sfruttate generalmente per risolvere problemi di clustering come quello della classificazione delle immagini o della segmentazione della clientela in ambito economico-finanziario. Nel caso di studio, la rete SOM è applicata, invece, all'anomaly detection dell'impianto sperimentale perché le mappe ottenibili risultano essere intuitive e di facile comprensione per l'identificazione delle anomalie.

Grazie alla scelta ottima dei parametri fondamentali della rete e della tipologia della griglia esagonale, nella mappa di output si preservano le relazioni topologiche di partenza e, per mezzo di diversi colori, le letture vengono proiettate in sei macroaree distinte: l'area di regime, quella del transitorio aria-acqua, quella dell'anomalia dell'acqua, dell'anomalia dell'aria, dell'anomalia fluido bifase aria-acqua e l'area dell'anomalia della perdita del serbatoio.

In caso di anomalia, dalla mappa di output si possono trarre due conclusioni: più l'intensità dell'ostruzione sarà maggiore, più la mappa proietterà la lettura verso centro dell'area di interesse, mentre se l'ostruzione sarà lieve il dato verrà mostrato lungo il bordo dell'area.

Una limitazione della rete è data dall'area del transitorio: l'algoritmo è in grado di avvertirci che il sistema non si trova più nelle condizioni di regime, ma non è capace di classificare il tipo di anomalia che si sta riscontrando. Per caratterizzare in maniera più approfondita l'area del transitorio, si potrebbe studiare come la rete classifica le letture registrate sul sistema quando su di esso vengono combinate più anomalie nello stesso istante. Tramite queste nuove prove, si potrebbero studiare in maniera più approfondita anche i singoli confini delle macroaree di output.

Per la gestione e l'individuazione delle anomalie sulla mappa, una possibile implementazione alla rete addestrata potrebbe essere quella di inserire un algoritmo decisionale in grado di restituire automaticamente il tipo di anomalia in base alla posizione del punto stesso.

Appendice

Tabella A. 1: Coordinate cartesiane dei cluster nella mappa di output

x	y	Regime	A. 3	A. 5	A. 6	A. 7	A. 8	A. 9	A. 10
5	10	338	15	3	8	10	3	0	0
6	10	19	20	0	12	0	13	12	0
5	12	77	0	0	0	0	0	0	0
6	9	11	0	0	0	1	0	0	8
6	11	0	5	0	0	0	2	0	0
5	11	0	5	0	0	0	0	0	7
4	11	0	4	0	0	0	0	0	0
0	10	0	4	0	0	0	0	0	34
4	12	0	6	0	0	0	0	0	45
5	13	0	2	0	0	0	0	0	0
5	14	0	36	0	0	0	0	0	0
3	11	0	7	0	0	0	0	0	0
2	11	0	7	0	0	0	0	0	0
1	11	0	1	0	0	0	0	0	0
3	12	0	1	0	0	0	0	0	0
4	13	0	1	0	0	0	0	0	0
4	14	0	9	0	0	0	0	0	0
0	11	0	1	0	0	0	0	0	0
4	15	0	2	0	0	0	0	0	0
3	16	0	2	0	0	0	0	0	0
3	17	0	3	0	0	0	0	0	0
1	12	0	10	0	0	0	0	0	0
2	12	0	12	0	0	0	0	0	0
3	13	0	47	0	0	0	0	0	0
1	14	0	28	0	0	0	0	0	0
0	15	0	2	0	0	0	0	0	0
2	17	0	7	0	0	0	0	0	0
3	15	0	11	0	0	0	0	0	0
0	12	0	1	0	0	0	0	0	0
2	15	0	14	0	0	0	0	0	0
2	14	0	25	0	0	0	0	0	0
0	13	0	8	0	0	0	0	0	0
0	14	0	2	0	0	0	0	0	0
0	16	0	3	0	0	0	0	0	0
0	17	0	0	5	0	0	0	0	0
0	2	0	0	47	0	0	0	0	0
1	1	0	0	24	0	0	0	0	0
1	0	0	0	11	0	0	0	0	0
4	3	0	0	17	0	0	0	0	0
2	1	0	0	5	0	0	0	0	0
2	2	0	0	4	0	0	0	0	0
3	2	0	0	3	0	0	0	0	0
4	2	0	0	14	0	0	0	0	0
4	0	0	0	1	0	0	0	0	0
5	3	0	0	4	0	0	0	0	0
6	2	0	0	1	0	0	0	0	0
6	0	0	0	2	0	0	0	0	0
7	3	0	0	10	0	0	0	0	0

8	3	0	0	1	0	0	0	0	0
7	1	0	0	6	0	0	0	0	0
8	0	0	0	4	0	0	0	0	0
7	2	0	0	4	0	0	0	0	0
8	1	0	0	1	0	0	0	0	0
8	2	0	0	1	0	0	0	0	0
7	0	0	0	0	6	0	0	0	0
6	7	0	0	0	10	0	0	0	0
0	3	0	0	0	12	0	0	0	0
5	8	0	0	0	47	0	0	0	0
5	9	0	0	0	10	0	0	0	0
3	10	0	0	0	25	0	0	0	0
3	4	0	0	0	10	0	0	0	0
0	4	0	0	0	90	0	0	0	0
5	7	0	0	0	29	0	0	0	0
4	9	0	0	0	9	0	0	0	0
4	5	0	0	0	9	0	0	0	0
4	8	0	0	0	14	0	0	0	0
4	6	0	0	0	14	0	0	0	0
0	5	0	0	0	4	0	0	0	0
3	9	0	0	0	5	0	0	0	0
1	9	0	0	0	2	0	0	0	0
4	7	0	0	0	3	0	0	0	0
3	8	0	0	0	3	0	0	0	0
2	8	0	0	0	1	0	0	0	0
0	7	0	0	0	2	0	0	0	0
3	7	0	0	0	1	0	0	0	0
2	5	0	0	0	1	0	0	0	0
0	6	0	0	0	1	0	0	0	0
1	6	0	0	0	0	1	2	9	0
6	6	0	0	0	0	2	0	0	0
6	14	0	0	0	0	18	0	0	0
7	14	0	0	0	0	5	0	0	0
8	14	0	0	0	0	53	0	0	0
11	17	0	0	0	0	22	0	0	0
6	15	0	0	0	0	75	0	0	0
5	16	0	0	0	0	14	0	0	0
7	15	0	0	0	0	26	0	0	0
6	17	0	0	0	0	5	0	0	0
6	16	0	0	0	0	2	0	0	0
7	17	0	0	0	0	2	0	0	0
10	17	0	0	0	0	1	0	0	0
7	16	0	0	0	0	1	0	0	0
8	16	0	0	0	0	0	19	0	0
9	13	0	0	0	0	0	11	0	0
12	16	0	0	0	0	0	6	0	0
13	17	0	0	0	0	0	19	0	0
10	13	0	0	0	0	0	1	0	0
16	11	0	0	0	0	0	128	0	0
10	14	0	0	0	0	0	7	0	0
15	15	0	0	0	0	0	4	0	0
12	15	0	0	0	0	0	11	0	0
12	14	0	0	0	0	0	26	0	0
13	14	0	0	0	0	0	118	0	0

17	5	0	0	0	0	0	95	0	0
17	6	0	0	0	0	0	47	0	0
17	7	0	0	0	0	0	30	0	0
16	6	0	0	0	0	0	1	0	0
15	17	0	0	0	0	0	65	0	0
17	8	0	0	0	0	0	13	0	0
14	17	0	0	0	0	0	12	0	0
17	11	0	0	0	0	0	121	0	0
16	10	0	0	0	0	0	38	0	0
17	10	0	0	0	0	0	3	0	0
16	13	0	0	0	0	0	19	0	0
15	14	0	0	0	0	0	11	0	0
11	14	0	0	0	0	0	9	0	0
14	14	0	0	0	0	0	4	0	0
13	15	0	0	0	0	0	27	0	0
16	14	0	0	0	0	0	20	0	0
16	8	0	0	0	0	0	25	0	0
17	9	0	0	0	0	0	4	0	0
17	12	0	0	0	0	0	2	0	0
17	13	0	0	0	0	0	2	0	0
16	12	0	0	0	0	0	0	3	0
7	5	0	0	0	0	0	0	26	0
9	12	0	0	0	0	0	0	5	0
15	10	0	0	0	0	0	0	97	0
15	7	0	0	0	0	0	0	5	0
8	5	0	0	0	0	0	0	24	0
7	6	0	0	0	0	0	0	18	0
7	8	0	0	0	0	0	0	16	0
13	12	0	0	0	0	0	0	3	0
11	12	0	0	0	0	0	0	49	0
10	5	0	0	0	0	0	0	78	0
9	5	0	0	0	0	0	0	10	0
8	6	0	0	0	0	0	0	9	0
8	7	0	0	0	0	0	0	24	0
9	6	0	0	0	0	0	0	52	0
8	8	0	0	0	0	0	0	32	0
9	9	0	0	0	0	0	0	15	0
9	8	0	0	0	0	0	0	16	0
10	7	0	0	0	0	0	0	17	0
10	9	0	0	0	0	0	0	17	0
11	7	0	0	0	0	0	0	6	0
10	8	0	0	0	0	0	0	24	0
11	11	0	0	0	0	0	0	14	0
10	10	0	0	0	0	0	0	23	0
12	7	0	0	0	0	0	0	8	0
12	11	0	0	0	0	0	0	1	0
13	7	0	0	0	0	0	0	6	0
13	11	0	0	0	0	0	0	8	0
11	9	0	0	0	0	0	0	7	0
11	10	0	0	0	0	0	0	5	0
12	8	0	0	0	0	0	0	1	0
12	10	0	0	0	0	0	0	115	0
13	9	0	0	0	0	0	0	0	3
6	12	0	0	0	0	0	0	0	1

7	12	0	0	0	0	0	0	0	229
12	5	0	0	0	0	0	0	0	94
13	5	0	0	0	0	0	0	0	79
14	5	0	0	0	0	0	0	0	68
15	4	0	0	0	0	0	0	0	103
11	4	0	0	0	0	0	0	0	18
13	4	0	0	0	0	0	0	0	11
14	4	0	0	0	0	0	0	0	36
12	4	0	0	0	0	0	0	0	10
14	3	0	0	0	0	0	0	0	3
12	3	0	0	0	0	0	0	0	8
11	3	0	0	0	0	0	0	0	19
13	3	0	0	0	0	0	0	0	4
10	2	0	0	0	0	0	0	0	1
10	1	0	0	0	0	0	0	0	2
11	1	0	0	0	0	0	0	0	34
11	0	0	0	0	0	0	0	0	13
11	2	0	0	0	0	0	0	0	5
12	2	0	0	0	0	0	0	0	4
12	1	0	0	0	0	0	0	0	14
12	0	0	0	0	0	0	0	0	10
13	2	0	0	0	0	0	0	0	31
13	1	0	0	0	0	0	0	0	27
14	2	0	0	0	0	0	0	0	12
13	0	0	0	0	0	0	0	0	2
15	2	0	0	0	0	0	0	0	13
17	1	0	0	0	0	0	0	0	34
15	1	0	0	0	0	0	0	0	2
14	0	0	0	0	0	0	0	0	49
15	0	0	0	0	0	0	0	0	15

Bibliografia

1. **UNI EN 13306:** *Manutenzione - Terminologia di manutenzione*. Italia, 2018.
2. **Sheldon M. Ross.** : *Probabilità e statistica per l'ingegneria e le scienze*. Apogeo, 2003.
3. **Marmo, Roberto:** *Algoritmi per l'intelligenza artificiale*. Milano : Ulrico Hoepli Milano, 2020. ISBN 978-88-203-9171-3.
4. **W. S. McCulloch, W. Pitts.:** *A Logical Calculus of the Ideas Immanent*. Nervous Activity. The bulletin of mathematical biophysics, 1943.
5. **Rosenblatt, F:** *The Perceptron, a Perceiving and Recognizing Automaton*. Cornell Aeronautical Laboratory, 1957.
6. **Raschka, Vahid Mirjalili Sebastian.** *Python Machine Learning SecondEdition*. Apogeo, 2017. ISBN 9788850333974.
7. **H. Abdi, L. J. Williams:** *Principal component analysis*. Wiley Interdisciplinary Reviews Computational Statistics, 2010, Vol. 2. ISSN 1939-5108.
8. **Hout, Michael, Papesh, Megan, Goldinger, Stephen:** *Multidimensional scaling*. WIREs Cognitive Science, 2013, Vol. 4. ISSN 1939-5078.
9. **Umut A., Secil E.** *Computational Intelligence Systems in Industrial Engineering: with Recent Theory and Applications*. Atlantis Press, 2012.
10. **Shalaginov, Andrii, Franke, Katrin:** *A New Method for an Optimal SOM Size Determination in Neuro-Fuzzy for the Digital Forensics Applications*. 2015. ISBN 978-3-319-19221-5.
11. **T. Kohonen:** *Self-Organizing Maps*. Berlino : Springer, 2001.
12. **G. Pözlbauer:** *Proceedings of the Fifth Workshop on Data Analysis*. Survey and Comparison of Quality Measures for Self-Organizing Maps. Slovacchia, 2004.