



UNIVERSITA' POLITECNICA DELLE MARCHE
FACOLTA' DI INGEGNERIA
Corso di laurea triennale in
INGEGNERIA INFORMATICA E DELL'AUTOMAZIONE

**Analisi e modellazione per forni di riscaldamento nell'industria
dell'acciaio**

Analysis and modelling for reheating furnaces in the steel industry

Relatore:

Prof. Silvia Maria ZANOLI

Correlatori:

Ing. Giacomo ASTOLFI

Ing. Crescenzo PEPE

Laureando:

Angelo DI CARLO

Sommario

Capitolo 1 Descrizione del processo.....	6
1.1. Impianto di lavoro	8
1.2. Sensoristica.....	14
1.1.1. Pirometri.....	14
1.1.2. Fotocellule:	14
1.1.3. Assorbimenti gabbie di laminazione	15
Capitolo 2 Scopo del lavoro.....	16
Capitolo 3 Concetti teorici.....	20
3.1. Regressione lineare	20
3.2. Bontà dell'adattamento	22
3.3. Coefficiente di determinazione	24
3.4. Scelta del modello	26
3.5. Indici e statistiche.....	29
3.5.1. P-value	29
3.5.2. t-Test.....	30
3.5.3. Livello di significatività.....	30
3.6. Analisi dei residui.....	30
Capitolo 4 Analisi assorbimento delle gabbie	32
4.1. Analisi di un modello lineare	32
4.2. Modello su un sottoinsieme di dati.....	33
4.3. Modello multivariato	37
4.3.1. Analisi residui.....	43
4.4. Confronto del modello su diversi periodi.....	45
4.5. Modello per stima degli assorbimenti della giornata.....	47
Capitolo 5 Problematica allineamento testa - coda	53
Capitolo 6 Conclusioni	62

Introduzione

L'energia è una componente sempre più presente in quasi tutti i modelli di innovazione tecnologica. Rappresenta un aspetto fondamentale per il moderno sviluppo ma, è doveroso ricordare che, anche l'energia è causa dell'attuale crisi ambientale. Per questo motivo, una politica aziendale consapevole deve cercare di ridurre al massimo gli sprechi e perseguire una "Green Economy energy policies".

Il termine automazione identifica la tecnologia che usa sistemi di controllo (come circuiti logici o elaboratori) per gestire macchine e processi, riducendo la necessità dell'intervento umano. A partire dalla rivoluzione industriale sono stati richiesti, agli ingegneri e alle industrie, degli enormi investimenti nell'automazione che hanno permesso il raggiungimento di notevoli progressi tecnologici.

Negli ultimi anni, si registra una crescente necessità di innovazione dei processi produttivi. Queste richieste sono dovute dalla necessità di garantire un compromesso ottimale tra gli obiettivi produttivi, le specifiche qualitative, l'efficienza energetica e, soprattutto, il rispetto dei rigorosi standard ambientali.

In questo elaborato, grazie alla collaborazione con l'azienda Alperia Bartucci S.p.A. è stato possibile analizzare un progetto di sviluppo di sistemi di controllo avanzati di forno di riscaldamento di semiprodotto. I forni di riscaldamento sono utilizzati, in questo ambito, per riscaldare billette o lastre provenienti da processi di colata continua e destinati alla produzione di prodotti come tondini di acciaio, travi e

altri profilati. Queste vengono lavorate secondo specifiche curve di riscaldamento per garantire il corretto funzionamento durante la fase di laminazione.

Il sistema di controllo avanzato APC (Advanced Process Control) è basato su tecniche di controllo predittivo (Model Predictive Control o MPC), che tengono conto della multivariabilità dei sistemi e della presenza di vincoli sulle variabili e sul funzionamento. Il sistema attualmente garantisce che la temperatura delle billette, in uscita dal forno, ricada in un certo range evitando malfunzionamenti, come la rottura delle gabbie stesse.

Nel particolare, in questa tesi ci siamo interessati ad applicazioni per forni di riscaldamento, riguardanti nel particolare la ricerca di una correlazione tra temperatura di uscita delle billette dal forno e gli assorbimenti delle gabbie di lavorazione. Inoltre, si è cominciato a valutare in che modo si potesse ottenere temperatura all'uscita della 3° gabbia uniforme, andando a controllare la differenza del valore dei bruciatori, nella zona finale del forno.

Il sistema di controllo avanzato, al momento, utilizza come set-point (o obiettivo di controllo) le temperature di uscita desiderate delle billette. Il valore di set-point viene settato in base ad informazioni che derivano dagli operatori degli impianti di laminazione.

I valori ricavati automaticamente dal sistema, sulla base degli assorbimenti delle gabbie di lavorazione, sono quelli di reale interesse e che sarebbe auspicabile sostituire alle informazioni di set-point.

Gli operatori fissano il valore dei set point in base alle loro conoscenze ed esperienze riguardo il forno, pertanto i set-point potrebbero essere impostati eccessivamente alti, in modo da essere conservativi, evitare il rischio di rottura delle gabbie, ma si consuma maggiormente.

In ogni caso, il controllore si pone l'obiettivo di ottimizzare la produzione, riducendo il più possibile i set-point dei bruciatori all'interno del forno, questo per minimizzare la quantità di combustibile, mantenendo comunque gli assorbimenti delle macchine sotto una certa soglia. Se fosse possibile lavorare direttamente con gli assorbimenti, si potrebbero migliorare notevolmente le prestazioni del sistema, soprattutto in ambito di risparmio energetico.

La tesi è così organizzata: nel primo capitolo viene descritto il processo di laminazione a caldo e la descrizione dell'impianto del quale sono stati analizzati i dati. Il secondo capitolo riguarda lo scopo di questo lavoro di tesi, e le principali problematiche affrontate. Nel terzo capitolo vengono

introdotti i concetti teorici sul quale si basa l'analisi del lavoro svolto, che è argomento del quarto e quinto capitolo; le conclusioni infine vengono esposte al capitolo .

Capitolo 1

Descrizione del processo

Il processo di ottenimento di alcuni tipici prodotti dell'industria siderurgica quali, ad esempio, lastre di acciaio, tondini, barre di acciaio con profili ad H,L, etc. sono ottenuti dalla laminazione di semilavorati quali *bramme (slabs)*, *billette (billets)* o *blumi (blooms)* tutti derivati da processi di colata continua che si differenziano per forma e dimensioni. La laminazione fa parte delle lavorazioni per deformazione plastica (deformazione permanente), alle quali è sottoposto il metallo che si trova in una opportuna forma e condizione.

Lo scopo della laminazione è quello di far assumere ai prodotti forma e dimensioni finali diverse ottenendo, ad esempio, lastre di acciaio, tondini, barre di acciaio con profili ad H,L, etc. (per approfondimenti, si veda [2],[3]). Questo procedimento avviene mediante uno sforzo di compressione del semilavorato fra due rulli contrapposti, rotanti in verso opposto tra di loro, che ne generano uno schiacciamento, allungamento e la forma desiderata.

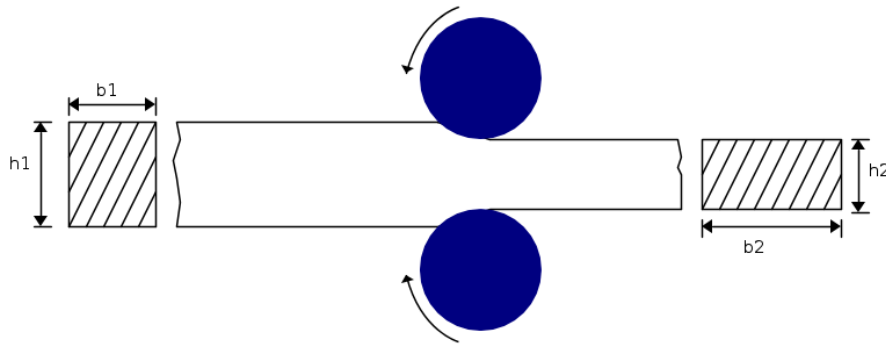


Figura-1 Rappresentazione del processo di laminazione

Ci sono due principali metodologie di laminazione: a caldo e a freddo. La laminazione è detta a caldo nel momento in cui il prodotto, prima di essere lavorato, viene riscaldato per essere portato alla temperatura di massima plasticità; quella a freddo invece è un processo di lavorazione meccanica che opera con il materiale a temperatura ambiente.

I rulli sono mantenuti in posizione di lavoro all'interno di una gabbia detta gabbia di laminazione. Le gabbie possono essere di tipo continuo (il materiale passa una sola volta) o di tipo reversibile (il materiale passa più volte, invertendo il senso di avanzamento).

Possono essere dotate di più tipi di rulli: i rulli di lavoro (sempre presenti) ed i rulli di appoggio. I rulli di lavoro sono quelli che deformano il materiale, mentre la funzione dei rulli di appoggio è quella di evitare l'incurvatura dei cilindri di lavoro nella laminazione delle bramme.

Le gabbie sono disposte a gruppi in stretta successione, a costituire un treno di laminazione. Gli impianti sono formati da diversi treni e possono essere dotati di vie a rulli, di dispositivi per la rotazione (ribaltatori), per la traslazione (manipolatori e trascinatori), per il taglio a caldo (cesoie, seghe o moli), circuiti d'acqua destinati al raffreddamento del prodotto (laminazione a temperatura controllata), placche o tappeti di raffreddamento[4].

Nel caso di laminazione a caldo, in genere, a monte del processo di laminazione si deve provvedere ad un opportuno trattamento di questi semilavorati derivanti dalla colata continua che consiste nel far transitare le *bramme*, le *billette* o i *blumi* in specifici forni di riscaldamento. In tali forni i semilavorati vengono riscaldati secondo precisi profili di temperatura.



Figura 2 Colata dell'acciaio impianto Feralpi

1.1. Impianto di lavoro

L'impianto, oggetto di studio in questa tesi è relativo allo stabilimento "Feralpi Siderurgica S.p.A." nel quale è attualmente installato un controllore APC, sviluppato con tecniche di controllo predittivo.

Nel presente impianto, le billette sono ottenute mediante colata continua e vengono preriscaldate (laminazione a caldo) in un forno di riscaldamento di tipo "Pusher type". Pusher type indica che l'avanzamento delle billette dentro il forno è dovuto alla spinta (push) di uno spintore, che spingendo le billette inserite per ultime nel forno, fa traslare le altre nelle zone successive, diversamente da come accade ad esempio nei forni "Walking Beam" dove lo spostamento delle billette è ottenuto mediante un meccanismo che alza le singole billette e le trasla in avanti.

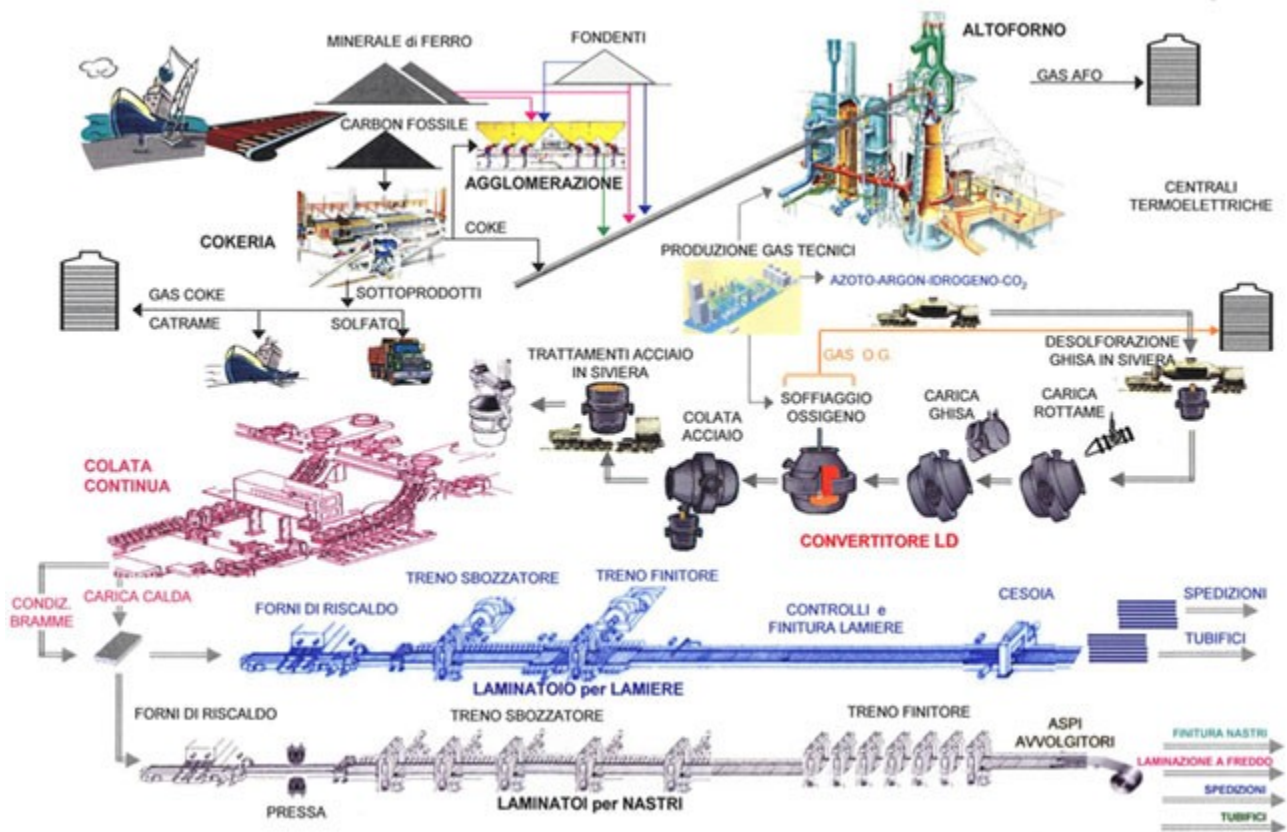


Figura 3 Stabilimento ILVA di Taranto – Schema parziale del processo produttivo: dalle materie prime alla laminazione a caldo con produzione coils e lamiere.
<http://astrolabio.amicidellaterra.it/node/1539>

Il forno dello stabilimento di Feralpi può contenere fino a 138 billette. Le billette possono entrare a coppia o singolarmente; mentre ne escono una per volta in quanto, alle gabbie, le billette vengono lavorate singolarmente. La billetta uscente dal forno può essere una tra quelle posizionate nelle ultime tre posizioni del forno “finali” del forno, in base alle esigenze.

L’ingresso delle billette nel forno è condizionato dalle billette in uscita: poiché è un forno Pusher type, nel forno non ci saranno zone senza billette (eccezion fatta nei momenti in cui viene tolta una billetta, e non ne sono state ancora reinserite altre), ma ha comunque una capienza limitata che ne vincola l’inserimento.

Prima di entrare nel forno le billette passano (a coppie o se è una sola, singolarmente) di fronte due pirometri, e ne viene acquisita la temperatura di ingresso.

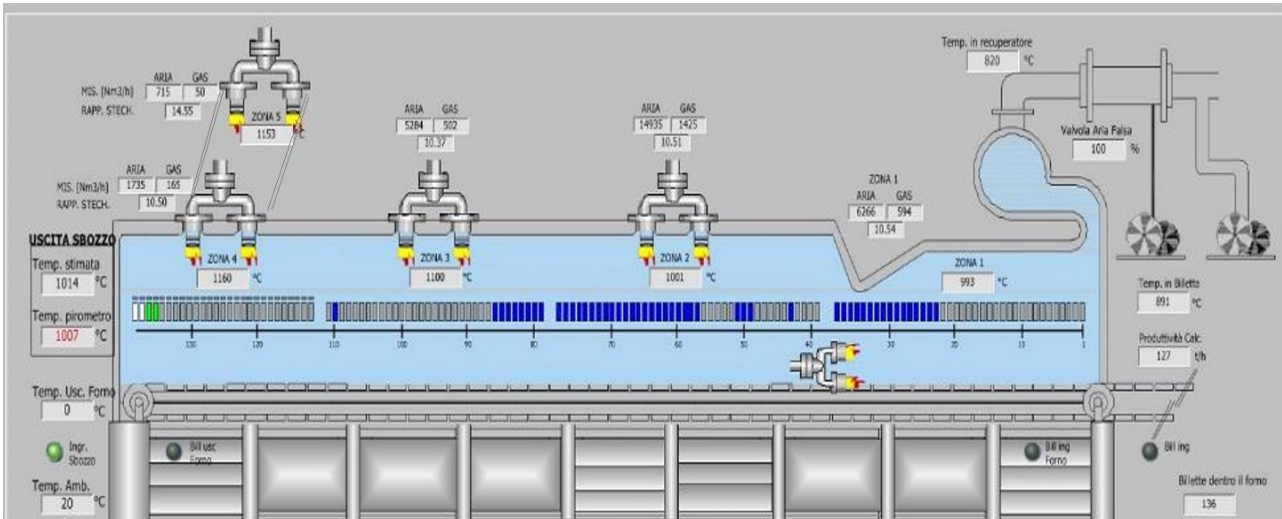


Figura 4 Rappresentazione del forno di riscaldamento

Le billette che vengono infornate sono tutte con le medesime caratteristiche, a prescindere dal tipo di prodotto che si andrà a realizzare: hanno sezione 150mm x 150 mm , lunghezza 1.2m e peso di 2155 Kg, il range di temperatura all'ingresso del forno invece è variabile e si trova tra gli 0° e 910°.

Il forno in questione è diviso in quattro regioni; la divisione è dovuta per contraddistinguere le caratteristiche che le caratterizza, principalmente relative ai differenti range dei set point e alla posizione dei bruciatori. Le informazioni sono riportate nella tabella seguente.

Regione del Forno	Numero di billette presenti	Lunghezza [m]	Range dei set point di temperatura [°C]
Regione 1	35	5495	825-1020
Regione 2	37	5735	925-1090
Regione 3	31	4710	1025-1180
Regione 4	25	3925	1115-1240

Tabella 1

Come si vede dalla figura 4, le prime tre regioni del forno sono caratterizzate da un'unica zona di riscaldamento, mentre la quarta ne ha due (chiamate Zona 4, Zona 5).

La prima regione è preposta al processo di preriscaldamento, nella seconda e terza avviene il processo di riscaldamento che si conclude nella quarta dove abbiamo la fase di equalizzazione/condizionamento (soaking) delle billette. Quest'ultima regione viene divisa in due zone definite Zona 4 e Zona 5: per ognuna di queste si ha un bruciatore, utilizzato per riscaldare la porzione sinistra della billetta e l'altro quella destra che rappresentano rispettivamente la testa e coda della billetta. Testa e coda della billetta fanno riferimento alla parte della billetta che, rispettivamente, viene processata per prima e per ultima dalle gabbie di laminazione posizionate in uscita al forno.

Viene utilizzata questa struttura poiché risulta necessario portare ad una temperatura maggiore la coda rispetto la testa, considerato che la coda, prima di arrivare alle gabbie di laminazione trascorrerà un tempo (variabile) fuori dal forno superiore a quello della testa che ne causerà una progressiva diminuzione di temperatura. Con una opportuna gestione delle due zone distinte, Zona 4 e Zona 5, si può cercare di uniformare la temperatura di lavoro della intera billetta.

La combustione del forno è regolata da regolatori di temperatura PID locali che agiscono sul rapporto aria/combustibile della relativa zona. Prima dell'installazione del sistema di controllo avanzato APC, i set-point di temperatura PID erano regolati dagli operatori di stabilimento sulla base della loro esperienza e capacità. A causa dell'ampia gamma di variabilità della temperatura delle billette in ingresso al forno (0-910 ° C) e del tasso di produzione del forno caratterizzato da forti oscillazioni (0-170 tonnellate all'ora), la conduzione manuale del forno risulta complessa e richiede operatori esperti: l'introduzione del sistema di controllo APC basato su metodologie di controllo predittivo ha permesso il miglioramento dell'efficienza energetica del forno[5].

All'uscita del forno è presente un pirometro ottico la cui misura non è attendibile per le impurità presenti sulla superficie esterna della billetta detta *scaglia*. Tale temperatura infatti non è utilizzata nel controllore APC e di conseguenza non è stata considerata nell'analisi[6],[7].

Spinte fuori dal forno, ogni billetta viene ruotata, traslata e, mediante un rullo, diretta verso il descagliatore che rimuove la scaglia di laminazione con getti d'acqua ad alta pressione sulla superficie. Successivamente la billetta viene diretta verso il treno sbozzatore, dove subirà il primo trattamento di laminazione a caldo.



Figura 5 Scaglia superficiale <http://www.siddikkardesler.com/en/index.php/production/rolling-mill>

Quando si riscalda l'acciaio, sulla sua superficie si forma la scaglia di laminazione che in sostanza è ossido di ferro, causata alla reazione del ferro con l'ossigeno nell'aria che oltre a comprometterne l'aspetto, ne compromette le proprietà in modo persistente. Pertanto, nella produzione di manufatti a caldo è inevitabile che tali impurità debbano essere rimosse dalla superficie metallica.

Il processo di laminazione è suddiviso in tre fasi differenti. La prima effettua la cosiddetta fase di sgrossatura o sbozzatura, attraverso il passaggio di sei differenti gabbie. La sbozzatura è la fase di laminazione che conferisce alla barra incandescente una forma grossolana ma simile a quella definitiva. Successivamente vi sono altre due unità di laminazione per un totale di ulteriori dodici gabbie, sei "intermedie" e sei di finitura. L'oggetto di studio di questa tesi ha riguardato, per quanto riguarda la fase di laminazione, la sola unità di sgrossatura.



Figura 6 Treno di laminazione <https://www.alfacciai.it/uploads/2018-11-16/1000-0/alfa-acciai-processi-laminazione-a-caldo-21.jpg.jpg>

Dopo che la billetta esce dalla terza gabbia di laminazione, viene acquisita la temperatura della stessa mediante un pirometro ottico situato tra la terza e la quarta gabbia.

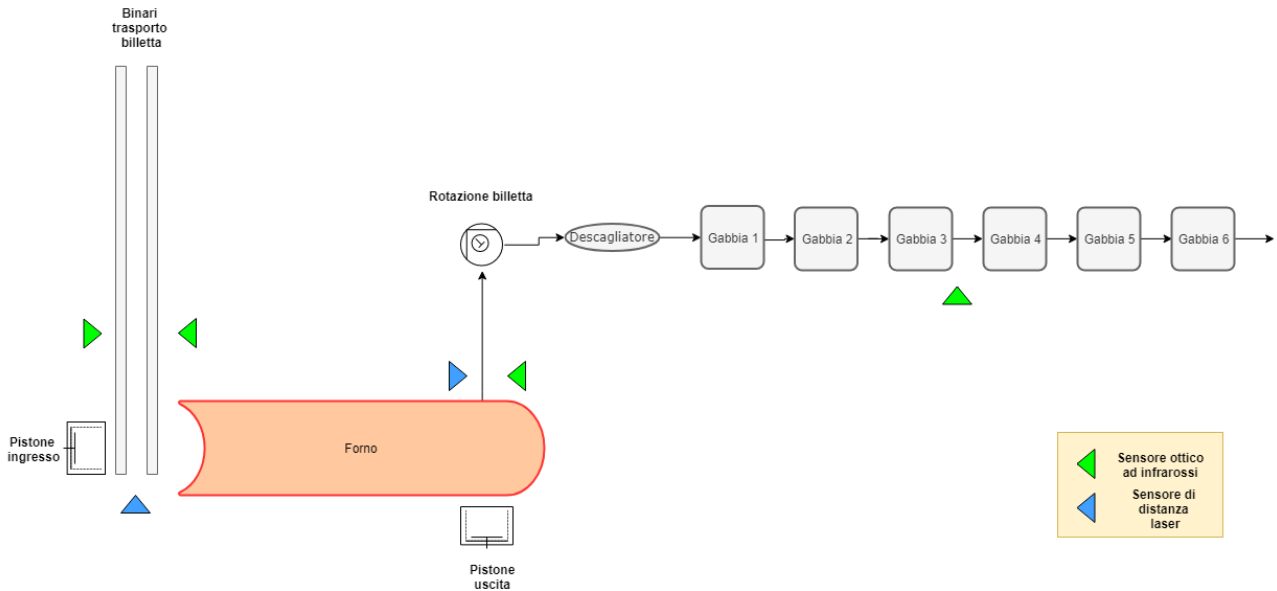


Figura 7 Schema del processo produttivo dello stabilimento Feralpi

1.2. Sensoristica

Nell'impianto sono presenti diversi sensori per monitorare e controllare il processo. Nel seguito sono riportate le loro caratteristiche generali.

1.2.1 Pirometri

Tipologie dei sensori utilizzati ed info su ognuno di essi

- Pirometri ingresso forno: Ottico ad infrarossi - SITEL, MINIRED SKS T14-09
- Pirometro uscita forno: Ottico ad infrarossi - SITEL, MINIRED STS-TE 5-12
- Pirometro dopo gabbia 3: Ottico ad infrarossi -LAND-SYSTEM, M2 300/1100

In Figura 8 e Figura 9 sono riportate le specifiche dei due pirometri, in ingresso e in uscita al forno.

	SKS-S-18	SKS-S-25	SKS-GA-13	SKS-GA-18	SKS-T4-10	SKS-T5-12	SKS-T5-25	SKS-T14-09
Campo di Misura	650°C...1800°C 1202°F...3272°F	800°C...2500°C 1472°F...4532°F	300°C...1300°C 572°F...2372°F	350°C...1800°C 662°F...3272°F	200°C...1000°C 392°F...1832°F	100°C...1200°C 212°F...2192°F	400°C...2500°C 752°F...4532°F	-32°C...900°C -25.6°F...1112°F
Risposta spettrale	0.8...1.1 μm	0.8...1.1 μm	1.45...1.8 μm	1.45...1.8 μm	3.9 μm	5.14 μm	5.14 μm	8...14 μm
Rapporto ottico	100:	150:	100:	100:	33:	50:	50:	50:
Tempo di risposta t90	20ms	20ms	20ms	20ms	100ms	100ms	100ms	100ms
Accuratezza	0.5% della lettura in °C+1K	0.5% della lettura in °C+1K	0.5% della lettura in °C+1K	0.5% della lettura in °C+1K	1% della lettura in °C+1K	1% della lettura in °C+1K	1% della lettura in °C+1K	1% della lettura in °C+1K
Ripetibilità	0.1% della lettura in °C+1K	0.1% della lettura in °C+1K	0.1% della lettura in °C+1K	0.1% della lettura in °C+1K	0.5% della lettura in °C+1K	0.5% della lettura in °C+1K	0.5% della lettura in °C+1K	0.5% della lettura in °C+1K
Emissività	5...100%	5...100%	5...100%	5...100%	20...100%	20...100%	20...100%	20...100%
Puntamento luminoso	Si	Si	Si	Si	-	-	-	-
Memorizzazione picco	Si	Si	Si	Si	-	-	-	-
Optica	I,II,III	I,II,III,IV	I,II,III	I,II,III	IV, V, VI	IV, V, VI	IV, V, VI	VII, VIII, IX
Uscita lineare	4-20 mA	4-20 mA	4-20 mA	4-20 mA	4-20 mA	4-20 mA	4-20 mA	4-20 mA
Carico max	700Ω (24V)	700Ω (24V)	700Ω (24V)	700Ω (24V)	700Ω (24V)	700Ω (24V)	700Ω (24V)	700Ω (24V)
Alimentazione	24V DC ± 25%	24V DC ± 25%	24V DC ± 25%	24V DC ± 5%	24V DC ± 25%	24V DC ± 25%	24V DC ± 25%	24V DC ± 25%
Protezione	P 65	P 65	P 65	P 65	P 65	P 65	P 65	P 65
Temperatura operativa	0...70°C 32...158°F	0...70°C 32...158°F	0...70°C 32...158°F	0...70°C 32...158°F	0...70°C 32...158°F	0...70°C 32...158°F	0...70°C 32...158°F	0...70°C 32...158°F
Temperatura di stoccaggio	-20...70°C -4...158°F	-20...70°C -4...158°F	-20...70°C -4...158°F	-20...70°C -4...158°F	-20...70°C -4...158°F	-20...70°C -4...158°F	-20...70°C -4...158°F	-20...70°C -4...158°F
Custodia	Acciaio	Acciaio	Acciaio	Acciaio	Acciaio	Acciaio	Acciaio	Acciaio
Peso	c.a. 450g.	c.a. 450g.	c.a. 450g.	c.a. 450g.	c.a. 450g.	c.a. 450g.	c.a. 450g.	c.a. 450g.

Figura 8 Pirometro SITEL, MINIRED SKS T14-09

Thermometer Description	Model N°	Wavelength (μm)	Range	Minimum Target Dia* (mm/in)
M2 Thermometers				
M2 thermometers use the latest generation of germanium detectors and operate at a wavelength of 1.6μm. They extend the measurement range of short wavelength thermometers down to 300°C/600°F and have a fast response time of 5ms.	M2 300/1100C M2 600/2000F	1.6	300 to 1100°C 600 to 2000°F	0.9/0.04 0.9/0.04

Figura 9 Pirometro LAND-SYSTEM, M2 300/1100

1.2.2 Fotocellule:

Una volta passate davanti ai pirometri posti sul piano a rulli, le billette vengono rilevate dal sensore di distanza laser posto davanti il forno [8]. Anche in uscita forno è presente una fotocellula che rileva il passaggio delle billette in uscita [9]. Si è rilevato che il segnale della fotocellula in uscita del forno era poco affidabile. Per porre rimedio è stato sfruttato il segnale del pirometro all'uscita del forno.

Non è stato possibile avere informazioni dettagliate sulle fotocellule.

1.2.3 Assorbimenti gabbie di laminazione

Le prestazioni delle gabbie di laminazione monitorate mediante le misure degli assorbimenti in corrente. Il nuovo sistema APC di controllo avanzato non ha accesso diretto alla misura del sensore ma questo dato viene fornito all'APC dal sistema di gestione dell'impianto di livello superiore.

Capitolo 2

Scopo del lavoro

Come già specificato nell'introduzione, in questa tesi ci siamo interessati a due problematiche del forno di riscaldamento:

- trovare un modello che metta in relazione gli assorbimenti delle prime sei gabbie, dopo il descagliatore, con le variabili del processo.
- Analizzare le variabili che influenzano le temperature di testa e coda delle billette rilevate all'uscita della terza gabbia

Per quanto riguarda il primo punto, in genere si osserva che all'aumentare della temperatura di uscita dei semilavorati si hanno assorbimenti decrescenti.

Attualmente, il sistema di controllo predittivo (MPC) controlla le temperature delle varie zone del forno [2], cercando di garantire alle billette un opportuno profilo di temperatura e di portare la loro temperatura all'uscita del forno ad un determinato valore di set-point.

Tale valore di set-point è imposto dagli operatori, in base alla loro esperienza e in funzione della tipologia di produzione da produrre; tutto allo scopo di mantenere bassi gli assorbimenti delle gabbie di laminazione.

Per il corretto funzionamento delle gabbie di laminazione e, in parte, per l'ottenimento di prodotti di qualità gli assorbimenti delle gabbie devono risultare limitati al di sotto di un certo valore.

Nel momento in cui le billette non sono scaldate adeguatamente, gli assorbimenti si alzano molto e, nel caso in cui superino una prefissata soglia, può determinare la rottura della gabbia stessa. Come conseguenza si ha che, per prevenire incidenti, evitando al massimo il rischio di rottura delle gabbie e limitando gli assorbimenti, spesso gli operatori preferiscono operare in maniera "conservativa", e impostano set-point eccessivamente alti, a discapito dei consumi -.Questo comporta degli sprechi dal punto di vista energetico con conseguenze sia sul lato economico che ambientale.

D'altra parte, il controllore si pone l'obiettivo di ottimizzare la produzione, riducendo il più possibile i set-point dei bruciatori all'interno del forno, questo per minimizzare la quantità di combustibile, mantenendo comunque gli assorbimenti delle macchine sotto una certa soglia.

Simili considerazioni avevano portato all'introduzione del sistema di controllo avanzato APC. Infatti, si ha che il processo è caratterizzato a un'ampia variabilità di condizioni operative: per la tipologia e per la temperatura in ingresso delle billette, per la velocità di produzione. Questo fatto, insieme all'elevata interconnessione tra le variabili che caratterizza questo processo, tutto ciò influisce negativamente sulla capacità dell'operatore di comprendere appieno il processo di riscaldamento delle billette, e quindi le implicazioni sugli effettivi assorbimenti in fase di laminazione.

A tale proposito si è pensato che se il controllore fosse in grado di bypassare l'informazione dei set point di temperatura alla laminazione derivante gli operatori e lavorare direttamente sui dati degli assorbimenti, si potrebbero migliorare notevolmente le prestazioni del sistema, soprattutto in ambito di risparmio energetico.

Per quanto riguarda il secondo punto invece, lo scopo è riuscire ad avere una lettura di temperatura del pirometro uniforme, "piatta" tra testa e coda a valle della terza gabbia. Il motivo che ci spinge a tale richiesta è che, in questo modo, risulta che la testa e la coda subiscono il processo di laminazione alla stessa temperatura.

Questo è importante perché, essendo la billetta molto lunga, il tempo che intercorre tra la lavorazione della testa e quello della coda può essere significativo. Questo implica che, una volta uscita dal forno, la temperatura della coda decresce progressivamente prima che venga lavorata, ovviamente aumentandone gli assorbimenti.

Il valore relativo agli assorbimenti è un valore unico per ogni billetta poiché opportunamente processato; in questa fase però non c'è stata la necessità dell'introduzione degli assorbimenti.

Si ottiene uno sbilanciamento della testa e della coda andando ad agire sui bruciatori nell'ultima zona del forno (vedi capitolo 2): si fissano i set point relativi alla coda più alti rispetto quelli della testa, in modo che pur essendo il tempo di raffreddamento della coda maggiore, la temperatura con la quale verrà lavorata sia quella desiderata.

Per la risoluzione di questa seconda problematica è necessario l'ottenimento di un modello che, in base alle caratteristiche della billetta e della produzione corrente, riesca a determinare automaticamente il valore del set point da impostare affinché si abbia il risultato desiderato.

Per risolvere le problematiche sopra discusse si sono analizzati i dati storici del funzionamento del forno di riscaldamento e delle prime sei gabbie di laminazione. Tali dati riguardano un periodo di circa tre mesi che va da febbraio 2019 ai primi di maggio 2019.

Molti dati presenti nel file log risultano allineati temporalmente con le singole billette. Ogni volta che una billetta esce dal primo treno di laminazione (sesta gabbia) in un certo set di dati di log vengono memorizzate le principali informazioni relative alla singola billetta quali ad esempio temperatura di uscita alle gabbie, temperatura in ingresso al forno, assorbimenti nelle singole gabbie.

Tali dati però risultano filtrati e spesso sono forniti solo i valori medi di segnali nel tempo. Ad esempio, per chiarire tale problematica la temperatura in ingresso memorizzata su tali file log associata alle singole billette è un valore filtrato in funzione delle temperature delle billette adiacenti. Questo per motivi legati ad esigenze di controllo che non è detto che siano in accordo con quelle legate alla problematica oggetto del presente lavoro di tesi. Un altro esempio è la temperatura di uscita alla terza gabbia: anche questo è un unico valore associato ad ogni billetta ottenuto a seguito di opportuni processamenti delle misure del pirometro. Si capisce chiaramente

che dovendo affrontare il problema dello sbilanciamento testa coda tale informazione non risulta adeguata.

Si è quindi dovuto procedere in maniera autonoma a ricreare tale allineamento ricostruendo il tracking della billetta partendo dal suo ingresso nel forno. Tale compito ha richiesto un certo dispendio di tempo richiedendo una più approfondita conoscenza del sistema di controllo sviluppato e della modalità di avanzamento delle billette.

Capitolo 3

Concetti teorici

3.1 Regressione lineare

La regressione lineare è un tipo di analisi predittiva volto a stimare il valore atteso di una variabile dipendente y condizionato da una o più variabili indipendenti x_1, \dots, x_m . La relazione tra le variabili indipendenti e quella dipendente può essere espressa come:

$$y = f(x_1, x_2, x_3, \dots, x_m) + \varepsilon = f(\mathbf{X}) \quad (1)$$

Che esprimendola in una relazione lineare risulta:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_m x_m + \varepsilon \quad (2)$$

dove abbiamo che le x_i rappresentano le variabili indipendenti, ε è l'**errore**, β_0 è l'intercetta mentre i $\beta_{i=1,m}$ sono i coefficienti della retta di regressione.

$$\varepsilon_i = y_i - \beta_0 - \beta_1 x_1 \text{ (caso con singolo regressore)} \quad (3)$$

Noi però non siamo in grado di ottenere con esattezza i valori dei coefficienti β_i per diversi motivi quali ad esempio errori di misurazioni, variabili non incluse nel modello, etc. Quello che otteniamo è una stima del valore dei coefficienti $\hat{\beta}_i$, ed il relativo modello è il seguente:

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 x_2 + \dots + \hat{\beta}_m x_m \quad (4)$$

Per ogni punto osservato andiamo a definire **residuo** la differenza tra valore reale e valore stimato per l'osservazione i-esima:

$$e_i = y_i - \hat{y}_i \quad (5)$$

Un esempio di valori reali e stimati sono riportati nella figura 10 dove si hanno i valori celesti che sono quelli misurati, mentre quelli rossi sono i valori ottenuti dalla regressione lineare; la distanza verticale dei valori reali dalla retta rappresenta il residuo i-esimo.

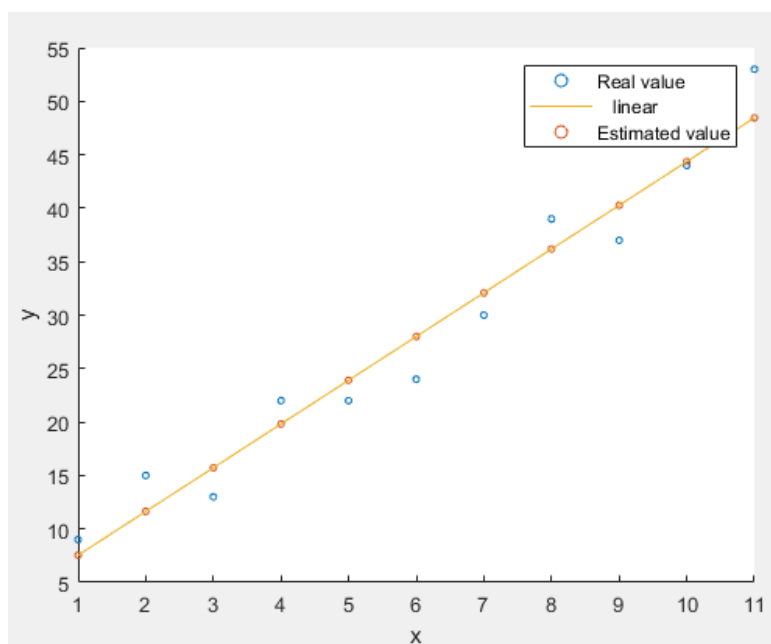


Figura 10 Esempio regressione lineare su insieme di dati

È detta retta dei minimi quadrati, quella che rende minima la somma dei residui al quadrato (Residual Sum of Squares)

$$RSS = \sum_{i=1}^n (y_i - \hat{y}_i)^2 = \sum_{i=1}^n e_i^2 \quad (6)$$

Considerando il caso con solo x_i come unica variabile predittiva, per ottenere la retta che minimizza l'errore quadratico esprimiamo e_i in funzione di $\hat{\beta}_0 + \hat{\beta}_1$

$$e_i = y_i - \hat{y}_i = y_i - \hat{\beta}_0 + \hat{\beta}_1 x_1 \quad (7)$$

$$RSS = \sum_{i=1}^n (e_i)^2 = \sum_{i=1}^n (y_i - \hat{\beta}_0 + \hat{\beta}_1 x_1)^2 \quad (8)$$

Quindi $\hat{\beta}_0, \hat{\beta}_1$ sono le quantità che minimizzano la formula somma, calcolate come segue:

$$\hat{\beta}_1 = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2} \quad (9)$$

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x} \quad (10)$$

di cui omettiamo la dimostrazione[10].

3.2 Bontà dell'adattamento

Dividendo SSR per il numero di campioni n , otteniamo MSE (Mean Square Error) che indica la discrepanza quadratica media fra i valori dei dati osservati ed i valori dei dati stimati. Più il valore MSE è vicino allo zero, più la nostra regressione è accurata, ma non abbiamo un valore soglia che ci certifica la "qualità" della regressione, è un indice più utile per comparare regressioni diverse. Per ovviare a questa limitazione introduciamo il **coefficiente di correlazione**.

Considerando due variabili (x,y) , introduciamo il coefficiente di correlazione che è una misura della forza dell'associazione lineare tra le due ed è indicato con r . Esso può assumere valori continui compresi tra -1 e +1; un valore positivo indica che le variabili (x,y) sono direttamente correlate (o correlate positivamente) e che quindi, ad un aumento della variabile x , aumenta la variabile y , nel caso negativo invece, inversamente correlate (o correlate negativamente) e all'aumento di x , la y diminuisce (e viceversa).

$$r = \frac{1}{n-1} \sum_{i=1}^n \frac{\sigma_{x,y}}{\sigma_x \sigma_y} = \sum_{i=1}^n \frac{(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}} \quad (11)$$

Dove σ_x, σ_y sono la deviazione standard relativa alla variabile x e y

$$\sigma_x = \sqrt{\frac{(x_i - \bar{x})^2}{n - 1}} \quad (12)$$

$$\sigma_y = \sqrt{\frac{(y_i - \bar{y})^2}{n - 1}} \quad (13)$$

mentre $\sigma_{x,y}$ è la covarianza tra x e y.

$$\sigma_{x,y} = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}) \quad (14)$$

In base al valore assunto da r, si ha una correlazione più o meno forte:

- $r = 0$ variabili incorrelate
- $0 < |r| < 0,3$ correlazione debole
- $0,3 < |r| < 0,7$ correlazione moderata
- $|r| > 0,7$ correlazione forte

Il segno di r indica se la correlazione è positiva o negativa

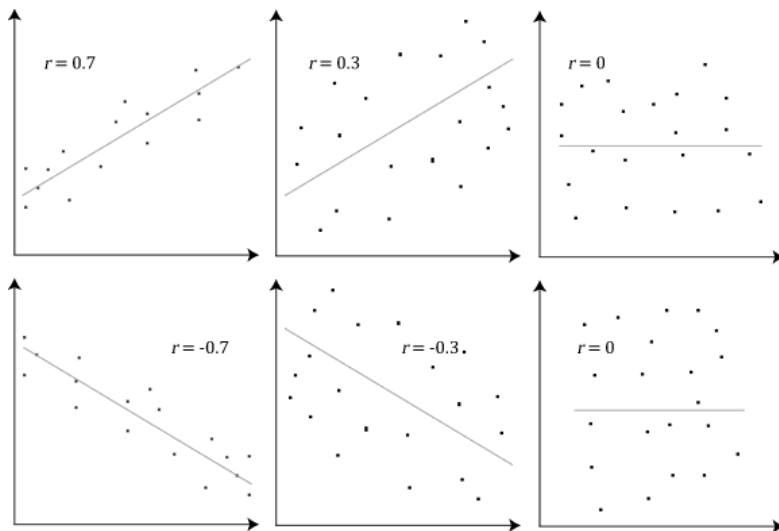


Figura 11-Esempi di correlazione positiva negativa e nulla

È importante ricordare che il coefficiente r non varia se

- Moltiplico ogni valore per una costante positiva

- Sommo una costante a ogni valore
- Scambio i valori di x con quelli di y

3.3 Coefficiente di determinazione

Una volta ottenuto il modello di regressione lineare dei dati è bene avere una misura della bontà della regressione; un modo per quantificare quanto la regressione giustifica i dati è il coefficiente di determinazione r^2 , più il valore r^2 è prossimo ad 1 e più il modello spiega correttamente i dati della variabile dipendente.

$$r^2 = \frac{\sum_{i=1}^n (y_i - \bar{y}) - \sum_{i=1}^n (y_i - \hat{y}_i)}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (15)$$

$$r^2 = \frac{ESS}{TSS} = 1 - \frac{RSS}{TSS} \quad (16)$$

dove:

- $ESS = \sum_{i=1}^n (\hat{y}_i - \bar{y})^2$ (17) devianza spiegata dal modello (Explained Sum of Squares)
- $TSS = \sum_{i=1}^n (y_i - \bar{y})^2$ (18) devianza totale (Total Sum of Squares)
- $RSS = \sum_{i=1}^n (y_i - \hat{y}_i)^2$ (8)
- y_i dato osservato, \bar{y} media dei dati osservati, \hat{y}_i dato stimato.

Poiché la regressione creerà un modello che minimizza i residui, sembra sensato confrontare i residui dei nostri modelli. Come si evince dall'ultima uguaglianza infatti, l'indice r^2 misura l'errore quadratico medio sia sul nostro modello lineare che su un modello definito "baseline" che ad ogni x risponde con un valore costante pari alla media dei valori assunti dalla variabile dipendente (non è un modello valido, ma si prende appunto come riferimento) e ad 1 sottrae il loro rapporto.

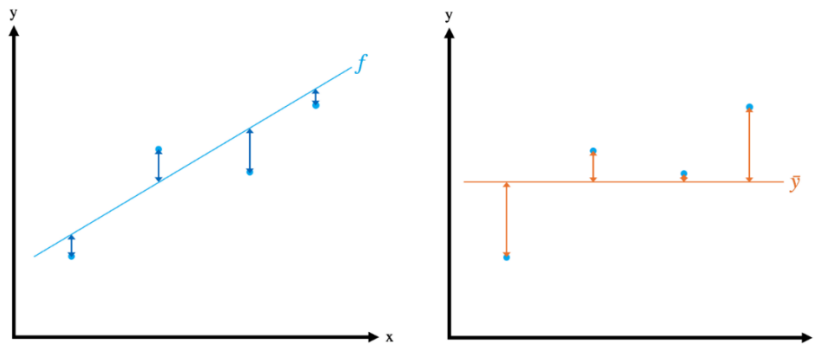


Figura 12 Confronto tra modello ottenuto dalla regressione ed il modello baseline

Nel momento in cui l'errore $RSS = TSS$, si ha che il nostro modello spiega allo stesso modo del modello baseline (ovvero male), ed $r^2 = 1 - 1 = 0$; se invece $RSS < TSS$ vuol dire che il nostro modello è migliore di quello "base" e, più ci avviciniamo ad 1 e più il nostro modello spiega correttamente i dati.

Nelle formule di RSS TSS ESS , la differenza tra valore reale e stimato è al quadrato: questo assicurerà che non dovremo affrontare valori negativi che porterebbero ad una diminuzione dell'indice; amplificherà inoltre l'effetto di grandi errori individuali, il che renderà ancora più evidente, quando un modello non sta facendo buone previsioni.

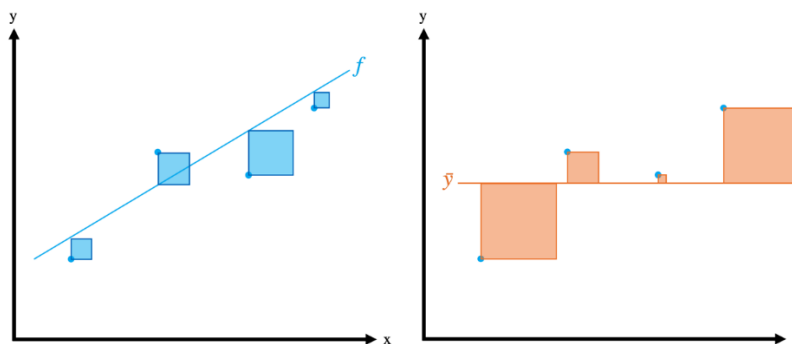


Figura 13

Nel momento in cui consideriamo una regressione multivariata però, il coefficiente di determinazione deve essere "corretto". Questo risulta necessario poiché aggiungendo il numero di predittori x_i indipendenti, aumenta anche il valore di r^2 , non diminuisce mai; inoltre può capitare che diventa eccessivamente "su misura", adattandosi alle peculiarità e al rumore casuale del set di dati, piuttosto che riflettere l'intera popolazione. Gli statistici lo definiscono *overfitting model* e produce valori R-quadrati ingannevolmente alti e una ridotta capacità di previsioni precise.

Introduciamo quindi il cosiddetto r^2 corretto detto anche adjusted:

$$r^2_{adj} = r^2 - \left(\frac{k}{n - k - 1} \right) (1 - r^2) \quad (19)$$

Dove abbiamo che k è il numero di variabili indipendenti preseti nel modello mentre n è il numero di osservazioni. Il valore di r^2_{adj} aumenta solo quando il nuovo predittore x_k migliora l'adattamento del modello maggiormente rispetto all'aumento "dovuto al caso"; diminuisce quando x_k non migliora l'adattamento del modello di una quantità sufficiente. Inoltre,

$$r^2_{adj} \leq r^2 \quad (20)$$

3.4 Scelta del modello

Nel caso di regressione multivariata, il modello ottenuto dipende dalle variabili indipendenti che scelgo; una buona selezione del modello si basa sul principio elementare noto **Rasoio di Occam**: Il miglior modello scientifico è il modello più semplice in grado di spiegare i fatti osservati. In termini di modelli lineari il rasoio di Occam implica il principio di parsimonia : un modello deve contenere il minor numero di variabili necessarie per adattarsi ai dati; i modelli che contengono solo le variabili necessarie per adattarsi ai dati sono detti **modelli parsimoniosi**[10]

Dato un modello lineare con un gran numero di variabili indipendenti, ci chiediamo quindi se esista un sottoinsieme di queste tali per cui l'accuratezza del modello non si riduca sensibilmente. Prendiamo come esempio il seguente modello

$$r^2_{adj} \leq r^2 \quad (21)$$

e supponiamo di sapere che è corretto e che esso rappresenti la vera relazione tra x e y . Vogliamo verificare l'ipotesi che

$$H_0: \beta_4 = 0$$

Se H_0 è vera, il modello completo può essere sostituito da quello ridotto

$$y_i = \beta_0 + \beta_1 * x_1 + \beta_2 * x_2 + \beta_3 * x_3 + \varepsilon_i$$

Tale ipotesi si verifica mediante il test di Fisher che mette a confronto le varianze delle due distribuzioni. L'analisi della varianza (in inglese: Analysis of variance, ANOVA) è utilizzata per testare le differenze tra medie campionarie e, per fare questo, si prendono in considerazione le rispettive varianze.

Il principio alla base di questo test è quello di stabilire se due o più medie campionarie possono derivare da popolazioni che hanno la stessa media parametrica (t test).

Quando le medie sono solamente due è indifferente usare il test di Fisher o il t-test, mentre dobbiamo necessariamente utilizzare l'ANOVA quando le medie sono più di due.

Il secondo di Fisher è dunque una derivazione del primo, anche dal punto di vista algebrico poiché

$$t^2 = f \quad (22)$$

ma non è limitato al dover confrontare solo due gruppi, difatti è quello maggiormente più usato

La statistica test F si ottiene dal rapporto tra la varianza di regressione e la varianza di dispersione del modello:

$$f = \frac{\frac{ESS_{ridotto} - ESS_{completo}}{p - k}}{\frac{ESS_{completo}}{n - p - 1}} \quad (23)$$

Abbiamo:

- Modello ridotto: modello che ha come variabili indipendenti $X_{ridotto} = [x_1, x_2, x_3]$
- Modello completo: modello che ha come variabili indipendenti $X_{completo} = [x_1, x_2, x_3, x_4]$
- n : numero delle osservazioni
- p : numero di variabili indipendenti del modello completo (nel nostro caso 4)
- k : numero di variabili indipendenti nel modello ridotto (nel nostro caso 3)

$ESS_{completo}/n - p - 1$ rappresenta una stima della varianza dell'errore σ^2 , poiché il modello completo è corretto; allora anche $ESS_{ridotto}/n - k - 1$ è una stima per la varianza dell'errore.

Se

- $ESS_{completo} \cong (n - p - 1)\sigma^2$

- H_0 è vera

Allora $ESS_{ridotto} \cong (p - k)\sigma^2$ e quindi $ESS_{ridotto} - ESS_{completo}/p - k$ sia vicino a σ^2 .

Se H_0 è vera sia il numeratore che il denominatore di f sono stime di σ^2 e quindi ci si aspetta che f assuma un valore prossimo ad 1, se H_0 non fosse vera il valore di f tenderebbe ad essere grande[10].

L'ipotesi nulla viene rigettata se, a un prefissato livello di significatività α , f così calcolata sui dati campionari è maggiore del valore della F di Fisher tabulato in corrispondenza di $p - k$ e $n - p - 1$ gradi di libertà:

$$f \geq f_{p-k, n-p-1} \quad (24)$$

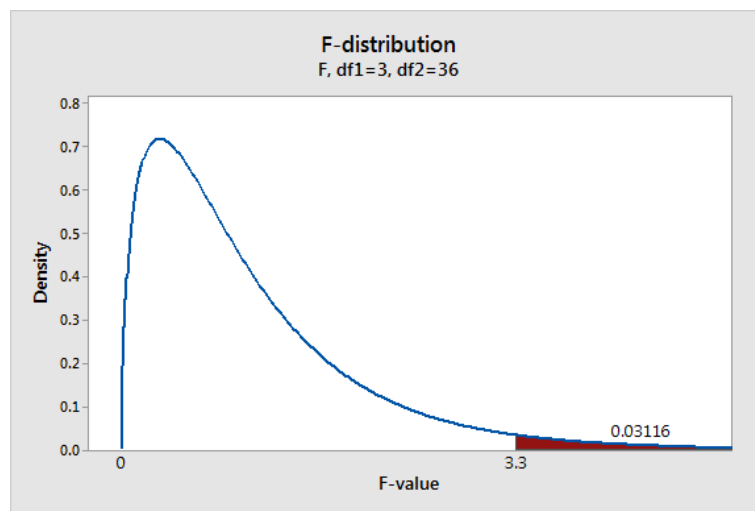


Figura 14

Graficamente la si può interpretare in questo modo: fissato un livello di significatività, mediante la tabella otteniamo il valore $f_{3,36}$ (in questo caso vale 3.3) e l'ipotesi nulla viene rigettata se f calcolato si trova nell'area rossa.

df2 / df1	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	22	24	26	28	30	35	40	45	50	60	70	80	100	200	500	1000	>1000	df1 / df2		
3	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.79	8.76	8.74	8.73	8.71	8.70	8.69	8.68	8.67	8.67	8.66	8.65	8.64	8.63	8.62	8.62	8.60	8.59	8.59	8.58	8.57	8.57	8.56	8.55	8.54	8.53	8.53	8.54	3		
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96	5.94	5.91	5.89	5.87	5.86	5.84	5.83	5.82	5.81	5.80	5.79	5.77	5.76	5.75	5.75	5.73	5.72	5.71	5.70	5.69	5.68	5.67	5.66	5.65	5.64	5.63	5.63	5.64	4	
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74	4.70	4.68	4.66	4.64	4.62	4.60	4.59	4.58	4.57	4.56	4.54	4.53	4.52	4.50	4.50	4.48	4.46	4.45	4.44	4.43	4.42	4.42	4.41	4.39	4.37	4.37	4.37	4.36	5	
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06	4.03	4.00	3.98	3.96	3.94	3.92	3.91	3.90	3.88	3.87	3.86	3.84	3.83	3.82	3.81	3.79	3.77	3.76	3.75	3.74	3.73	3.72	3.71	3.69	3.68	3.67	3.67	3.67	6	
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.64	3.60	3.57	3.55	3.53	3.51	3.49	3.48	3.47	3.46	3.44	3.43	3.41	3.40	3.39	3.38	3.36	3.34	3.33	3.32	3.30	3.29	3.29	3.27	3.25	3.24	3.23	3.23	3.23	7	
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.35	3.31	3.28	3.26	3.24	3.22	3.20	3.19	3.17	3.16	3.15	3.13	3.12	3.10	3.09	3.08	3.06	3.04	3.03	3.02	3.01	2.99	2.99	2.97	2.95	2.94	2.93	2.93	2.93	8	
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.14	3.10	3.07	3.05	3.03	3.01	2.99	2.97	2.96	2.95	2.94	2.92	2.90	2.89	2.87	2.86	2.84	2.83	2.81	2.80	2.79	2.78	2.77	2.76	2.73	2.72	2.71	2.71	2.71	9	
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98	2.94	2.91	2.89	2.86	2.85	2.83	2.81	2.80	2.79	2.77	2.75	2.74	2.72	2.71	2.70	2.68	2.66	2.65	2.64	2.62	2.61	2.60	2.59	2.56	2.55	2.54	2.54	2.54	10	
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90	2.85	2.82	2.79	2.76	2.74	2.72	2.70	2.69	2.67	2.66	2.65	2.63	2.61	2.59	2.58	2.57	2.55	2.53	2.52	2.51	2.49	2.48	2.47	2.46	2.43	2.42	2.41	2.41	2.41	11	
12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80	2.75	2.72	2.69	2.66	2.64	2.62	2.60	2.58	2.57	2.56	2.54	2.52	2.51	2.49	2.48	2.47	2.44	2.43	2.41	2.40	2.38	2.37	2.36	2.35	2.32	2.31	2.30	2.30	2.30	12	
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71	2.67	2.63	2.60	2.58	2.55	2.53	2.51	2.50	2.48	2.47	2.46	2.44	2.42	2.41	2.39	2.38	2.36	2.34	2.33	2.31	2.30	2.28	2.27	2.26	2.23	2.22	2.21	2.21	2.21	13	
14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65	2.60	2.57	2.53	2.51	2.48	2.46	2.44	2.43	2.41	2.40	2.39	2.37	2.35	2.33	2.32	2.31	2.28	2.27	2.25	2.24	2.22	2.21	2.20	2.19	2.16	2.14	2.14	2.14	14		
15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59	2.54	2.51	2.48	2.45	2.42	2.40	2.38	2.37	2.35	2.34	2.33	2.31	2.29	2.27	2.26	2.25	2.22	2.20	2.19	2.18	2.16	2.15	2.14	2.12	2.10	2.08	2.07	2.07	2.07	15	
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54	2.49	2.46	2.42	2.40	2.37	2.35	2.33	2.32	2.30	2.29	2.28	2.25	2.24	2.22	2.21	2.19	2.17	2.15	2.14	2.12	2.11	2.09	2.08	2.07	2.04	2.02	2.02	2.01	2.01	16	
17	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.49	2.45	2.41	2.38	2.35	2.33	2.31	2.29	2.27	2.26	2.24	2.23	2.21	2.19	2.17	2.16	2.15	2.12	2.10	2.09	2.08	2.06	2.05	2.03	2.02	1.99	1.97	1.97	1.96	1.96	17	
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46	2.41	2.37	2.34	2.31	2.29	2.27	2.25	2.23	2.22	2.20	2.19	2.17	2.15	2.13	2.12	2.11	2.08	2.06	2.05	2.04	2.02	2.00	1.99	1.98	1.95	1.93	1.92	1.92	1.92	18	
19	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42	2.38	2.34	2.31	2.28	2.26	2.23	2.21	2.20	2.18	2.17	2.16	2.13	2.11	2.10	2.08	2.07	2.05	2.03	2.01	2.00	1.98	1.97	1.96	1.94	1.91	1.89	1.88	1.88	1.88	19	
20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39	2.35	2.31	2.28	2.25	2.23	2.20	2.18	2.17	2.15	2.14	2.12	2.10	2.08	2.07	2.05	2.04	2.01	1.99	1.98	1.97	1.95	1.93	1.92	1.91	1.88	1.86	1.85	1.84	1.84	20	
22	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.34	2.30	2.26	2.23	2.20	2.17	2.15	2.13	2.11	2.10	2.08	2.07	2.05	2.03	2.01	2.00	1.98	1.96	1.94	1.92	1.91	1.89	1.88	1.86	1.85	1.82	1.80	1.79	1.78	1.78	1.78	22
24	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30	2.25	2.22	2.18	2.15	2.13	2.11	2.09	2.07	2.05	2.04	2.03	2.00	1.98	1.97	1.95	1.94	1.91	1.89	1.88	1.86	1.84	1.83	1.82	1.80	1.77	1.75	1.74	1.73	1.73	1.73	24
26	4.23	3.37	2.98	2.74	2.59	2.47	2.39	2.32	2.27	2.22	2.18	2.15	2.12	2.09	2.07	2.05	2.03	2.02	2.00	1.99	1.97	1.95	1.93	1.91	1.90	1.87	1.85	1.84	1.82	1.80	1.79	1.78	1.76	1.73	1.71	1.70	1.69	1.69	26	
28	4.20	3.34	2.95	2.71	2.56	2.45	2.36	2.29	2.24	2.19	2.15	2.12	2.09	2.06	2.04	2.02	2.00	1.99	1.97	1.96	1.93	1.91	1.90	1.88	1.87	1.84	1.82	1.80	1.79	1.77	1.75	1.74	1.73	1.69	1.67	1.66	1.66	1.66	28	
30	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21	2.16	2.13	2.09	2.06	2.04	2.01	1.99	1.98	1.96	1.95	1.93	1.91	1.89	1.87	1.85	1.84	1.81	1.79	1.77	1.76	1.74	1.72	1.71	1.70	1.66	1.64	1.63	1.62	1.62	30	
35	4.12	3.27	2.87	2.64	2.49	2.37	2.29	2.22	2.16	2.11	2.08	2.04	2.01	1.99	1.96	1.94	1.92	1.91	1.89	1.88	1.85	1.83	1.82	1.80	1.79	1.76	1.74	1.72	1.70	1.68	1.66	1.65	1.63	1.60	1.57	1.57	1.56	1.56	35	
40	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12	2.08	2.04	2.00	1.97	1.95	1.92	1.90	1.89	1.87	1.85	1.84	1.81	1.79	1.77	1.76	1.74	1.72	1.69	1.67	1.66	1.64	1.62	1.61	1.59	1.55	1.53	1.52	1.51	1.51	1.51	40
45	4.06	3.20	2.81	2.58	2.42	2.31	2.22	2.15	2.10	2.05	2.01	1.97	1.94	1.92	1.89	1.87	1.86	1.84	1.82	1.81	1.78	1.76	1.74	1.73	1.71	1.68	1.66	1.64	1.63	1.60	1.59	1.57	1.55	1.51	1.49	1.48	1.47	1.47	45	
50	4.03	3.18	2.79	2.56	2.40	2.29	2.20	2.13	2.07	2.03	1.99	1.95	1.92	1.89	1.87	1.85	1.83	1.81	1.80	1.78	1.76	1.74	1.72	1.70	1.69	1.66	1.63	1.61	1.60	1.58	1.56	1.54	1.52	1.48	1.46	1.45	1.44	1.44	50	
60	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	2.04	1.99	1.95	1.92	1.89	1.86	1.84	1.82	1.80	1.78	1.76	1.75	1.72	1.70	1.68	1.66	1.65	1.62	1.59	1.57	1.56	1.53	1.52	1.50	1.48	1.44	1.41	1.40	1.39	1.39	60	
70	3.98	3.13	2.74	2.50	2.35	2.23	2.14	2.07	2.02	1.97	1.93	1.89	1.86	1.84	1.81	1.79	1.77	1.75	1.74	1.72	1.70	1.67	1.65	1.64	1.62	1.59	1.57	1.55	1.53	1.50	1.49	1.47	1.45	1.40	1.37	1.36	1.35	1.35	70	
80	3.96	3.11	2.72	2.49	2.33	2.21	2.13	2.06	2.00	1.95	1.91	1.88	1.84	1.82	1.79	1.77	1.75	1.73	1.72	1.70	1.68	1.65	1.63	1.62	1.60	1.57	1.54	1.52	1.51	1.48	1.46	1.45	1.43	1.38	1.35	1.34	1.33	1.33	80	
100	3.94	3.09	2.70	2.46	2.31	2.19	2.10	2.03	1.97	1.93	1.89	1.85	1.82	1.79	1.77	1.75	1.73	1.71	1.69	1.68	1.65	1.63	1.61	1.59	1.57	1.54	1.52	1.49	1.48	1.45	1.43	1.41	1.39	1.34	1.31	1.30	1.28	1.28	100	
200	3.89	3.04	2.65	2.42	2.26	2.14	2.06	1.98	1.93	1.88	1.84	1.80	1.77	1.74	1.72	1.69	1.67	1.66	1.64	1.62	1.60	1.57	1.55	1.53	1.52	1.48	1.46	1.43	1.41	1.39	1.36	1.35	1.32	1.26	1.22	1.21	1.19	1.19	200	
500	3.86	3.01	2.62	2.39	2.23	2.12	2.03	1.96	1.90	1.85	1.81	1.77	1.74	1.71	1.69	1.66	1.64	1.62	1.61	1.59	1.56	1.54	1.52	1.50	1.48	1.45	1.42	1.40	1.38	1.35	1.32	1.30	1.28	1.21	1.16	1.14	1.12	1.12	500	
1000	3.85	3.00	2.61	2.38	2.22	2.11	2.02	1.95	1.89	1.8																														

L'ipotesi alternativa è la diversa motivazione per spiegare meglio il fenomeno oggetto di indagine. Come tale, deve essere l'unica o la migliore possibile spiegazione alternativa. In questo modo, se il valore p convalida un rifiuto dell'ipotesi nulla, allora l'ipotesi alternativa può essere considerata credibile.

3.5.2 t-Test

Il test t (o, dall'inglese, t -test) è un test statistico di tipo parametrico con lo scopo di verificare se, il valore medio di una distribuzione (paired t -test) o due distribuzioni (unpaired t -test), si discostano significativamente da un certo valore di riferimento.

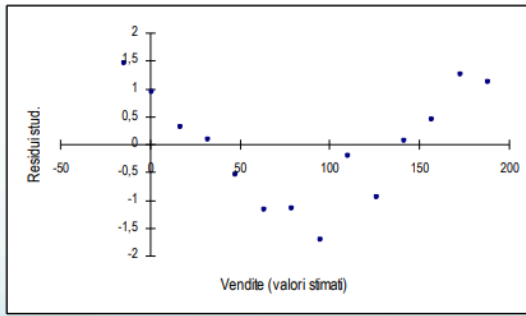
3.5.3 Livello di significatività

Il livello di significatività, rappresentato con α , di un test è dato solitamente da una verifica del test d'ipotesi. Nel caso più semplice è definita come la probabilità di accettare o rigettare l'ipotesi nulla. Se il test considerato da un valore $p < \alpha$ l'ipotesi nulla è rifiutata. Più basso è il livello di significatività stabilita e maggiore è l'evidenza

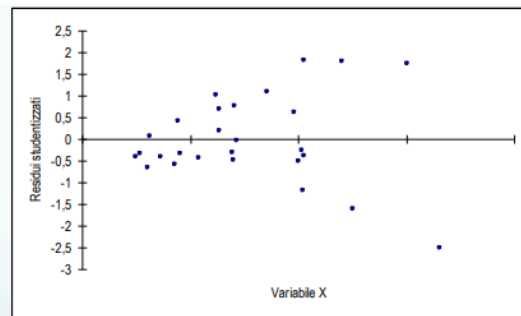
3.6 Analisi dei residui

Oltre valutare il coefficiente r^2 è opportuno analizzare il grafico dei residui per stabilire se il modello ipotizzato è corretto. Prenderemo in considerazione in futuro il diagramma di dispersione dei residui che ha in ascissa il valore stimato della variabile dipendente \hat{y} mentre in ordinata il valore dei residui. Quello che dobbiamo avere è che:

- i residui devono disporsi in modo casuale, senza nessuna struttura riconoscibile; se ciò non accade si può ipotizzare ad esempio di non avere una relazione lineare tra i dati.
- la varianza dei residui deve essere costante, ci deve essere omoschedasticità (valutabile tramite test F)



Violazione dell'ipotesi di linearità



Violazione dell'ipotesi di omoschedasticità

Figura 16- <https://www.docenti.unina.it/webdocenti-be/allegati/materiale-didattico/578826>

Capitolo 4

Analisi assorbimento delle gabbie

Sono stati elaborati i dati ottenuti dal sistema SCADA che fa da interfaccia tra il sistema di controllo avanzato APC e gli apparati reali posti in campo.

Ci sono stati forniti i file di log relativi al periodo febbraio - aprile. Dopo aver estratto i dati, ci siamo accorti di quanta varietà di produzioni ci fossero in questo periodo; l'elaborato è incentrato sulla produzione **Green Plus 12mm** poiché è quella che viene prodotta maggiormente nel periodo fornitoci e della quale avevamo intervalli di produzione lunghi e continui da poter analizzare.

4.1 Analisi di un **modello** lineare

È ragionevole ipotizzare, per quanto riguarda la relazione tra assorbimenti alle gabbie e temperatura misurate in uscita, un andamento lineare: maggiore è la temperatura con cui arrivo alle gabbie e minore è l'assorbimento.

Andando a graficare il grafico di dispersione (o scatter-plot), delle due variabili, ci siamo accorti di come, fissata una temperatura, i valori degli assorbimenti di billette differenti non fossero contenuti in un piccolo intervallo, ma fossero invece dispersi in un ampio range.

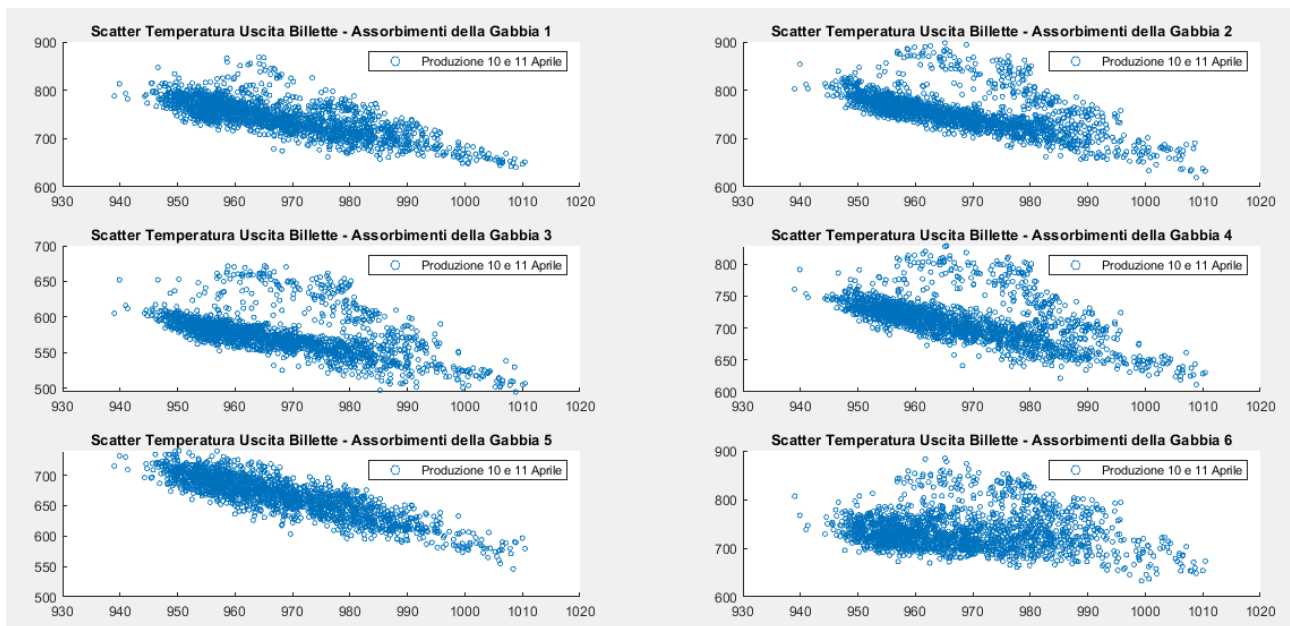


Figura 17 Assorbimento delle 6 gabbie in funzione della temperatura di uscita

Questo ci indirizza su due diverse strade:

1. l'insieme totale della produzione deve essere suddiviso in base a "condizioni simili", in modo da raggruppare billette che hanno la stessa storia, e subito la stessa lavorazione
2. valutare un modello, sempre lineare, ma che tenga conto anche di altri fattori oltre la temperatura di uscita

4.2 Modello su un sottoinsieme di dati

Iniziamo dalla prima e dividiamo l'insieme totale dei dati in funzione del tempo di permanenza nel forno.

Zona 1	Tempo in forno < 2 h
Zona 2	2 h < Tempo in forno < 2 h 30 min
Zona 3	Tempo in forno > 2 h 30 min

Tabella 2

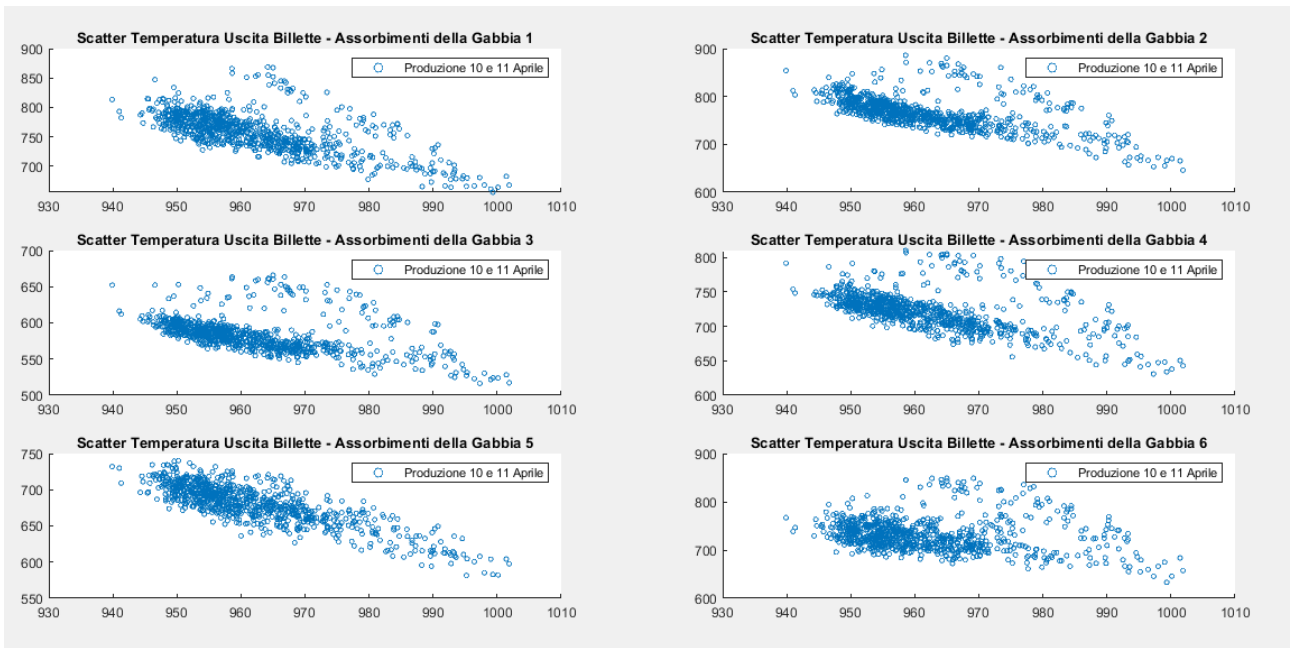


Figura 18 Zona 1: Assorbimenti billette con tempo di permanenza in forno inferiore a 2 ore

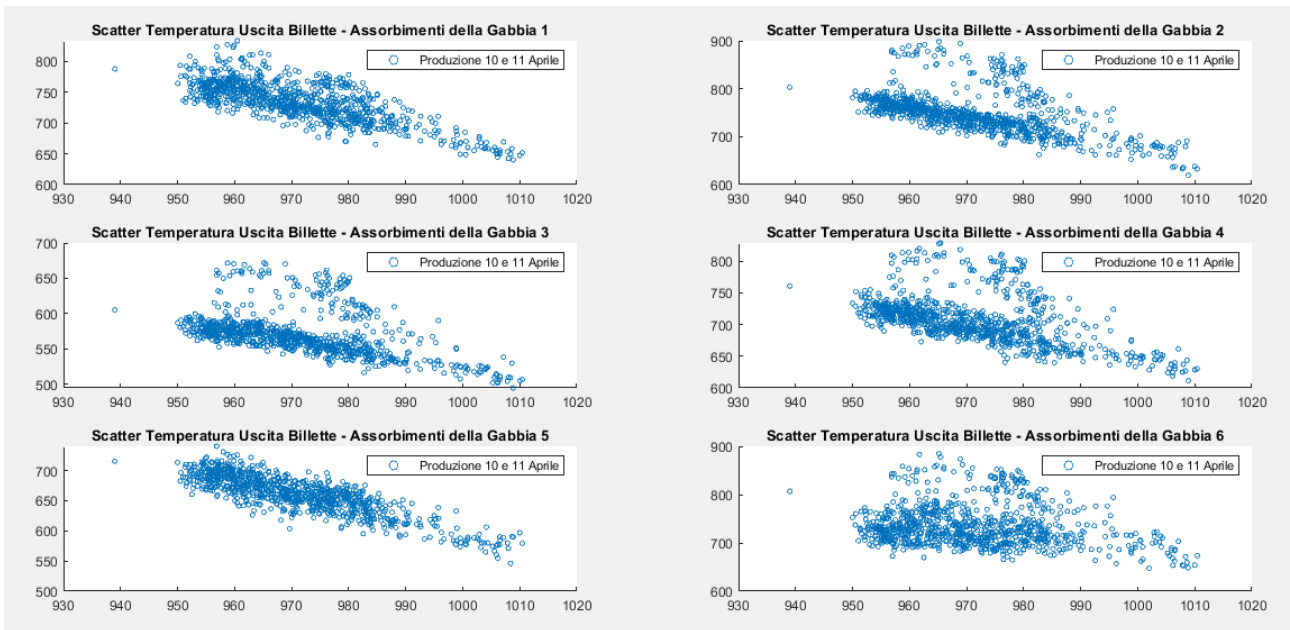


Figura 19 Zona 2: Assorbimenti billette con tempo di permanenza in forno compreso tra 2 ore e 2 ore e mezza

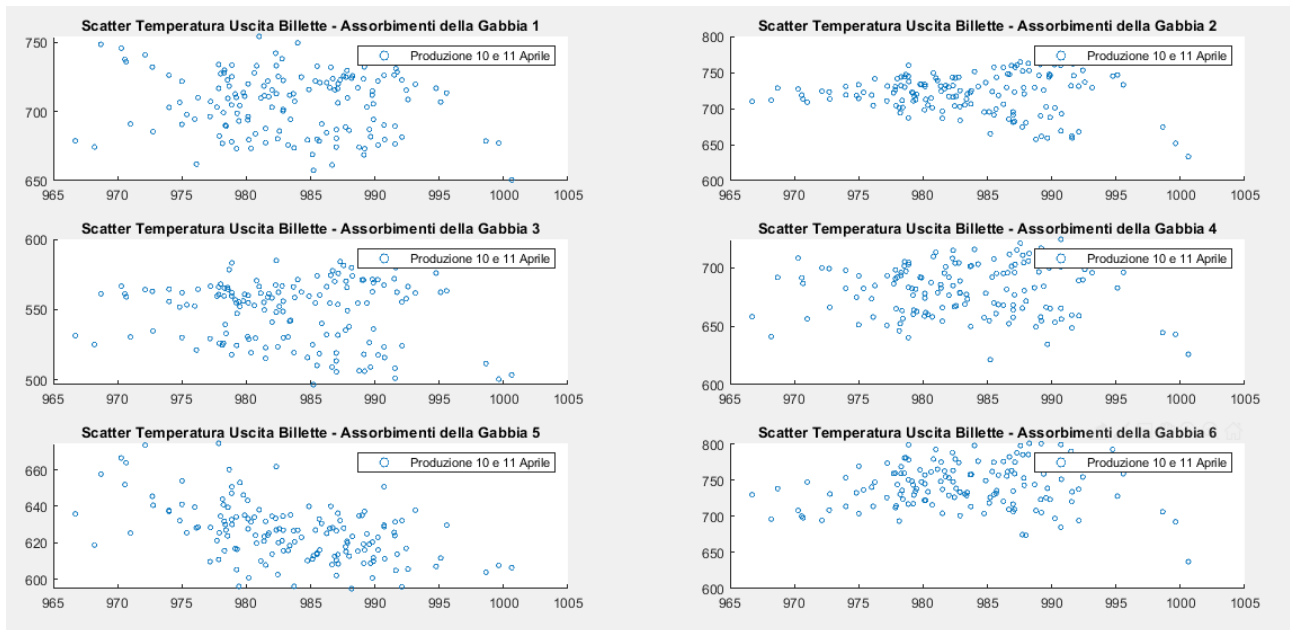


Figura 20 Zona 3: Assorbimenti billette con tempo di permanenza in forno superiore alle 2h 30 min

Otteniamo che nonostante la suddivisione, in ogni cluster si ha un andamento molto “disperso” della relazione assorbimenti-temperatura .

Consideriamo ora come variabile discriminante la temperatura di ingresso (reale) delle billette nel forno.

Zona 1	Temperatura di ingresso < 400 °C
Zona 2	400 °C < Temperatura di ingresso < 650 °C
Zona 3	Temperatura di ingresso > 650 °C

Tabella 2

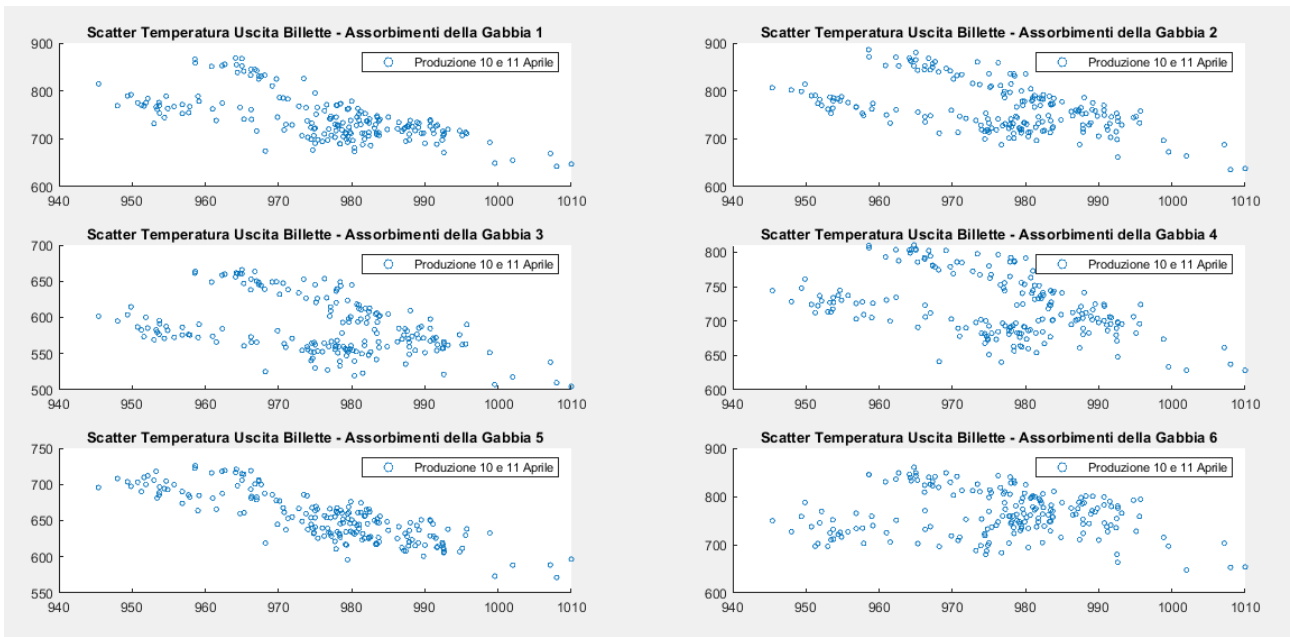


Figura 21 Zona 1: Temperatura di ingresso < 400°C

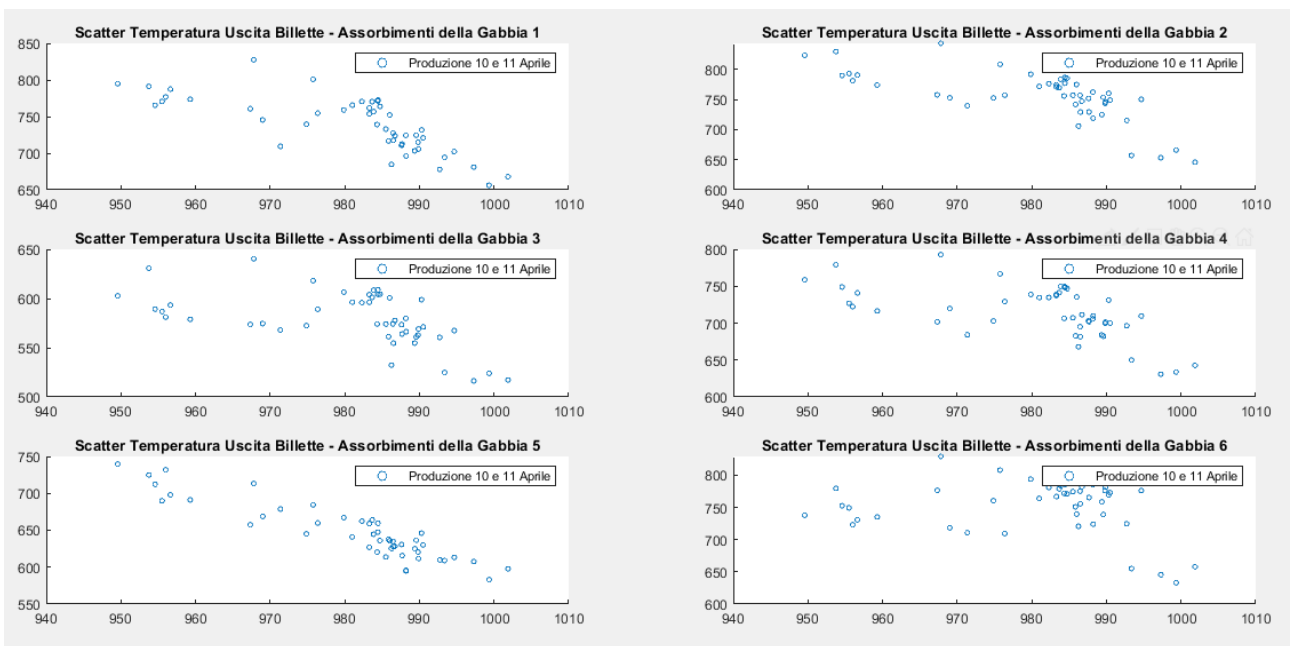


Figura 22 Zona 2: Temperatura di ingresso compresa tra 400°C e 650°C

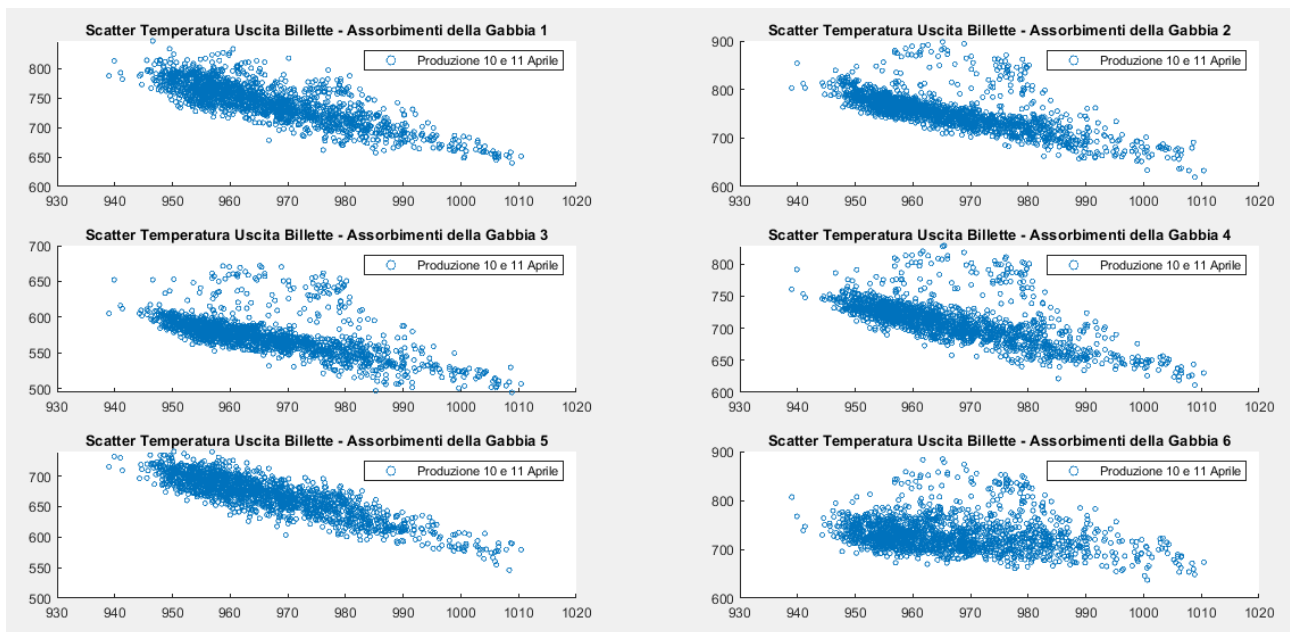


Figura 23 Zona 3: Temperatura di ingresso > 650°C

Anche mediante questa seconda suddivisione otteniamo risultati simili che non chiariscono l'andamento degli assorbimenti. Questo ci fa capire che non è questo il metodo corretto per procedere.

Pur raggruppando billette con caratteristiche simili, ci sono altri fattori, non ancora considerati, che vanno ad influire sugli assorbimenti e che ne determinano la loro variabilità. Di fatto, la dinamica del nostro sistema fisico non è affatto banale, e per questo, andremo ad introdurre invece di una regressione lineare semplice, una multivariata.

4.3 Modello multivariato

Prendiamo ad esempio la produzione di Green Plus 12mm relativa al 9 – 10 Aprile 2019; e facciamone un modello lineare, sulla totalità della produzione.

L'equazione iniziale che andremo a considerare quindi sarà:

$$\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_{1i} * x_1 + \hat{\beta}_{2i} * x_2 + \hat{\beta}_{3i} * x_3 + \hat{\beta}_{4i} * x_4 + \hat{\beta}_{5i} * x_5$$

dove

- y_i è l'assorbimento della Gabbia i-esima
- x_1 è la temperatura uscita
- x_2 è la temperatura filtrata in ingresso
- x_3 è il tempo che la billetta trascorre nel forno
- x_4 è la temperatura reale ingresso

Distinguiamo la temperatura di ingresso “filtrata” e “reale” poiché per motivi di controllo, alla coppia di billette in ingresso viene assegnata, ad entrambe, la temperatura di quella più alta; come temperatura reale appunto indichiamo la temperatura estrapolata dal treno di letture di pirometri, opportunamente processata.

Mediante la funzione integrata di matlab “fitlm” operiamo una regressione multivariata, dove le variabili indipendenti o regressori sono le x_i e quelle dipendenti sono le y_i . Otteniamo quindi per il vettore X di regressori presenti nella figura 24 il seguente modello:

```

Command Window
X = [x1, x2, x3, x4]

mdl1 =

Linear regression model:
y ~ 1 + Temperaturauscita + Temperaturadiingresso + Tempodipermanenzaaforno + Temperaturarealeingresso

Estimated Coefficients:

                Estimate      SE      tStat      pValue
                _____  _____  _____  _____
(Intercept)          3261      34.709      93.95         0
Temperatura uscita   -2.4232     0.035773    -67.74         0
Temperatura di ingresso -0.15499    0.0034676   -44.698    4.6281e-299
Tempo di permanenza forno -16.962     0.83243    -20.377    1.1632e-83
Temperatura reale ingresso 0.0031814   0.0027118    1.1732     0.24088

Number of observations: 1911, Error degrees of freedom: 1906
Root Mean Squared Error: 18.5
R-squared: 0.795, Adjusted R-Squared: 0.794
F-statistic vs. constant model: 1.85e+03, p-value = 0
fx >>

```

Figura 24

Matlab ci restituisce diverse informazioni, per quanto riguarda l’intercetta ed i regressori si ha:

Estimate	Stima dei coefficienti per ciascun termine corrispondente al modello
SE	Errore standard, stima della deviazione standard dello stimatore
tStat	<p>Statistica t per ciascun coefficiente per verificare l’ipotesi nulla che il coefficiente corrispondente sia nullo, rispetto l’alternativa che è diverso da zero. È definito come</p> $\frac{\text{Estimate}}{SE}$

pValue	valore p relativo alla statistica t del test di ipotesi, secondo cui il coefficiente corrispondente è uguale a zero o no
---------------	--

Tabella 3

Per quanto riguarda il modello complessivo, invece abbiamo:

Number of observations	n , Numero dei campioni analizzati.
Error degrees of freedom	$n - p$, con p che rappresenta il numero di coefficienti nel modello, inclusa l'intercetta.
Root mean squared error	Radice quadrata dell'errore quadratico medio, che stima la deviazione standard della distribuzione dell'errore.
R-squarede Adjusted R - squared	Coefficiente di determinazione e coefficiente di determinazione corretto, rispettivamente.
F-statistic vs. constant model	Prova statistica per il test F sul modello di regressione, che verifica se il modello si adatta significativamente meglio di un modello degenerato costituito solo da un termine costante.
p-value	p -value per il test F sul modello.

Tabella 4

Fissiamo come livello di significatività $\alpha = 0.05$; il p -value associato al test F è < 0.05 , quindi rifiuto l'ipotesi nulla H_0 per la quale tutti i coefficienti sono nulli: il modello ha capacità esplicativa.

Valutiamo adesso la significatività dei singoli coefficienti del modello, in modo da individuare quali variabili predittive possono essere rimosse dal nostro modello, dal momento che ne complicano unicamente la struttura e non ne apportano nessun miglioramento significativo.

Fissato un livello di significatività $\alpha = 0.05$, per ogni regressore valuto il p -value associato al test- t ; nel caso in cui sia inferiore del valore α , posso rifiutare l'ipotesi H_0 di coefficiente nullo: il regressore corrispondente è rilevante per la spiegazione della variabile dipendente. Nel caso in cui quindi sia maggiore, posso rimuovere il regressore dal modello.

Eseguendo questi test, le variabili predittive scelte sono date dal vettore $X = [x_1, x_2, x_3]$

Quello che si ottiene, per le 6 gabbie, mediante fitlm è:

mdl_Gabbia1 =

	Estimate	SE	tStat	pValue
Intercetta	3072.2	32.121	95.644	0
Temperatura uscita	-1.1224	0.047335	-69.331	0
Temperatura di ingressi	-0.15922	0.0032493	-40.301	2.0192e-257
Tempo ne forno	-4.8268	1.1004	-4.3862	1.3481e-73

Number of observations: 1911, Error degrees of freedom: 1907

Root Mean Squared Error: 17.1

R-squared: 0.774, Adjusted R-Squared: 0.773

F-statistic vs. constant model: 2.17e+03, p-value = 0

mdl_Gabbia2 =

	Estimate	SE	tStat	pValue
Intercetta	3261.5.2	34.71	93.963	0
Temperatura uscita	-2.4236	0.35775	-67.747	0
Temperatura di ingressi	-0.15212	0.0024558	-61.945	0
Tempo ne forno	-17.006	0.83169	-20.447	3.5427e-85

Number of observations: 1911, Error degrees of freedom: 1907

Root Mean Squared Error: 18.5

R-squared: 0.795, Adjusted R-Squared: 0.794

F-statistic vs. constant model: 2.46e+03, p-value = 0

mdl_Gabbia3 =

	Estimate	SE	tStat	pValue
Intercetta	2133	24.4	87.416	0
Temperatura uscita	-1.4702	0.025149	-58.458	0
Temperatura di ingressi	-0.11924	0.0017264	-69.072	0
Tempo ne forno	-17.396	0.58466	-29.754	3.965e-160

Number of observations: 1911, Error degrees of freedom: 1907

Root Mean Squared Error: 13

R-squared: 0.794, Adjusted R-Squared: 0.794

F-statistic vs. constant model: 2.45e+03, p-value = 0

mdl_Gabbia4 =

	Estimate	SE	tStat	pValue
Intercetta	2650.7	30.881	85.836	0
Temperatura uscita	-1.8511	0.031828	-58.161	0
Temperatura di ingressi	-0.13074	0.0021848	-59.841	0
Tempo ne forno	-18.515	0.73993	-25.022	9.858e-120

Number of observations: 1911, Error degrees of freedom: 1907

Root Mean Squared Error: 16.4

R-squared: 0.767, Adjusted R-Squared: 0.767

F-statistic vs. constant model: 2.09e+03, p-value = 0

mdl_Gabbia5 =

	Estimate	SE	tStat	pValue
Intercetta	2828.1.2	27.984	101.06	0
Temperatura uscita	-2.185	0.028842	-75.759	0
Temperatura di ingressi	-0.033826	0.0019799	- 17.085	5.2483e-61
Tempo ne forno	-8.5952	0.67052	-12.819	3.81711e-36

Number of observations: 1911, Error degrees of freedom: 1907

Root Mean Squared Error: 14.9

R-squared: 0.795, Adjusted R-Squared: 0.794

F-statistic vs. constant model: 2.46e+03, p-value = 0

mdl_Gabbia6 =

	Estimate	SE	tStat	pValue
Intercetta	1961.1	45.927	42.701	3.6768e-280
Temperatura uscita	-1.1224	0.047335	-23.712	3.8068e-109
Temperatura di ingressi	-0.15922	0.0032493	-49.001	0
Tempo ne forno	-4.8268	1.1004	-4.3862	1.2162e-05

Number of observations: 1911, Error degrees of freedom: 1907

Root Mean Squared Error: 24.5

R-squared: 0.582, Adjusted R-Squared: 0.581

F-statistic vs. constant model: 884, p-value = 0

Facendo un modello lineare, multivariato, sulla produzione Green Plus 12mm riusciamo ad avere un modello molto buono per le 6 gabbie.

Andando a valutare i residui, si vede come essi abbiano media approssimabile a zero (figura 25)

```
Command Window
Il valor medio dei residui, per la gabbia 1 è 1.461212e-12
Il valor medio dei residui, per la gabbia 2 è 1.148529e-12
Il valor medio dei residui, per la gabbia 3 è 8.046720e-13
Il valor medio dei residui, per la gabbia 4 è 9.247838e-13
Il valor medio dei residui, per la gabbia 5 è 1.026691e-12
Il valor medio dei residui, per la gabbia 6 è 4.750932e-13
```

Figura 25

4.3.1 Analisi residui

Andiamo adesso a valutare il grafico dei residui. Essendo una regressione multivariata è bene farlo, oltre in funzione della variabile stimata, anche in funzione dei predittori per vedere se ci sono delle violazioni di ipotesi.

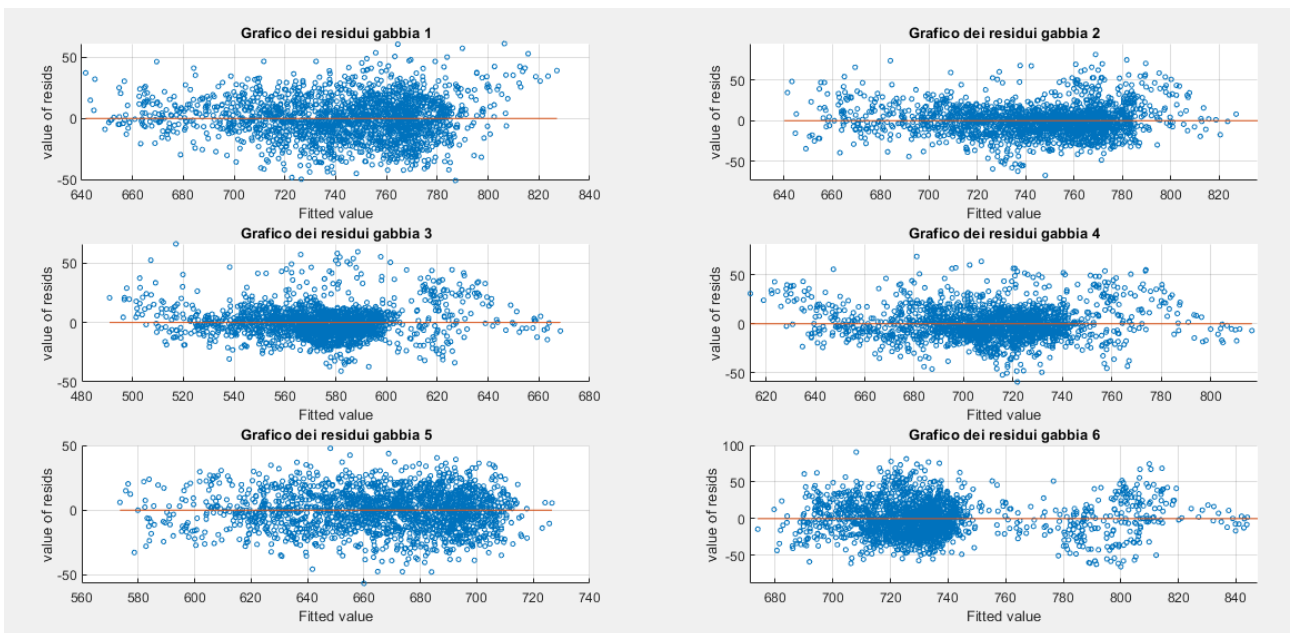


Figura 26 Grafico residui con assorbimento reale in ascissa

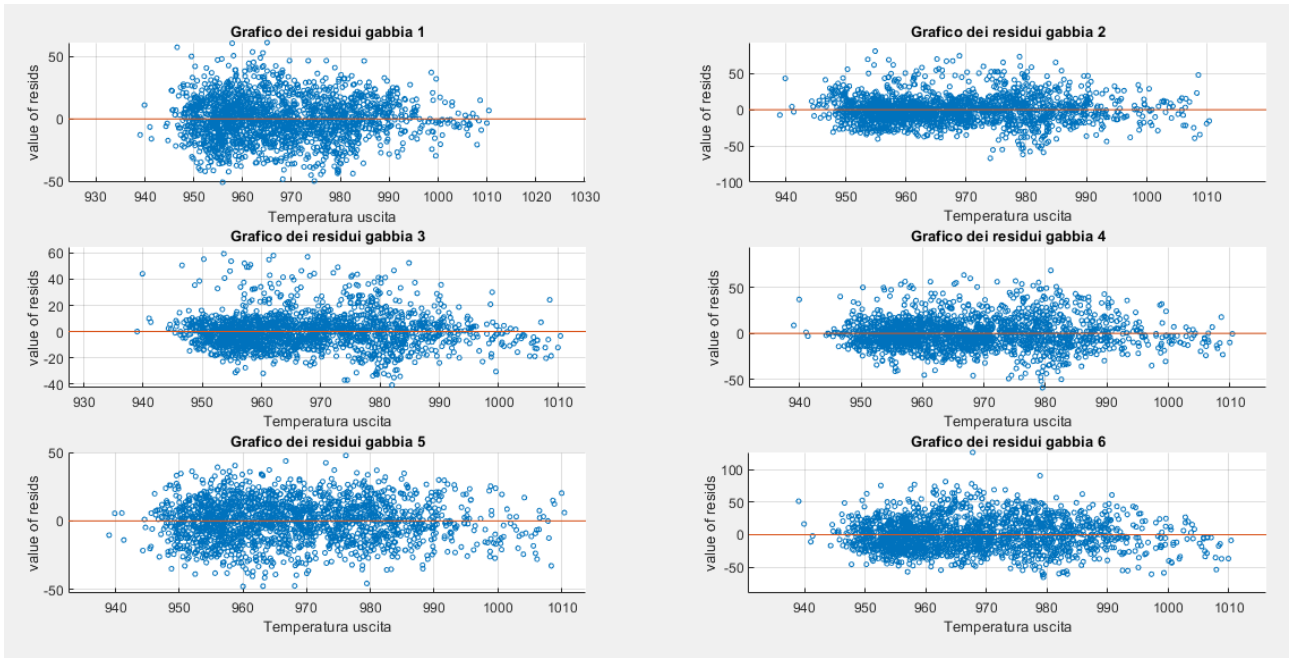


Figura 27 Grafico residui con temperatura di uscita in ascissa

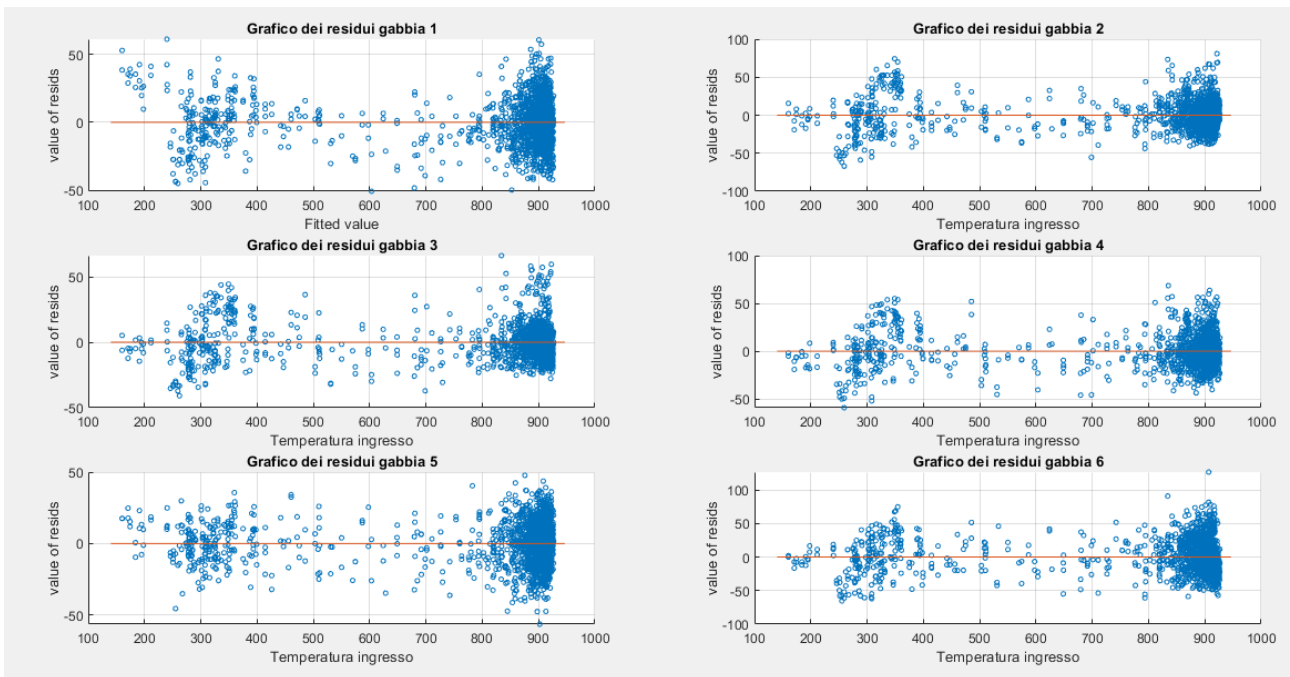


Figura 28 Grafico residui con temperatura di ingresso in ascissa

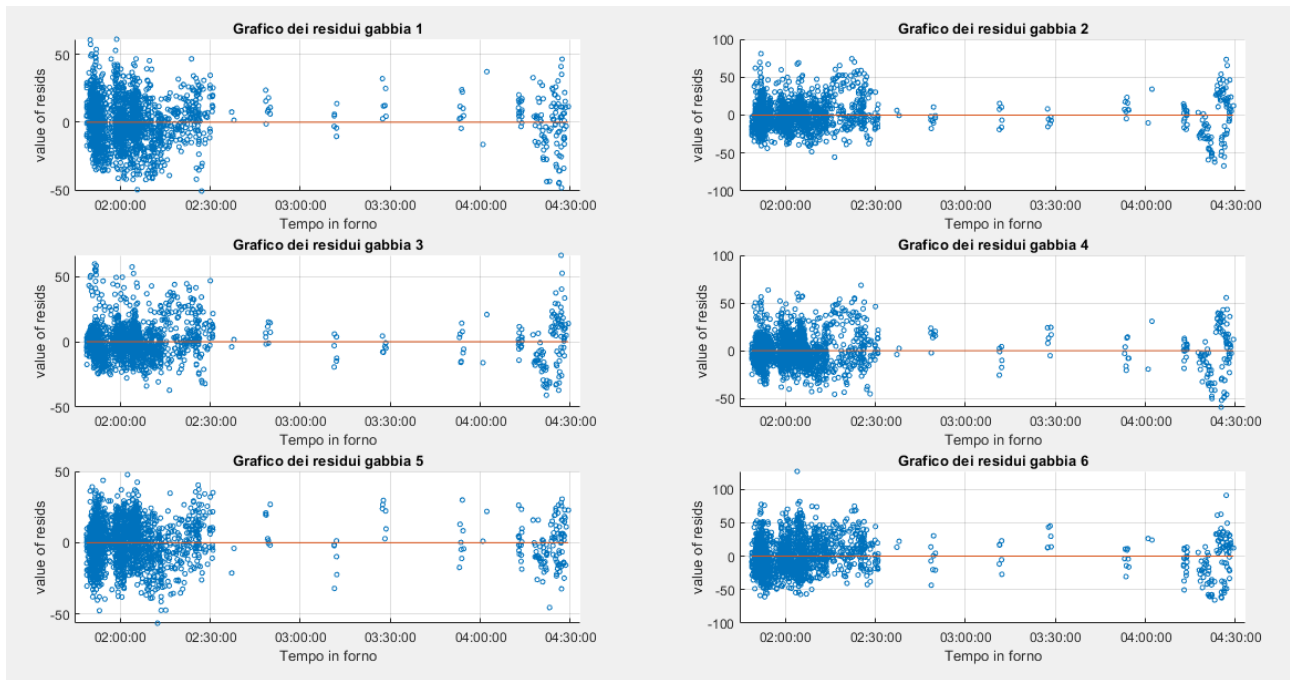


Figura 29 Grafico residui con tempo di permanenza nel forno in ascissa

Poiché i grafici non mostrano un andamento patologico, i modelli ottenuti possono essere ritenuti validi.

Andiamo quindi ad analizzare la stessa produzione, in un periodo diverso da quello considerato. Lo scopo di ciò è capire se i risultati ottenuti (i 6 modelli, uno per ogni gabbia) possono essere considerati dei modelli generali e validi sempre, per questa specifica produzione.

4.4 Confronto del modello su diversi periodi

Prendiamo in analisi quindi la produzione (Green Plus 12mm) relativa al 2 Marzo.

Operiamo come prima una regressione lineare multivariata con i medesimi predittori. Otteniamo dei nuovi modelli che, seppur stimino in modo soddisfacente (si veda figura 30) gli assorbimenti delle 6 gabbie, confrontandoli con quelli relativi alla produzione di aprile non emergono somiglianze.

```
Command Window
PRODUZIONE COMPLETA GREEN PLUS 12

ASSORBIMENTO GABBIA 1:

y1 = 3.345e+03 -2.541e+00*x1 -7.198e-02*x2 -3.769e-02*x3
Il valore del coefficiente di determinazione R^2 adj. vale 6.126210e-01
La deviazione standard vale 1.93e+01

ASSORBIMENTO GABBIA 2:

y2 = 3.345e+03 -3.337e+00*x1 -5.658e-02*x2 -7.584e-02*x3
Il valore del coefficiente di determinazione R^2 adj. vale 6.256686e-01
La deviazione standard vale 2.36e+01

ASSORBIMENTO GABBIA 3:

y3 = 2.746e+03 -2.102e+00*x1 -5.870e-02*x2 -4.250e-02*x3
Il valore del coefficiente di determinazione R^2 adj. vale 6.547215e-01
La deviazione standard vale 1.50e+01

ASSORBIMENTO GABBIA 4:

y4 = 3.345e+03 -2.489e+00*x1 -8.779e-02*x2 -4.111e-02*x3
Il valore del coefficiente di determinazione R^2 adj. vale 6.737252e-01
La deviazione standard vale 1.76e+01

ASSORBIMENTO GABBIA 5:

y5 = 3.345e+03 -2.167e+00*x1 -7.700e-02*x2 -1.730e-02*x3
Il valore del coefficiente di determinazione R^2 adj. vale 6.022200e-01
La deviazione standard vale 1.71e+01

ASSORBIMENTO GABBIA 6:

y6 = 3.345e+03 -1.803e+00*x1 -9.814e-02*x2 -5.755e-02*x3
Il valore del coefficiente di determinazione R^2 adj. vale 5.245402e-01
La deviazione standard vale 2.26e+01

fx >> |
```

Figura 30 Modelli ottenuti dalla produzione del 2 marzo

Molto probabilmente si hanno queste differenze poiché, nonostante in questi due periodi si produca lo stesso prodotto finale, i settaggi delle gabbie, non incluse nei modelli, rendono la stima incompleta e non “generale”. Tali settaggi sono frutto dell’esperienza del singolo operatore che spesso opera piccoli personali aggiustamenti, diversi quindi da un operatore ad un altro. Poiché vengono inseriti manualmente, il sistema automatizzato non tiene traccia e per questo non sono reperibili dai file di log fornitici.

Un altro fattore da considerare può essere l'usura delle gabbie stesse: poiché le stesse gabbie vengono usate per tutte le produzioni e sono sempre in funzione, non è da escludere che il rendimento delle stesse sia diminuisca, causando il progressivo aumento degli assorbimenti.

Il fatto che i modelli vengano molto differenti rende impossibile l'utilizzo di un modello stimato in un periodo di funzionamento per la predizione degli assorbimenti in un diverso periodo, non consecutivo a quello analizzato.

Questo è evidenziato nella figura 31 dove si è utilizzato il modello ricavato dai dati di aprile per stimare i dati di marzo.

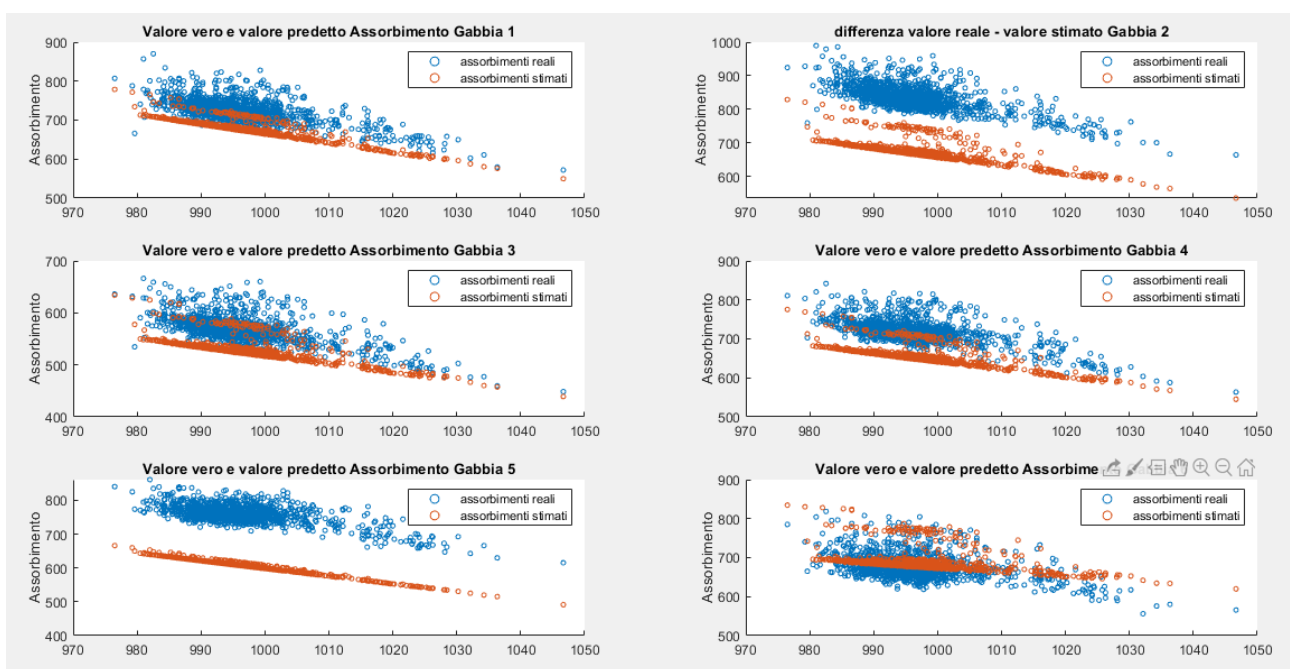


Figura 31 Assorbimenti stimati relativi al periodo di marzo mediante modello ottenuto sui dati di aprile

Quello che si può fare a questo punto è considerare un "breve" intervallo temporale nel quale acquisirò informazioni da usare per il modello (breve, ma sufficiente affinché il modello riesca a prevedere gli assorbimenti), mediante il quale potrò restituire una stima sugli assorbimenti giornalieri.

4.5 Modello per stima degli assorbimenti della giornata

Considerando la figura 32, nel plot in alto abbiamo messo in ascissa il numero della billetta prodotta e in ordinata la temperatura di ingresso; nel secondo invece l'ascissa è relativa tempo nel forno,

mentre il terzo rappresenta la temperatura in ingresso. Le billette rosse sono quelle che consideriamo nel modello.

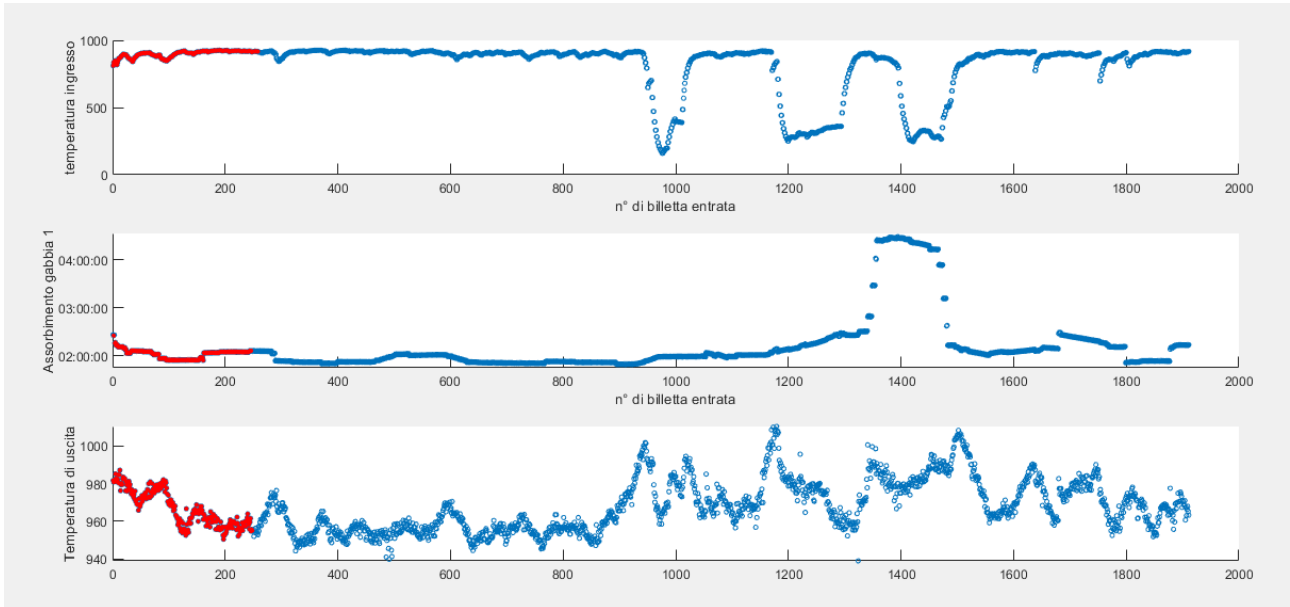


Figura 32

Utilizziamo le prime 250 billette prodotte per questa tipologia per ottenere il modello andiamo a stimare i valori degli assorbimenti per le successive.

La previsione degli assorbimenti è riportata nella seguente figura:

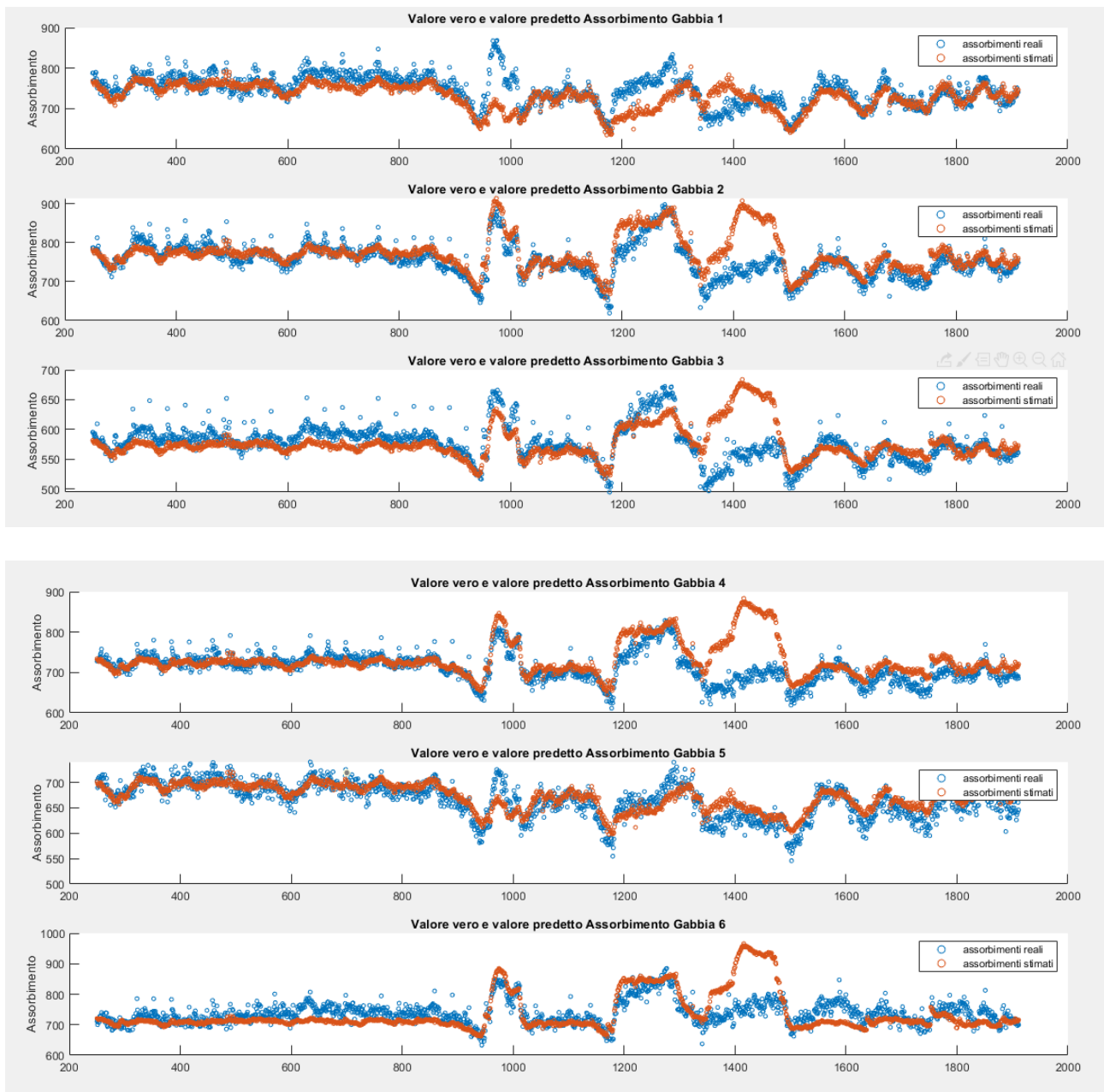


Figura 33

Due considerazioni sono doverose: ovviamente la stima dipende dalla bontà del modello ottenuto: se ho un modello che stima male i valori fornitigli (nel nostro caso le 250 billette iniziali prodotte), ovviamente le previsioni saranno poco coerenti con ciò che aspettiamo. Inoltre, la stima produrrà una buona previsione sugli assorbimenti, nel caso in cui le condizioni di lavoro delle billette utilizzate per stimare il modello, siano simili a quelle di cui ho interesse nel prevedere gli assorbimenti.

Come possiamo vedere nella figura 32, dalla 900° billetta circa, vengono prodotte billette con temperatura in ingresso molto basse, nel nostro modello non ne erano presenti ed infatti (figura 33) da lì, gli assorbimenti iniziano ad essere sbagliati.

Per questo motivo ci deve essere una nuova fase di training per il modello, ogni qualvolta si presenta un “cambio di condizioni” quali ad esempio una sosta o se ad esempio stanno per entrare billette tanto fredde etc.

Prendendo infatti come dati per il modello tutte quelle billette che entrano in forno con una temperatura bassa (si veda figura 34), la stima degli assorbimenti è simile a quella reale (si ha però una variabilità dovuta al tempo nel forno che dalla 1350° prodotta aumenta, visibile in figura 32)

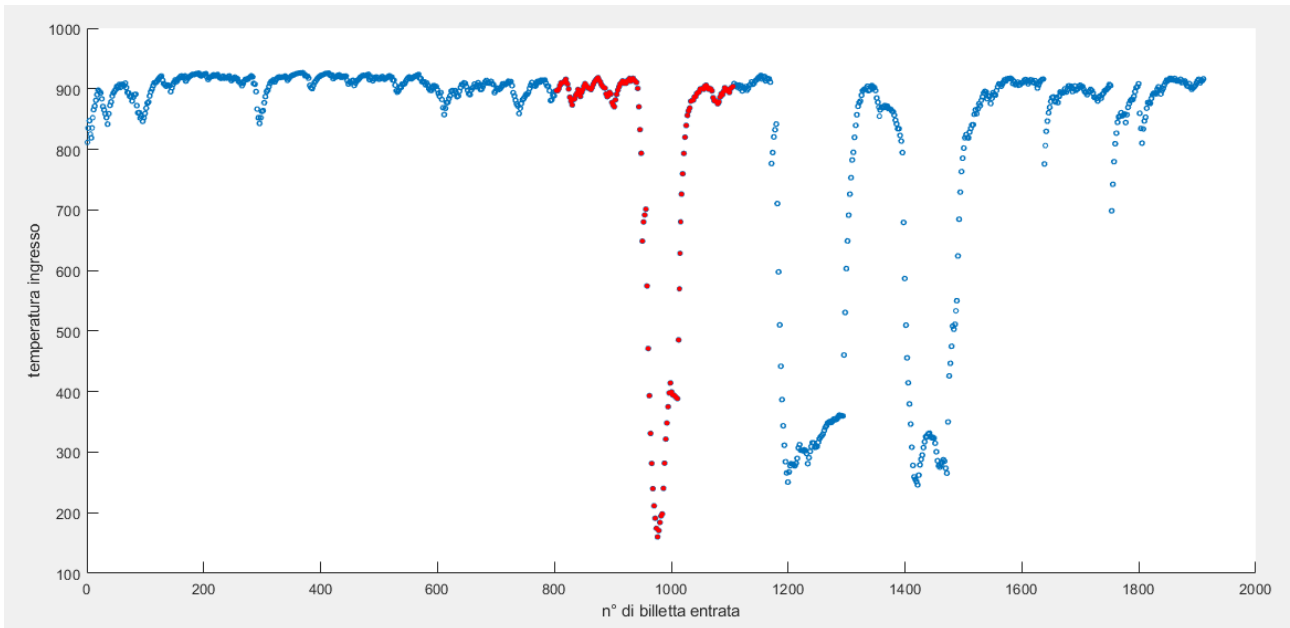


Figura 34

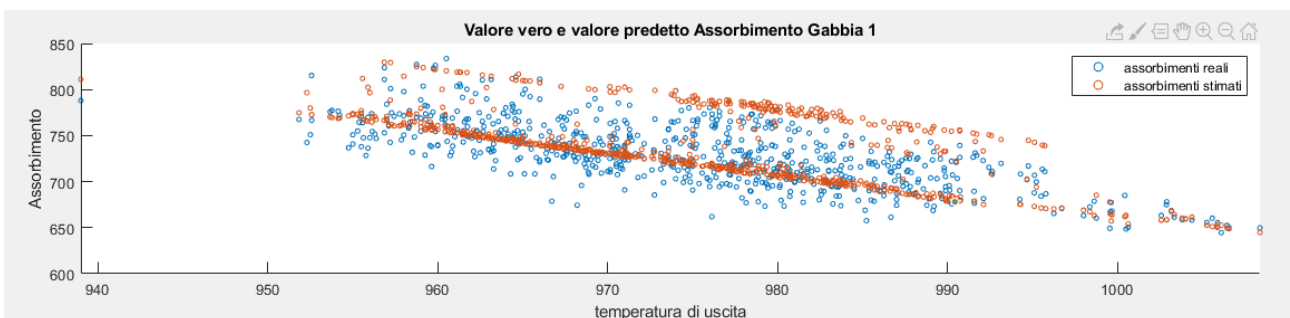


Figura 34

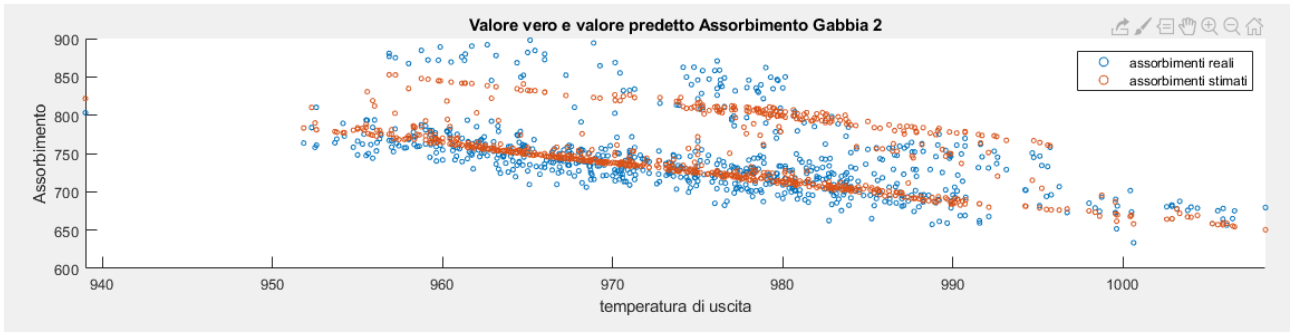


Figura 35

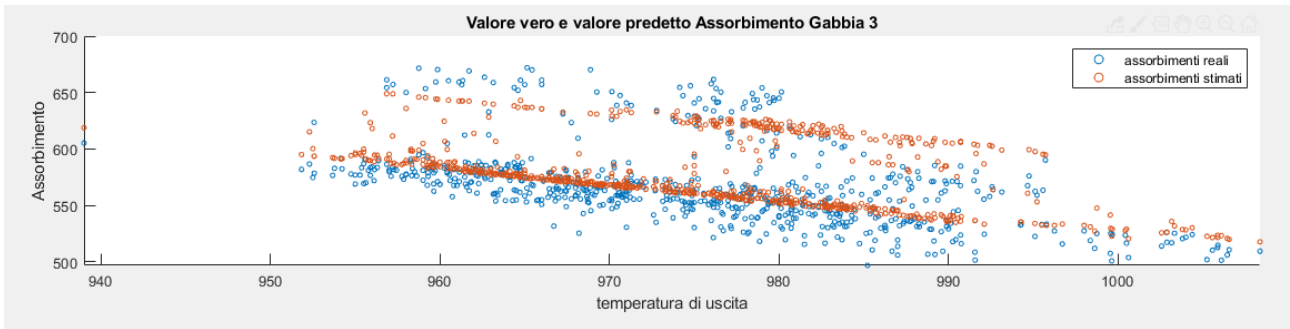


Figura 36

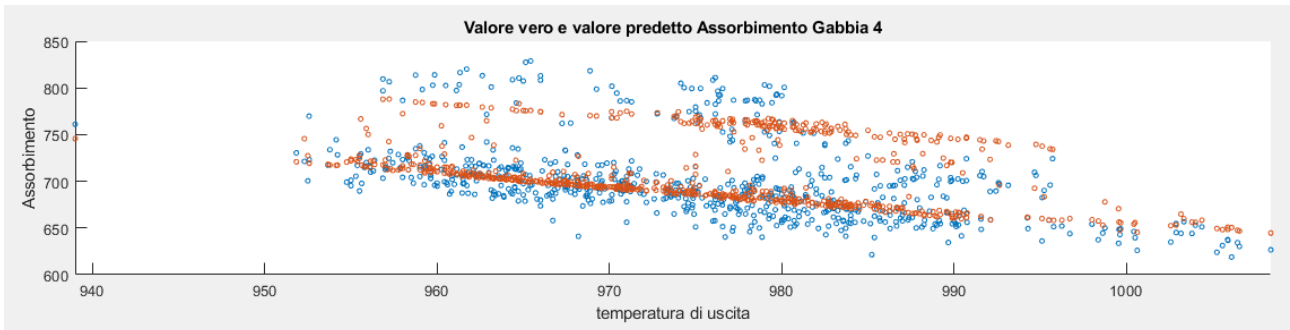


Figura 37

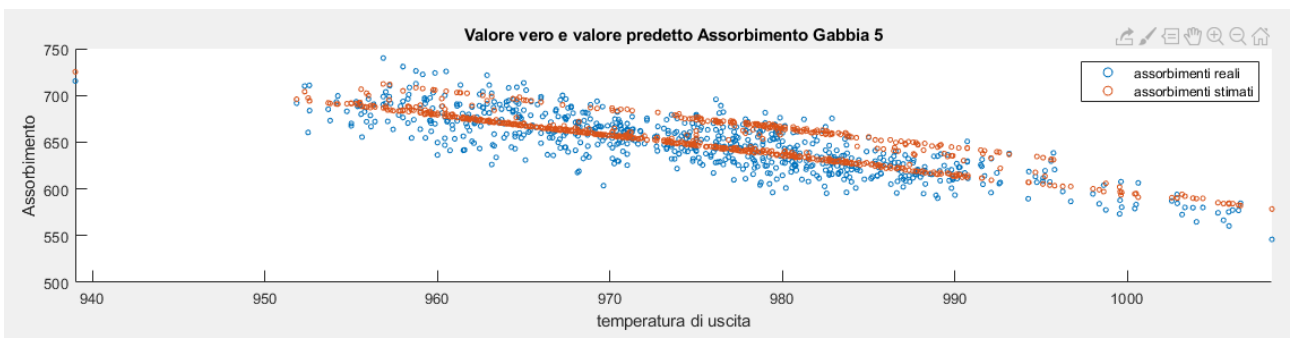


Figura 38

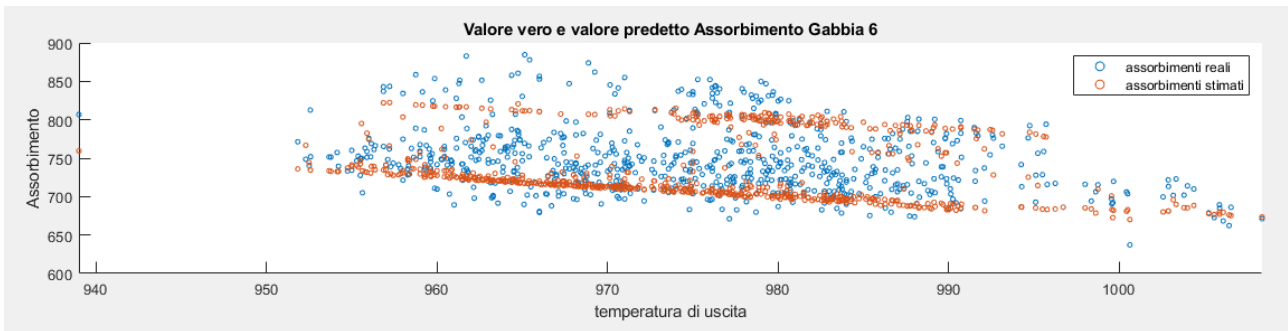


Figura 39

Una cosa particolare si è riscontrata quando, si sono messi a confronto due modelli fatti su un numero diverso di osservazioni (200 e 950): notiamo come per le gabbie 2-4-5, gli assorbimenti sono molto simili. Questo sottolinea ulteriormente quanto non è tanto la quantità dei campioni considerati, quanto le loro caratteristiche che caratterizzano un buon modello. Sul perché proprio quelle gabbie risultino molto simili, verrà chiesto con gli ingegneri del progetto.

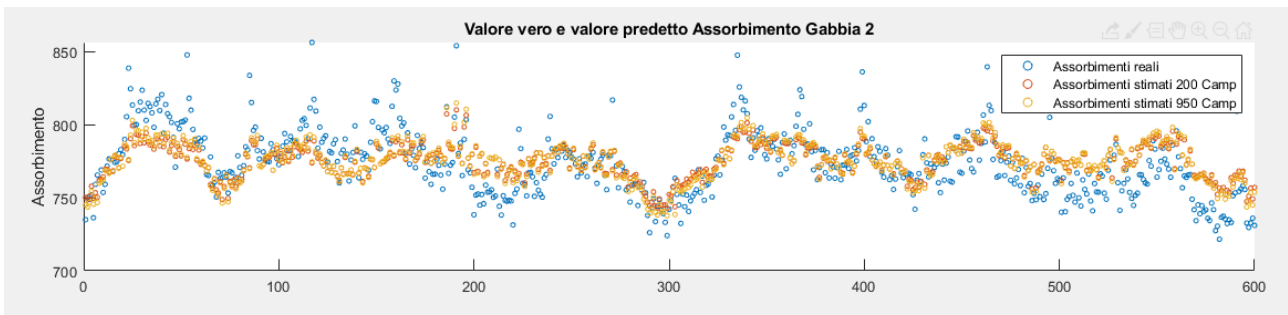


Figura 40

Capitolo 5

Problematica allineamento testa - coda

Per quanto riguarda la seconda problematica, ovvero ottenere la temperatura della testa e della coda all'uscita della terza gabbia uniforme, lo studio è stato il lavoro è stato limitato per manca . Verranno comunque esposte le problematiche affrontate, le considerazioni e i passi futuri dal quale partire.

Come già chiarito in precedenza è indispensabile che la billetta, nella fase di laminazione rispetti predefiniti range di temperatura affinché si abbia una lavorazione soddisfacente in termini qualitativi ed economici.

Nei file di log fornitici dall'azienda si vede che ,mediamente, il tempo di processamento di una billetta è 1min 40 sec; la lavorazione della cosiddetta coda non avviene dunque immediatamente dopo l'uscita della billetta dal forno ma trascorre un tempo variabile.

Come prima cosa, abbiamo raccolto le informazioni riguardanti le billette quali ad esempio data e ora di ingresso e di uscita della billetta nel forno, temperatura di ingresso, etc. per sapere a quale billetta stessimo facendo riferimento in ogni momento. Fatto ciò, poiché Per quanto riguarda la misura della temperatura di uscita a valle della terza gabbia, il valore fornito dal controllore risulta

un singolo valore ottenuto dal il processamento del segnale relativo alla temperatura in uscita durante tutta la fase di laminazione e quindi a partire dalla testa e a finire dalla coda della billetta. Questo dato non risultata utile allo scopo della presente analisi, avendo bisogno dell'intero treno di misure e non di un solo valore. Per questo motivo è stato necessario andare a leggere i valori dei pirometri istante per istante, dal "campo".

Ciò che risulta, andando a leggere i valori dal "field data" è segnale di questo tipo (figura 41) dove in ascissa è riportato l'asse temporale, mentre in ordinata si ha il valore della temperatura (°C) letto in quell'istante di tempo.

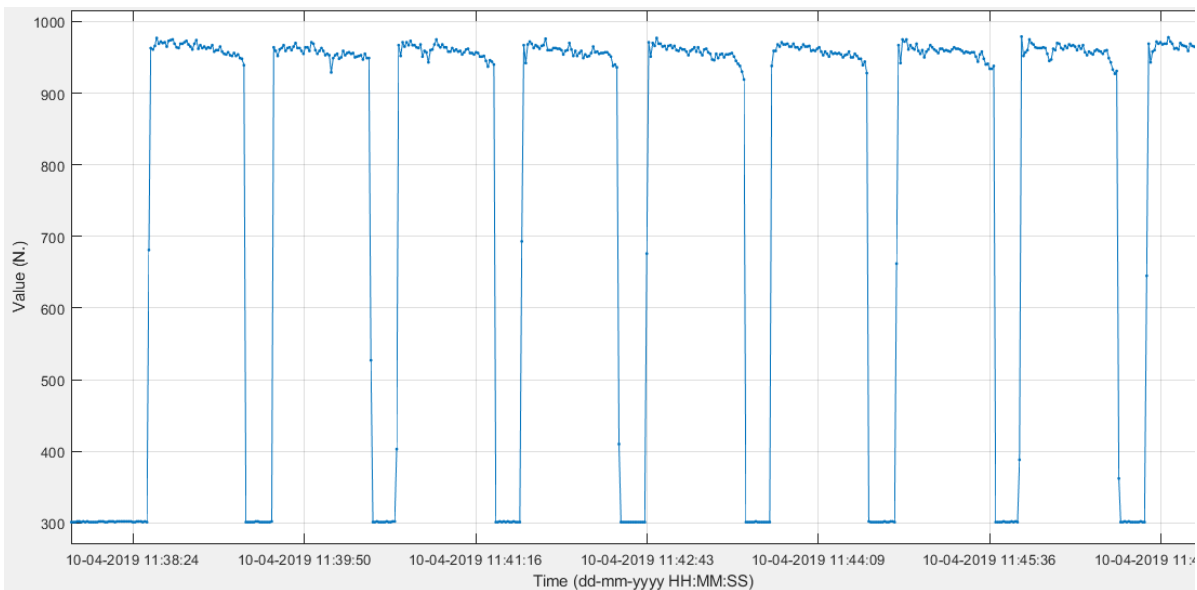


Figura 41 Valori di temperatura letti dai pirometri

In figura 41 possiamo distinguere due principali " fasi di lettura: una relativa alla billetta che passa davanti al pirometro, ed una relativa alla lettura "a riposo" del pirometro, quando non c'è nessuna billetta.

L'andamento non "stabile" delle letture ha richiesto di mettere a punto una procedura di filtraggio per poter associare valori corretti alla testa e alla coda delle billette.

Analizzando i dati si è scelto di discriminare sia l'inizio che la fine della lettura in funzione del valore assunto dalla temperatura: nel momento in cui si ha una transizione da temperatura bassa, ad una temperatura $\geq 780^{\circ}\text{C}$ (fronte di salita del segnale), si inizia ad acquisire il treno di misure, che verrà interrotto nel momento in cui si ha il primo valore di temperatura $< 780^{\circ}\text{C}$ (fronte di discesa).

Si è scelto di eliminare i primi e gli ultimi valori del treno poiché relativi a letture (di temperatura calda), prima che arrivi la billetta davanti al pirometro e dopo che la billetta ha superato il pirometro; infatti, essendo le billette molto calde, riscaldano l'aria vicina al pirometro e ne falsano i primi e gli ultimi valori della lettura.

Un esempio di valori acquisiti del treno di misure è presentato nella figura 42 dove il segnale rosso è quello associato alla temperatura salvata, mentre quello blu è il segnale di lettura continua del pirometro.

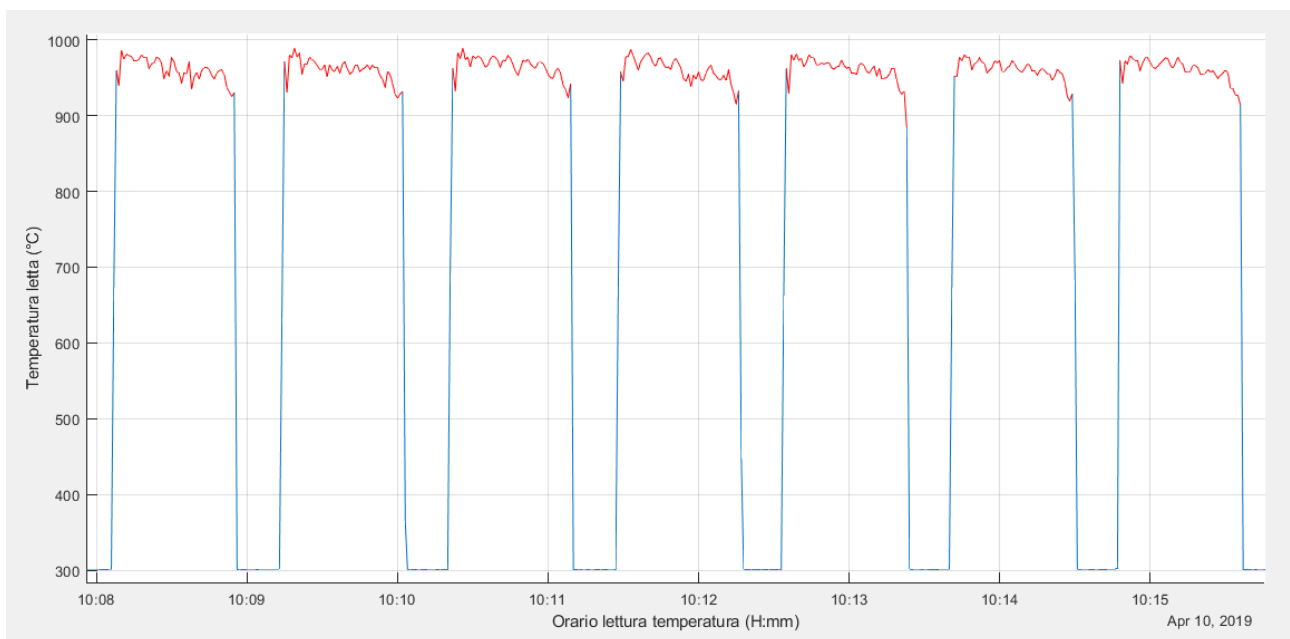


Figura 42 Esempio treno temperatura acquisito

Dovendo definire una temperatura per la testa della billetta ed una per la coda, dobbiamo elaborare ogni treno ottenuto.

Inizialmente, si è ipotizzato di utilizzare i primi k campioni per la testa e gli ultimi k per la coda e farne la media.

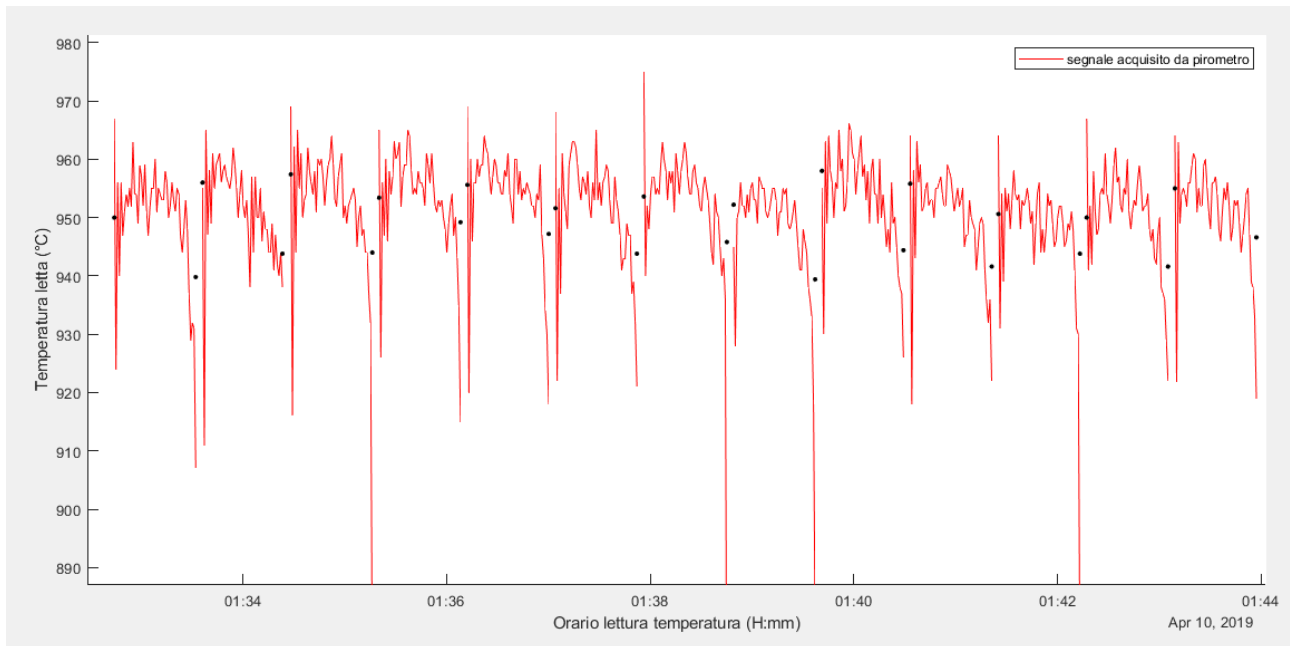


Figura 43 Valori temperatura testa coda ottenuti, rispettivamente mediante **media dei primi ed ultimi 5 istanti di misura**

Attraverso questo semplice metodo si ottiene un valore abbastanza veritiero per i punti analizzati ma, considerando la totalità di letture del treno, sembra che l'oscillazione iniziale dei valori, porti ad una sottostima della temperatura per la testa.

Questa variabilità si presenta poiché ogni billetta che passa di fronte il pirometro lo fa molto rapidamente; questo ci porta a dover filtrare ogni treno di letture.

Per ovviare a questa problematica si sono applicate diverse tecniche e i risultati sono stati analizzati e messi a confronto

Sono stati valutati le seguenti tecniche di filtraggio:

- a media mobile
- Savitzky-Golay

In figura 30, 31, 32 sono riportati 3 segnali:

- grigio: misura di temperatura letta dal pirometro,
- blu: misura filtrata mediante la tecnica Savitzky-Golay,
- rosso: misura filtrata mediante filtro a media mobile.

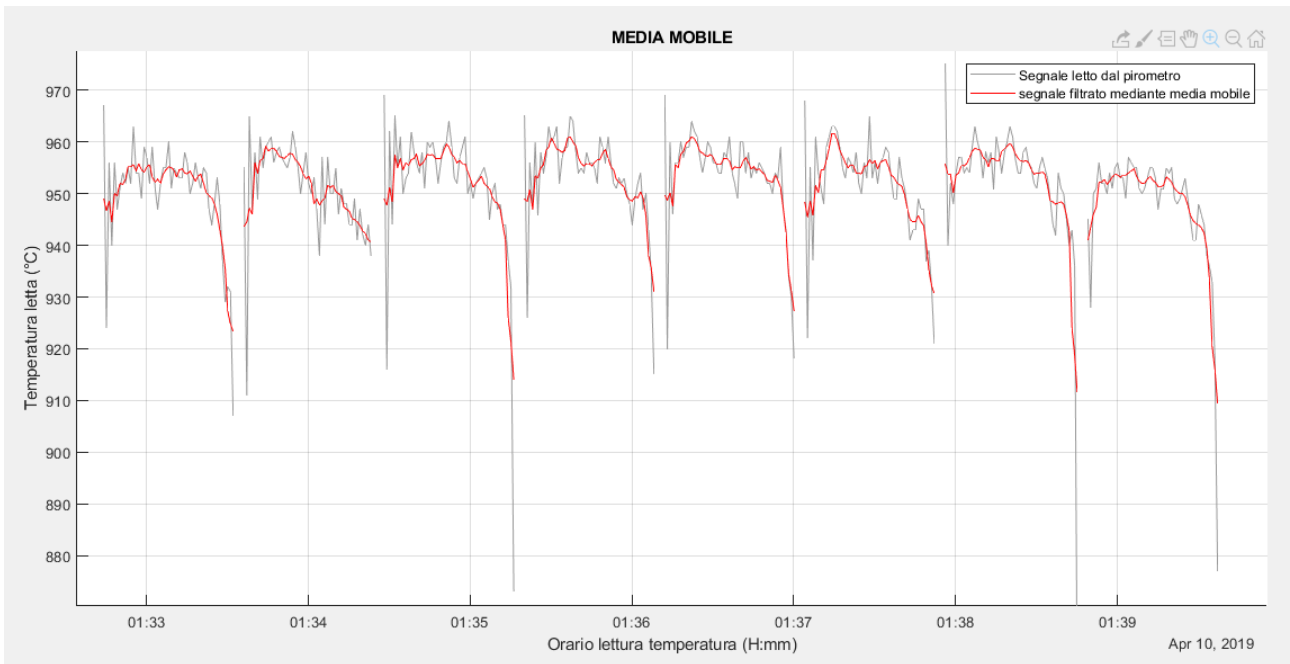


Figura 44-Filtro mediante media mobile

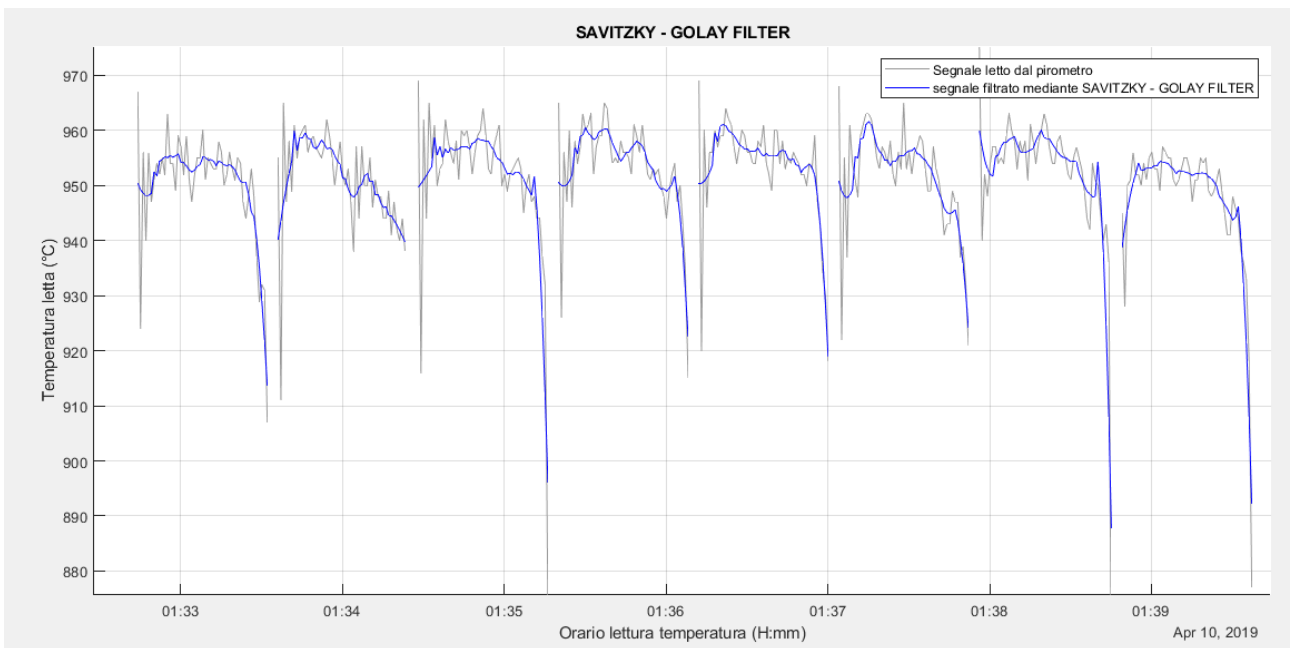


Figura-45 Filtro mediante Savitzky - Golay filter

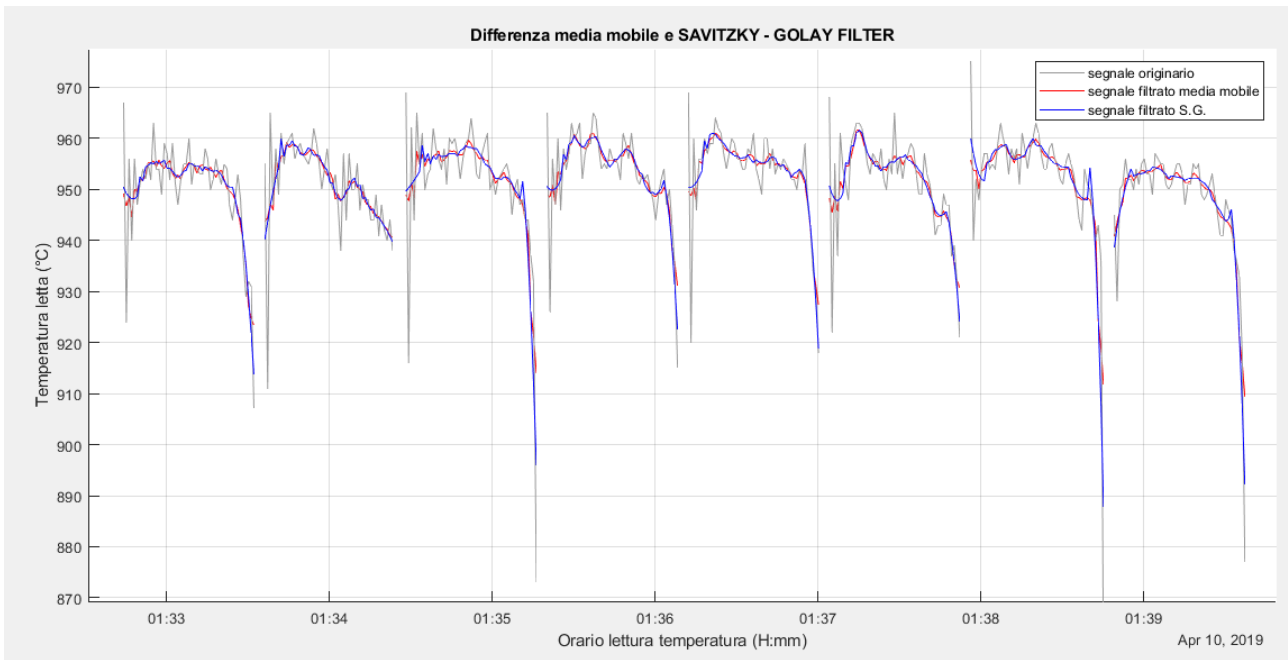


Figura-46 Segnali sovrapposti

Dall'analisi dei risultati si può notare innanzitutto che le prestazioni dei due metodi di filtraggio, relativamente al caso di nostro interesse, siano sostanzialmente simili. Inoltre, si nota che nonostante il filtraggio, non si riesce a definire un valore soddisfacente di temperatura per la testa e la coda poiché l'andamento della temperatura rimane altalenante.

Ciò che è stato fatto quindi è approssimare il segnale, sia filtrato che non, con una retta e una parabola.

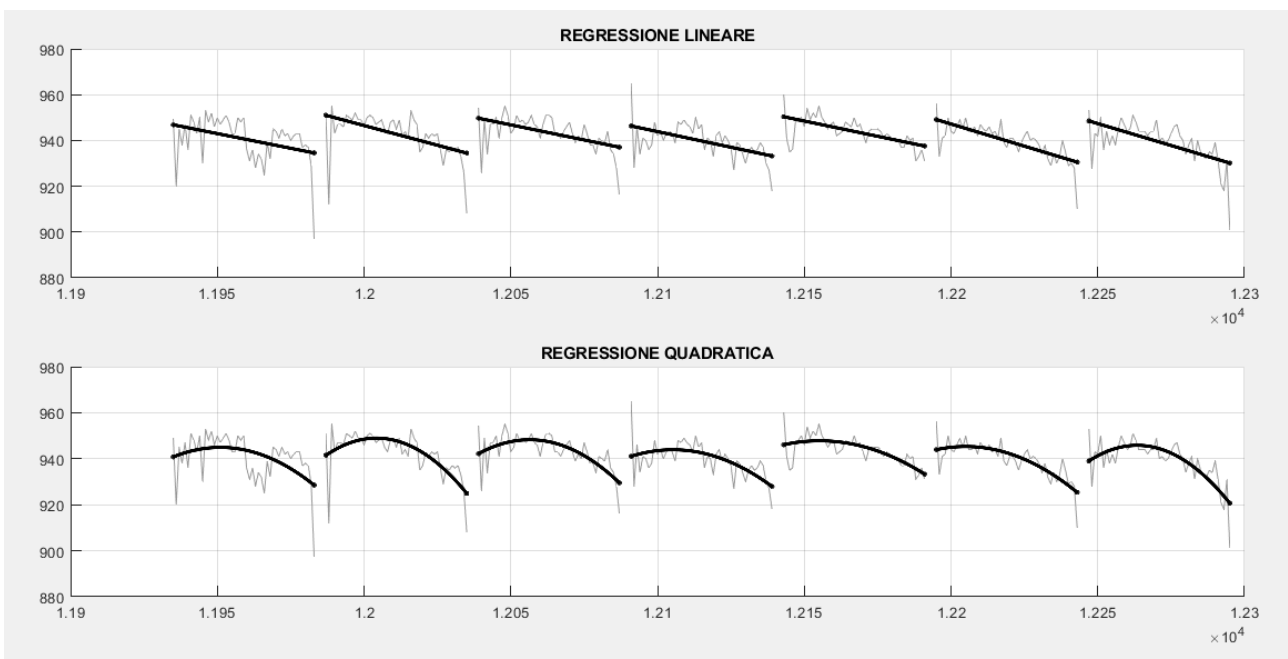


Figura 47 Regressione sul segnale originale

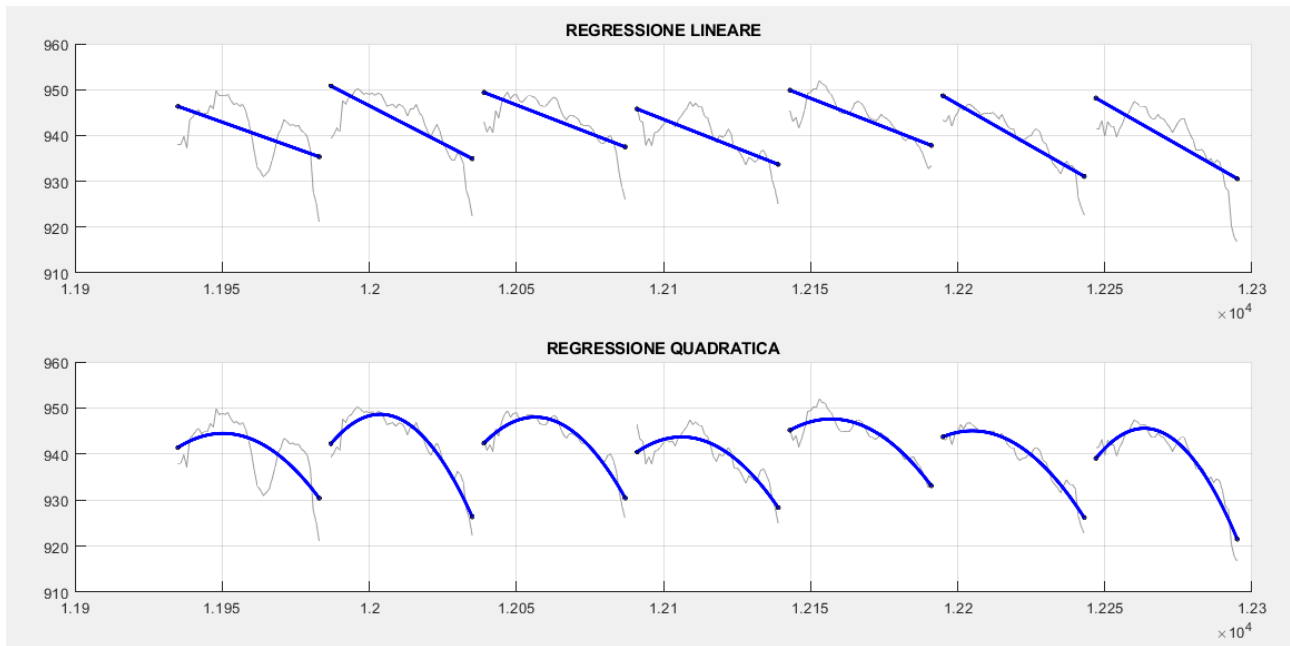


Figura 48 Filtro sul segnale a media mobile

Si è ritenuto migliore come metodo per stimarne la temperatura di testa e di coda (non per l'andamento complessivo) il metodo che approssima il segnale non filtrato con una retta. Non sono state rilevate sostanziali differenze applicando la regressione al segnale filtrato.

È opportuno, una volta scelto il metodo per stabilire la temperatura per la testa e per la coda, operare un ulteriore controllo i valori ottenuti, poiché, capitano treni di lettura che danno luogo, mediante regressione, ad una temperatura di coda inferiore a quella di testa. Questi casi, come si può notare dalla figura 35 sono relativi ad andamenti anomali del segnale del pirometro per quella singola billetta. Pertanto, tali billette sono state escluse dall'analisi. Si è quindi messa a punto una logica per l'eliminazione di tali dati.

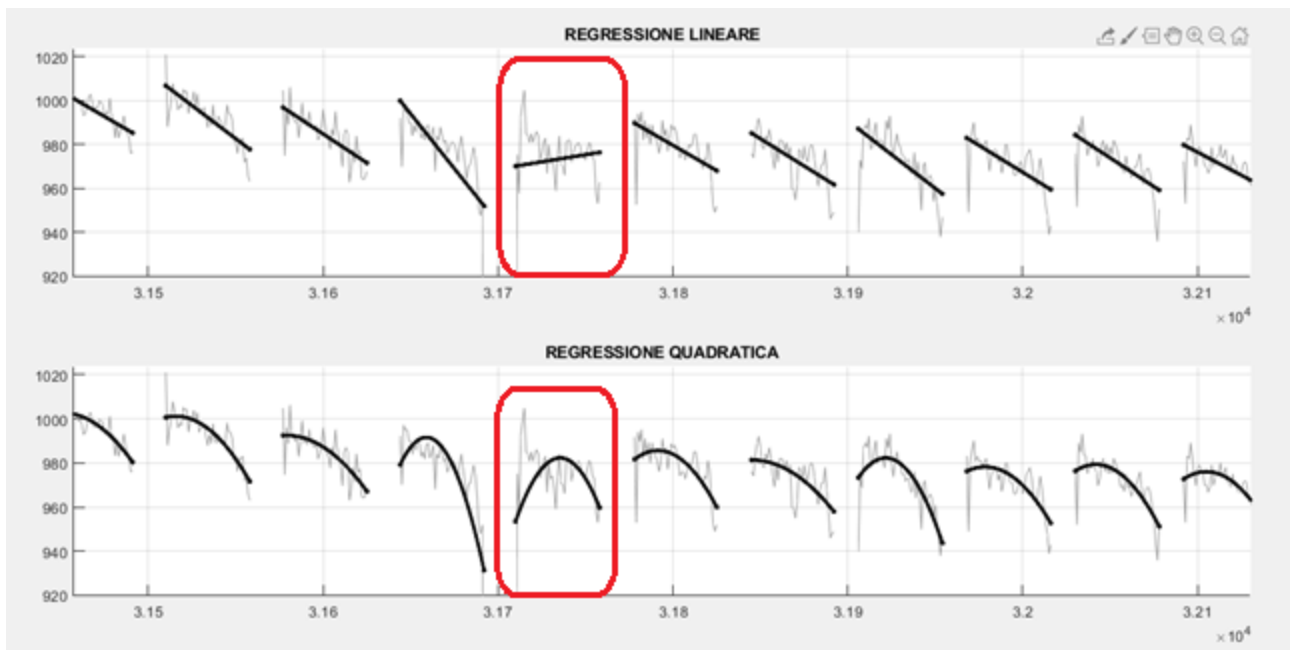


Figura 49 Anomalia assegnazione temperatura testa coda

Altra variabile di interesse ai fini della valutazione dello sbilanciamento testa-coda sono i valori dei bruciatori relativi alle Zone 4 e 5 che, come spiegato nel paragrafo 2, influenzano rispettivamente la testa e la coda della billetta. A tal fine, si sono acquisiti i valori dei set point e delle misure di temperatura dei bruciatori nella regione finale del forno, nel periodo in cui la billetta ne era all'interno.

Dal file di log fornitoci, è possibile ricavare l'istante temporale in cui ogni billetta esce dal forno. Da questo, si è deciso con gli ingegneri dell'azienda, di andare a considerare i cinque minuti precedenti all'uscita dal forno della billetta ed acquisire i valori dei bruciatori. Anche in questo caso si avrà un treno di misure per ogni billetta che dovrà essere elaborato per ottenere un singolo valore per ogni bruciatore.

Oltre alle misure di temperature di zona però è da tenere in considerazione il valore dei set point dei relativi bruciatori poiché sono i valori a cui i bruciatori tendono.

Come è mostrato in figura 36, il segnale relativo ai bruciatori (quello colorato) è distante un Δt (variabile nel tempo) al segnale di set point (segnale grigio) appunto perché il set point è l'obiettivo, ma la traiettoria con cui tali set point vengono raggiunti è legata alle dinamiche interne del sistema. Inoltre, tali set point non sono costanti ma variano nel tempo.

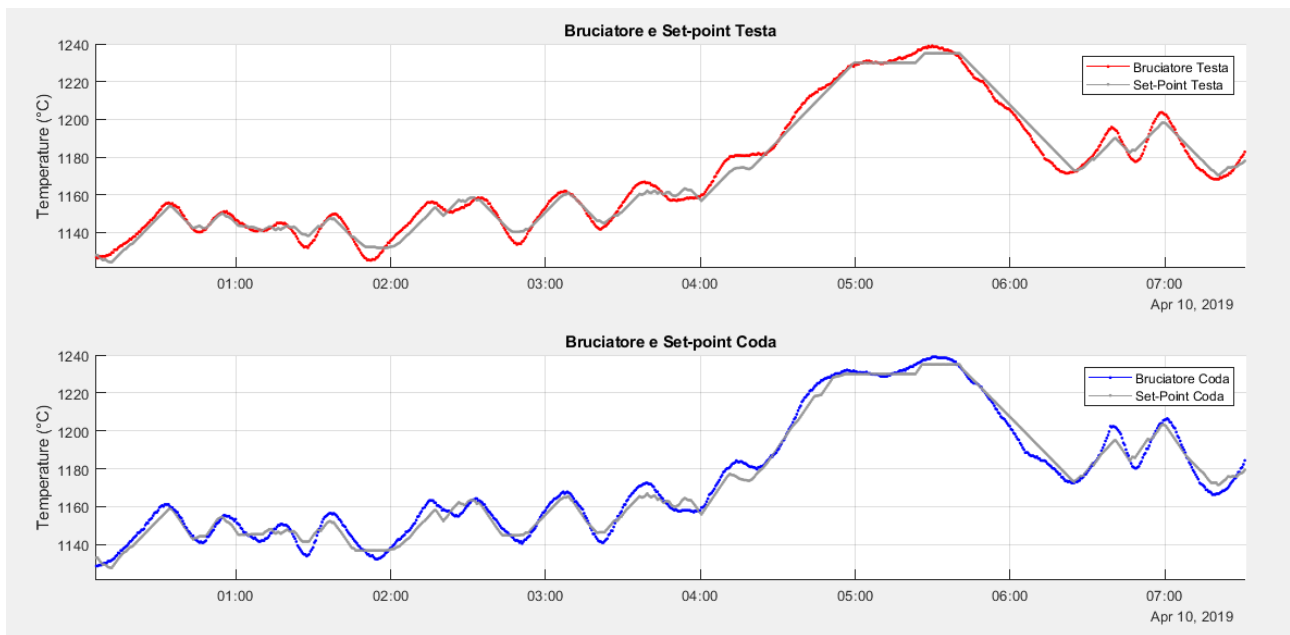


Figura 50 Segnali set point e bruciatori nel tempo

Tale problematica per il momento non è stata ulteriormente approfondita per la necessità di ottenere ulteriori dati e per mancanza di tempo. Nel capitolo seguente sono espone le considerazioni relative.

Capitolo 6

Conclusioni

Nel presente lavoro di tesi sono stati trattati problemi riguardante processi legati all'industria dell'acciaio; in particolare sia l'aspetto legato ai processi di laminazione che gli aspetti legati ai forni di riscaldamento, fase che precede la fase di laminazione.

Il primo problema affrontato ha riguardato la stima degli assorbimenti delle gabbie nella fase di sbizzo. Dopo aver analizzato i dati storici di funzionamento dell'impianto, a tale riguardo, si è adottato un approccio data-based, applicando tecniche di regressione multivariata. I risultati ottenuti hanno evidenziato che la stima degli assorbimenti è strettamente legata al tipo di produzione considerata ad esempio il diametro del prodotto finito. Anche nella stessa produzione però sono state notate differenze negli assorbimenti, al variare del periodo produttivo, mentre su periodi prolungati ma continuativi, tali differenze non sono state riscontrate.

La grande variabilità di condizioni operative, pur nella stessa tipologia di prodotto, comporta talvolta, una limitata capacità di predizione del modello stimato.

Come discusso nei risultati, se nei dati di addestramento sono presenti più condizioni operative, le prestazioni del modello, in termini predittivi, saranno migliori, adattando la predizione alla

specifica condizione operativa. Altre soluzioni, limitate al caso di produzioni contigue, prevedono un opportuno addestrando del modello.

Dai dati a disposizione non è stato possibile ottenere modelli soddisfacenti che vadano bene per produzioni, su periodi non contigui. Dalle analisi svolte ci sembra di poter attribuire questo problema al fatto che i settaggi delle gabbie possano differire nei diversi periodi, nonostante siano la stessa produzione. Queste informazioni sono impostate manualmente dagli operatori e non sono reperibili nei file di log forniti. Eventualmente, in seguito ad un colloquio con gli ingegneri del progetto, le informazioni saranno richieste e in caso possibile introdotte per le produzioni future.

Per quanto riguarda invece la seconda problematica affrontata ovvero rendere minimo lo sbilanciamento di temperatura tra la testa e la coda della billetta, all'uscita della terza gabbia, l'analisi è stata limitata allo sviluppo di algoritmi per il trattamento dei segnali grezzi, non presenti nel data log, al fine di ottenere le informazioni utili all'identificazione del modello richiesto. Dopo varie analisi è stato definito un metodo efficiente per la stima del valore di temperatura di testa e di coda e per i valori di temperatura della billetta nel forno che influenzano lo sbilanciamento testa – coda.

Per lo sviluppo del modello richiesto, dall'analisi effettuata è emerso che sarebbe utile approfondire i seguenti punti:

- considerando una media statistica dei valori assunti dai bruciatori in quel lasso di tempo, quanta differenza di temperatura si ottiene a valle della terza gabbia tra “testa e coda”?
- C'è, e in caso affermativo, che tipo di correlazione lega il valore dei set point con la differenza “testa – coda”?
- C'è variazione nel momento in cui vengono acquisiti, non gli ultimi 5 minuti, ma l'intero intervallo di tempo che la billetta trascorre nella quarta regione (dove abbiamo i due bruciatori per la testa e per la coda), magari introducendo un modello termodinamico che ne modelli le caratteristiche in funzione di quanto e in che modo percorre quella regione?

Per quanto riguarda l'ultimo punto, l'informazione relativa alla regione del forno in cui si trova la singola billetta in un certo istante non è ottenibile dai dati che ci sono stati messi a disposizione (se

non in modo approssimato), ma dovrebbe essere possibile modificare i file di log per memorizzare anche questa informazione che risulterebbe molto utile.

BIBLIOGRAFIA E SITOLOGIA

1. <https://it.wikipedia.org/wiki/Automazione>
2. Kalpakjian S., Schmid S. R. "Tecnologia Meccanica; Pearson Prentice Hall, 2012
3. Trinks, W., Mawhinney, M. H., Shannon, R. A., Reed R. J., Garvey, J. R.: Industrial Furnaces. John Wiley & Sons, New York, USA , 2004.
4. <http://www.fiom.brescia.it/site/home/servizi/ufficio-ambiente-e-sicurezza/profili-di-rischio-in-metallurgia/articolo25003164.html>
5. Zanolli, S. M., Pepe, C., Barboni, L.: Application of Advanced Process Control techniques to a pusher type reheating furnace. Journal of Physics: Conference Series 659(1) (2015).
6. Pepe C., Model Predictive Control aimed at energy efficiency improvement in process industries, Ph. D. thesis, 2017.
7. Astolfi, G., Barboni, L., Cocchioni, F., Pepe, C.: Metodo per il controllo di forni di riscaldamento. Italian Patent n. 0001424136 awarded by Ufficio Italiano Brevetti e Marchi, 2016.
8. www.sitelcontrol.it
9. www.ametek-land.com
10. William Navidi Probabilità e statistica per l'ingegneria e le scienze McGraw-Hill 2006
11. <https://statistics.laerd.com/statistical-guides/pearson-correlation-coefficient-statistical-guide.php>
12. <https://towardsdatascience.com/an-ode-to-r-squared-804d8d0ed22c>