

**UNIVERSITÀ DEGLI STUDI DI PADOVA**  
DIPARTIMENTO DI INGEGNERIA INDUSTRIALE  
CORSO DI LAUREA MAGISTRALE IN INGEGNERIA CHIMICA E DEI PROCESSI INDUSTRIALI

**Tesi di Laurea Magistrale in  
Ingegneria Chimica e dei Processi Industriali**

**Formulazione di schiume poliuretatiche rigide mediante  
tecniche multivariate di machine learning**

*Relatore: Prof. Pierantonio Facco*  
*Correlatore: Ing. Nicola Mela*

*Laureando: MASSIMO MATTIELLO*

ANNO ACCADEMICO 2022 – 2023



# **Prefazione**

Questa Tesi è stata svolta in collaborazione con Europoliuretani s.r.l. (Trebaseleghe, PD, Italia), azienda leader nella produzione di sistemi per la produzione di schiume poliuretatiche rigide, che ha fornito i dati su cui è stata elaborata questa Tesi e le strutture e la strumentazione analitica per lo svolgimento della campagna sperimentale e della raccolta dati.



# Abstract

Rigid polyurethane foams are polymeric materials, produced through simultaneous expansion and polymerization reactions, mainly used for thermal insulation purposes. The formulation of chemicals with which these foams are realized has a great impact on the final product quality, determined through time profiles of foam growth velocity and temperature gradients.

The aim of the Thesis is the development of multivariate machine learning methodologies for both the characterization of the foams and the formulation of the polyurethane systems. First of all, methodologies for the study of the expansion process are developed by means of a multi-way principal components analysis (PCA, Wise e Gallagher, 1996, Camacho *et al.*, 2008) and a pattern recognition technique based on similarity indices (Ottavian *et al.* 2012) applied to time profiles of growth velocity and temperature gradient. These methodologies guarantee a in-depth understanding of the polyurethane systems and a science-based method which demonstrate to be much more efficient than the analytical state-of-the-art methods, which are mostly based on operators' experience. Furthermore, a reduction of 50-75% of the duration of the product quality laboratory analysis is proposed. Finally, based on the inversion (Tomba *et al.*, 2016) of a partial least squares regression methods (PLS, De Jong, 1993) calibrated on data from designed experimental campaigns, a tool is developed to suggest the most appropriate formulation which guarantees the desired final product quality in terms of growth velocity and foam's height time profiles.



# Riassunto

Le schiume poliuretatiche rigide sono materiali polimerici, prodotti attraverso reazioni simultanee di espansione e polimerizzazione, principalmente usati nel campo dell'isolamento termico. La formulazione delle sostanze chimiche con le quali queste schiume vengono prodotte ha un impatto notevole nella qualità del prodotto finale, determinata attraverso profili temporali di velocità di crescita e gradiente di temperatura. Lo scopo della Tesi è l'introduzione di tecniche multivariate di machine learning per la caratterizzazione e la formulazione dei sistemi poliuretatici.

Innanzitutto, sono state sviluppate metodologie per lo studio del processo di espansione attraverso l'analisi delle componenti principali (PCA, Wise e Gallagher, 1996, Camacho *et al.*, 2008) e una tecnica di *pattern recognition* basato su fattori di similarità (Ottavian *et al.* 2012) applicate ai profili temporali di velocità di crescita e gradiente di temperatura. Queste metodologie garantiscono una profonda comprensione dei sistemi poliuretatici e un metodo scientifico molto più efficiente dei metodi analitici allo stato dell'arte, quasi esclusivamente basati sull'esperienza dell'operatore. Inoltre, viene proposta una riduzione del 50-75% della durata dell'analisi sperimentale della qualità del prodotto. Infine, basandosi sull'inversione (Tomba *et al.*, 2016) di metodi di regressione lineare ai minimi quadrati parziali (PLS, De Jong, 1993), calibrato su dati provenienti da campagne sperimentali pianificate, è stato sviluppato uno strumento in grado di suggerire la formulazione più appropriata per garantire la qualità del prodotto finale desiderata in termini di profili temporali di velocità di crescita e altezza della schiuma.



# Indice

<b>INTRODUZIONE.....</b>	<b>1</b>
<b>CAPITOLO 1.....</b>	<b>3</b>
<b>METODOLOGIE DI ANALISI MULTIVARIATA DEI DATI E PIANIFICAZIONE DEGLI ESPERIMENTI .</b>	<b>3</b>
1.1 ANALISI MULTIVARIATA DEI DATI.....	3
1.1.1 <i>Approccio matematico-statistico</i> .....	3
1.1.2 <i>Pretrattamento dei dati</i> .....	5
1.1.3 <i>Analisi delle componenti principali</i> .....	6
1.1.3.1 Formulazione matematica del modello PCA.....	6
1.1.3.2 Rappresentazione geometrica del modello PCA.....	8
1.1.3.3 Selezione del numero ottimale di componenti principali.....	9
1.1.3.4 Diagnostica dell'osservazione.....	10
1.1.3.5 Organizzazione dei dati.....	11
1.1.3.6 Metodologie di <i>Machine Learning</i> .....	12
1.1.3.7 Analisi attraverso <i>similarities</i> .....	13
1.1.4 <i>Proiezione su strutture latenti</i> .....	13
1.1.4.1 Criticità dei comuni metodi di regressione:.....	14
1.1.4.2 Formulazione matematica del modello PLS.....	14
1.1.4.3 Interpretazione geometrica di PLS.....	16
1.1.4.4 Inversione del modello PLS.....	17
1.2 PIANIFICAZIONE DEGLI ESPERIMENTI.....	18
1.2.1 <i>Central Composite Design (CCD)</i> .....	19
1.2.2 <i>Algoritmo di Kennard-Stone</i> .....	20
<b>CAPITOLO 2.....</b>	<b>23</b>
<b>TECNOLOGIE DEI SISTEMI POLIURETANICI RIGIDI .....</b>	<b>23</b>
2.1 SCHIUME POLIURETANICHE RIGIDE.....	23
2.1.1 <i>Storia, mercato, tipologie ed utilizzi dei poliuretani</i> .....	23
2.1.2 <i>Chimica delle schiume poliuretaniche rigide</i> .....	24
2.1.2.1 Reazione di gel.....	25
2.1.2.2 Reazione di <i>blow</i> .....	26
2.1.2.3 Reazioni di reticolazione e trimerizzazione.....	27
2.1.3 <i>Proprietà fisiche delle schiume poliuretaniche</i> .....	28
2.1.3.1 Morfologia.....	28
2.1.3.2 Densità e proprietà meccaniche.....	29
2.1.4 <i>Formulazione di una schiuma poliuretanic</i> a.....	31
2.1.4.1 Isocianato.....	31
2.1.4.2 Polieteri e Poliesteri.....	33
2.1.4.3 Catalizzatori.....	35
2.1.4.4 Agenti Espandenti.....	35
2.1.4.5 Antifiamma e tensioattivi.....	36
2.2 PROCESSO DI PRODUZIONE E TECNICHE DI ANALISI.....	36
2.2.1 <i>Produzione di una schiuma poliuretanic</i> a rigida.....	36
2.2.2 <i>Stato dell'arte su analisi e formulazione del processo</i> .....	37
2.2.3 <i>Campagna sperimentale</i> .....	39
2.2.3.1 Procedura di analisi sperimentale.....	39
2.2.4 <i>Dati sperimentali disponibili e formulati di riferimento</i> .....	41
2.2.4.1 Dati di calibrazione.....	42
2.2.4.2 Dati di convalida.....	43
<b>CAPITOLO 3.....</b>	<b>45</b>
<b>METODOLOGIE DI MACHINE LEARNING PER LO STUDIO DELLA CINETICA DI ESPANSIONE DI SISTEMI POLIURETANICI RIGIDI.....</b>	<b>45</b>
3.1 METODOLOGIA MULTIVARIATA DI ANALISI DI SISTEMI POLIURETANICI RIGIDI.....	45

3.2	APPLICAZIONE DELLA METODOLOGIA .....	46
3.2.1	<i>Pretrattamento e visualizzazione dei dati</i> .....	46
3.2.1.1	Filtrazione dei dati.....	46
3.2.1.2	Trasformazione dei dati.....	48
3.2.1.3	Visualizzazione dei profili .....	50
3.2.2	<i>Analisi esplorativa</i> .....	52
3.2.2.1	Costruzione del modello globale PCA .....	52
3.2.2.2	Risultati dell'analisi esplorativa .....	53
3.2.2.3	Caratterizzazione della cinetica dei formulati .....	58
3.2.2.3	Utilizzo dell'analisi esplorativa per la classificazione dei formulati .....	59
3.2.3	<i>Analisi approfondita di un singolo formulato</i> .....	61
3.2.4	<i>Analisi comparativa dei formulati</i> .....	63
3.2.5	<i>Conclusioni dell'applicazione dell'analisi</i> .....	64
3.3	CONFRONTO TRA LA NUOVA METODOLOGIA E L'ANALISI ALLO STATO DELL'ARTE.....	65
3.3.1	<i>Riduzione dei tempi di analisi</i> .....	66
3.4	CONCLUSIONI E POSSIBILI SVILUPPI .....	68
<b>CAPITOLO 4.....</b>		<b>69</b>
<b>SVILUPPO DI UN MODELLO PREDITTIVO PER LA FORMULAZIONE DI SISTEMI POLIURETANICI RIGIDI.....</b>		<b>69</b>
4.1	CAMPAGNE SPERIMENTALI.....	69
4.1.1	<i>Campagna sperimentale per lo studio dell'effetto della catalisi</i> .....	70
4.1.2	<i>Campagna sperimentale per lo studio dell'effetto congiunto di tutti i componenti</i> .....	72
4.2	COSTRUZIONE DEL MODELLO PREDITTIVO .....	76
4.2.1	<i>Predizione e studio dell'effetto della catalisi</i> .....	76
4.2.1.1	Costruzione del modello a superficie di risposta mediante <i>PLS</i> .....	77
4.2.1.2	Interpretazione del modello <i>PLS</i> .....	80
4.2.1.3	Risultato della convalida del modello .....	83
4.2.2	<i>Predizione e studio dell'effetto di tutte le componenti</i> .....	85
4.2.2.1	Costruzione del modello <i>PLS</i> e problema di <i>aliasing</i> .....	85
4.3	CONCLUSIONI E POSSIBILI SVILUPPI .....	87
<b>CONCLUSIONI.....</b>		<b>89</b>
<b>APPENDICE A .....</b>		<b>91</b>
<b>PROVE SPERIMENTALI ED ULTERIORI DATI .....</b>		<b>91</b>
<b>NOMENCLATURA .....</b>		<b>101</b>
<b>RIFERIMENTI BIBLIOGRAFICI .....</b>		<b>105</b>

# Introduzione

Il poliuretano è un polimero che si ottiene combinando tra loro sostanze diverse, di cui le principali sono polioli e poliisocianati. Si tratta di un materiale molto versatile che trova largo impiego in molteplici settori di applicazione industriale: dalle forme compatte ad alta densità per *wood imitation*, agli adesivi, alle guarnizioni sigillanti fino ad arrivare agli espansi flessibili (imbottiture nel settore arredamento) e agli espansi rigidi la cui principale funzione è l'isolamento termico.

Le forme espanse si ottengono realizzando un materiale la cui morfologia potrebbe essere definita bifasica, dove una fase solida contiene del gas intrappolato nella matrice polimerica. Le zone contenenti il gas, vengono chiamate *celle*. La particolare struttura *a celle* si ottiene attraverso l'espansione del materiale, che avviene parallelamente alla polimerizzazione, attraverso reazioni (*blowing*) innescate in fase di costruzione del reticolo polimerico e che permettono lo sviluppo di gas. Lo studio e la corretta modulazione della fase di espansione del materiale, così come la valutazione della cinetica delle reazioni chimiche collegate, risultano quindi di importanza fondamentale per il processo di produzione dello stesso, in quanto andranno a determinare la resa in termini di produttività e la qualità del prodotto finale.

Le tecniche di caratterizzazione del processo di espansione e della cinetica chimica attualmente utilizzate nella pratica industriale, sono basate su metodi empirici e sull'esperienza dell'operatore deputato al controllo qualità. Naturalmente queste analisi risultano poco efficienti ed oggettive.

In letteratura, sono presenti numerosi modelli matematici il cui obiettivo è simulare l'intero processo di reazione e di produzione del materiale espanso (Al-Moameri *et al.*, 2016); la complessità del sistema rende però molto complicato il loro utilizzo a livello industriale, in quanto essi tendono ad essere molto specifici oppure eccessivamente approssimativi per via delle numerose assunzioni semplificative.

Lo sviluppo di un nuovo metodo oggettivo, basato su analisi obiettive, ripetibili e riproducibili, su metodologie matematiche e statistiche, potrebbe rappresentare un notevole passo in avanti per quanto riguarda la tecnologia e la formulazione dei materiali poliuretanicici che permetterebbe il distacco dai metodi attuali di analisi.

Le tecniche analitiche che usualmente vengono utilizzate prevedono la raccolta di larghe quantità di dati, in particolare, vengono misurati profili temporali di altezza e temperatura del materiale in fase di espansione poi successivamente analizzati solamente attraverso un confronto grafico. Inoltre, il controllo della qualità del prodotto viene eseguito principalmente attraverso delle misurazioni dei tempi caratteristici del processo di produzione del materiale, che forniscono una rapida interpretazione della scala temporale su cui avviene il fenomeno ma sono puramente basati su misurazioni empiriche non oggettive e difficilmente ripetibili.

In questa Tesi si propone una procedura innovativa basata su tecniche multivariate di *machine learning* per ottenere una caratterizzazione dei sistemi poliuretanicici obiettiva, ridurre i tempi (e quindi i costi) di analisi legati alla fase di convalida e al contempo di distaccarsi da un approccio di tipo

prettamente empirico. Inoltre, le tecniche proposte sono in grado di aiutare la formulazione di nuovi prodotti con prestazioni desiderate.

La Tesi è quindi strutturata in diverse sezioni. Nel Capitolo 1 si introducono i concetti matematici generali delle metodologie di *machine learning* e delle tecniche di pianificazione ed analisi degli esperimenti. Nel Capitolo 2 si descrive la chimica dei poliuretani, inquadrando inoltre il caso preso in esame. Successivamente, nel Capitolo 3 verranno presentati i risultati del metodo di analisi dell'espansione di diversi formulati mediante analisi delle componenti principali (PCA, Wise e Gallagher, 1996), estesa a metodiche multi-way per lo studio dei profili temporali. Infine, nel Capitolo 4 saranno illustrati i risultati dell'estensiva campagna sperimentale condotta per lo sviluppo di una tecnica di formulazione di nuovi prodotti che sia in grado di proporre il dosaggio dei componenti della schiuma basato sull'inversione di un modello di regressione ai minimi quadrati parziali (PLS, De Jong, 1993) al fine di garantire la qualità desiderata finale al sistema poliuretano.

# Capitolo 1

## Metodologie di analisi multivariata dei dati e pianificazione degli esperimenti

Questa Tesi è stata svolta utilizzando tecniche di analisi statistica multivariata, nello specifico metodologie di machine learning, sia supervisionato che non, e campagne sperimentali pianificate mediante tecniche di pianificazione degli esperimenti.

In questo Capitolo, verranno presentate le basi teoriche su cui queste tecniche si fondano.

### 1.1 Analisi multivariata dei dati

L'analisi multivariata dei dati consiste nella manipolazione di grandi quantità di informazioni con l'obiettivo di rimuovere il rumore e tutta l'informazione che non è legata ai fenomeni studiati.

Queste tecniche vengono anche utilizzate come modelli matematici semi-empirici descrittivi e predittivi.

Le due metodologie utilizzate per lo svolgimento della Tesi sono l'Analisi delle Componenti Principali (*Principal Components Analysis, PCA*; Wise e Gallagher, 1996) e la Proiezione su Strutture Latenti (*Projection to Latent Structure, PLS*; De Jong, 1993). Nell'ingegneria di processo le due tecniche si differenziano per lo scopo: PCA ha come scopo una rappresentazione più chiara e facilmente comprensibile di osservazioni e fenomeni racchiusi in grandi quantità di dati, la cui comprensione è spesso complessa e limitata. PLS permette di analizzare allo stesso tempo due tipologie di dati: variabili indipendenti, cioè le condizioni di processo, caratteristiche di catalizzatori e materie prime, e fattori di un esperimento, e variabili dipendenti, cioè i parametri di qualità del prodotto

#### 1.1.1 Approccio matematico-statistico

L'analisi statistica consiste nello studio del comportamento di una o più variabili e di come queste siano fra loro correlate quando caratterizzate da incertezza. Una variabile  $x$  [u. d. m.] può essere definita come segue:

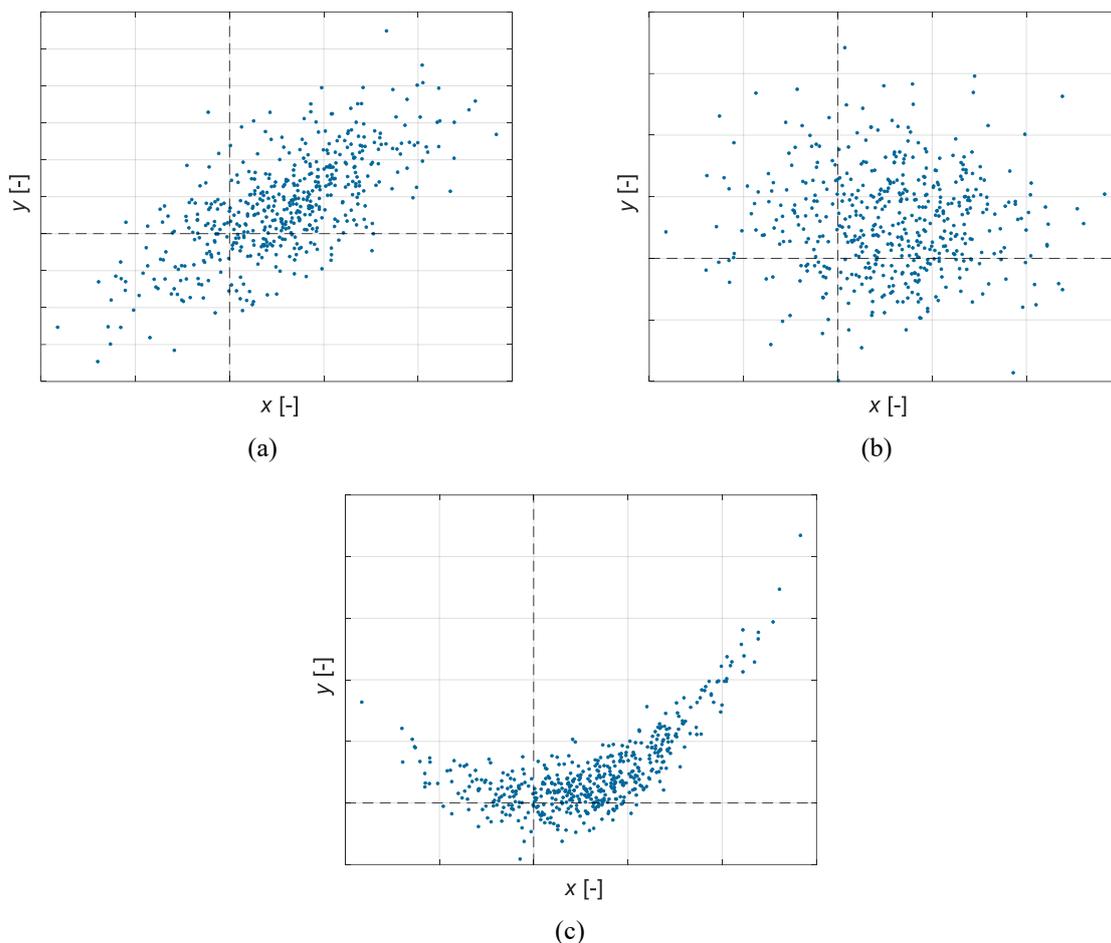
$$x = \eta + \varepsilon,$$

$$\eta \text{ [u. d. m.]} = \text{parte sistematica dell'informazione}, \quad (1.1)$$

$$\varepsilon \text{ [u. d. m.]} = \text{rumore}.$$

Misurando il valore di  $x$  più volte, a seconda delle condizioni che interferiscono con la misura stessa, si ottiene una distribuzione di valori. Ogni diversa misurazione  $i$  è chiamata osservazione ed è identificata come  $x_i$  [u. d. m.].

L'insieme di più variabili deve essere studiato attraverso un approccio multivariato, perché gran parte delle variabili in un sistema altamente multivariato sono autocorrelate e correlate fra di loro, ovvero seguono precisi profili e dipendono le une dalle altre.



**Figura 1.1.** Rappresentazione di diverse possibilità di correlazioni fra variabili: (a) variabili non correlate; (b) variabili linearmente correlate; (c) variabili con correlazione non lineare.

È quindi necessario studiarle assieme per coglierne la natura. Come è possibile osservare dalla Figura 1.1, prendendo in considerazione due variabili  $x$  [u. d. m.] e  $y$  [u. d. m.], esse possono essere fra loro indipendenti oppure possono essere correlate secondo un preciso profilo. Ad esempio, Figura 1.1 (a) mostra due variabili non correlate tra loro, il caso di correlazione lineare è invece rappresentato in Figura 1.1 (b): all'aumentare del valore di una variabile, anche l'altra aumenta allo stesso modo. La correlazione fra due variabili può anche essere non lineare ma seguire andamenti diversi, come in Figura 1.1 (c).

La covarianza  $\sigma_{xy}$  [u. d. m.], parametro che indica come le due variabili variano assieme, è definita come segue:

$$\sigma_{xy} = \text{covarianza} = \frac{1}{N} \sum_i (x_i - \mu_x)(y_i - \mu_y). \quad (1.2)$$

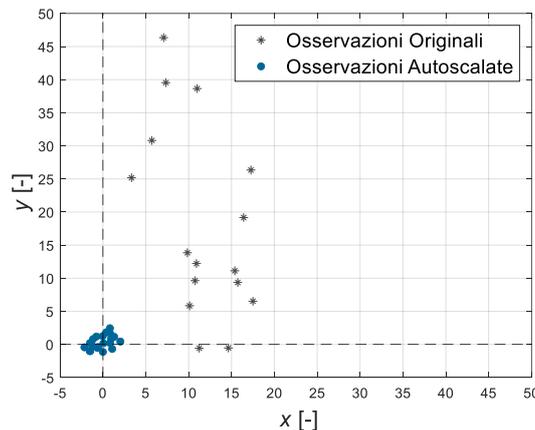
Numerosi problemi sorgono, dunque, quando vengono studiate più variabili allo stesso tempo:

- dimensionalità elevata: all'aumentare del numero di variabili, aumenta la difficoltà nell'interpretazione dei risultati;
- multicollinearità: la presenza di un gran numero di correlazioni fra diverse variabili genera molta confusione in fase di comprensione delle informazioni;
- rumore;
- incompletezza dei dati: talvolta i dataset possono essere parzialmente incompleti.

### 1.1.2 Pretrattamento dei dati

Considerando che, in un sistema di più variabili, esse possono assumere valori su scale di grandezza diverse, esistono numerose tecniche per pretrattare i dati in modo da renderli esplicativi e farli pesare allo stesso modo sull'analisi. La tecnica più utilizzata è l'*autoscaling* (riferimento): consiste nel centrare tutte le osservazioni di ogni variabile sulla loro media  $\mu_x$  (*mean-centering*) e scalarle sulla loro deviazione standard, la radice della varianza  $\sigma_x$ :

$$x_{\text{autoscalata}} [-] = \frac{[x - \mu_x]}{\sqrt{\sigma_x}}. \quad (1.3)$$



**Figura 1.2.** Effetto dell'*autoscaling* su variabili generiche caratterizzate da un diverso ordine di grandezza.

L'effetto di questa tecnica è visibile in Figura 1.2: entrambe le variabili autoscalate variano sullo stesso ordine di grandezza e la loro media è zero.

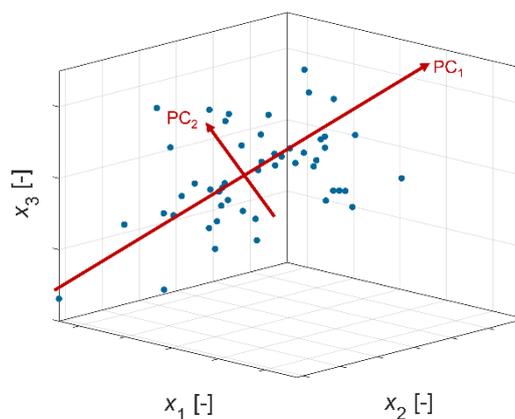
Una ulteriore metodologia di pretrattamento dei dati che può essere applicata, viene definita *Standard Normal Variate (SNV; riferimento!)*. La tecnica *SNV* è molto simile all'*autoscaling*, in quanto vengono eseguite le medesime operazioni, tuttavia, invece di applicarle ad ogni variabile, vengono

applicate ad ogni osservazione. La tecnica è particolarmente utile nei casi in cui le variabili da utilizzare sono profili altamente autocorrelati, quali spettri o profili temporali.

### 1.1.3 Analisi delle componenti principali

PCA (Wise e Gallagher,1996) è un metodo di riduzione della dimensionalità di un dataset che permette di studiare una grande quantità di informazioni eliminando il rumore.

Il metodo consiste nel generare, a partire dallo spazio delle osservazioni di partenza, un nuovo spazio virtuale le cui direzioni sono le direzioni di massima variabilità dei dati, catturando solamente le informazioni utili. Queste nuove direzioni sono chiamate componenti principali (*Principal Components, PC*).



**Figura 1.3.** Costruzione dello spazio delle PC da uno spazio di tre variabili misurate: la prima componente principale  $PC_1$  rappresenta la direzione di variabilità più elevata, mentre la  $PC_2$  rappresenta la direzione di variabilità più elevata ortogonale alla precedente.

Proiettando i dati nel nuovo spazio virtuale sono facilmente individuabili:

- raggruppamenti fra osservazioni simili;
- correlazioni fra le variabili;
- correlazioni fra le osservazioni;
- andamenti e profili delle variabili;
- osservazioni anomale;
- relazioni fra variabili e osservazioni.

#### 1.1.3.1 Formulazione matematica del modello PCA

Al fine di costruire il modello, la prima operazione consiste nel ricavare i cosiddetti *loadings*  $\mathbf{P}$ , i quali sono i coseni direttori delle PC. Dato un dataset autoscalato  $\mathbf{X} [N \times V]$ , dove  $N$  rappresenta il numero di osservazioni e  $V$  il numero di variabili, viene costruita la matrice di autocorrelazione definita come:

$$\mathbf{X}^T\mathbf{X} = \text{matrice di autocorrelazione} . \quad (1.4)$$

Essa rappresenta anche la covarianza fra le diverse variabili del dataset  $\mathbf{X}$ .

Le direzioni delle PC devono soddisfare due requisiti: catturare la massima variabilità del dataset  $\mathbf{X}$  ed essere ortogonali e indipendenti fra loro, in modo da rappresentare fenomeni diversi (cioè ortogonali e quindi indipendenti):

$$\text{condizione di massima variabilità: } \max_{\mathbf{P}}(\mathbf{P}^T\mathbf{X}^T\mathbf{X}\mathbf{P}), \quad (1.5)$$

$$\text{condizione di ortogonalità: } \mathbf{P}\mathbf{P}^T = \mathbf{1}. \quad (1.6)$$

Le dimensioni della matrice  $\mathbf{P}$  è  $[V \times A]$ , dove  $A$  è il numero di PC.

La proiezione di ogni osservazione sullo spazio delle PC è chiamata *score*  $\mathbf{t}_{n,v}$ , dove  $n$  è una osservazione del dataset, mentre  $v$  è una PC. L'insieme di queste  $N$  proiezioni consiste in una matrice di dimensione  $[N \times A]$  ed è calcolata come:

$$\mathbf{T} = \mathbf{X}\mathbf{P}. \quad (1.7)$$

La formulazione di PCA può essere riformulata in un problema di decomposizione ai valori singolari che per la  $a$ -esima componente principale è:

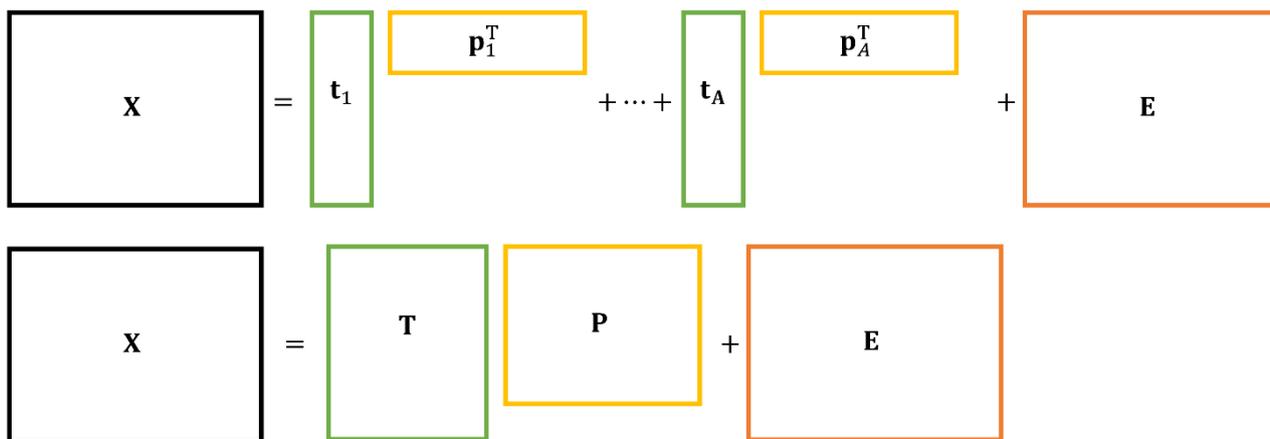
$$\mathbf{X}^T\mathbf{X}\mathbf{p}_a = \lambda_a\mathbf{p}_a \quad , \quad (1.8)$$

dove  $\mathbf{p}_a$  è l'autovettore della matrice di covarianza per l' $a$ -esima PC e rappresenta la direzione dell' $a$ -esima PC, mentre  $\lambda_a$  è l'autovettore associato. L'operazione viene ripetuta più volte per ogni  $a = 1, 2, \dots, A$ , applicandola ad ogni iterazione all'informazione residua alla precedente. In questo modo viene rimossa ad ogni step una parte dell'informazione, finché rimarrà soltanto il rumore e la parte di informazione non sistematica, definita come residuo  $\mathbf{E}$   $[N \times V]$ .

Quindi, il modello PCA che viene ottenuto può essere rappresentato nel modo seguente (Wise e Gallagher, 1996):

$$\mathbf{X} = \mathbf{T}\mathbf{P}^T + \mathbf{E} = \sum_i^A \mathbf{t}_i\mathbf{p}_i^T + \mathbf{E}, \quad (1.9)$$

$$\hat{\mathbf{X}} = \mathbf{T}\mathbf{P}^T = \text{approssimazione di } \mathbf{X} \text{ sulle } A \text{ PC}. \quad (1.10)$$



**Figura 1.4.** Rappresentazione schematica del modello PCA: la matrice  $X [N \times V]$  può essere rappresentata dalla somma dei prodotti di score e loading (che forniscono l'approssimazione nello spazio  $A$  dimensionale delle PC) e la matrice dei residui  $E [N \times V]$ .

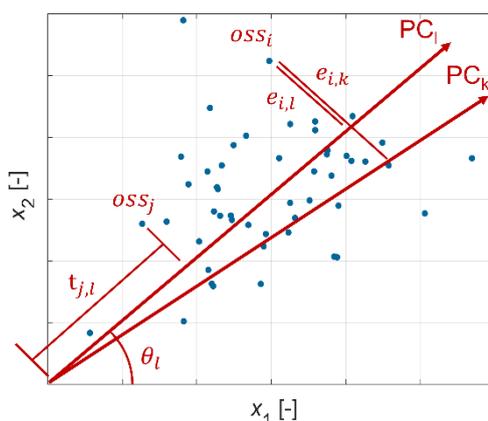
### 1.1.3.2 Rappresentazione geometrica del modello PCA

Come descritto precedentemente, il *loading* di una variabile latente rappresenta la direzione della PC, cioè l'angolo fra la PC e indica come la  $v$ -esima variabile viene rappresentata nella  $a$ -esima PC. Si può, quindi, dire che i *loading* rappresentano le correlazioni fra le variabili.

Gli *score*, invece, sono definiti come proiezioni delle osservazioni nello spazio delle PC; perciò, è possibile affermare che rappresentano le relazioni fra le diverse osservazioni. Di conseguenza, le osservazioni la cui proiezione è vicina nello spazio degli score saranno simili e avranno valori simili delle variabili.

La distanza di una osservazione dallo spazio delle PC è il residuo.

Graficamente, è possibile dimostrare come prendendo la direzione di variabilità massima, si ha intrinsecamente una riduzione dell'errore di approssimazione.



**Figura 1.5.** Rappresentazione geometrica del modello PCA: la distanza dal centro dello spazio virtuale è lo score; l'angolo  $\theta$ , descritto dal loading rappresenta la direzione della PC, la distanza fra l'osservazione e il nuovo spazio della PC è il residuo. Importante notare come la scelta della direzione la cui variabilità non è massima ( $PC_k$ ), comporta un aumento dell'errore  $e_i$ .

In Figura 1.5, il modello PCA è rappresentato in modo grafico, a dimostrazione dei concetti precedentemente espressi.

### 1.1.3.3 Selezione del numero ottimale di componenti principali

Dal punto di vista matematico ed operativo, possono essere trovate tante PC quante il rango della matrice di autocorrelazione con la quale il modello è costruito. All'aumentare del numero di PC incluse nel modello, aumenta la varianza catturata dal modello.

Tuttavia, un numero eccessivo di variabili latenti causerebbe l'inclusione di parte del rumore che si vuole eliminare. Quindi serve una metodologia per la scelta del numero più appropriato di PC da includere nel modello.

Esistono diversi test che aiutano la scelta del numero adatto di PC ai fini di costruire il modello. Quelli utilizzati nella Tesi sono:

- *scree test* (Hartmann *et al.*,2018): è una procedura visuale per cui nel grafico della varianza catturata cumulativa vs. numero di PC viene selezionato il numero di PC che corrisponde al gomito della curva; superato il numero ideale di PC da includere, la varianza introdotta aumenta in modo poco significativo (perciò la curva tenderà ad appiattirsi);
- *eigenvalue test* (K. Hartmann *et al.*, 2018): l'autovalore corrispondente ad ogni PC indica indirettamente la quota di varianza che quest'ultima rappresenta e, secondo una regola pratica, il numero di variabili originali che vengono rappresentate in una singola variabile latente. Se l'autovalore è minore di 1, allora la PC corrispondente rappresenterà meno informazioni di una delle variabili originali, perciò non viene ritenuta utile ai fini della rappresentazione fornita dal modello;
- *convalida incrociata mediante test sul RMSECV (Root of Minimum Squared Error in Cross Validation, Wise e Gallagher,1996)*: si sceglie il numero di PC che minimizza l'errore in fase di previsione; il numero di PC viene però scelto prima di utilizzare il modello, cioè quando si ha a disposizione solamente i dati di calibrazione. Perciò, si possono utilizzare diverse strategie per selezionare una parte di dati; questa viene poi rimossa, e viene costruito il modello sui dati rimanenti; l'errore di previsione viene calcolato utilizzando l'informazione eliminata precedentemente come convalida del modello ricavato con i dati parziali. Iterando più volte la procedura, in modo da provare tutte le combinazioni con la matrice di calibrazione e in modo da utilizzare ogni volta un numero diverso di PC, calcolando ad ogni iterazione il valore di RMSECV come descritto in 1.11(Wise e Gallagher,1996):

$$RMSECV = \sqrt{\frac{\mathbf{E}^T \cdot \mathbf{E}}{\text{numero di osservazioni}}} \quad (1.11)$$

### 1.1.3.4 Diagnostica dell'osservazione

Molto spesso i modelli PCA vengono utilizzati come tecniche di controllo statistico del processo (*Statistical Process Control, SPC*) per il monitoraggio di processo, al fine di determinare se esso segua un comportamento normale o se qualche anomalia si stia verificando mediante test d'ipotesi. In particolare, mediante due indici statistici è possibile capire se una nuova osservazione del processo sia conforme alla norma o irregolare.

Il primo indice è il cosiddetto *Hotelling  $T^2$*  e misura la distanza fra la proiezione della nuova osservazione nel modello e il centro dello spazio latente (Wise e Gallagher, 1996):

$$T^2 = \mathbf{t}_n^T \mathbf{\Lambda}^{-1} \mathbf{t}_n = \sum_a \frac{\mathbf{t}_a^2}{\lambda_a}, \quad (1.12)$$

$$\mathbf{\Lambda} = \text{matrice diagonale degli autovalori} = \begin{pmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_A \end{pmatrix}.$$

Ogni qualvolta un'osservazione presenti una anomalia di  $T^2$  significa che la sua proiezione è lontana dalle condizioni medie; perciò, quell'osservazione avrà qualche variabile che presenta dei valori estremi e quindi, anomali.

Il secondo parametro utilizzato è lo *Squared Prediction Error SPE* che rappresenta il residuo della nuova osservazione (Wise e Gallagher, 1996):

$$Q = \mathbf{e}_n^T \mathbf{e}_n, \quad (1.13)$$

$$\mathbf{e}_n = (\mathbf{x}_n - \hat{\mathbf{x}}_n).$$

Quando un'osservazione ha un valore anomalamente alto di  $Q$ , essa non viene rappresentata in modo adeguato dal modello; perciò, le variabili non seguono la correlazione descritta dai dati analizzati in fase di calibrazione.

Questi indici permettono di fare un test d'ipotesi per verificare la conformità dell'osservazione con il set di calibrazione ricavato da dati storici che definiscono le condizioni operative normali. Il test d'ipotesi viene fatto mediante carte di controllo. Se nelle carte di controllo almeno uno dei due indici uscirà dal proprio limite di confidenza (*Confidence Limit, CL*) con una predeterminata probabilità, significherebbe che l'osservazione è anomala e non conforme a quelle utilizzate per la calibrazione del modello PCA. Generalmente, viene considerato il valore limite con una confidenza del 95%.

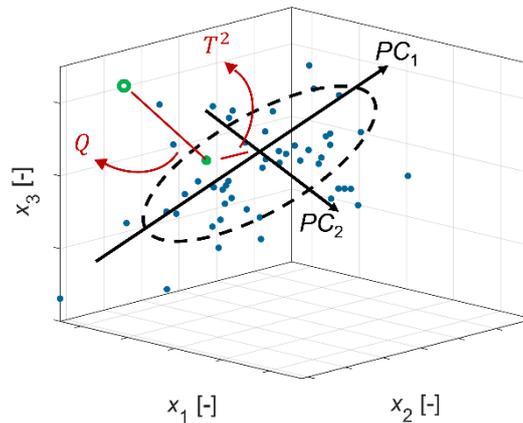
Basandosi sulle distribuzioni statistiche che rappresentano la distribuzione statistica di questi due indici, è possibile determinare il limite di confidenza. Considerando che l'indice  $T^2$  è una somma di valori normalmente distribuiti che viene divisa per una varianza, perciò, seguirà una distribuzione  $F$   $A$  gradi di libertà al numeratore e  $N - A$  al denominatore. Allo stesso modo,  $Q$  seguirà una distribuzione  $\chi^2$ , in quanto somma di residui, che, in linea di principio sono normalmente distribuiti. Quindi, i limiti dei due indici sono (Wise e Gallagher, 1996):

$$T^2_{95\%CL} = \frac{A \cdot (N - 1)}{N - A} F_{A, N-A, 1-0.95}, \quad (1.14)$$

$$Q_{95\%CL} = \theta_1 \left[ \frac{\chi^2_{1-0.95}}{\theta_1} \sqrt{2\theta_2 h_0^2} + 1 + \frac{\theta_2 h_0 (h_0 - 1)}{\theta_1^2} \right]^{1/h_0} \quad (1.15)$$

$$\theta_i = \sum_j^N \lambda_j^i, \quad i = 1, 2, 3 \quad (1.16)$$

$$h_0 = 1 - \frac{2\theta_1 \theta_3}{3\theta_2^2} \quad (1.17)$$



**Figura 1.6.** Rappresentazione geometrica dei due indici per il controllo statistico dei processi: in blu, le osservazioni di calibrazione; la distanza fra la proiezione dell'osservazione (cerchio verde pieno), e il centro del modello PCA rappresenta l'indice  $T^2$ , mentre la distanza normale fra l'osservazione (cerchio verde vuoto) e la sua proiezione nel modello rappresenta l'indice  $Q$ .

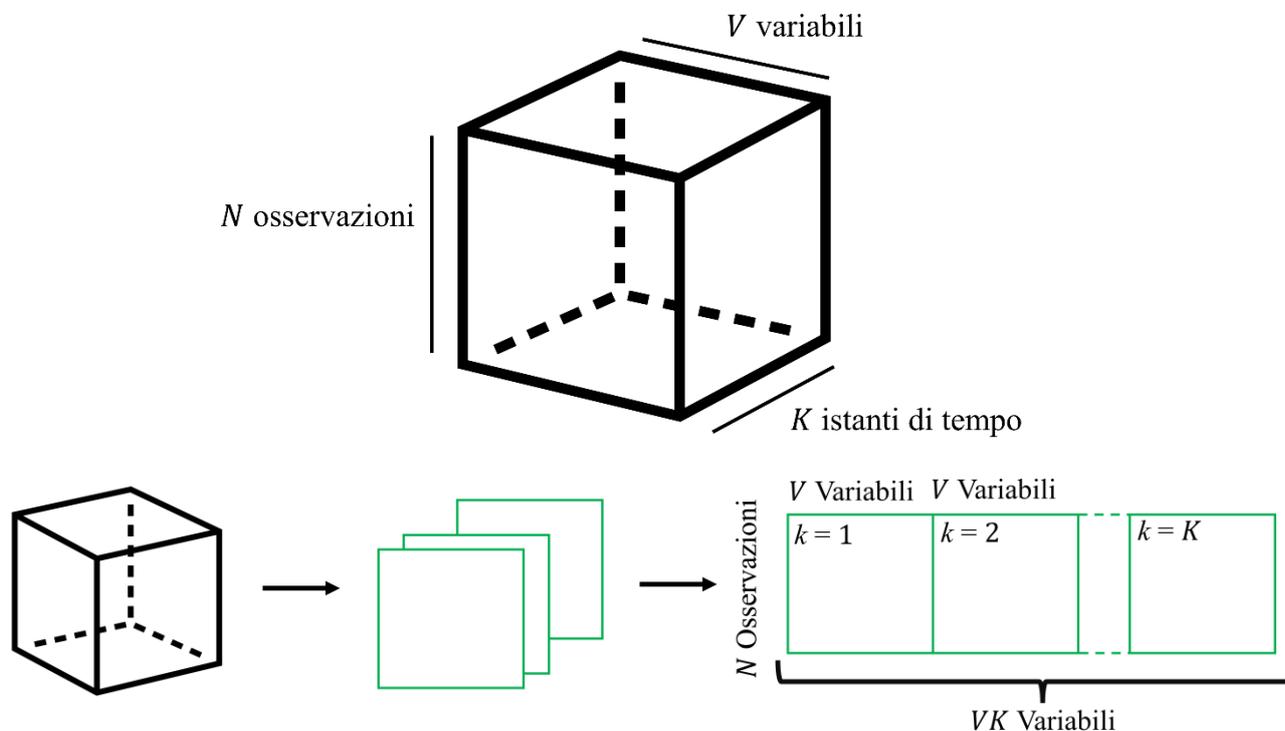
In Figura 1.6 è possibile osservare il limite di confidenza nello spazio delle PC per il  $T^2$  di Hotelling, entro il quale il 95% delle osservazioni è racchiuso (Wise e Gallagher, 1996).

### 1.1.3.5 Organizzazione dei dati

La prima operazione da svolgere in fase di costruzione del modello PCA è l'organizzazione e l'ordinamento dei dati in modo tale da procedere alla decomposizione ai valori singolari. Nel caso di processi *batch* la matrice  $\mathbf{X}$  possiede una terza dimensione, cioè, ad esempio, il tempo, ed ha dimensioni  $[N \times V \times K]$ , dove  $K$  sono gli istanti temporali, come visualizzabile in Figura 1.7.

Perciò, è necessaria una riorganizzazione dei dati mediante *batch-wise unfolding* (Camacho *et al.*, 2008). Per riportare la matrice tridimensionale  $\mathbf{X}$  ad una forma bidimensionale che sia trattabile con PCA, occorre "affettarla" in  $K$  matrici formate di  $N$  osservazioni e  $V$  variabili, ognuna delle quali

rappresenta un istante di tempo. Mettendo queste nuove matrici  $[N \times V]$  una a fianco all'altra, si otterrà una struttura  $[N \times VK]$ , dove  $VK$  è il nuovo numero di colonne. Ognuna delle  $VK$  colonne rappresenta quindi una variabile originale in un determinato istante di tempo.



**Figura 1.7.** Nel caso di profili temporali, la matrice ha tre dimensioni, di cui la terza è il tempo: metodologia di batch-wise unfolding applicata ad una matrice  $[N \times V \times K]$ .

### 1.1.3.6 Metodologie di *Machine Learning*

I modelli PCA possono essere accoppiati ad altre metodologie di *machine learning* supervisionato. In particolare, in questa Tesi sono state utilizzate delle metodologie di classificazione, per cercare di costruire un modello in grado di assegnare ad ogni *score* una classe di appartenenza. Il metodo utilizzato viene chiamato *Linear Discriminant Analysis* (Fisher, 1936) e permette di ottenere una classificazione attraverso la massimizzazione di una funzione obiettivo  $\delta_c [-]$ ; questa funzione descrivere la probabilità di un vettore di appartenere ad una determinata classe.

La funzione discriminante viene calcolata in questo modo:

$$\delta_c = \text{funzione discriminante} = x^T \Sigma^{-1} \mu_m - 1/2 \Sigma^{-1} \mu_m^T \mu_m + \ln \pi_m \quad (1.18)$$

$$\Sigma = \frac{\sum \sum (x_i - \mu_c)(x_i - \mu_c)^T}{N_t - c} \quad (1.19)$$

$\Sigma =$  matrice di covarianza fra le osservazioni di una classe

$$\mu_m = \sum \frac{x_i}{N_m} = \text{media dei valori delle osservazioni della classe } m \quad (1.20)$$

$$\pi_m = N_m/N_t$$

$$N_m = \text{numero di osservazioni appartenenti alla classe } m \quad (1.21)$$

$$N_t = \text{numero totale di osservazioni}$$

Una volta definite quindi tutte le classi di osservazioni, il metodo calcola la probabilità che ogni nuova osservazione appartenga ad una determinata classe, andando poi ad assegnare ogni osservazione alla categoria la cui funzione discriminante è massima.

### 1.1.3.7 Analisi attraverso similarities

Una metodologia che permette il confronto tra due diversi modelli PCA utilizza dei parametri definiti come parametri di similarità (*similarity factors*,  $S_{i,k} [-]$ ).

Questo parametro viene definito in questo modo (Ottavian *et al.*, 2012):

$$S_{i,j} [-] = \frac{\text{traccia} \left[ (\mathbf{P}_i^W)^T (\mathbf{P}_j^W) (\mathbf{P}_j^W)^T (\mathbf{P}_i^W) \right]}{\sum_a \lambda_{ia} \lambda_{ja}} \quad (1.22)$$

$$\mathbf{p}_{ia}^W [-] = \text{Weighted Loadings} = \mathbf{p}_{ia} \lambda_{ia} \quad (1.23)$$

$$\lambda_{ia} = \text{autovalore della } PC_a \text{ del modello costruito su } i$$

Essi sono dei parametri che quantificano quanto due modelli PCA si assomigliano, in altre parole, sono degli indici il cui valore può andare da zero (nessuna somiglianza) a uno (modelli uguali) valutando quanto i *loadings*, ovvero le direzioni di massima variabilità dei modelli costruiti sono fra loro simili.

La valutazione viene pesata attraverso l'autovalore corrispondente ad ogni PC, in modo tale da dare più importanza alle direzioni maggiormente significative.

Chiaramente, per poter utilizzare gli indici di similarità è necessario costruire un modello diverso per ogni classe di osservazioni. Non è quindi possibile utilizzare il metodo per confrontare due diverse tipologie di cui si dispone solamente una osservazione ciascuna.

### **1.1.4 Proiezione su strutture latenti**

La tecnica di analisi proiezione su strutture latenti (*Projection to Latent Structure, PLS*), detta anche regressione ai minimi quadrati parziali (*Partial Least Squares Regression*), è una metodologia di

riduzione della dimensionalità molto simile a PCA. In particolare, PLS è una tecnica di regressione, quindi di machine learning supervisionato, in cui vengono studiate le correlazioni di una serie di variabili indipendenti, detti anche fattori, con le variabili dipendenti, dette anche risposte. Questa tecnica prevede la costruzione di un modello di regressione lineare fra effetti e risposte grazie al quale è possibile poi predire come una data combinazione di fattori determini i parametri di qualità.

#### 1.1.4.1 Criticità dei comuni metodi di regressione:

Considerando i fattori  $\mathbf{X}$  e le risposte  $\mathbf{Y}$ :

$$\mathbf{X} = \text{Fattori} = \text{Variabili indipendenti} = [N \times V], \quad (1.24)$$

$$\mathbf{Y} = \text{Risposte} = \text{Variabili dipendenti} = [N \times M], \quad (1.25)$$

è possibile legarli attraverso diversi modelli di regressione. Il comune metodo di regressione multipla lineare (equazione 1.25) prevede che i parametri del modello (cioè i coefficienti di regressione  $\mathbf{B}$ ) vengano stimati in modo algebrico dall'inversione della matrice di covarianza di  $\mathbf{X}$  (equazione 1.26):

$$\mathbf{Y} = \mathbf{B}\mathbf{X} + \mathbf{E}, \quad (1.25)$$

$$\mathbf{B}_r = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{Y}. \quad (1.26)$$

Alcune criticità emergono però all'applicazione di questi modelli: potrebbero non esserci soluzioni (nel caso il numero delle osservazioni sia minore del numero di variabili), le soluzioni potrebbero essere infinite, nel caso il numero di osservazioni sia maggiore del numero di variabili), oppure la matrice  $\mathbf{X}^T\mathbf{X}$  potrebbe non essere invertibile. Quindi, un modello di regressione multipla lineare non può essere stimato nel caso in cui le variabili indipendenti siano correlate tra loro.

Il modello PLS supera questo problema e permette di stimare i coefficienti di regressione nel caso di variabili indipendenti correlate.

Questo viene fatto riducendo la dimensionalità del problema mediante un sistema di variabili latenti, come vedremo di seguito.

#### 1.1.4.2 Formulazione matematica del modello PLS

Analogamente a PCA, PLS consente di estrarre linee, piani, spazi o iperspazi nei quali la massima variabilità dei dati è catturata, sia per quanto riguarda le risposte, sia per i fattori. Inoltre, allo stesso tempo consente di legare le due tipologie di variabili fra loro. In particolare, PLS trova le direzioni di massima variabilità delle  $\mathbf{X}$  che sono ruotate in modo da essere più predittive per la  $\mathbf{Y}$ . Queste direzioni si chiamano variabili latenti (*Latent Variables*, LV).

In PLS le matrici  $\mathbf{X}$  e  $\mathbf{Y}$  vengono scomposte in questo modo (De Jong, 1993):

$$\mathbf{X} = \mathbf{T}\mathbf{P}^T + \mathbf{E} = \sum_{a=1}^A \mathbf{t}_a \mathbf{p}_a^T + \mathbf{E}, \quad (1.27)$$

$\mathbf{T}$  = Scores di  $\mathbf{X}$ ,       $\mathbf{P}$  = Loadings di  $\mathbf{X}$ ,       $\mathbf{E}$  = Residui di  $\mathbf{X}$ ,

$$\mathbf{Y} = \mathbf{U}\mathbf{Q}^T + \mathbf{F} = \sum_{a=1}^A \mathbf{u}_a \mathbf{q}_a^T + \mathbf{F}, \quad (1.28)$$

$\mathbf{U}$  = Scores di  $\mathbf{Y}$ ,       $\mathbf{Q}$  = Loadings di  $\mathbf{Y}$ ,       $\mathbf{F}$  = Residui di  $\mathbf{Y}$ ,

$$\mathbf{u}_a = b_a \mathbf{t}_a, \quad (1.29)$$

$b_a$  = coefficienti di regressione interni del modello PLS .

In PLS sono inoltre introdotti i pesi  $\mathbf{W}$  (*Weights*), che similmente ai *Loadings*, rappresentano le direzioni delle LV.

SIMPLS è un algoritmo introdotto da De Jong (1993), utilizzato per la costruzione dei modelli PLS. È il più veloce e robusto e consiste in una serie di operazioni in cui, innanzi tutto, si stima la matrice di covarianza tra  $\mathbf{X}$  e  $\mathbf{Y}$ :

$$\mathbf{S} = \mathbf{X}^T \mathbf{Y}. \quad (1.30)$$

Una volta definita la matrice di covarianza, si utilizza un algoritmo iterativo costruito sui seguenti passaggi:

$$a = \text{generica latent variable}, \quad (1.31)$$

$$\mathbf{w}_a = \text{weights di } LV_a = (\mathbf{S}^a) \cdot (\text{autovettore dominante di } \mathbf{S}^{aT} \mathbf{S}^a),$$

$$\mathbf{t}_a = \text{Score di } \mathbf{X} \text{ su } LV_a = \mathbf{X} \cdot \mathbf{w}_a, \quad (1.32)$$

$$\mathbf{t}_a = \frac{(\mathbf{t}_a - \text{media}(\mathbf{t}_a))}{\|\mathbf{t}_a\|}, \quad \text{normalizzazione di } \mathbf{t}_a, \quad (1.33)$$

$$\mathbf{p}_a = \text{Loading di } X \text{ su } LV_a \text{ di } \mathbf{X} = \mathbf{X}^T \cdot \mathbf{t}_a, \quad (1.34)$$

$$\mathbf{q}_a = \text{Loading di } Y \text{ su } LV_a = \mathbf{Y}^T \cdot \mathbf{t}_a, \quad (1.35)$$

$$\mathbf{u}_a = \text{Score di } \mathbf{Y} \text{ su } LV_a = \mathbf{Y} \cdot \mathbf{q}_a, \quad (1.36)$$

$$\mathbf{v}_a = \mathbf{p}_a - \mathbf{V}(\mathbf{V}^T \mathbf{p}_a), \quad \text{ortogonalizzazione del Loading di } \mathbf{X}, \quad (1.37)$$

$$\mathbf{v}_a = \frac{(\mathbf{v}_a - \text{media}(\mathbf{v}_a))}{\|\mathbf{v}_a\|}, \quad \text{normalizzazione di } \mathbf{v}_a, \quad (1.38)$$

$$\mathbf{u}_a = \mathbf{u}_a - \mathbf{T}(\mathbf{T}^T \mathbf{u}_a), \quad \text{ortogonalizzazione fra gli score di } \mathbf{X} \text{ ed } \mathbf{Y}, \quad (1.39)$$

$$\mathbf{S}^{a+1} = \mathbf{S}^a - \mathbf{v}_a(\mathbf{v}_a^T \mathbf{S}^a), \quad \text{deflazione della matrice di covarianza.} \quad (1.40)$$

La necessità di utilizzare i loading ortogonali fra loro ( $\mathbf{v}_a$ ) in fase di deflazione della matrice di covarianza sta nel fatto che le informazioni della suddetta, rimosse una volta proiettate, devono essere indipendenti fra loro. Non è però più necessario che i *loadings* siano ortogonali fra loro, in quanto non rappresentano più le direzioni del nuovo spazio latente, che sono invece rappresentati dai *weights*.

È poi possibile calcolare i coefficienti di regressione complessivi del modello PLS, che legano  $\mathbf{X}$  e  $\mathbf{Y}$ :

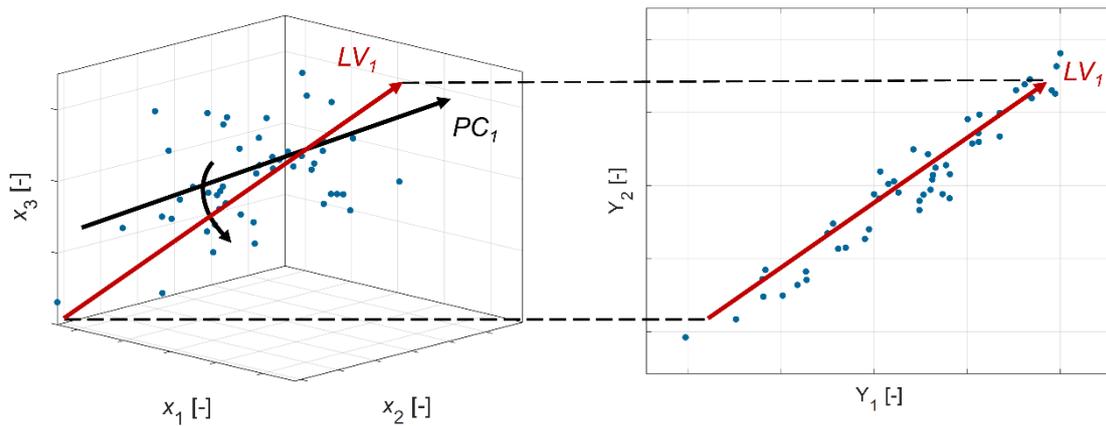
$$\mathbf{B} = \mathbf{WQ}^T \quad (1.41)$$

$$\mathbf{B} = \text{coefficienti di regressione complessivi del modello PLS}$$

$$\hat{\mathbf{Y}} = \mathbf{XB} \quad (1.42)$$

### 1.1.4.3 Interpretazione geometrica di PLS

Il modello PLS è geometricamente rappresentabile in modo simile al modello PCA, in questo caso però, le variabili latenti vengono ruotate dalla direzione di massima variabilità che permettono la migliore predizione di  $\mathbf{Y}$ , come è possibile osservare in Figura 1.8.



**Figura 1.8.** Rappresentazione geometrica del modello PLS. È possibile notare come la rotazione della PC nello spazio dei fattori vada a rappresentare al meglio le risposte.

Il numero di LV viene scelto in modo analogo a quanto presentato per la PCA: si possono utilizzare lo *Scree Test*, con il quale si analizzano sia la variabilità catturata di  $\mathbf{Y}$ , o la *Cross Validation*, se lo scopo è di stimare la  $\mathbf{Y}$ . Nei problemi di inversione (che verranno presentati nel Sottoparagrafo successivo) si dà invece peso anche alla rappresentazione della  $\mathbf{X}$ .

Per quanto riguarda la diagnosi statistica, ancora una volta sono utilizzabili concetti simili a quelli introdotti per PCA e visti al Sottoparagrafo 1.1.3.4.

#### 1.1.4.4 Inversione del modello PLS

Nella sua forma diretta il modello PLS permette la predizione di valori di qualità  $\mathbf{Y}$  a partire da osservazioni di determinati fattori  $\mathbf{X}$ . È però possibile invertire il modello (Tomba *et al.*, 2016) al fine di capire come a partire da una  $\mathbf{Y}$  desiderata si calcoli la combinazione dei fattori che garantiscano la  $\mathbf{Y}$  desiderata.

Siccome nell'inversione di PLS si passa allo spazio delle  $\mathbf{X}$  da uno spazio  $\mathbf{Y}$ , che usualmente è di dimensione molto inferiore a quello delle  $\mathbf{X}$  stesse, abitualmente non esiste una sola combinazione dei fattori che garantisce la  $\mathbf{Y}$  desiderata. Bensì esiste uno spazio nullo che raccoglie le infinite combinazioni di fattori che corrispondono alla  $\mathbf{Y}$  desiderata. Al fine di trovare la migliore combinazione di fattori che assicuri la  $\mathbf{Y}$  desiderata si risolve un problema di ottimizzazione nel quale è possibile inserire vincoli sia di tipo modellistico che di tipo operativo sulle variabili in gioco.

La soluzione del problema nello spazio degli score è:

$$\min_{\mathbf{t}_{calc}} \|(\mathbf{t}_{des} - \mathbf{t}_{calc}) \cdot \mathbf{\Lambda}^{0.5}\|, \quad (1.43)$$

in cui:

$$\mathbf{t}_{des} = (\mathbf{Q}^T \mathbf{Q})^{-1} \mathbf{Q}^T \mathbf{Y}_{des}^T, \quad (1.44)$$

e:

$$\mathbf{t}_{calc} = \mathbf{X}_i \cdot \mathbf{W}. \tag{1.45}$$

L'equazione 1.43 descrive la funzione obiettivo, e non è altro che la radice della somma dei quadrati della differenza fra lo *Score* desiderato e lo *Score* calcolato con il valore di  $\mathbf{x}$  tentativo, pesata secondo la varianza spiegata da ogni variabile latente ( $\Lambda$ ), in modo da favorire la minimizzazione degli *score* sulle *LV* più importanti.

Infine, viene inserito un vincolo non lineare che assicura l'appartenenza dello *Score* calcolato al limite del 95% di confidenza, in modo da forzare il modello di inversione a predire valori all'interno del dominio con il quale il modello è stato costruito, evitando estrapolazioni:

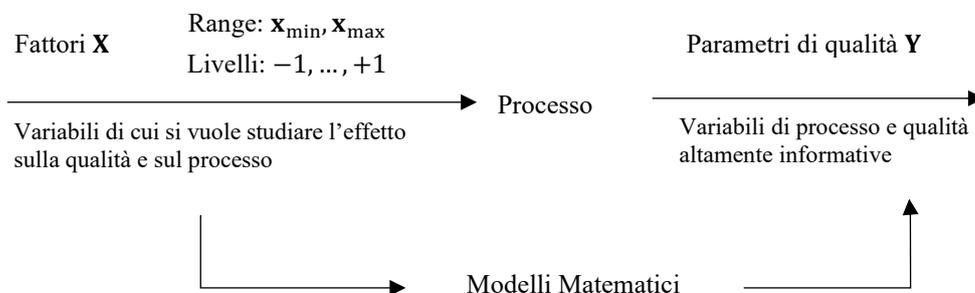
$$T_{score\ calcolato}^2 < T_{95\%CL}^2. \tag{1.46}$$

L'efficienza predittiva dell'inversione del modello *PLS* viene valutata attraverso una metodologia di *Leave One Out (LOO)*.

## 1.2 Pianificazione degli esperimenti

Le tecniche di pianificazione degli esperimenti (*Design of Experiments, DoE*, Ryan, 2007) sono delle metodologie utili ad organizzare, effettuare e analizzare prove sperimentali nella maniera più efficiente ed informativa possibile. Viene definito esperimento una serie di prove nelle quali vengono misurate le variabili di qualità  $\mathbf{Y}$  variando ogni volta in modo sistematico ognuno dei fattori della matrice  $\mathbf{X}$ .

Una procedura sistematica di organizzazione degli esperimenti è il cosiddetto *Full-Factorial Design* (Ryan, 2007). Nel *Full-Factorial Design*, i  $G$  fattori vengono fatti variare in modo discreto su  $L$  livelli predeterminati. Tipicamente i fattori vengono normalizzati e le variabili codificate in modo da assumere un valore  $+1$  al massimo del proprio campo di estensione, mentre il livello minimo assume valore  $-1$ . Il numero totale di combinazioni possibili dei fattori e quindi degli esperimenti da fare nel *Full-Factorial Design* è  $L^G$ . In Figura 1.9, la pianificazione degli esperimenti è schematicamente rappresentata.



**Figura 1.9.** Rappresentazione schematica della raccolta di dati attraverso pianificazione degli esperimenti.

La procedura per eseguire un *DoE* è la seguente:

1. inquadramento del problema e dell'obiettivo della campagna sperimentale;
2. selezione delle risposte da misurare;
3. scelta dei fattori, dei livelli e dei range su cui essi variano;
4. scelta della tipologia di campagna sperimentale;
5. esecuzione degli esperimenti;
6. analisi dei dati;
7. conclusioni.

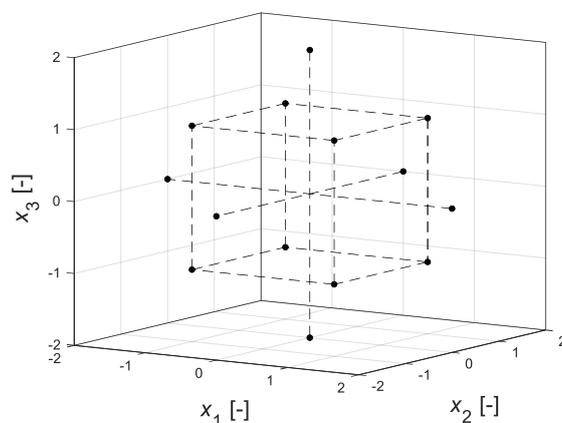
Esistono diverse tecniche di pianificazione di esperimenti, che possono essere ricondotte a quattro tipologie: oltre al *Full Factorial Desing*, sono comunemente utilizzati anche *Fractional Factorial Design*, *Central Composite Desing*, *Irregular Desings* (Ryan, 2007). Questi si differenziano nella scelta delle combinazioni di fattori su cui devono essere fatti gli esperimenti.

In questo Capitolo, andremo a descrivere le tipologie di *DoE* che sono state utilizzate in questa Tesi.

### 1.2.1 Central Composite Design (CCD)

Secondo questa tecnica di pianificazione degli esperimenti, i fattori possono variare il loro valore su due livelli; tutte le combinazioni sono quindi studiate. In aggiunta, vengono studiati anche i punti centrali, ovvero delle combinazioni in cui  $K - 1$  fattori sono tenuti costanti e uno è invece mantenuto in un livello centrale del range. I punti centrali permettono di introdurre nel modello eventuali non linearità della risposta al variare del fattore all'interno del range di misura.

Il vantaggio dell'utilizzo di questa tecnica è la capacità di studiare tutti gli effetti combinati dei fattori, ma anche avere informazioni sull'effetto dei punti centrali, mantenendo però un numero di esperimenti relativamente limitato. Esistono diverse tipologie di CCD, in base a dove vengono presi i punti centrali questo studio è stato utilizzato un Circumnscribed Central Composite Design (Xu e Ting, 2004): come rappresentato in Figura 1.10.



**Figura 1.10.** Rappresentazione grafica della pianificazione degli esperimenti Circumnscribed Central Composite Design.

### 1.2.2 Algoritmo di Kennard-Stone

Nel caso in cui il numero di fattori sia invece molto elevato, non è possibile effettuare una campagna sperimentale completa, in quanto richiederebbe troppo tempo e investimenti.

Per questo motivo, si utilizzano delle strategie di *Fractional Factorial Design* (FFD), che consistono nella scelta di un numero limitato di combinazioni tra tutte quelle possibili di un *Full Factorial Design*.

In questo caso, si è deciso di utilizzare l'algoritmo di *Kennard-Stone* (Gomes *et al.*, 2018); esso, appartiene alla categoria dei design a riempimento spaziale: queste tipologie di design sono particolarmente utilizzate nei problemi di formulazione, quando il numero di fattori e livelli sono molto elevati. Prendendo in considerazione lo spazio formato da tutte le e combinazioni possibili che possono essere studiate, considerando anche alcuni vincoli fisici o imposti dal formulatore, le combinazioni migliori vengono scelte in modo iterativo, ottenendo una serie di punti nello spazio che descrivano e riempiano lo spazio di design nel modo più omogeneo possibile.

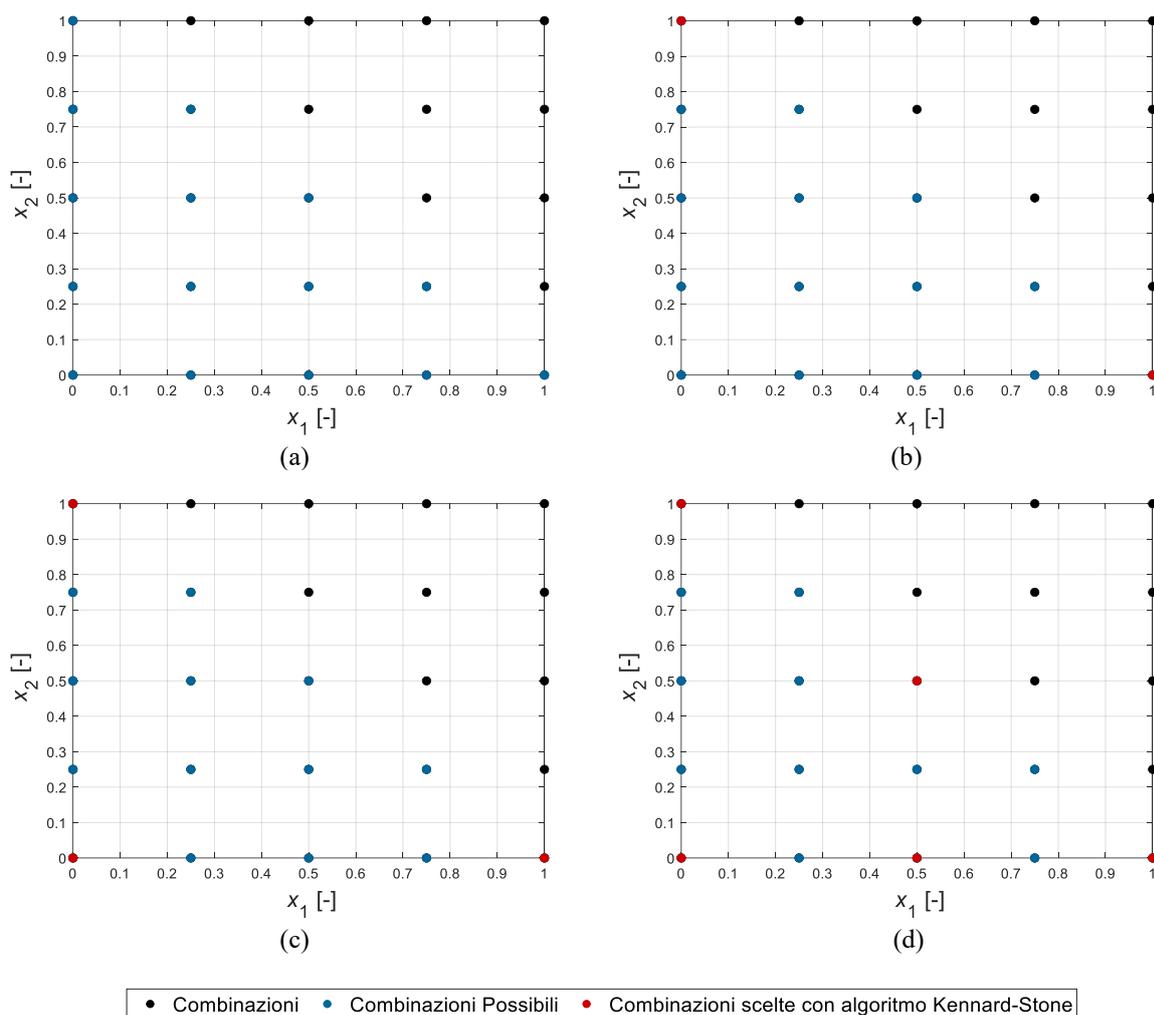
Il numero (sovrabbondante) di possibili combinazioni dello spazio viene definito  $N_c$ , mentre  $N_e$  è il numero di esperimenti da eseguire, scelto arbitrariamente. L'algoritmo proposto da *Kennard-Stone* (Gomes *et al.*, 2018) esegue la procedura descritta nel seguito:

- si calcola la distanza tra tutti gli  $N_c$  punti scelti:

$$d_{i,j} = \|(\mathbf{x}_i - \mathbf{x}_j)\| = \text{distanza fra le combinazioni } \mathbf{x}_i \text{ e } \mathbf{x}_j; \quad (1.47)$$

- si scelgono  $n_c [-] = 2$  punti,  $\mathbf{x}_1$  ed  $\mathbf{x}_2$ , in modo tale che la loro distanza sia la massima tra tutte quelle calcolate;
- viene definito come  $\Delta_i(N) [-]$  il minimo valore di distanza tra il punto  $x_i$  e ognuno fra gli  $n_c$  scelti.
- il punto  $n_c + 1$ , che verrà aggiunto all'insieme di combinazioni scelte, sarà il punto con il valore massimo di  $\Delta_i(N)$ . In questo modo, il nuovo punto  $n + 1$  sarà il punto più distante possibile dall'insieme degli  $n_c$  già scelti.
- La procedura viene ripetuta fino a che  $n_c$  non sarà uguale a  $N_e$ , aggiungendo via via combinazioni sempre più distanti fra loro nello spazio.

L'algoritmo è rappresentato schematicamente in Figura 1.11.



**Figura 1. 11.** *Algoritmo Kennard Stone applicato a due fattori che possono variare su cinque livelli; (a) divisione del dominio in osservazioni possibili e impossibili; (b) Primo step,  $n = 2$ ; (c) secondo step,  $n = 3$ ; (d) ultimo step,  $n = N_e = 5$ ;*

Questo algoritmo permette di esportare i punti dello spazio fra loro più distanti, in modo da ottenere informazioni in tutto il dominio e lo studio degli effetti di tutti i fattori con un limitato numero di esperimenti.

L'algoritmo di *Kennard Stone* per la pianificazione degli esperimenti soffre però di una problematica tipica dei *FFD*. In particolare, lo studio di parte del dominio genera un fenomeno che viene definito confondimento (*aliasing*): gli effetti di fattori diversi vengono confusi ed interpretati allo stesso dai modelli matematici costruiti sui dati raccolti dalla campagna sperimentale, a causa della mancanza di esplorazione di tutte le combinazioni dei fattori ai livelli scelti. Questo problema non può essere risolto a priori, il modello potrebbe essere migliorato solamente aggiungendo nuovi dati e nuove campagne sperimentali all'insieme di esperimenti eseguiti.



# Capitolo 2

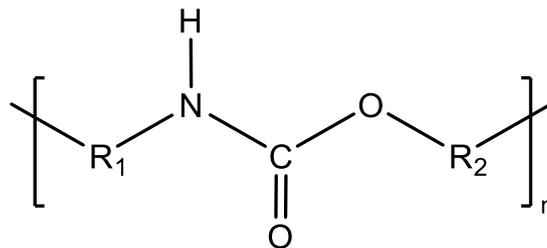
## Tecnologie dei sistemi poliuretanicici rigidi

I polimeri sono materiali utilizzati in tutto il mondo per la loro versatilità e adattabilità in ogni tipologia di applicazione. La caratteristica che li contraddistingue è la particolare struttura chimica formata da lunghe catene di atomi (principalmente di carbonio). Tra questi, spiccano i prodotti poliuretanicici, la cui cinetica che ne regola e descrive la trasformazione è oggetto di questo studio. In questo capitolo verranno dunque illustrate le principali caratteristiche del materiale considerato. Più in specifico, i materiali studiati sono schiume poliuretanicche rigide.

### 2.1 Schiume poliuretanicche rigide

Il poliuretano è un materiale polimerico utilizzato principalmente in forma termoindurente; ciò significa che a differenza della maggior parte dei comuni polimeri è infusibile e insolubile. Queste proprietà derivano dalla particolare struttura che viene definita reticolata, ovvero, nella quale le lunghe catene molecolari tipiche dei polimeri sono connesse fra loro attraverso legami chimici forti. Queste connessioni generano una fitta rete, che rende il materiale impossibile da fondere e solubilizzare in quanto l'unico modo di rompere le connessioni fra le maglie della rete è la degradazione dei legami primari, cosa che però porterebbe al collasso dell'intera struttura.

Il legame che caratterizza questo polimero è il legame uretanico: l'unità ripetitiva del poliuretano è rappresentata in Figura 2.1:



**Figura 2.1.** Unità ripetitiva del poliuretano. Il legame uretanico è dato da un particolare gruppo funzionale che caratterizza i poliuretani.

#### 2.1.1 Storia, mercato, tipologie ed utilizzi dei poliuretani

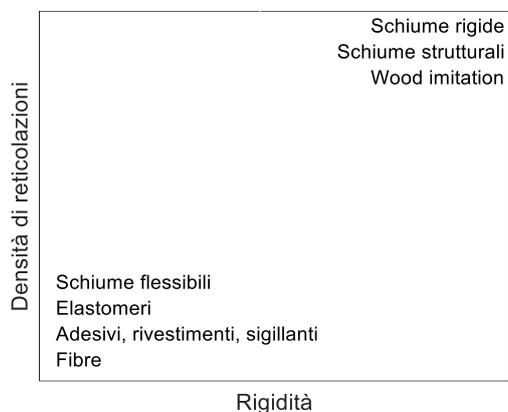
Il primo composto contenente un legame uretanico fu sintetizzato nel 1849 dal chimico Wurtz; solo nel 1937 in Germania, fu sintetizzato il primo poliuretano presso l'azienda tedesca IG Farben, attraverso la reazione di un diisocianato e un poliesteri. Fu in questa occasione che venne coniato il termine poliaddizione, come categoria di reazione di polimerizzazione, che si contraddistingue per la produzione del polimero attraverso una cinetica non radicalica, e senza la generazione di molecole di condensazione (Ionescu, 2016). I dettagli in merito alle reazioni chimiche verranno discussi nel

sottoparagrafo 2.1.2. Il poliuretano in forma di schiuma rigida (a densità medio bassa) viene utilizzato nell'isolamento termico di edifici, sistemi refrigeranti, isolamento di tubi e trasporto di materiale a temperature ridotte.

Circa un terzo della produzione di poliuretani viene impiegato nella realizzazione di materassi e nel settore automobilistico; in questi casi il materiale è utilizzato in forma di schiuma flessibile.

Applicazioni minori spaziano dall'utilizzo di poliuretano elastomerico nell'industria calzaturiera o nella produzione di pneumatici alla produzione di adesivi poliuretanici, guarnizioni e fibre al riempimento di materiali galleggianti come boe o tavole da surf (Ionescu, 2016).

Tutte queste forme e applicazioni possono essere divise in due tipologie principali: sistemi elastici e sistemi rigidi. In Figura 2.2 viene mostrato come le due categorie si differenziano per densità di reticolazioni e rigidità.



**Figura 2.2.** Tipologie di utilizzi dei poliuretani in base a densità di reticolazioni e rigidità. (adattata da Ionescu, 2016).

Questa Tesi si concentra sulla produzione delle schiume poliuretatiche rigide, sulla loro formulazione<sup>1</sup> e la loro cinetica in fase di trasformazione.

Queste schiume vengono ampiamente utilizzate nel campo dell'isolamento termico per l'elevata capacità isolante per unità di spessore. Inoltre, la possibilità di produrre la schiuma direttamente *in loco* rende il prodotto versatile ed efficiente in numerosi utilizzi e permette il trasporto del prodotto in forma liquida anziché solida, riducendo volumi e quindi costi di trasporto. Altre proprietà peculiari di questa schiuma sono l'eccellente adesione ai materiali da costruzione, come metallo, ceramica, vetro e legno, l'ottima resistenza al calore e il compromesso fra proprietà meccaniche e peso ridotto.

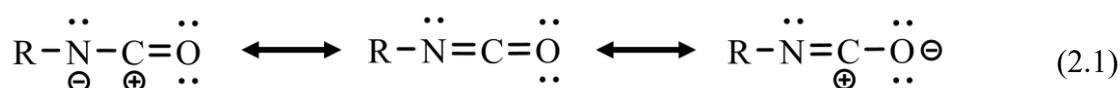
### 2.1.2 Chimica delle schiume poliuretatiche rigide

La sintesi di una schiuma poliuretatica, sia essa rigida o flessibile, avviene attraverso due composti principali: poliolo e poliisocianato. La reazione avviene tra il gruppo funzionale isocianato e quello

<sup>1</sup> Quando si parla di formulazione si indica il dosaggio delle sostanze chimiche utilizzate per la produzione di una schiuma poliuretatica

idrossile, tipico dei polioli. La funzionalità dei due composti, ovvero il numero di gruppi funzionali che ogni molecola possiede, deve essere maggiore o uguale a due, in modo tale da permettere la costruzione di una catena attraverso una reazione di poliaddizione.

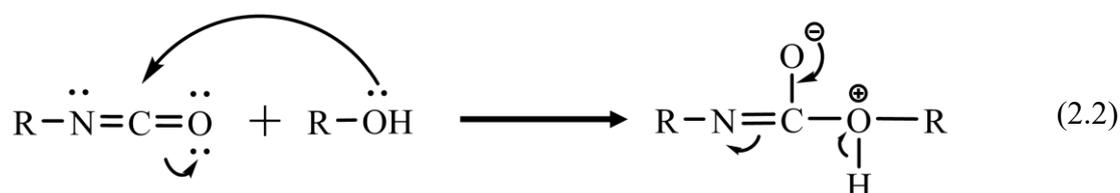
Il gruppo funzionale isocianato è costituito da un atomo di azoto, uno di carbonio e uno di ossigeno, legati fra loro in questo esatto ordine attraverso doppi legami. Possiede una particolare forma di risonanza (i legami chimici del gruppo funzionale esistono in modo instabile e in continua oscillazione di forma e combinazioni) che lo rende particolarmente reattivo verso le molecole contenenti un idrogeno attivo o in altre parole facilmente cedibile:



Come rappresentato in equazione 2.1, la densità elettronica del gruppo funzionale è spostata verso l'atomo di ossigeno, più elettronegativo, questo genera la presenza di una carica positiva sull'atomo di carbonio e una intermedia negativa sull'azoto. La reazione si sviluppa, dunque, quando il carbonio dell'isocianato con carattere elettrofilo viene attaccato dall'ossigeno presente nel gruppo ossidrilico del poliolo, in quanto si comporta da nucleofilo. I gruppi elettron-attrattori presenti in *R*, ovvero quei gruppi funzionali organici la cui caratteristica è di attrarre a sé la nuvola elettronica della molecola, rendono più reattiva la molecola di isocianato. Per questo motivo, gli isocianati aromatici tendono ad avere una reattività più elevata di quelli alifatici. Un eventuale ingombro sterico, dovuto alla presenza di gruppi ad elevato peso molecolare in *R*, per quanto riguarda entrambi i monomeri, va invece a ridurre la reattività (Ionescu, 2016).

### 2.1.2.1 Reazione di gel

La reazione di polimerizzazione viene descritta in equazione 2.2, in cui viene rappresentato il meccanismo di reazione (Hart *et al.*, 1985) fra un gruppo isocianato ed un idrossile.



La reazione di polimerizzazione viene industrialmente chiamata *reazione di gel*, ed è caratterizzata da una notevole esotermia ( $\Delta H_r^0 = -100.4 \text{ kJ/eq}_{\text{NCO}}$ , G.Woods<sup>9</sup>), con formazione di numerosi legami chimici.

In questo particolare tipo di poliaddizione, le specie reattive sono i monomeri, i dimeri, gli oligomeri e i polimeri stessi. Questo permette di ottenere un elevato peso molecolare solamente a conversioni molto elevate di poliolo ed isocianato. La reazione complessiva è riportata in equazione 2.3.



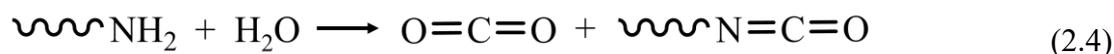
Come in tutte le reazioni di poliaddizione, il rapporto fra le quantità di poliolo ed isocianato ha un impatto notevole sul peso molecolare del composto finale, che poi influenzerà le proprietà della schiuma; perciò, deve essere controllato e sarà un fattore chiave in fase di formulazione (Ionescu, 2016). Come vedremo, per ottenere un polimero molto reticolato, la formazione del legame uretanico avverrà anche tra gruppi funzionali già all'interno delle molecole polimeriche e gruppi funzionali liberi, in modo tale da ottenere la connessione fra le diverse catene.

### 2.1.2.2 Reazione di *blow*

Oltre alla reazione di polimerizzazione è necessaria una ulteriore reazione per permettere al materiale di generare la sua caratteristica schiuma. Questa reazione viene detta *reazione di blow* e consiste nella formazione di una quantità di gas, contemporaneamente alla polimerizzazione, responsabile del considerevole aumento di volume del polimero nella fase di espansione.

La reazione avviene fra l'isocianato e un agente espandente chimico. La necessità di specificare la natura chimica di questo tipo di espansione nasce dall'esistenza di due modalità che permettono l'ottenimento della schiuma poliuretanic: essa può essere infatti generata da una espansione di tipo chimico o di tipo fisico. Come suggerisce la denominazione, l'espansione di tipo fisico avviene invece attraverso l'evaporazione di un espandente volatile a seguito della generazione di calore della reazione di gel.

Uno degli espandenti chimici più utilizzati è l'acqua, che porta alla produzione di anidride carbonica a seguito della reazione con l'isocianato, come descritto dall'equazione 2.4.

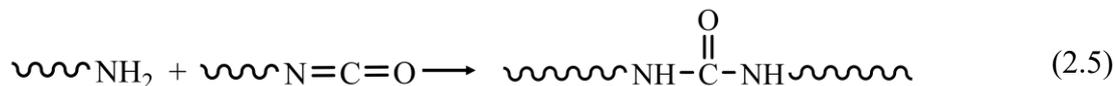


La reazione 2.4 comprende anche un intermedio instabile; infatti, prima di formare l'anidride carbonica e l'ammina primaria, si forma l'acido carbammico, che però si decompone subito nei prodotti finali, per questo motivo la sua presenza e la reazione intermedia non verranno considerati.

Un secondo espandente chimico che viene spesso utilizzato è l'acido formico. La reazione è molto simile alla 2.4, tuttavia, viene prodotto anche monossido di carbonio (Ionescu, 2016).

L'espansione chimica è una reazione esotermica ( $\Delta H_r^0 = -94.14 \text{ kJ/eq}_{\text{NCO}}$ , Woods, 1987), questa è una delle altre importanti differenze rispetto alla espansione fisica sopracitata, in quanto quest'ultima, essendo una evaporazione, assorbe calore.

L'ammina reagisce successivamente in modo molto rapido con l'isocianato residuo o terminale in una catena polimerica, viene quindi prodotta urea bisostituita (Hart *et al.*, 1985), come descritto in equazione 2.5.



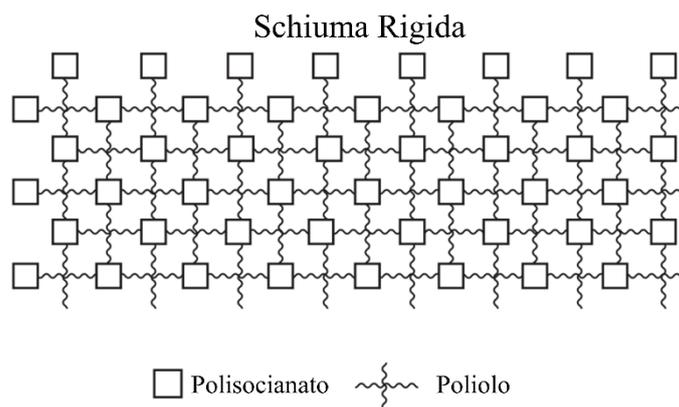
### 2.1.2.3 Reazioni di reticolazione e trimerizzazione

La struttura finale del prodotto è termoindurente. Questa viene generata attraverso numerose reticolazioni che legano fra loro le catene polimeriche formatesi durante la reazione di gel. Le reticolazioni possono essere generate attraverso diverse reazioni chimiche, ognuna però ha come reagente principale l'isocianato.

La reazione è possibile solamente se si utilizza un monomero polisostituito, in quanto sono necessari dei gruppi isocianati non terminali che saranno liberi all'interno della catena, e capaci di produrre la reticolazione.

Altre reazioni di reticolazione, consistono nella formazione di gruppi detti allofanati e biuretici. I primi, si formano quando il gruppo isocianato di un monomero ancora libero, reagisce con il gruppo uretanico appena formatosi all'interno di una catena polimerica già prodotta, mentre il secondo avviene fra una molecola di urea bisostituita prodotta durante il *blow* e un gruppo isocianato terminale (Iounescu, 2016).

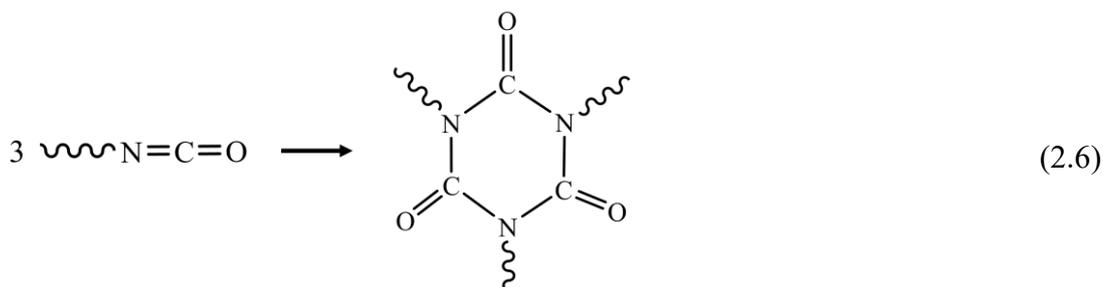
In Figura 2.3 è rappresentata l'elevata reticolazione di una schiuma poliuretanic rigida:



**Figura 2.3.** Struttura reticolata in una schiuma poliuretanic. (Adattata da M. Szycher <sup>5</sup>)

Infine, un'ultima, ma non meno importante, forma di reticolazione, soprattutto per quanto riguarda le schiume rigide, è la formazione di una complessa struttura ciclica: la reazione di *trimerizzazione*. Essa porta alla formazione di gruppi ciclici all'interno della struttura polimerica, questi vengono chiamati isocianurati.

Questi particolari gruppi sono la causa dell'elevato indice di reticolazione in una schiuma poliuretanicica rigida e si formano secondo la reazione 2.6.



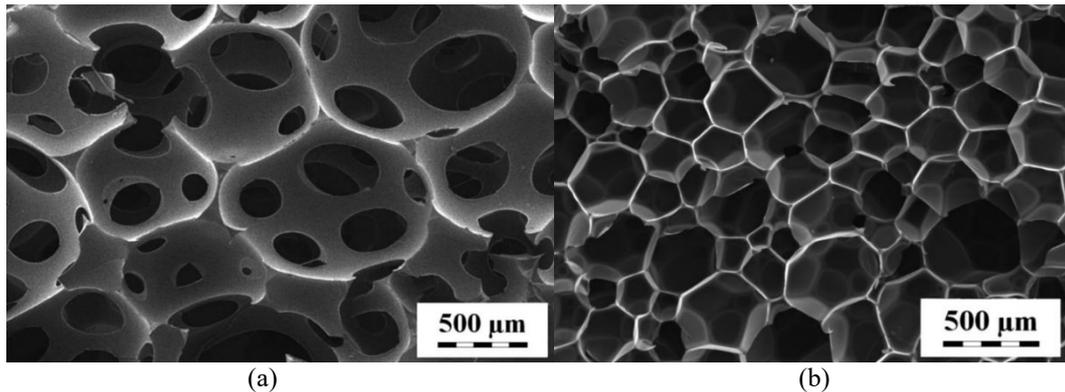
La reazione avviene quando è presente un eccesso di isocianato. Queste strutture cicliche possiedono una elevata resistenza al fuoco e rendono il materiale finale molto più rigido (Lee e Ramesh, 2004). Anche la reazione di *trimerizzazione* è esotermica ( $\Delta H_r^0 = -64.85 \text{ kJ/eq}_{\text{NCO}}$ , Woods, 1987).

### 2.1.3 Proprietà fisiche delle schiume poliuretanicche

Durante la produzione di una schiuma poliuretanicca la simultaneità delle reazioni di *gel* e *blow* permette al materiale di espandersi, ovvero di aumentare il suo volume, contemporaneamente alla crescita delle catene polimeriche. Questo significa che, partendo da un volume iniziale liquido molto ridotto si ottiene tipicamente una struttura finale bifasica meccanicamente stabile, costituita da una matrice solida contenente gas intrappolato al suo interno. La morfologia del materiale, la sua densità, la scelta del gas di espansione, andranno a determinare le proprietà fisiche della schiuma.

#### 2.1.3.1 Morfologia

Il gas si trova all'interno di alcune zone della matrice polimerica che vengono chiamate celle. Le celle possono essere aperte o chiuse; le celle aperte caratterizzano le schiume flessibili, e permettono un flusso di fluidi dall'interno all'esterno o viceversa, cosa che conferisce ad esse la loro tipica elasticità e resistenza alle deformazioni. Per quanto riguarda le schiume rigide, si parla generalmente di celle chiuse. In questo caso, il gas prodotto durante la reazione di *blow* rimane isolato all'interno della struttura, che risulterà relativamente fragile e non elasticamente deformabile. È proprio questa particolare morfologia che conferisce al materiale le sue notevoli capacità isolanti. In Figura 2.4 è mostrata la differenza fra la struttura a celle aperte, Figura 2.4 (a) e celle chiuse, Figura 2.4 (b).



**Figura 2.4.** Diverse morfologie dei prodotti poliuretanic: (a) struttura a celle aperte; (b) struttura a celle chiuse (K. Skleničková et al<sup>8</sup>).

L'elevata reticolazione è quindi un fattore chiave per permettere la formazione delle schiume rigide in quanto rende il materiale resistente alla contrazione di volume in fase di raffreddamento, permettendo lo stiramento senza rottura delle pareti delle celle.

La percentuale di celle chiuse in una schiuma poliuretanic dipende dal grado di reticolazione, dal tipo di tensioattivo utilizzato e dalla massa molare del poliolo. Per la maggior parte degli utilizzi è richiesto un elevato contenuto di celle chiuse. La dimensione delle celle dipende fortemente dal tipo di monomero utilizzato in fase di polimerizzazione (Szycher, 2013).

### 2.1.3.2 Densità e proprietà meccaniche

La densità di una schiuma poliuretanic è un parametro di qualità molto importante. Essa è un indice della quantità di gas presente all'interno della struttura bifasica. Chiaramente, essendo il materiale costituito da una fase solida ed una fase gassosa, quando si discuterà di questa proprietà ci si riferirà alla densità apparente o di *bulk*:

$$\rho_{bulk} [kg/m^3] \equiv \frac{m_{PU}}{Vol_{PU}},$$

$$m_{PU} = \text{massa totale di materiale espanso [kg]}, \quad (2.7)$$

$$Vol_{PU} = \text{volume totale di materiale espanso [m}^3\text{]}.$$

A parità di sistema poliolic, tutte le proprietà meccaniche, come il modulo di elasticità o i parametri di resistenza, la conducibilità termica, la tendenza ad assorbire umidità dipendono fortemente dalla densità del materiale (Szycher, 2013). L'obiettivo per quanto riguarda il controllo di questo parametro è ottenere un compromesso fra stabilità dimensionale e isolamento termico: una ridotta densità significa un elevato contenuto di gas nel materiale, questo comporterebbe proprietà termiche molto buone per il suo utilizzo, ma andrebbe a comprometterne la resistenza ed integrità in quanto si ottiene una diminuzione di resistenza meccanica.

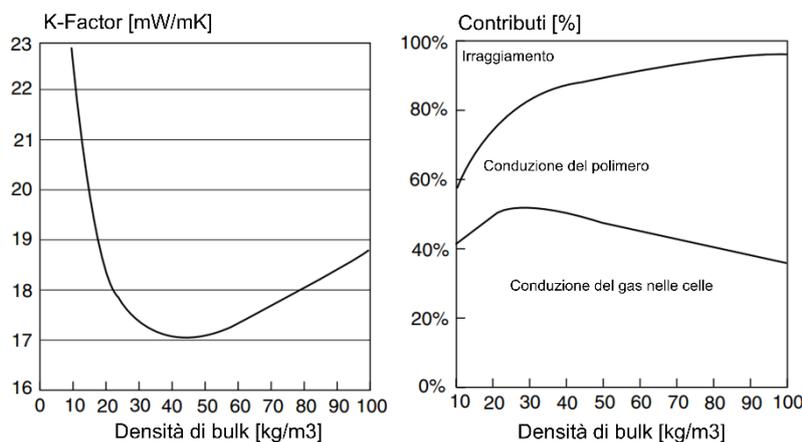
La densità delle schiume rigide può variare dai  $10 \text{ kg/m}^3$  ai  $300 \text{ kg/m}^3$  (Szycher, 2013); schiume con densità inferiore ai  $24 \text{ kg/m}^3$  sono dimensionalmente instabili (l'elevata diminuzione di volume in fase di raffreddamento ne causa il collasso).

Il parametro che più influisce sulla densità della schiuma finale è la quantità di espandente inserita. Infatti, all'aumentare della quantità di gas prodotto si otterranno celle di dimensioni maggiori e di conseguenza una densità più elevata.

È necessario ora discutere dell'isolamento termico, e della sua correlazione con la densità del prodotto. Il trasferimento di calore avviene attraverso conduzione del polimero, conduzione del gas ed irraggiamento. La somma dei contributi va a determinare il coefficiente globale di scambio termico. Esso dipende dalla densità del polimero, ma il comportamento è difficile da prevedere.

Assumendo costante la dimensione delle celle, l'irraggiamento diventa importante a densità ridotte, dove si ha un minor numero di *finestre cellulari* per unità di volume nelle quali la radiazione impatta. Perciò, a basse densità, quando questo effetto è determinante, si ha una scarsa capacità di isolamento. Man mano che la densità cresce, la conduzione del polimero diventa il fenomeno dominante, perciò, si ha un aumento dello scambio termico ad elevate densità (Lee e Ramesh, 2004).

Il miglior isolamento termico si ottiene a densità tali per cui la condizione del polimero e l'irraggiamento sono entrambi limitati, questo avviene a circa  $30 \text{ kg/m}^3$ . In modo analogo, al diminuire della dimensione delle celle diminuisce l'irraggiamento ma aumenta il fenomeno di conduzione del polimero. Essendo però questi due fattori fra loro collegati, generalmente si cerca di controllare solamente la densità (Lee e Ramesh, 2004). L'effetto precedentemente descritto è illustrato in Figura 2.5.



**Figura 2.5.** Effetto della densità sulla conducibilità termica: essa deve variare in un range limitato per ottenere il minimo coefficiente di scambio termico globale rappresentato dal K-factor nel grafico a sinistra. (Adattata da Lee, Ramesh, 2004).

Oltre all'isolamento termico, anche la stabilità dimensionale della schiuma è importante, in quanto si ha un notevole sforzo di compressione in fase di raffreddamento. La stabilità termica dipende dalla formulazione del polimero, dalla densità, dal punto di ebollizione dal tipo di espandente fisico (all'aumentare della volatilità, quest'ultimo tenderà a condensare meno durante l'utilizzo della

schiuma a temperature medio basse, generando meno sforzo di compressione). Inoltre, altri fattori che vanno ad influenzare la stabilità termica sono: lo scorrimento della schiuma (se la schiuma non scorre in modo adeguato, lo sforzo di compressione non si distribuisce e sarà più focalizzato in alcuni punti critici, generando rotture e difetti nel materiale); la tendenza dei prodotti di espansione fisica di diffondere all'esterno della schiuma; l'abilità degli agenti espandenti fisici di diffondere all'interno delle celle (Lee e Ramesh, 2004). La scelta dell'agente espandente è quindi un fattore chiave che verrà discusso successivamente.

#### **2.1.4 Formulazione di una schiuma poliuretanica**

La scelta dei monomeri e degli additivi necessari alla produzione di una schiuma poliuretanica è di importanza fondamentale per la sua cinetica di reazione/espansione. Le miscele chimiche per la produzione di una schiuma poliuretanica rigida sono due e vengono generalmente chiamati componente A e componente B.

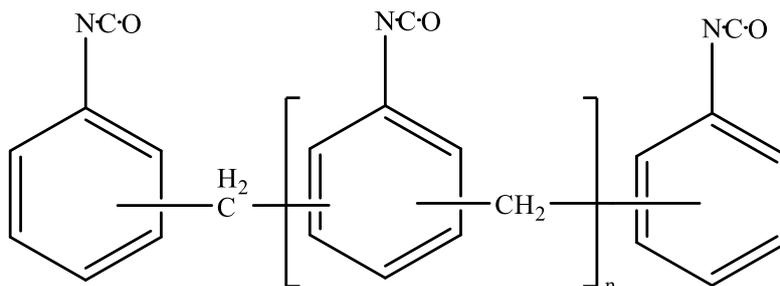
Il componente A contiene la miscela poliolica, assieme agli additivi: catalizzatori, agenti espandenti chimici e fisici, tensioattivi e antifiamma. Il componente B contiene il poliisocianato. Vedremo dunque tutte le caratteristiche di ogni tipologia di ingrediente chiave.

##### **2.1.4.1 Isocianato**

La prima categoria di sostanze che devono essere analizzate è il poli-isocianato generalmente definito come "Componente B" nella terminologia tecnica dei sistemi.

Il termine poli-isocianato si riferisce ad una categoria di molecole organiche contenenti più di un gruppo funzionale isocianato, essenziale alla produzione del poliuretano. I poli-isocianati, vengono più semplicemente chiamati isocianati. La funzionalità dei poli-isocianati può variare notevolmente, il valore minimo per ottenere una polimerizzazione è 2. All'aumentare della funzionalità, aumenta il grado di reticolazione, in quanto nella catena polimerica saranno presenti gruppi liberi laterali che potranno reagire con i polioli e formare le connessioni fra le catene stesse, come illustrato nel Sottoparagrafo 2.1.2.3.

Industrialmente, due isocianati sono utilizzati per la gran parte delle applicazioni: Toluene Di-Isocianato (TDI) o Difenil-Metano Di-Isocianato (MDI). Questi due prodotti hanno in comune una struttura organica aromatica, a cui sono legati i gruppi funzionali designati; il composto impiegato per questo studio è l'MDI. In particolare, viene utilizzata una versione alternativa a questo, ovvero l'MDI polimerico, la cui struttura è rappresentata in Figura 2.6.



**Figura 2.6.** Struttura chimica dell'isocianato polimerico MDI. Si nota come la struttura polimerica del composto generi catene di poliuretano con moltissimi gruppi liberi in catena, adatti ad ottenere una struttura altamente reticolata.

L'utilizzo di MDI polimerico è preferibile al composto comune, in quanto possiede una funzionalità molto elevata e permette il suo utilizzo per la produzione di composti altamente reticolati (Lee e Ramesh, 2004).

La presenza di un gruppo aromatico direttamente legato all'azoto del gruppo  $-NCO$  rende molto più reattivo il composto, in quanto l'anello benzenico è un gruppo elettrone-attrattore, che attrae la carica negativa o nuvola elettronica della molecola a sé, rendendo più attaccabile il carbonio dall'ossigeno dei polioli e dagli agenti espandenti chimici.

L'MDI è preferito invece al TDI in quanto ha una minor pressione di vapore, perciò è meno volatile, riducendo i rischi dovuti alla tossicità dei vapori di isocianato. Inoltre, l'MDI polimerico produce schiume con una stabilità meccanica e termica superiore (Lee e Ramesh, 2004).

Per quanto riguarda questo composto, importante è il contenuto di isocianato che lo caratterizza, esso viene calcolato in questo modo:

$$PE_{NCO} [g/eq] = \frac{MW_{Gruppo-NCO}}{\%NCO \text{ nella molecola}/100} \quad (2.8)$$

Ogni poli-isocianato possiede un valore di peso equivalente, perciò, è possibile utilizzare questo parametro per calcolare quanti equivalenti di poliolo vanno inseriti in miscela per ottenere una miscela stechiometrica:

$$\begin{aligned} \text{Indice di NCO} [-] &= \frac{\text{Equivalenti di Isocianato}}{\text{Equivalenti di poliolo}} \\ \text{Indice di NCO} [-] &= \frac{\text{Equivalenti di Isocianato reali}}{\text{Equivalenti di poliolo stechiometrici}} \end{aligned} \quad (2.9)$$

L'indice di NCO misura la quantità di gruppi  $-NCO$  rispetto alla quantità di gruppi  $-OH$  (sia di polioli che di espandenti chimici) e viene mantenuto costante, in modo tale da ottenere schiume poliuretaniche aventi densità, proprietà termiche, proprietà meccaniche e morfologiche accettabili, in quanto, esse dipendono dal contenuto di isocianato in eccesso. In genere, si lavora con un

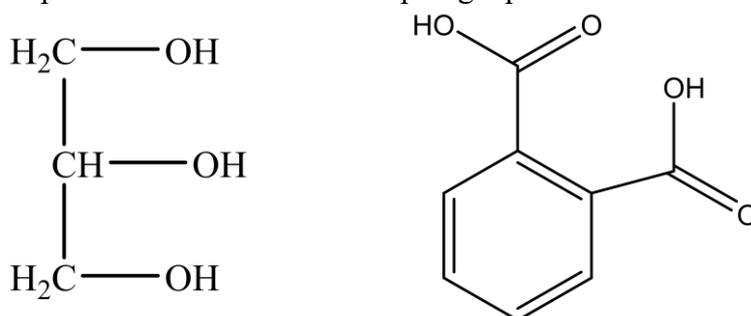
Indice di NCO oscillante attorno a 1.2 [–]. L'eccesso di isocianato viene inserito per stimolare la reazione di *trimerizzazione* e la produzione di poliisocianurati.

Infine, l'isocianato è l'unico componente della miscela di formulazione ad essere particolarmente pericoloso, in quanto è irritante e pericoloso per la salute.

#### 2.1.4.2 Polieteri e Poliesteri

I polioli sono le sostanze principali per quanto riguarda il Componente A. Come precedentemente descritto costituiscono la seconda parte di monomeri necessaria alla poliaddizione. Si distinguono fondamentalmente in due categorie: polieteri e poliesteri.

I polieteri sono generalmente alcoli a funzionalità elevata e sono basati su molecole di partenza come glicerina, sorbitolo, saccarosio e altri derivati aromatici. I poliesteri invece sono basati su acidi organici, come ad esempio l'acido ftalico. Le due tipologie possono essere viste in Figura 2.7.



**Figura 2.7.** Polioli tipicamente utilizzati nella formulazione dei poliuretani: (a) molecola di glicerina, il polietere più semplice; (b) acido adipico, comune poliesteri impiegato.

Il vantaggio dei polieteri sta nel poter disporre di una varietà molto elevata di composti, cosa che permette al formulatore di ottenere più facilmente le proprietà richieste; portano anche alla produzione di schiume più rigide e stabili. I poliesteri invece, possono essere recuperati da materiali di scarto (ad esempio dal riciclo di bottiglie in PET), oppure possono essere prodotti da starter di origine vegetale, inoltre a seconda dell'andamento delle materie prime possono avere un costo inferiore. Spesso possiedono gruppi aromatici che possono migliorare la risposta al fuoco. I poliesteri tendono ad avere cinetiche eccessivamente veloci e generano schiume più morbide in quanto tendono a rompere le celle chiuse. L'elevata reattività della molecola di poliesteri può essere spiegata analizzando la capacità delle molecole di delocalizzare le cariche negative: essendo presente un carbonile ( $C = O$ ) legato al carbonio del gruppo idrossile, una eventuale perdita dell'idrogeno genera la presenza di una carica negativa sull'ossigeno, che può essere delocalizzata appunto nell'ossigeno del carbonile, stabilizzando la forma ionica, e rendendo l'ossigeno più nucleofilo. Questo non può succedere nei polieteri, in quanto l'ossidrile è legato ad un carbonio che stabilizza più difficilmente una eventuale carica negativa dell'ossigeno, rendendo più difficile l'indebolimento del legame tra ossigeno ed idrogeno (Lee e Ramesh, 2004).

Il peso equivalente dei polioli è una delle caratteristiche più importanti, il composto infatti costituisce poi la componente principale delle reticolazioni fra le catene; perciò, una reticolazione più o meno fitta va a modificare la rigidità o la flessibilità della struttura, oltre che della catena stessa. Inoltre, esso andrà ad influire su altre caratteristiche importanti tra cui: la capacità di assorbire l'umidità, la percentuale di celle chiuse nella schiuma, la resistenza al fuoco e sulla quantità di gas perso durante l'espansione. Per quanto riguarda le schiume rigide, i polioli utilizzati hanno un ridotto peso molecolare rispetto alle altre applicazioni poliuretaniche.

Ogni gruppo ossidrilico presente nei polioli ha una reattività diversa, in base alle molecole organiche a cui sono legati. In ordine dal più al meno reattivo, abbiamo idrossili primari, secondari, terziari e infine fenolici. La differenza di reattività è notevole: in un range da 25 a 50°C, la reattività degli alcoli secondari va da uno a tre ordini di grandezza in meno.

Per quantificare il numero di gruppi idrossili presenti nella molecola poliolica, viene utilizzato il seguente parametro:

$$N^{\circ} OH_i [mgKOH/g \text{ poliolo}] = MW_{KOH} \cdot 1000 \cdot \frac{\text{funzionalità}_{\text{poliolo}}}{MW_{\text{poliolo}}}. \quad (2.10)$$

Il numero di idrossido è un parametro empirico specifico per ogni poliolo e rappresenta la concentrazione di gruppi ossidrilici (o gruppi in grado di reagire con il gruppo funzionale  $-NCO$ ) per unità di massa di poliolo, espressi in milligrammi equivalenti di potassio idrossido.

L'indice viene dato dall'azienda produttrice della sostanza, come accade per il peso di isocianato equivalente descritto in precedenza. Questo parametro è utilizzato non solo per i polioli, ma anche per gli agenti espandenti chimici, in quanto anch'essi reagiscono con l'isocianato e ne consumano una quota parte. È quindi necessario tenere conto anche della loro presenza in fase di formulazione. È possibile, perciò, dosare la massa di poliolo che deve essere utilizzata nella formulazione della schiuma. Comunemente, la massa di miscela poliolica viene tenuta come riferimento costante:

$$eq \text{ di } OH = \sum_i \frac{m_i \cdot N^{\circ} OH_i}{MW_{KOH} \cdot 1000}, \quad i = \text{poliolo, esp. chimici}, \quad (2.11)$$

$$eq.NCO_{stech} = eq.OH, \quad (2.12)$$

$$eq.NCO_{reali} = \text{Indice di } NCO \cdot eq.NCO_{stech}, \quad (2.13)$$

$$m_{isocianato} = eq.NCO_{reali} \cdot \text{Peso Equivalente di Isocianato} \quad (2.14)$$

Fisicamente, i polioli si presentano come sostanze altamente dense e viscosi. Non ci sono particolari rischi legati loro utilizzo, occorre comunque prestare attenzione in quanto composti chimici organici caratterizzati da una particolare acidità (Lee e Ramesh, 2004).

### 2.1.4.3 Catalizzatori

La catalisi delle reazioni coinvolte nella produzione di una schiuma poliuretanic è uno dei fattori più importanti per la cinetica di produzione di questo tipo di materiale. Essa, infatti, permette di modulare la velocità di espansione di una schiuma, di regolarizzare la forma finale e di rendere più veloce il processo di produzione.

Esistono moltissimi tipi di molecole che catalizzano la reazione di formazione del legame uretanico. Tuttavia, il meccanismo di azione è soggetto a diversi dibattiti e non è ancora propriamente definito. I meccanismi proposti sono specifici per il tipo di catalizzatore ed è possibile identificarne diverse tipologie, che verranno qui descritte.

- catalizzatori generici: ammine terziarie, il cui effetto è l'accelerazione di molteplici reazioni, oltre a quella di formazione dell'uretano. Il meccanismo più probabile è l'azione polarizzante dovuta al doppietto elettronico dell'atomo di azoto, che va a polarizzare le molecole di isocianato ed idrossido, rendendo più elettrofilo il carbonio in  $-NCO$  e più stabile la rimozione dell'idrogeno legato all'ossigeno. Le ammine terziarie possono essere chimicamente sviluppate per favorire in modo più efficace una fra le reazioni di *gel* o di *blow* (Ionescu, 2004, Sonnenschein, 2015);
- catalizzatori ad effetto ritardato: sali di ammonio quaternario i quali si dissociano solamente in seguito allo sviluppo di calore generato dalla reazione di *blow*, liberando una ammina, che va a catalizzare la reazione di *gel* e un acido, che reagisce con l'isocianato producendo ulteriori anidride carbonica e monossido di carbonio. L'effetto ritardato è necessario quando si desidera una rapida espansione, innescando solo successivamente la poliaddizione, in modo da mantenere il materiale scorrevole (Lee e Ramesh, 2004);
- catalizzatori di *trimerizzazione*: in genere, i catalizzatori di questo tipo sono molto basici e nucleofili, ipoteticamente essi operano attaccando il carbonio elettrofilo di un isocianato in modo da stabilirne un legame, questo rompe il doppio legame azoto-carbonio e rende l'azoto carico negativamente. L'azoto può quindi attaccare un secondo isocianato allo stesso modo e a sua volta un terzo. Una volta che si è formata la stabile struttura ad anello con tre gruppi isocianato, il catalizzatore si allontana dal primo carbonio attaccato. Questo tipo di catalizzatori è generalmente costituito da sali di metalli alcalini carbossilati (Sonnenschein, 2015). La loro struttura permette che vengano facilmente solvatati dall'acqua e perciò agiscono solamente al termine della reazione di *blow* quando quest'ultima viene del tutto consumata.

### 2.1.4.4 Agenti Espandenti

Come anticipato al Sottoparagrafo 2.3.2, esistono due tipologie di sostanze utilizzate nella creazione della struttura a *celle* di una schiuma poliuretanic: espandenti chimici o espandenti fisici.

Gli espandenti chimici, producono anidride carbonica o monossido di carbonio a seguito di una reazione chimica esotermica con l'isocianato. Acqua e acido formico sono spesso utilizzati per questo scopo. Gli espandenti fisici invece sono sostanze volatili che evaporano in seguito allo sviluppo del calore di reazione e generano un aumento della dimensione delle *celle* portando all'espansione della schiuma, per via endotermica. Inizialmente, gli espandenti fisici erano composti Cloro-Fluoro-Carbonati (CFC), ma sono stati successivamente rimossi a causa del loro impatto negativo sull'ozonosfera (Lee e Ramesh, 2004). Il dosaggio di questi due componenti permette, in modo analogo alla catalisi, seppur con un effetto ridotto, la modulazione del profilo di crescita di una schiuma.

#### 2.1.4.5 Antifiamma e tensioattivi

Come suggerisce il nome, gli additivi antifiamma conferiscono al polimero una resistenza al fuoco necessaria al loro impiego. Lo scopo è la riduzione del calore e dell'emissione di gas durante la combustione del materiale. Generalmente, gli antifiamma sono organofosfati o alogenuri organici inerti, che agiscono mediante meccanismi diversi: diluiscono i prodotti di combustione, riducono la densità ottica dei fumi, stabilizzano i radicali liberi e producono un *char*, ovvero una sostanza voluminosa a base di carbonio che si forma a seguito di una combustione, in grado di separare la fiamma dal residuo polimerico non combusto e di isolare la propagazione del calore.

Alcuni polioli utilizzati possono contenere in catena sostanze antifiamma, in modo da ottenere intrinsecamente questi effetti. I tensioattivi sono invece sostanze a base siliconica (dimetilpolisilossani), aggiunti in percentuale non superiore all'1%, che stabilizzano la schiuma immediatamente dopo al mixing dei componenti A e B e successivamente anche durante la fase di crescita della schiuma. Essi hanno l'effetto di diminuire la tensione superficiale all'interfaccia gas liquido e permettere alle celle di espandersi e non collassare, assicurando inoltre la formazione di una struttura cellulare più fine e regolare (Lee e Ramesh, 2004).

## **2.2 Processo di produzione e tecniche di analisi**

Il processo di produzione di una schiuma poliuretana rigida è il caso studio considerato in questa Tesi. In particolare, l'obiettivo è l'introduzione di metodologie sistematiche, oggettive e basate su tecniche scientifiche per analizzare i processi di formulazione e di crescita. In questo paragrafo, si introduce la tecnica di produzione di una schiuma rigida e le metodologie di analisi del processo e di formulazione attualmente impiegate.

### ***2.2.1 Produzione di una schiuma poliuretana rigida***

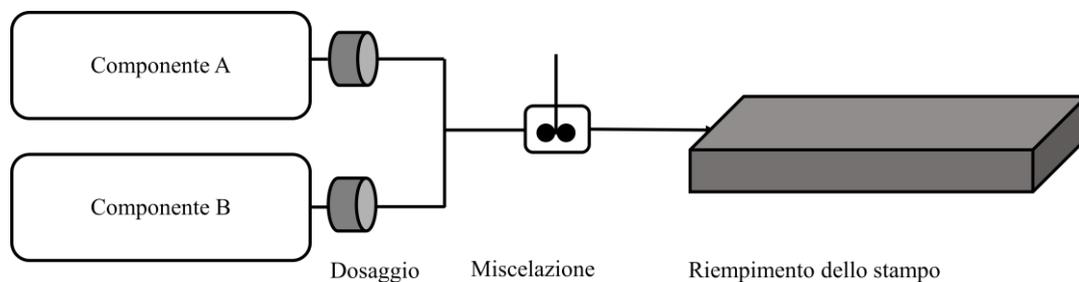
Le schiume poliuretane rigide analizzate in questo lavoro di tesi, vengono prodotte per iniezione. Essa viene svolta direttamente in loco, il che significa che non viene commercializzata successivamente alla polimerizzazione, ma è necessario produrre la schiuma allo stesso tempo della

produzione del manufatto che la contiene. La tecnica prevede la rapida miscelazione di componente A e componente B in un brevissimo tempo, e l'iniezione della miscela risultante all'interno dello stampo ad opera di macchine cosiddette *schiumatrici*.

Il rapporto fra le due parti è costantemente controllato attraverso misuratori di portata e deve mantenere un rapporto fisso per rispettare l'*indice di NCO*:

$$\text{Rapporto } [-] = B/A = \frac{\text{Massa di componente B}}{\text{Massa di componente A}} \quad (2.15)$$

In precedenza, si è affermato che lo scorrimento della schiuma e la sua velocità di espansione sono cruciali per la qualità del prodotto; questo è dovuto alla necessità di riempire lo stampo in modo efficace ed uniforme. Il processo è rappresentato in modo schematico in Figura 2.8.



**Figura 2.8.** Rappresentazione schematica del processo di produzione di un pannello in poliuretano espanso attraverso iniezione.

È quindi in merito a questo processo che si sviluppa il problema di formulazione di una schiuma poliuretanicica, in quanto, la cinetica della sua espansione determina il modo in cui essa riempirà lo stampo una volta iniettata. In altre parole, la formulazione determina la qualità del prodotto finale. Inoltre, si deve notare che se la velocità di espansione di una schiuma rispetto alla velocità di polimerizzazione non è abbastanza elevata, allora essa andrà a riempire solamente parzialmente lo stampo, generando un composito difettoso e disfunzionale. Se invece la velocità di *gel* è troppo ridotta, allora si andrà incontro a problemi di allungamento dei tempi di stampaggio e conseguente diminuzione della produttività. Occorre quindi trovare il giusto compromesso per ottenere una cinetica adeguata<sup>2</sup>.

### 2.2.2 Stato dell'arte su analisi e formulazione del processo

I tempi caratteristici di una schiuma poliuretanicica sono dei parametri empirici che contraddistinguono ogni prodotto ed in modo molto semplice e veloce ne caratterizzano il comportamento.

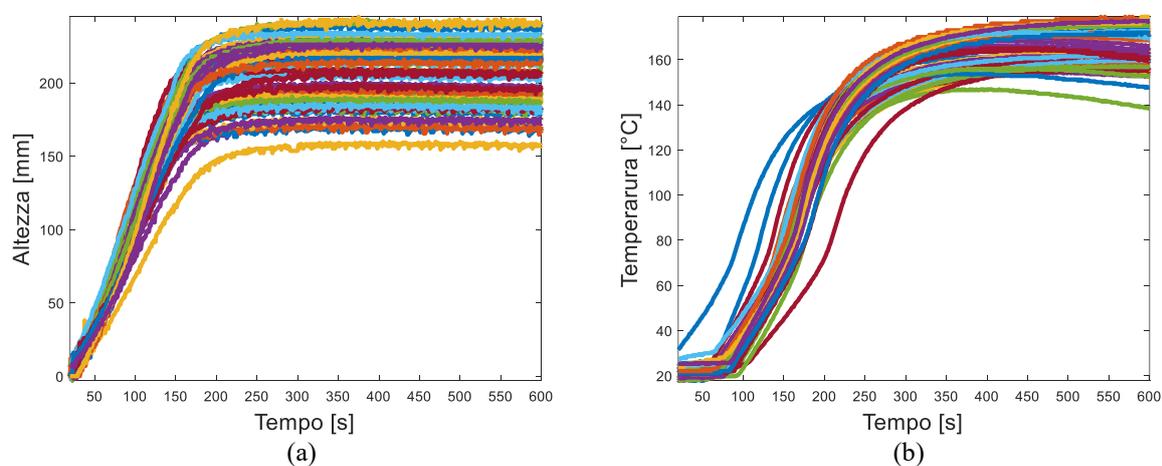
<sup>2</sup> La cinetica di un sistema poliuretanicico in espansione, ovvero la sua velocità di crescita e di aumento di temperatura, è rappresentata dai profili di crescita di una schiuma e viene utilizzata come indicatore per determinare la qualità del prodotto.

Essi, assieme alla densità del prodotto, sono dati di qualità che vengono riportati nella scheda tecnica di un formulato, e permettono all'utilizzatore finale della schiuma di capire quale sarà il suo comportamento una volta iniettato nello stampo. Come è possibile notare in Tabella 3.1, l'analisi risulta molto empirica: è necessaria la presenza di un operatore esperto per essere compiuta in modo adeguato e riproducibile.

**Tabella 3.1.** Descrizione schematica dei tempi caratteristici di una schiuma poliuretanica, della loro misurazione e dei valori comuni per i formulati standard

	Indicazione	Misurazione	Valori Comuni [s]
<b>Tempo di crema</b>	Inizio della reazione di <i>blow</i>	Misura del tempo al quale la miscela ottiene un aspetto cremoso	1-15
<b>Tempo di gel/filo</b>	Inizio della reazione di polimerizzazione	Misura del tempo al quale una volta inserita una bacchetta nella schiuma, essa trascina a sé dei filamenti all'estrazione	100-150
<b>Tempo di tocco</b>	Fine della reazione di polimerizzazione e inizio indurimento	Misura del tempo al quale toccando il polimero esso risulta solido e non più vischioso	250-350

Tuttavia, essa rappresenta uno strumento veloce ed efficace per valutare, in prima approssimazione, il comportamento della schiuma. Non è però adeguata quando è necessario approfondire la conoscenza del processo, ad esempio, quando è necessario riformulare la miscela per creare un nuovo prodotto. Un'analisi più accurata viene invece svolta attraverso misurazioni dei profili di crescita della schiuma nel tempo. Come verrà descritto al Sottoparagrafo 2.2.3, i profili possono essere raccolti attraverso un dispositivo ad ultrasuoni che misura la distanza fra lo strumento e la schiuma in ogni istante temporale. Successivamente alla manipolazione preliminare, il confronto e l'analisi del profilo di crescita di ogni singolo formulato; questi permettono al formulatore uno studio con un'accuratezza e una quantità di informazioni decisamente maggiori. Anche i profili di temperatura della schiuma in espansione possono essere raccolti attraverso l'utilizzo di una semplice termocoppia. In Figura 2.9, sono riportati alcuni esempi dei profili temporali di altezza e temperatura delle schiume.



**Figura 2.9.** Esempi dei profili ricavati attraverso l'analisi della crescita delle schiume poliuretaniche. A sinistra, profili di altezza, a destra, profili di temperatura

Tuttavia, il procedimento rimane molto soggettivo: l'analisi viene svolta visivamente e i confronti fra due curve vengono fatti ad opera dell'analista, che può scorgere i dettagli in base alla sua esperienza, senza quindi utilizzare un approccio sistematico.

Inoltre, una analisi completa di tutte le curve diverse e della correlazione che le lega risulta essere molto complessa, in quanto, ogni modifica causa al formulato una diversificazione dei profili; come risultato, il numero di profili raccolti è molto elevato e vario.

In questo studio, questi dati verranno analizzati attraverso metodologie di analisi multivariata dei dati che permette di studiare i profili temporali di queste variabili, come questi siano correlati tra loro e come siano correlati con la formulazione del prodotto.

Attualmente, infatti, anche il processo di formulazione avviene per mezzo di un'analisi dei profili e dei tempi caratteristici, attuando una procedura iterativa di tipo *trial and error*: viene proposta una formulazione il cui comportamento viene ritenuto essere simile a quello desiderato dagli esperti del sistema; il dosaggio è quindi basato sulle conoscenze degli operatori del settore in merito all'effetto che ogni sostanza dei componenti A e B ha sulla qualità della schiuma; tipicamente si effettua l'analisi sperimentale e si confrontano i risultati con i profili già raccolti per dei prodotti ritenuti standard e, infine, viene deciso se sia necessario correggere la formulazione o se il prodotto ottenuto soddisfa le specifiche desiderate.

### **2.2.3 Campagna sperimentale**

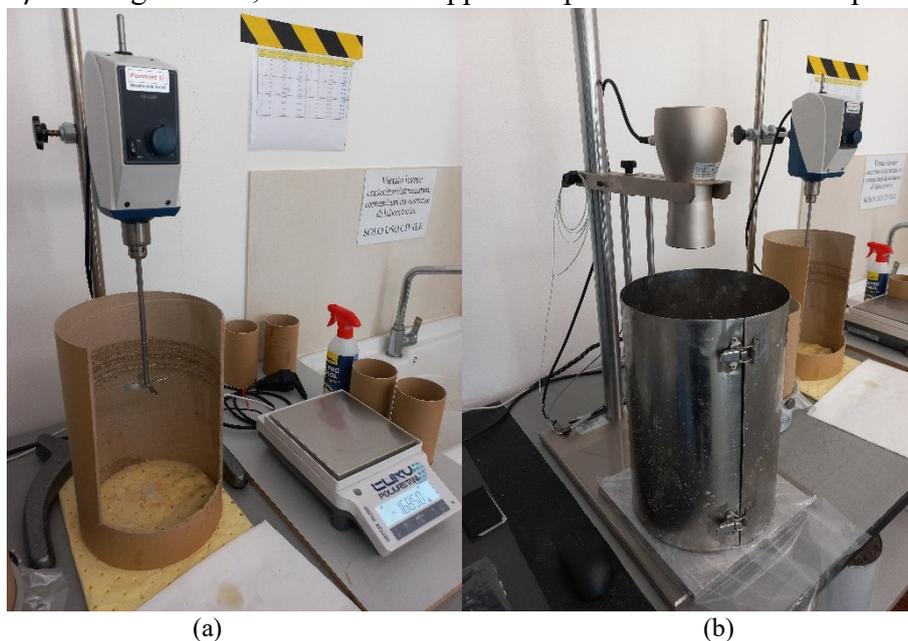
Essendo stata svolta una campagna sperimentale per la raccolta di dati, vengono illustrate le metodologie sperimentali utilizzate per eseguire le prove che permettono di raccogliere dati sui tempi caratteristici e sui profili di altezza e temperatura. Inoltre, è stata misurata la densità finale del prodotto; questa misura viene effettuata andando a pesare un volume standard di materiale espanso ricavato dal nucleo interno del provino al termine della prova.

#### **2.2.3.1 Procedura di analisi sperimentale**

Generalmente, l'analisi si svolge per una durata totale di 615 s, i quali comprendono 15 s necessari alla miscelazione ed al versamento delle due componenti e 600 s nei quali la schiuma si espande.

L'attrezzatura si compone di: una bilancia tecnica (Mettler Toledo® ME 3002E, Giessen, Germania, precisione 0.01 g, portata massima 3200 g), due contenitori del volume di 250 mL, un agitatore a flusso assiale (WiseStir® HS-120A, Witeg Labortechnik GmbH, Wertheim, Germania, potenza motore di 120W, frequenza di rotazione massima 3000 rpm, dimensioni della girante 70 mm a 4 rante), un contenitore cilindrico in metallo ( $D = 230 \text{ mm}$ ,  $H = 320 \text{ mm}$ ), una bacchetta di metallo, un cronometro, un sensore di distanza ad ultrasuoni (Ultrasonic Fan Sensor LR4® Format Labortechnik GmbH, Karlsruhe, Germania), una termocoppia (tipo K, NiCr/Ni, temperatura massima di 1000°C) e infine un software (CDC – Curve Di Crescita®, HVT Technologies, Bangalore, India, sviluppato per raccogliere i dati del sensore ad ultrasuoni e della termocoppia tramite Datalogger TC-

08 della Pico-Technology, a 8 canali ed interfaccia USB) che permette l'elaborazione dei dati raccolti dai misuratori. Il sensore di distanza e la termocoppia misurano i dati con una frequenza di 1 misurazione/s. In Figura 2.10, è illustrato l'apparato sperimentale utilizzato per l'analisi.



**Figura 2.10.** Apparecchiatura di analisi: (a) bilancia tecnica utilizzata per il dosaggio, (b) il contenitore cilindrico di metallo in cui espanderà la schiuma, il sensore di distanza e la termocoppia; (b) girante utilizzata per la fase di miscelazione.

La procedura di analisi sperimentale consiste in:

- preparazione del componente A, dosaggio di tutti i componenti della miscela;
- preparazione del componente B;
- miscelazione dei componenti A e B per 15 s;
- inizio delle misurazioni di distanza e temperatura;
- miscelazione delle componenti A e B;
- versamento della miscela all'interno del contenitore cilindrico;
- misurazione del tempo di crema andando ad osservare quando la miscela ottiene un aspetto cremoso;
- una volta che la miscela ha raggiunto i 70°C, inizio della fase di misurazione del tempo di *gel*: si inizia ad inserire una bacchetta di metallo all'interno della schiuma in espansione, non appena all'estrazione della bacchetta si noteranno dei filamenti residui, il tempo di crema viene misurato;
- a partire dai 200 s, si inizia a toccare la superficie libera della schiuma all'interno del contenitore, nel momento in cui la superficie ottiene caratteristiche solide al tocco, viene misurato il tempo di tocco.
- misurazione della distanza finale tra il sensore e la calotta della schiuma;
- misurazione dell'altezza del cilindro di poliuretano prodotto;

- una volta raggiunti 615 s dal versamento, la raccolta dei profili è terminata; il pezzo prodotto viene infine tagliato e pesato, in modo da ottenere la misura della densità di bulk. Le fasi di raccolta dei tre tempi caratteristici sono illustrate in Figura 2.11.



**Figura 2.11.** Fasi di misurazione dei tempi caratteristici di espansione della schiuma: (a) misurazione visiva del tempo di crema; (b) misurazione del tempo di gel; (c) misurazione del tempo di tocco; (d) materiale espanso finale.

Esistono in commercio apparecchiature avanzate per la raccolta dei suddetti profili, che permettono una lettura approfondita e un controllo qualità dei dati più accurato, in grado di raccogliere inoltre altri dati significativi come, ad esempio, i profili di viscosità della schiuma; questi strumenti, tuttavia, tendono ad essere poco accessibili dal punto di vista economico e vanno a sostituire completamente i vecchi strumenti.

#### 2.2.4 Dati sperimentali disponibili e formulati di riferimento

Ogni profili temporale dispone di 600 valori di distanza ed altrettanti valori di temperatura. Vanno distinti i dati di calibrazione dai dati di convalida: i primi sono dati di formulati riferimento, mentre i secondi provengono da nuove formulazioni.

### 2.2.4.1 Dati di calibrazione

Nel corso degli anni è stato misurato un numero molto elevato di curve di crescita dei diversi formulati<sup>3</sup> con il fine di analisi attraverso il confronto grafico fra profili diversi, di cui si è parlato precedentemente. Tutti questi dati sono utilizzati in questa Tesi per la calibrazione di una PCA esplorativa, i cui risultati e dettagli sono riportati al Capitolo 3.

Il dataset di calibrazione di cui si dispongono i dati è descritto in Tabella 2.1

**Tabella 2.1.** *Dettagli in merito al dataset di calibrazione disponibile per l'analisi.*

<b>Variabili <math>V</math> [-]</b>	2
<b>Istanti temporali <math>K</math> [-]</b>	600
<b>Osservazioni <math>N</math> [-]</b>	1013

Essendo il numero di formulati diversi all'interno di questo dataset molto elevato, sono state scelte otto tipologie in base alla loro rappresentatività, al numero di osservazioni di cui dispongono, alla quantità che ne viene generalmente prodotta dall'azienda, cercando di spaziare all'interno di tutto il dominio di misura, in modo da sceglierne alcuni di standard e alcuni il cui comportamento è estremo, che sono prese come riferimento per i dati standard, al fine di studiarne il comportamento in modo più chiaro.

Per motivi di sicurezza dei dati aziendali di Europoliuretani s.r.l., la composizione dei formulati analizzati non verrà divulgata, inoltre, verranno criptati tramite una sigla alfanumerica progressiva, in sostituzione della loro denominazione effettiva;

Questi sono riportati in Tabella 2.2, assieme al numero di osservazioni disponibile per ognuno di essi:

**Tabella 2.2.** *Classificazione dei formulati presi come riferimento, sui quali si concentrerà l'analisi PCA.*

<b>Formulato</b>	F1	F2	F3	F4	F5	F6	F7	F8	Totale Campioni
<b>Numero di Osservazioni [-]</b>	32	25	30	33	5	23	19	26	193

Per ogni categoria di formulati di calibrazione sono disponibili i valori dei tempi caratteristici e di densità presenti in scheda tecnica, in Tabella 2.2 sono riportati i valori dei formulati presi come riferimento.

<sup>3</sup> Il termine formulato indica una categoria di componente A con una specifica composizione chimica, ovvero con la stessa quantità e tipologia di polioli, agenti espandenti, catalizzatori ed additivi.

**Tabella 2.2.** Valori di densità e tempi caratteristici, provenienti dalle schede tecniche degli otto formulati presi in esame.

<b>Formulato</b>	<b>Tempo di Crema [s]</b>	<b>Tempo di Gel [s]</b>	<b>Tempo di tocco [s]</b>	<b>Densità [kg/m<sup>3</sup>]</b>
<b>F1</b>	10	145	240	29
<b>F2</b>	13	135	265	30
<b>F3</b>	9	92	155	28,5
<b>F4</b>	13	240	420	30
<b>F5</b>	17	60	75	40
<b>F6</b>	12	127	210	29,5
<b>F7</b>	45	230	420	40
<b>F8</b>	10	45	70	37

#### 2.2.4.2 Dati di convalida

I dati di convalida rappresentano dei nuovi formulati prodotti con il fine di sostituire il formulato F1, in quanto un catalizzatore ed uno dei polioli per la sua formulazione sono divenuti indisponibili e hanno raggiunto un costo eccessivo.

**Tabella 2.3.** Dettagli in merito al dataset di convalida disponibile per l'analisi.

<b>Variabili V [-]</b>	2
<b>Istanti temporali K [-]</b>	600
<b>Osservazioni N [-]</b>	29

Le 29 osservazioni si dividono in tre versioni del nuovi formulato, le caratteristiche dettagliate del dataset di convalida sono riportate in Tabella 2.3.

**Tabella 2.3.** Rappresentazione schematica delle nuove formulazioni di convalida.

<b>Nuove formulazioni</b>	NF1v1	NF1v2	NF1v3
<b>Numero di osservazioni [-]</b>	8	9	9

La nuova metodologia di analisi dei sistemi poliuretanic, introdotta nel Capitolo successivo, potrebbe essere validata valutando le prestazioni attraverso l'analisi di convalida di formulati dal comportamento medio ed altri dal comportamento estremo. Per motivi di tempistiche e necessità aziendali al momento dello sviluppo del nuovo sistema, i dati di convalida trattati nel capitolo hanno un comportamento medio.

Si analizzano tre diverse versioni di formulazione adatte alla sua sostituzione, e decretare quale sia la più indicata alla sostituzione.



# Capitolo 3

## Metodologie di machine learning per lo studio della cinetica di espansione di sistemi poliuretanici rigidi

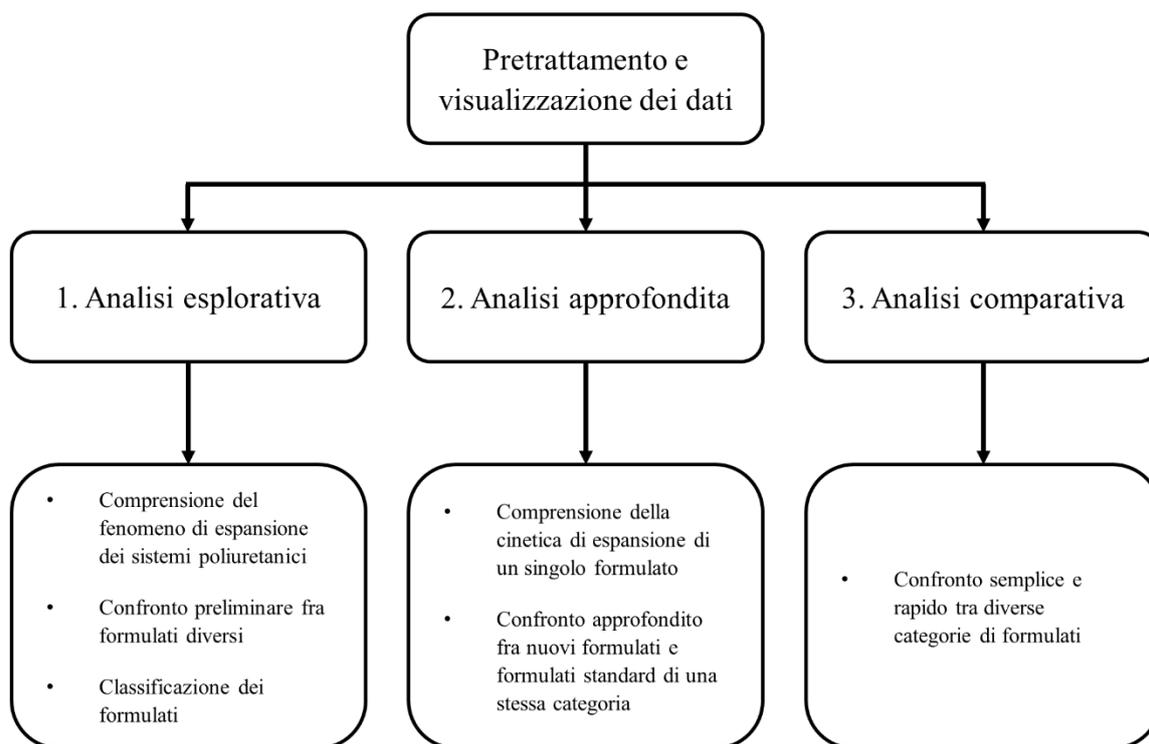
In questo Capitolo viene proposto un nuovo metodo di analisi dei dati per lo studio della cinetica di espansione di schiume rigide. L'obiettivo è proporre una modalità di analisi delle cinetiche d'espansione basata su metodi scientifici dell'Industria 4.0, quindi più complete, oggettive ed accurate delle tecniche allo stato dell'arte presentate nel Capitolo 2.

### 3.1 Metodologia multivariata di analisi di sistemi poliuretanici rigidi

Mentre le tecniche di analisi che vengono tipicamente utilizzate nella pratica industriale (§2) e che costituiscono attualmente lo stato dell'arte si limitano alla visualizzazione dei dati grezzi di profili di altezza e temperatura del poliuretano in espansione, ovvero ad un confronto non oggettivo i diversi profili di crescita, in questo Capitolo si propone una nuova metodologia di analisi della cinetica di crescita di un sistema poliuretano mediante tecniche multivariate di analisi dei dati. Innanzitutto, i dati devono essere opportunamente pretrattati e trasformati, in modo da renderli più descrittivi e facilmente trattabili. Si propongono, in particolare, 3 tecniche diverse di analisi, a seconda delle informazioni di cui si dispone e dell'obiettivo:

1. **analisi esplorativa:** tutte i profili di crescita raccolti ed analizzabili vengono utilizzati per calibrare un modello PCA al fine di ottenere: una comprensione accurata del fenomeno studiato rispetto ai formulati di cui si possiedono i profili; un confronto preliminare tra la cinetica di espansione di nuovi formulati con i formulati standard utilizzati per la calibrazione; l'assegnazione ad un nuovo formulato di cui non si conosce la composizione alla categoria a cui il suo comportamento è riconducibile;
2. **analisi approfondita sullo standard di un formulato di riferimento:** l'analisi PCA è calibrata utilizzando le osservazioni standard di un solo formulato allo scopo di ottenere un riscontro oggettivo tra eventuali variazioni di un formulato e lo standard dello stesso;
3. **analisi comparativa tra formulati:** è possibile confrontare diversi formulati fra loro per ottenere interpretazioni veloci e facilmente interpretabili di nuovi formulati.

Lo schema di analisi è rappresentato in Figura 3.1.



**Figura 3.1.** Schema a blocchi della metodologia proposta per l'analisi della cinetica di espansione di sistemi poliuretanic rigidi.

## 3.2 Applicazione della metodologia

In questo Paragrafo si mostra l'applicazione della metodologia di analisi della cinetica proposta per tutti i formulati di riferimento con l'obiettivo di trovare quale formulazione NF1 sia la più adatta alla sostituzione di F1, come anticipato nel Sotto-sottoparagrafo 2.2.4.2 attraverso ogni blocco dello schema in Figura 3.1.

### 3.2.1 Pretrattamento e visualizzazione dei dati

In questo Sottoparagrafo vengono descritte le tecniche di pretrattamento e trasformazione dei dati preliminari all'analisi, necessarie a renderli analizzabili in modo ottimale nelle successive fasi. Vengono poi visualizzati i profili standard dei formulati di riferimento con i quali il modello viene costruito, per ottenere alcune informazioni preliminari sul caso studio.

#### 3.2.1.1 Filtrazione dei dati

Il misuratore di distanza e la termocoppia raccolgono i profili a partire da 15 s successivi all'avvio dell'analisi sperimentale, che inizia con la miscelazione delle due componenti ed il versamento della all'interno dello stampo. Il sensore misura la distanza che lo separa dalla testa della schiuma

$D_{t_i}$  [mm], dove  $t_i$  indica l'istante di tempo a cui il dato viene misurato;  $t_0$  corrisponde al primo istante temporale.

Si osserva che spesso i primi valori raccolti dallo strumento presentano un rumore molto elevato e valori anomali durante i secondi iniziali di analisi; ad esempio, l'operatore potrebbe non avere ancora finito di versare la miscela, oppure essa potrebbe non avere ancora una forma ben definita all'interno dello stampo per cui la misura risulta indefinita. Per questo motivo si è deciso di rimuovere i dati raccolti dallo strumento nei primi 20 s. Il tempo di analisi andrà quindi da 20 s a 600 s successivi all'inizio della raccolta delle misure.

Occorre effettuare anche una procedura di filtrazione dei profili: un problema che riguarda principalmente i profili temporali di altezza è la misurazione errata di alcuni valori di distanza, ad esempio, quando l'operatore misura i tempi caratteristici della schiuma, può succedere che posizioni la mano fra il sensore e la schiuma, introducendo un errore di misura e la registrazione di un valore anomalo (*outlier*). Per rimuovere questi valori anomali, viene introdotto un filtro che agisce in 3 stadi successivi:

1. se  $D_{t_i}$  possiede un valore eccessivamente diverso dalla misura al precedente istante temporale, allora esso verrà sostituito da  $D_{t_{i-1}}$ . Viene definita come eccessivamente diversa una differenza fra valori ad istanti temporali adiacenti maggiore di  $a_f$  [-] volte la deviazione standard dei  $b_f$  [-] valori attorno ad  $D_{t_i}$ ;
2. se  $D_{t_i}$  si differenzia dal precedente valore per più di  $c_f$  [mm], allora esso verrà sostituito da  $D_{t_{i-1}}$ ;
3. se la distanza ha un valore inferiore a  $d_f$  [mm], allora esso verrà sostituito da  $D_{t_{i-1}}$ .

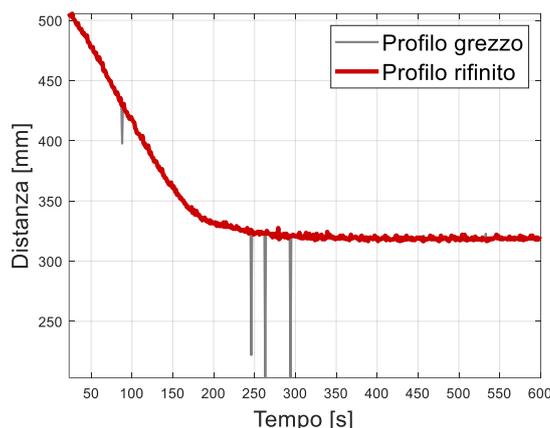
L'unica eccezione riguarda  $D_{t_0}$ , che nel caso di misurazioni anomale viene sostituito da  $D_{t_1}$ .

I parametri  $a, b, c, d$  sono stati tarati per ottenere una filtrazione ottimale e sono riportati in Tabella 3.1.

**Tabella 3.4.** Valori dei parametri utilizzati per il filtraggio dei dati sperimentali.

$a_f$ [-]	$b_f$ [-]	$c_f$ [mm]	$d_f$ [mm]
4	10	60	200

Il risultato della filtrazione dei dati è mostrato in Figura 3.2. In rosso è rappresentata la curva filtrata, in grigio i dati grezzi. È importante notare come le interferenze attorno ai 50,250 e 300 secondi dall'inizio della prova vengano rimosse.



**Figura 3.2.** Esempio di filtrazione dei dati per un formulato standard sul profilo temporale della distanza.

A seguito della filtrazione alcuni campioni sono eliminati in quanto le misure non sono affidabili.

**Tabella 3.5:** Risultato della rimozione dei profili non analizzabili dal dataset disponibile inizialmente.

Osservazioni	Prima della filtrazione	A seguito della filtrazione
Numero di osservazioni [-]	1013	648
Numero di formulati [-]	66	53

In Tabella 3.2 vengono riportati i dettagli in merito al numero dei dati inutilizzabili rimossi.

### 3.2.1.2 Trasformazione dei dati

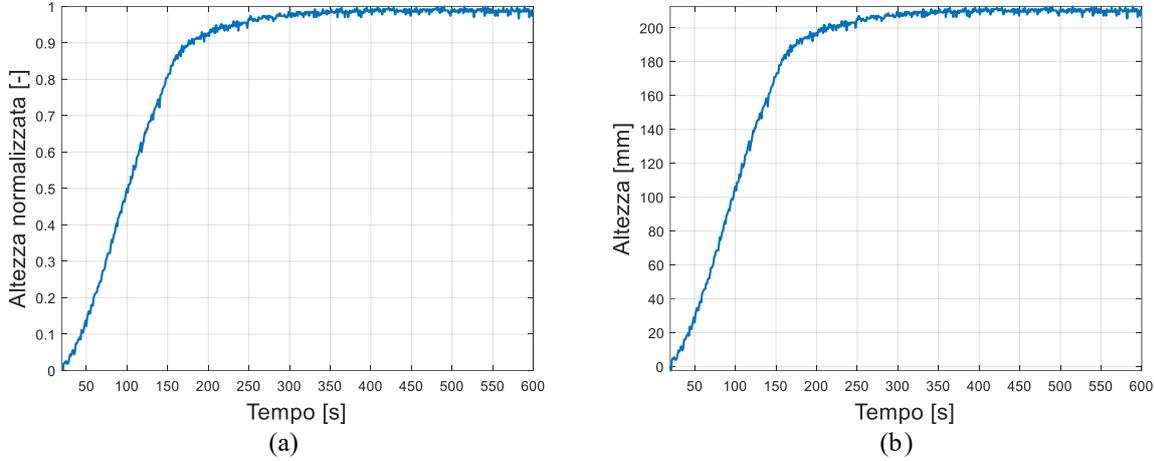
Una trasformazione che rende più esplicitiva la misura è la conversione dei profili temporali di distanza a profili temporali di altezza del poliuretano espanso mediante l'equazione 3.1:

$$H_{t_i} [mm] = D_{t_o} - D_{t_i} . \tag{3.1}$$

Una ulteriore trasformazione prevede di normalizzare l'altezza: essa, infatti, dipende dalla densità della schiuma finale, in quanto schiume meno dense, saranno generalmente più alte di altre la cui densità è maggiore, e quindi dalla massa di formulato che viene testata. Nonostante, come descritto nel Sotto-sottoparagrafo 2.1.3.2, si cerchi di mantenere costante la densità a  $30 \text{ kg/m}^3$ , è considerata accettabile la variabilità tra  $28$  e  $32 \text{ kg/m}^3$ . La normalizzazione è utilizzata per rimuovere l'effetto dovuto alla massa di formulato impiegata e concentrare l'interesse nella dinamica effettiva dell'espansione:

$$HN_{t_i} [-] = \frac{H_{t_i}}{\max (H_{t_i})} . \tag{3.2}$$

Si parla quindi di altezza normalizzata; l'operazione di normalizzazione ed il profilo di altezza sono visibili in Figura 3.3.



**Figura 3.3.** (a) Profilo di altezza ricavato dalle misure di distanza; (b) trasformazione a profilo di altezza normalizzata.

È poi possibile ricavare altre variabili importanti ai fini dell'analisi:

$$v_{HN_{t_i}} [s^{-1}] = \text{velocità di crescita} = \frac{dHN}{dt} \cong \frac{HN_{t_i} - HN_{t_{i-1}}}{t_i - t_{i-1}}, \quad (3.3)$$

$$v_{T_{t_i}} [^{\circ}C \cdot s^{-1}] = \text{gradiente di temperatura} = \frac{dT}{dt} \cong \frac{T_{t_i} - T_{t_{i-1}}}{t_i - t_{i-1}}, \quad (3.4)$$

$$mv_{HN_{t_j}} [s^{-1}] = \text{velocità di crescita media} = \frac{\sum_i^{j=i+5} v_{HN_{t_i}}}{5}, \quad (3.5)$$

$$mv_{T_{t_j}} [^{\circ}C \cdot s^{-1}] = \text{gradiente medio di temperatura} = \frac{\sum_i^{j=i+5} v_{T_{t_i}}}{5}. \quad (3.6)$$

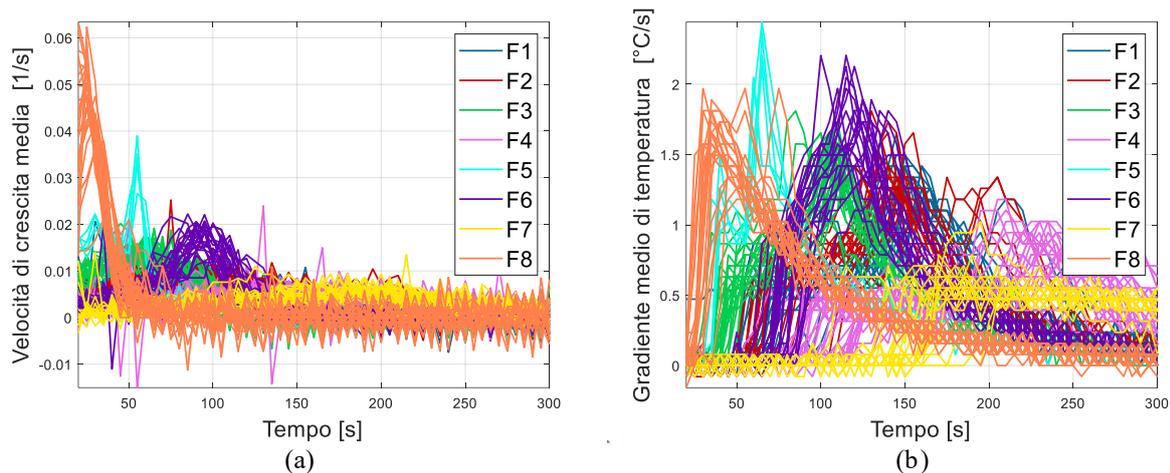
In particolare, la velocità di crescita media ed il gradiente medio di temperatura, sono le media di valori di gradienti di altezza e temperatura, eseguita ogni 5 s, e sono importanti variabili ai fini di analizzare la cinetica di espansione della schiuma. Esse, rivelano molti dettagli sull'espansione dei poliuretani e permettono un confronto più accurato fra i formulati diversi. Per questo motivo, sono state selezionate per la costruzione dei modelli attraverso analisi PCA.

L'analisi verrà applicata alla velocità di crescita media ed al gradiente medio di temperatura in quanto si è osservato che i profili temporali di velocità di crescita media e gradiente medio di temperatura forniscono indicazioni più accurate per descrivere la cinetica di un sistema poliuretanic. Una seconda scelta, fatta ai fini di migliorare l'analisi, è lo studio limitato ai primi 300 istanti temporali,

in quanto successivamente ad essi è possibile notare una stabilizzazione dei profili ad un valore costante. Le velocità sono intrinsecamente molto rumorose, vengono quindi filtrate attraverso una media mobile.

### 3.2.1.3 Visualizzazione dei profili

In Figura 3.4, sono riportati i profili temporali di velocità di crescita media e gradiente medio di temperatura di tutte le osservazioni degli otto formulati presi come riferimento.

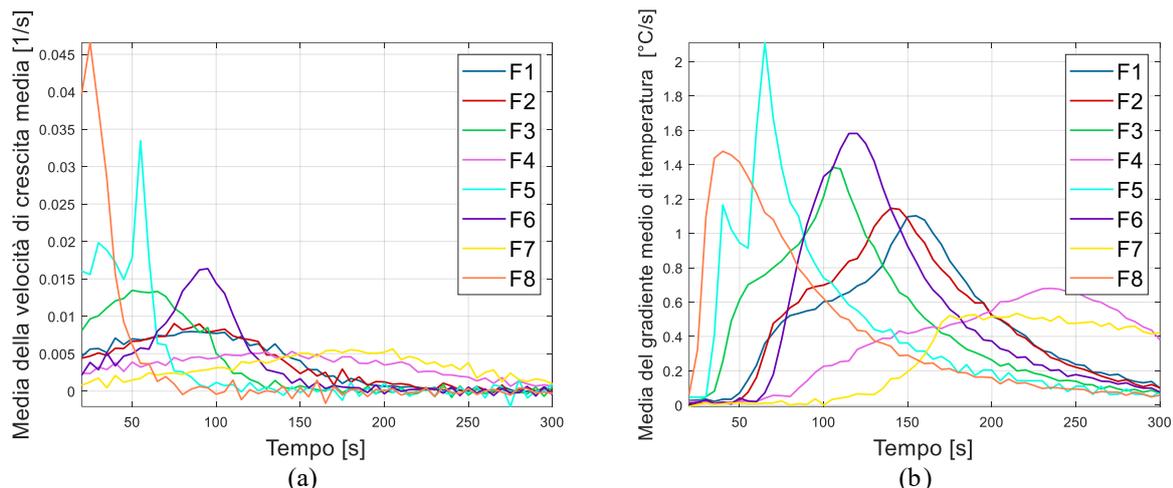


**Figura 3.4.** Profili temporali delle osservazioni degli otto formulati di riferimento: (a) profili di velocità di crescita media; (b) profili di gradiente medio di temperatura.

È possibile osservare che:

- I profili sono molto rumorosi; questo è un risultato atteso, considerando che derivano da profili di altezza normalizzata che presentano molte oscillazioni di misura, nonostante il filtro a media mobile e la rimozione degli *outlier*. Tuttavia, il rumore residuo può essere trattato mediante PCA;
- quasi tutte le osservazioni di uno stesso formulato seguono profili simili; alcune invece, si comportano in maniera sostanzialmente diversa fra loro, probabilmente a causa di piccole disomogeneità sulla formulazione oppure di errori di analisi.

Riportando invece la media di tutte le osservazioni di uno stesso formulato di riferimento, Figura 3.5, è possibile osservare il comportamento medio del formulato stesso, in modo da poter visualizzare i profili in modo più chiaro.

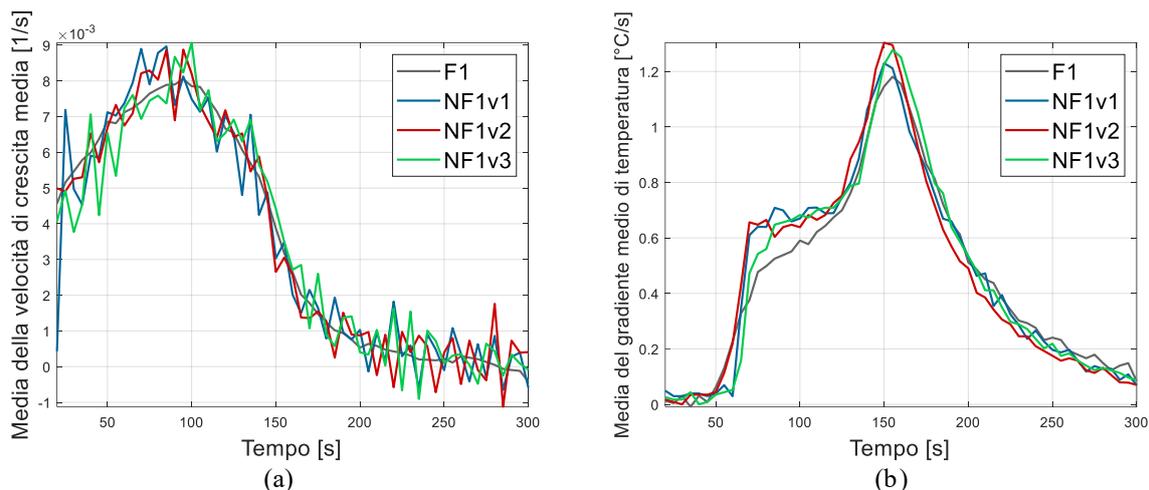


**Figura 3.5.** Profili temporali medi degli otto formulati di riferimento di: (a) velocità di crescita; (b) gradiente di temperatura.

È possibile individuare i formulati che si comportano in modo simile, come ad esempio F1 ed F2, o i formulati con una cinetica più veloce o più lenta. In particolare, F8 ed F5 sembrano essere molto veloci, invece F4 ed F6 sono molto lenti. Da di gradiente medio di temperatura, è possibile inoltre osservare i due contributi termici delle reazioni chimiche: spesso, sono presenti due scalini: il primo è dovuto ad alla reazione di *blow*, il secondo invece è causato dalla reazione di polimerizzazione. La velocità presenta, invece, un solo picco: la reazione di *blow* è il contributo fondamentale alla crescita del polimero, infatti, il massimo di altezza è raggiunto in prossimità del primo scalino della curva di variazione di temperatura.

Nonostante possano essere ricavate numerose informazioni analizzando i profili di velocità e gradiente di temperatura, non è possibile definire in maniera oggettiva il comportamento di ogni formulato, e non è facile capirne le correlazioni in modo esaustivo e completo. Inoltre, questa non solo è una rappresentazione della media delle osservazioni dei formulati, ma non comprende tutti i 53 formulati presenti nell'intero dataset.

Sono riportate anche le medie dei profili delle nuove formulazioni NF1 che dovrebbero sostituire F1, comparate con lo standard, in Figura 3.6.



**Figura 3.6.** Comportamento medio dei nuovi formulati NF1 rispetto allo standard desiderato F1; (a) profili di velocità di crescita media; (b) profili di gradiente medio di temperatura.

Considerando che, allo stato dell'arte, l'analisi si ferma a questo step, si può capire da Figura 3.6 che la scelta del migliore formulato adatto a sostituire F1 attraverso questa metodologia risulta molto complessa; sembra infatti che i nuovi profili replichino bene i profili standard e che le differenze siano minime. Tuttavia, una lieve differenza nei profili di crescita può determinare grandi differenze in termini di riempimento degli stampi e la determinazione puramente visiva della differenza fra i profili non è esaustiva e utilizzabile come criterio per la valutazione della qualità di un formulato.

### 3.2.2 Analisi esplorativa

L'obiettivo è lo studio della cinetica di espansione di una schiuma poliuretanicca e il confronto preliminare tra tutti i diversi formulati, in modo da ottenere una informazione generale sui prodotti di riferimento dell'azienda. È poi possibile utilizzare la prima parte della metodologia per ottenere un riscontro preliminare in merito al comportamento delle nuove versioni di F1 rispetto allo standard.

#### 3.2.2.1 Costruzione del modello globale PCA

Come discusso nel Sotto-sottoparagrafo 1.1.3.3, per costruire il modello è necessario selezionare il numero ottimale di PC. Questo, viene fatto grazie a *Scree Test*, *Eigenvalue Test* e *RMSECV Test*. I risultati di questi test, per quanto riguarda questa particolare analisi, sono riportati graficamente di seguito in Figura 3.7.

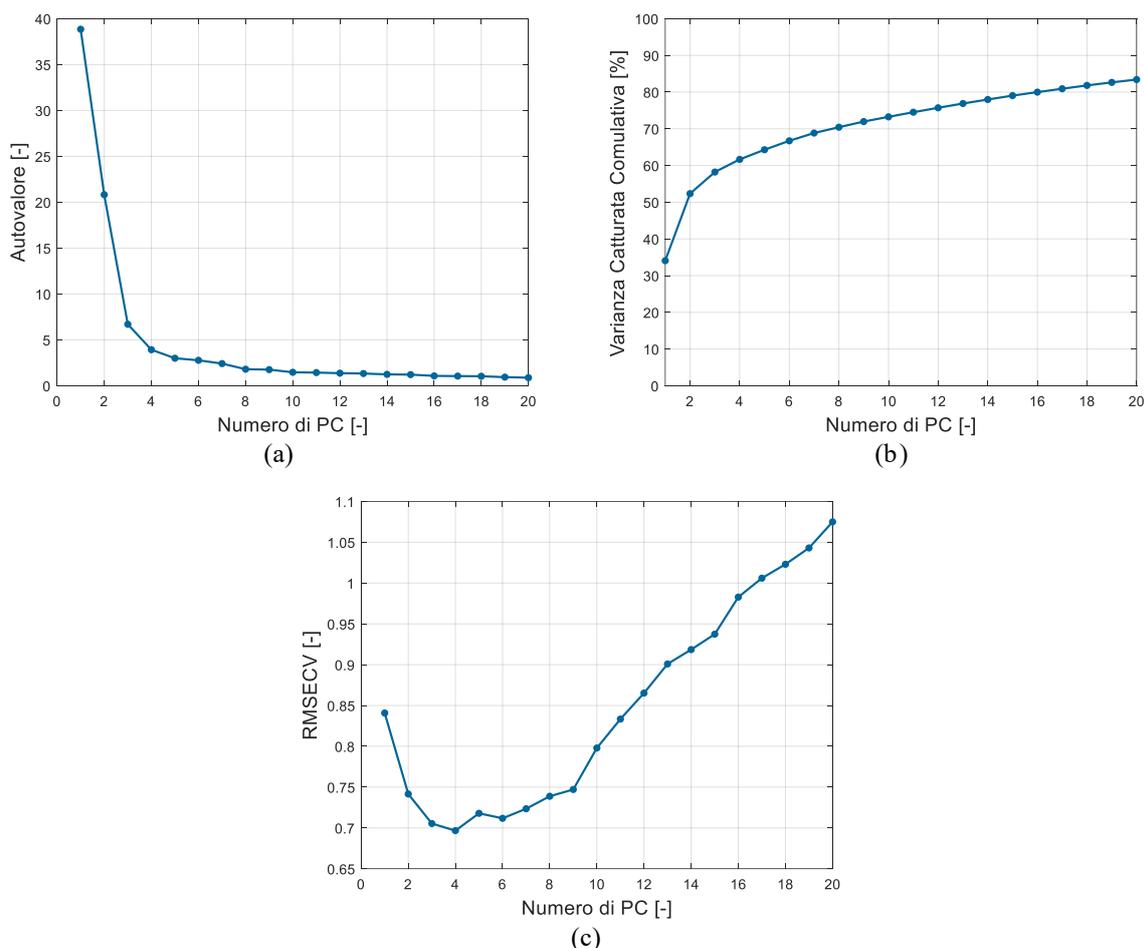
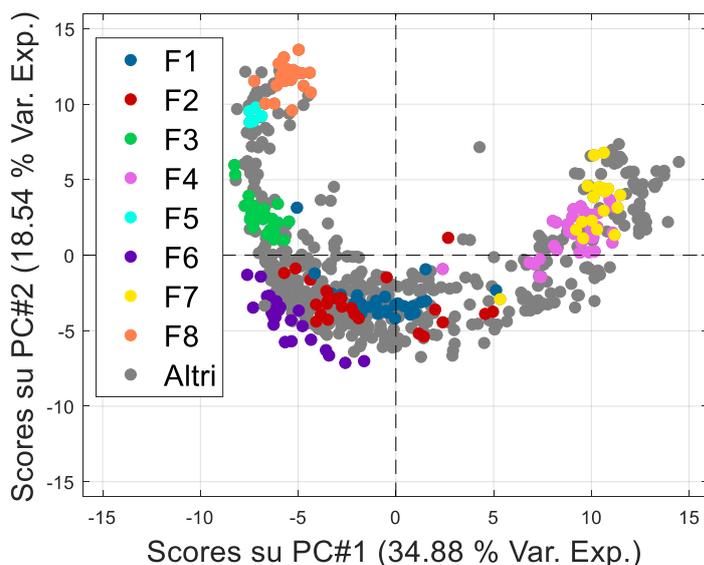


Figura 3.7. (a) Scree Test; (b) Eigenvalue Test; (c) RMSECV Test. I primi due test suggeriscono un numero elevato di PC, tuttavia, analizzando anche RMSECV, il minimo errore di approssimazione si ottiene con quattro PC.

Il numero adatto di PC scelto per costruire il modello è 4. Questa decisione è stata determinata principalmente dal RMSECV: l'errore di approssimazione minimo si ottiene con quattro PC. Inoltre, scegliere un numero di PC più alto potrebbe portare ad *overfitting*.

### 3.2.2.2 Risultati dell'analisi esplorativa

Con il modello PCA si studiano le relazioni fra osservazioni e variabili mediante *score* e *loading*. In Figura 3.8, ogni punto rappresenta lo *score* di una osservazione proiettata nel modello costruito con tutti i formulati.

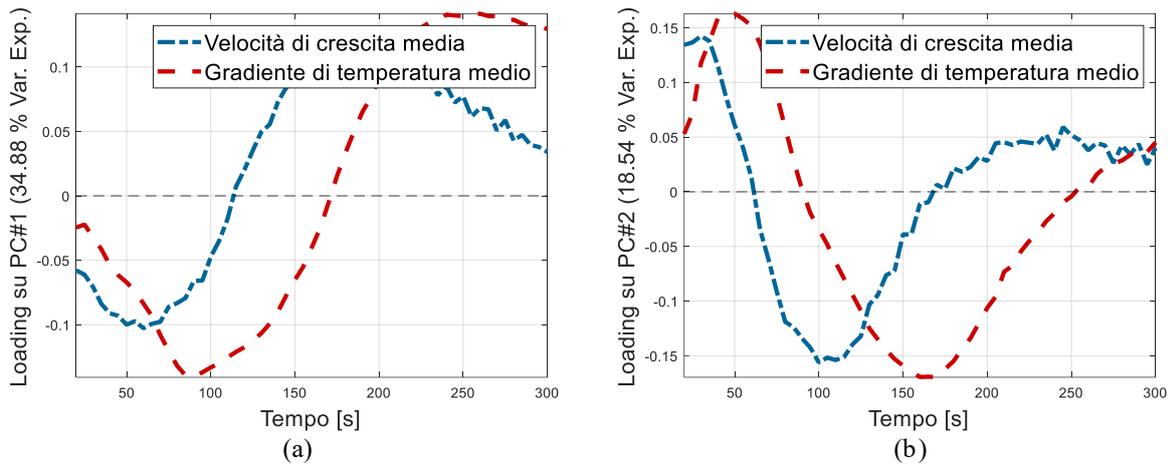


**Figura 3.8.** Score plot del modello PCA costruito con tutti i formulati. Ogni punto rappresenta la proiezione di un'osservazione nel nuovo spazio virtuale della PC. Sono riportate solamente le proiezioni sulle prime due PC, in quanto sono le più rappresentative del modello.

Analizzando lo *score plot* di Figura 3.8 si possono ricavare diverse informazioni:

- la varianza spiegata da ogni PC è piuttosto elevata per entrambe PC1 e PC2 a significare che il modello PCA globale rappresenta bene il comportamento delle schiume in espansione;
- le osservazioni appartenenti ad uno stesso formulato si comportano allo stesso modo, in quanto, le loro proiezioni sono vicine fra loro. Il modello riconosce la differenza fra tutti i diversi formulati e la differenza fra osservazioni diverse dello stesso formulato o variazioni leggere di un composto. Questo verrà sfruttato per lo sviluppo di metodologie di classificazione dei prodotti;
- la distribuzione degli *score* non è casuale, ma essi si distribuiscono secondo una particolare forma ad arco. Ciò significa che entrambe le PC rappresentano dei fenomeni importanti che avvengono in fase di espansione della schiuma;
- i formulati, che a primo impatto erano stati definiti lenti, F4 ed F7, si trovano a destra, cioè hanno valori positivi lungo PC#1. I formulati invece veloci, F5 ed F8, presentano valori negativi lungo PC#1, ma molto positivi lungo PC#2. I formulati con un comportamento medio si trovano vicino all'origine dello spazio delle LV, tuttavia, sono leggermente spostati verso il III quadrante.

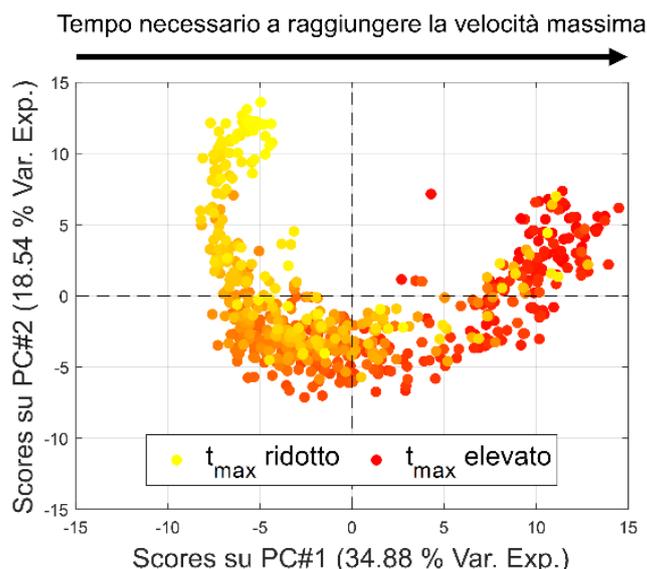
Dallo studio dei profili dei loading, è possibile andare più in dettaglio in merito al significato delle PC.



**Figura 3.9.** Loading plot: rappresentazione dei profili dei loading delle due variabili studiate. (a) i profili dei loading su PC#1; (b) i profili dei loading su PC#2.

In Figura 3.9 si osserva una consistente similitudine fra i *loading* di velocità di crescita media e gradiente termico di temperatura. Questo significa che le due variabili sono estremamente correlate: se la velocità aumenta, dopo circa cinquanta secondi, aumenta anche il gradiente medio di temperatura. Questo effetto ritardato potrebbe essere dovuto alla conduzione del calore nella schiuma in espansione: la reazione di *blow* è la responsabile della crescita della schiuma, e produce anche molto calore. Perciò inizialmente la schiuma aumenta di volume a seguito della produzione di anidride carbonica, poi, man mano che il calore generato raggiunge la termocoppia, anche la temperatura misurata aumenta. Il ritardo della variazione di temperatura è dovuto alla caratteristica isolante che il materiale sta assumendo man mano che cresce e le celle si formano, come è possibile confermare anche considerando l'aumento progressivo della differenza fra le due curve nel tempo. Interessante è l'osservazione dei *loading plot* assieme allo *score plot*: tutte le osservazioni la cui proiezione sulla prima PC è negativa, sviluppano il massimo di velocità entro i primi 100 secondi. Questo significa che proiezioni la cui PC1 è negativa presentano un massimo nella prima parte del profilo. Un ragionamento opposto può essere fatto analizzando i profili delle velocità delle osservazioni il cui *score* nella prima PC è positivo. Possiamo dire quindi che la prima PC descrive il momento nel quale una osservazione raggiunge la velocità massima, come si può notare in Figura 3.10.

$$t_{max,j} [s] = \text{tempo al quale l'osservazione } j \text{ ha velocità massima} \quad (3.7)$$



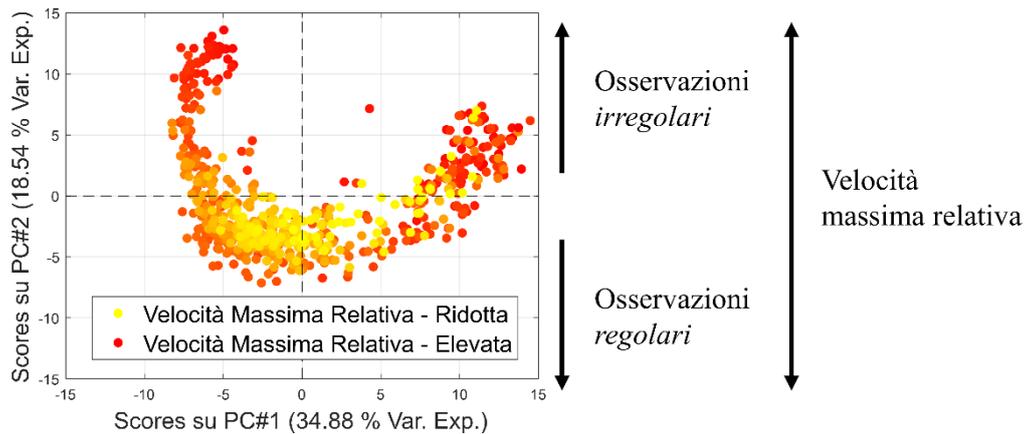
**Figura 3.10.** Score plot in cui ogni score è colorato in base al tempo necessario a raggiungere il massimo di velocità, dal minore, giallo, al maggiore, rosso. È possibile notare come la variazione di colore si sposti dal giallo al rosso, seguendo esattamente il profilo della PC#1.

Il profilo del *loading* della PC2 mostra che valori positivi rappresentano tutte quelle osservazioni il cui comportamento è anomalo, troppo presto o troppo tardi. Osservando infatti i profili originali di velocità di temperatura e velocità di crescita, è possibile ottenere una importante differenziazione:

- i formulati normali sono quei formulati la cui temperatura aumenta più lentamente rispetto all'altezza;
- i formulati molto veloci, presentano un aumento di temperatura quasi contemporaneo alla velocità di crescita, poiché la conduzione di calore nei primi istanti è molto più semplice in quanto le celle non sono ancora completamente formate, perciò, l'aumento di temperatura non è ostacolato;
- per i formulati molto lenti lo step determinante non è la conduzione di temperatura, ma la crescita, perciò, il divario fra le due variabili non è così elevato poiché in questo modo i due profili si sviluppano su scale temporali comparabili.

Inoltre, viene osservato che man mano che ogni proiezione si allontana dalla media degli *score* lungo la PC2, aumenta la distanza tra il massimo di velocità di un formulato e la media delle velocità di tutti gli altri formulati, calcolata nell'istante temporale corrispondente al massimo. È stata quindi introdotta una nuova variabile:

$$velocità\ massima\ relativa\ osservazione\ j = \frac{\max(v_{HNj})}{media(v_{HNt_{max,j}})} \tag{3.8}$$



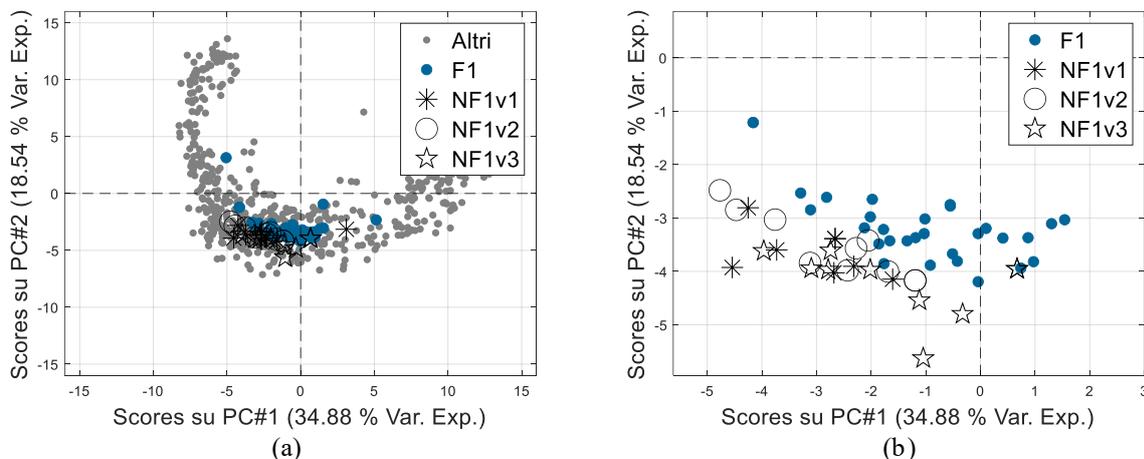
**Figura 3.11.** Score plot in cui ogni score è colorato in base al suo valore di velocità di crescita massima relativa, si va dal giallo, dove i valori sono più bassi, al rosso, dove i valori sono maggiori, man mano che ci si allontana dal centro.

In figura 3.11 si può avere la conferma di quanto discusso finora, in quanto gli *score* sono stati colorati in base al valore di *velocità massima relativa*.

Riassumendo, possiamo dire che:

- PC1: rappresenta il tempo necessario a raggiungere il massimo di velocità; è quindi una indicazione della scala temporale del processo di espansione per ogni formulato.
- PC2: essa indica se il tempo al quale il formulato raggiunge il massimo di velocità di crescita è *regolare* o *irregolare*; Il valore assoluto dello score rappresenta quanto un formulato al momento del massimo avrà una velocità maggiore rispetto agli altri.
- i formulati si muovono in direzione circolare in senso antiorario lungo lo *score plot*, dal più lento al più veloce.

Le informazioni ricavate attraverso l'analisi esplorativa possono essere utilizzate anche per una valutazione del comportamento delle nuove versioni di F1.

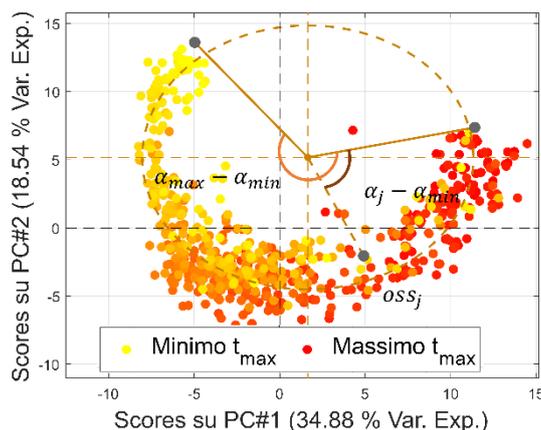


**Figura 3.12.** Score plot delle osservazioni delle nuove versioni di F1 proiettate nel modello globale; (a) score plot con tutti i formulati di calibrazione e convalida; (b) score plot del modello globale concentrato nella zona prevalente di F1 e NF1.

In Figura 3.12 si può notare che, proiettando le osservazioni di convalida nello *score plot*, il modello riconosce l'appartenenza dei nuovi formulati alle classi che essi cercano di imitare. Le nuove versioni di F1 sono leggermente spostate verso *score* sulla PC2 più negativi. Questo porta a concludere che le nuove formulazioni raggiungono il massimo in un tempo comparabile con F1, in quanto sono posizionati allo stesso modo lungo la prima componente principale, ma la loro velocità massima relativa sarà leggermente più elevata. Osservando i profili medi delle osservazioni delle diverse versioni, Figura 3.6, questo viene confermato, e va a convalidare inoltre l'interpretazione della seconda componente principale del modello generale. Solamente attraverso una analisi esplorativa preliminare è possibile ottenere una valutazione molto più accurata dei formulati F1.

### 3.2.2.3 Caratterizzazione della cinetica dei formulati

L'arco in cui le osservazioni si dispongono nel diagramma degli *score* può essere fittato<sup>4</sup> per trovare il centro di questa particolare forma geometrica; in questo modo l'angolo tra ogni *score* e l'asse orizzontale del cerchio che fitta l'arco delle osservazioni (Figura 3.13) può essere legato ad un parametro di velocità di crescita delle schiume.



**Figura 3. 13.** *Fitting delle osservazioni nel diagramma degli score: l'angolo secondo cui si dispone una osservazione determina un parametro di velocità di crescita della schiuma.*

A questo punto viene definito come *parametro di velocità*:

$$PV_j [-] = \left[ \frac{(\alpha_j - \alpha_{min})}{(\alpha_{max} - \alpha_{min})} \right] \cdot 100, \tag{3.9}$$

$\alpha$  [°] = *angolo del formulato rispetto al centro.*

Questo parametro permette quindi di valutare in modo semplice ed oggettivo la velocità di crescita di un formulato in relazione a tutti gli altri. Se il parametro di velocità è elevato per una osservazione *j*,

<sup>4</sup> Circle fit. (1987). V. Pratt. "Direct least-squares fitting of algebraic surfaces", Computer Graphics, Vol. 21, pages 145-152

allora il formulato avrà una cinetica veloce. Al contrario, se  $PS_j$  possiede valori limitati sarà caratterizzato da una cinetica lenta.

Ai fini di una valutazione preliminare della cinetica delle nuove versioni di F1 viene calcolato il parametro di velocità dei formulati NF1 e confrontato con il valore medio dei formulati standard.

**Tabella 3.6.** Risultati del calcolo del parametro di scorrimento medio per ogni nuova versione, comparata con il parametro di scorrimento standard.

Nuove formulazioni	F1	NF1v1	NF1v2	NF1v3
Parametro di velocità medio [-]	53.61	55.92	57.44	54.73

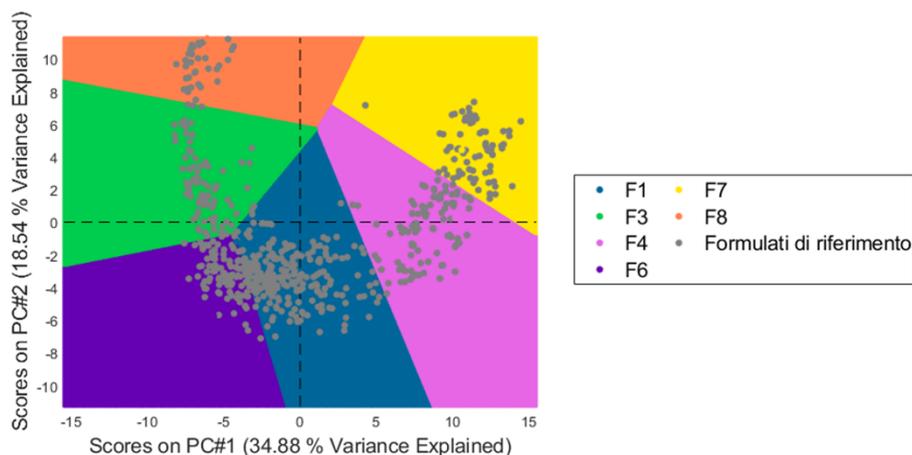
Le nuove versioni NF1 tendono ad essere più veloci rispetto alla versione standard F1, tuttavia, la terza versione si avvicina di più al valore richiesto.

### 3.2.2.3 Utilizzo dell'analisi esplorativa per la classificazione dei formulati

Come illustrato nel Sotto-sottoparagrafo precedente, tutti i formulati di una stessa classe, hanno uno *score* simile: lo *score plot* può essere quindi suddiviso in alcune zone, ognuna delle quali corrispondente ad un diverso formulato con delle caratteristiche cinetiche peculiari. Così facendo, se ci fosse bisogno di classificare una nuova formulazione, di cercarne una che si comporti allo stesso modo di una nota, o capire quali azioni siano necessarie per spostare una formulazione da un certo tipo di caratteristiche della schiuma a caratteristiche differenti, basterebbe analizzare il nuovo profilo riducendo la dimensionalità del problema attraverso PCA e verificare in quale zona dello spazio degli *score* esso si posizioni. La classificazione delle diverse zone dello spazio degli *score* che differenziano le classi di prodotto viene fatta mediante *Linear Discriminant Analysis*.

La tecnica per dividere lo *score plot* in zone è stata applicata agli *score* di sei fra gli otto formulati presi come riferimento e analizzati nei precedenti paragrafi; non tutti i formulati, quindi, sono stati utilizzati per costruire la mappa nello *score plot*: in particolare, F5 possiede troppe poche repliche di dati per rappresentare una intera zona, mentre F1 ed F2 sono molto simili fra loro, viene quindi scelto come riferimento F1, in quanto è un formulato più utilizzato e prodotto.

In Figura 3.14 è rappresentato il risultato: il diagramma degli *score* viene diviso in sei zone, ogni osservazione può quindi essere ricondotta ad uno dei sei formulati di riferimento con cui sono state ricavate le zone. Chiaramente, man mano che ci si allontana dal centro del cluster di osservazioni dei formulati con cui le zone sono state ricavate, l'accuratezza di classificazione diminuisce.



**Figura 3.14.** Divisione in classi delle formulazioni mediante Linear Discriminant Analysis nello spazio degli score di una PCA.

L'accuratezza del modello di classificazione può essere determinata analizzando l'errore di assegnazione:

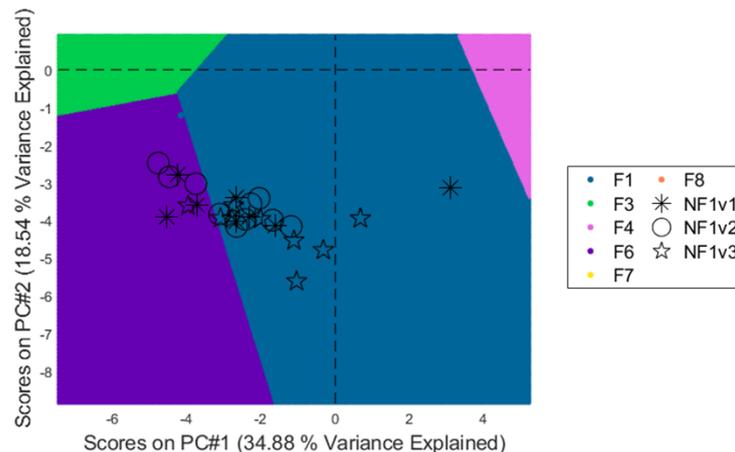
$$E_{ass} [-] = \frac{\text{Numero di assegnazioni errate}}{\text{Numero totale di osservazioni}} \cdot 100 = 10.43 \% \quad (3.10)$$

Come si può notare, l'errore di assegnazione è relativamente basso e principalmente dovuto alla similitudine tra le caratteristiche delle cinetiche dei formulati F4 ed F7, il cui comportamento è molto simile. Le altre classi vengono assegnate in maniera molto più precisa.

Questo, quindi, risulta essere uno strumento preliminare ed oggettivo di analisi, basato su una mappa nello spazio degli *score* di una PCA sulle variabili di qualità del prodotto in termini dei profili temporali delle variabili che identificano la cinetica di espansione.

In Figura 3.15 sono riportati i risultati di classificazione dei nuovi formulati NF1, ed è possibile osservare come alcune formulazioni di NF1, sono più veloci dello stesso F1 e si posizionano nella zona appartenente a F6.

Questo, è un risultato inatteso, tuttavia, dimostra come una classificazione preliminare di alcune formulazioni permette anticipatamente l'eliminazione di uno o più nuovi formulati tra i candidati adatti a sostituire quelli standard.



**Figura 3.15.** Risultato della classificazione di nuovi formulati proiettati nello spazio degli score della PCA (ingrandimento sullo standard della classe che identifica il formulato F1).

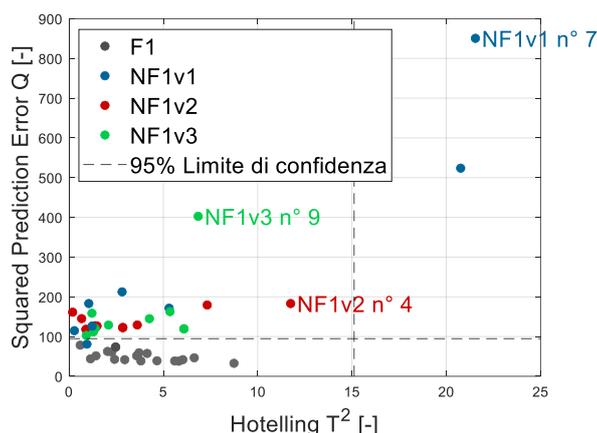
È necessario però considerare che le nuove formulazioni appartenenti ad una regione diversa da quella del formulato che si vuole sostituire, sono tutte molto vicine al confine fra le due, dove la divisione tende ad essere meno marcata.

### 3.2.3 Analisi approfondita di un singolo formulato

Con le medesime tecniche di modellizzazione è possibile costruire modelli specifici utilizzando le osservazioni di un'unica classe di formulati. L'obiettivo è capire più approfonditamente se le nuove versioni di F1 sono adatte alla sua sostituzione di questo formulato e quale di queste sia la più indicata. La procedura di costruzione dei modelli calibrati con i singoli formulati sono sostanzialmente le medesime applicate al modello globale. Per questo motivo non verranno nuovamente discusse. Il numero di PC scelto per questi due modelli è mantenuto pari a quattro.

Si ottiene un riscontro in merito alla differenza dei nuovi formulati NF1 con quello F1 di riferimento, proiettandoli nel modello PCA calibrato sul riferimento F1. In particolare, analizzando gli indici statistici  $T^2$  ed  $SPE$  è possibile capire quale fra le nuove formulazioni si avvicini in misura maggiore al profilo ricercato, considerando gli interi profili temporali di velocità di crescita media e gradiente medio di temperatura.

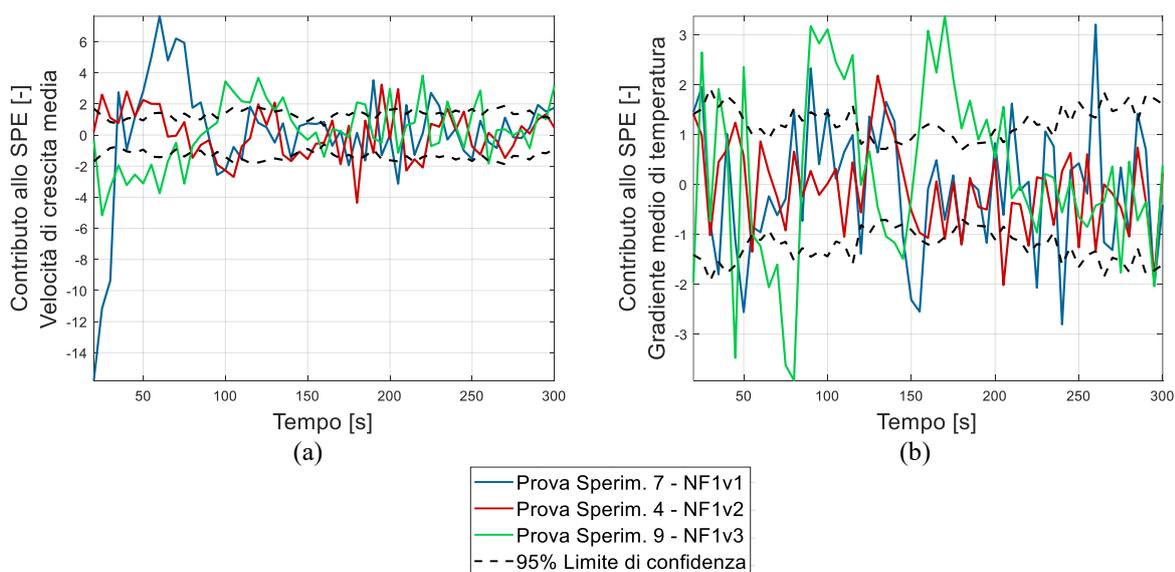
Innanzitutto, si nota da Figura 3.16, che quasi tutti i nuovi formulati sono all'interno del limite di confidenza di  $T^2$ : essi hanno dei valori molto simili ai profili di F1 per la media dei profili temporali.



**Figura 3.16.** Analisi degli indici statistici delle nuove formulazioni proiettate nel modello specifico costruito con il formulato standard F1. Sono indicate le prove anomale, che verranno analizzate in dettaglio di seguito.

Solamente due osservazioni della prima versione escono sia dai limiti di  $T^2$  che di  $SPE$ : i profili di velocità crescita media e di gradiente medio di temperatura presentano dei valori ed una struttura di correlazione anomala. Tutte le altre osservazioni sono invece posizione vicino al limite di confidenza di  $SPE$ ; ciò significa che imitano in modo soddisfacente il comportamento dei profili del formulato standard, ma è sempre presente una differenza non irrisoria nella forma del profilo. Per andare più nel dettaglio è necessario osservare Figura 3.17.

Per capire però a cosa sono dovute le differenze, si possono osservare infatti i contributi di ciascuna variabile nel tempo sullo  $SPE$ . Particolarmente interessanti sono le analisi dei contributi delle osservazioni che escono di molto dal limite di confidenza di  $Q$ : la settima prova sperimentale della prima versione, la quarta della seconda versione e la nona della terza versione (Figura 3.16).



**Figura 3.17.** Contributi allo  $SPE$  rispetto al formulato standard di: (a) velocità di crescita media e (b) gradiente medio di temperatura.

Si possono notare da Figura 3.17 diverse deviazioni: la prima prova sperimentale di NF1, all’inizio cresce lentamente rispetto allo standard, ed esce dai limiti di confidenza inferiori, per poi accelerare molto attorno ai 75 s. La quarta prova sperimentale della seconda formulazione presenta una velocità leggermente maggiore alla media di F1 nei primi istanti mentre successivamente è di poco inferiore. La prova sperimentale 9 della terza versione invece non solo esce dal limite di confidenza della velocità di crescita media ma anche del gradiente medio di temperatura: il formulato risulta troppo “lento” fino al centesimo istante temporale, ma troppo “rapido” attorno ai 120 s. Per quanto riguarda il gradiente termico, non sono presenti anomalie molto rilevanti; tuttavia, i profili sembrano oscillare molto attorno ai limiti di confidenza, inoltre, l’osservazione 9 della terza versione esce più volte dai limiti, questo però è dovuto nuovamente alla velocità di espansione ridotta, in quanto anche il profilo di temperatura si sviluppa con ritardo rispetto allo standard.

### 3.2.4 Analisi comparativa dei formulati

Per poter confrontare rapidamente due diversi formulati, si procede con la creazione di modelli PCA calibrati utilizzando i profili temporali di velocità di crescita media e gradiente medio di temperatura di ogni singolo formulato, per poi confrontare i *loading* come descritto nel Sotto-sottoparagrafo 1.1.3.7.

Innanzitutto, l’analisi è stata applicata agli otto formulati presi come riferimento; i risultati sono riportati graficamente in Figura 3.18.

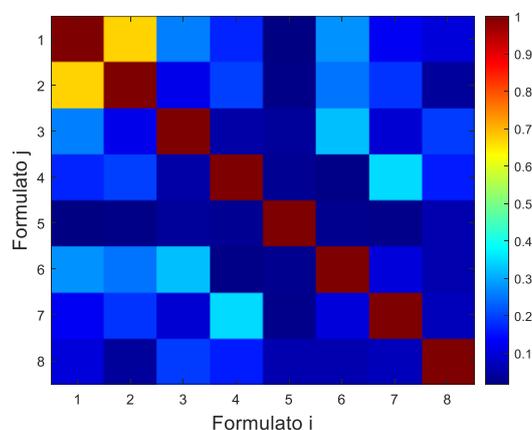
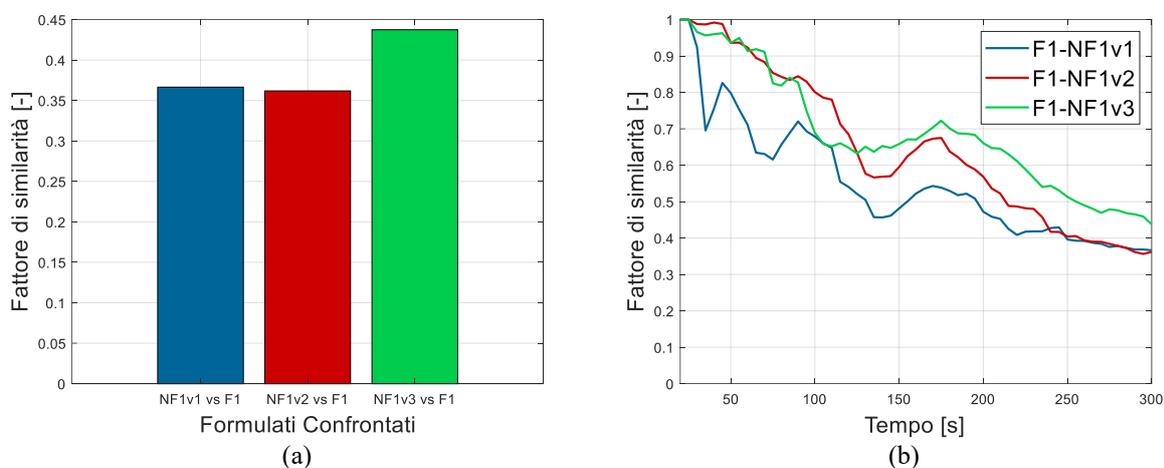


Figura 3.18. Mappa di calore dei fattori di similarità degli otto formulati F1-F8.

Come si può vedere, l’analisi riconosce molto bene i formulati che fra loro sono diversi per le caratteristiche cinetiche dei profili temporali di velocità di crescita media e gradiente medio di temperatura. Gli unici formulati fra loro simili sono F1 ed F2. La differenza fra F4 ed F7 viene osservata in modo molto marcato, formulati il cui errore di classificazione attraverso *Linear Discriminant Analysis* era molto elevato. Si può quindi dire che questo strumento funziona molto bene e che sia un valido supporto alla classificazione dei formulati. Una cosa molto interessante che si può vedere dalla Figura 3.18 è che questa analisi tende ad essere uno strumento molto sensibile alle

variazioni di formulazione; infatti, la massima somiglianza tra due formulati diversi si ha tra due che risultavano molto vicini nello *score plot*, con un fattore di somiglianza attorno allo 0.7 su una scala da 0 (profili totalmente diversi) a 1 (profili uguali).

Inoltre, non solo è possibile applicare l'analisi attraverso gli  $S_{i,j}$  a tutti gli istanti temporali del profilo temporale, ma è anche possibile ottenere una informazione in merito ai singoli istanti di tempo e la storia passata della prova sperimentale. Si calcola in modo iterativo il valore di  $S_{i,j}$  all'istante temporale  $k$ , si costruiscono i modelli PCA su matrici in ingresso incomplete, ovvero contenenti solamente i valori di velocità di crescita media e gradiente medio di temperatura degli istanti da 1 a  $k$  per tutte le osservazioni. Questo si ripete per ogni istante  $t_k$ . I risultati sono presentati in Figura 3.19.



**Figura 3.19.** (a) confronto fra i fattori di similarità tra le nuove versioni di F1 e lo standard in tutto il dominio temporale; (b) andamento dei fattori di similarità nel tempo.

Innanzitutto, risalta subito che la terza versione di NF1 è la più simile a F1.

Nel tempo, Figura 3.19 (b), la somiglianza fra i nuovi formulati e lo standard ha una diminuzione lineare, questo è dovuto principalmente alla presenza del rumore che viene costantemente introdotto nel modello costruito. Quello che però è importante notare sono i picchi di diminuzione del fattore di similarità, in quanto è in quel momento che i nuovi formulati si vanno a differenziare dallo standard per motivi non legati al rumore introdotto. Vediamo quindi che la prima versione si differenzia da F1 in maniera considerevole per tutti i primi 150 s, mentre le altre due versioni sembrano avere solamente dei picchi tra 100s e 150s. La diminuzione della seconda versione aumenta di intensità anche nella seconda parte dell'analisi, suggerendo una diminuzione della velocità di crescita media diversa rispetto allo standard.

### 3.2.5 Conclusioni dell'applicazione dell'analisi

Ogni fase della nuova analisi ha permesso di ottenere risultati circa i comportamenti dei nuovi formulati. È importante notare che tutte le diverse fasi di analisi hanno permesso di ottenere risultati fra loro coerenti, rendendo altresì possibile pensare ad una riduzione i tempi di analisi, ottenendo allo

stesso tempo risultati più robusti. In Tabella 3.4 sono presentati schematicamente tutti i risultati ottenuti:

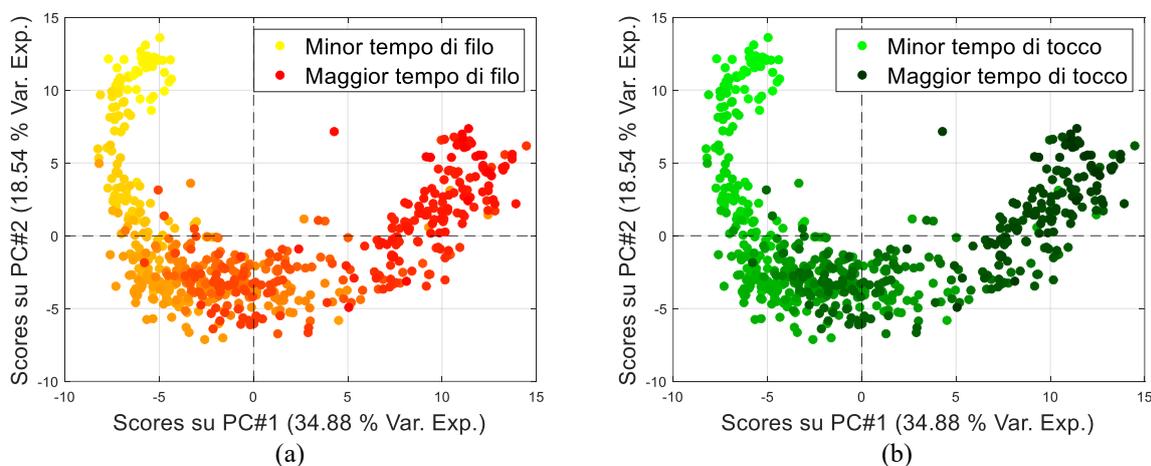
**Tabella 3.7.** Riassunto dei risultati dell'applicazione della nuova metodologia di analisi alla sostituzione del formulato F1.

Fase di analisi	Risultato
Esplorativa – analisi degli score	Tutte e tre le formulazioni presentano un massimo di velocità più elevato rispetto alla formulazione standard. Il tempo impiegato a raggiungere il massimo rimane tuttavia molto simile.
Esplorativa – parametro di velocità	Il parametro di scorrimento medio che più si avvicina allo standard appartiene alla terza versione del nuovo formulato.
Esplorativa – classificazione e monitoraggio	Le nuove formulazioni con una percentuale più elevata di assegnazione a F1 appartengono alla terza versione.
Approfondita – confronto con lo standard	Le variazioni sono dovute sia ad un comportamento anomalo dei formulati, sia ad anomalie dei valori delle variabili. La prima versione sembra essere la più anomala.
Comparativa – analisi delle similarità	La versione che più è riconducibile al formulato standard è la terza.

La versione scelta, a seguito dell'analisi, è la terza. Questo è un ottimo risultato in quanto l'analisi era stata parallelamente condotta dall'azienda attraverso le metodologie allo stato dell'arte ed ottenendo lo stesso risultato, confermando la validità del sistema di analisi introdotto.

### 3.3 Confronto tra la nuova metodologia e l'analisi allo stato dell'arte

In questa parte conclusiva del Capitolo 3 si confrontano la metodologia di analisi proposta l'attuale sistema di analisi industriale. Interpretando gli *score plot* secondo i valori dei tempi caratteristici (che attualmente rivestono il sistema di analisi allo stato dell'arte) è possibile osservare che i risultati ottenuti con la metodologia proposta sono pienamente concordi: la prima componente principale è un indicatore della scala temporale del processo e, allo stesso modo, dei tempi di gel e di tocco.

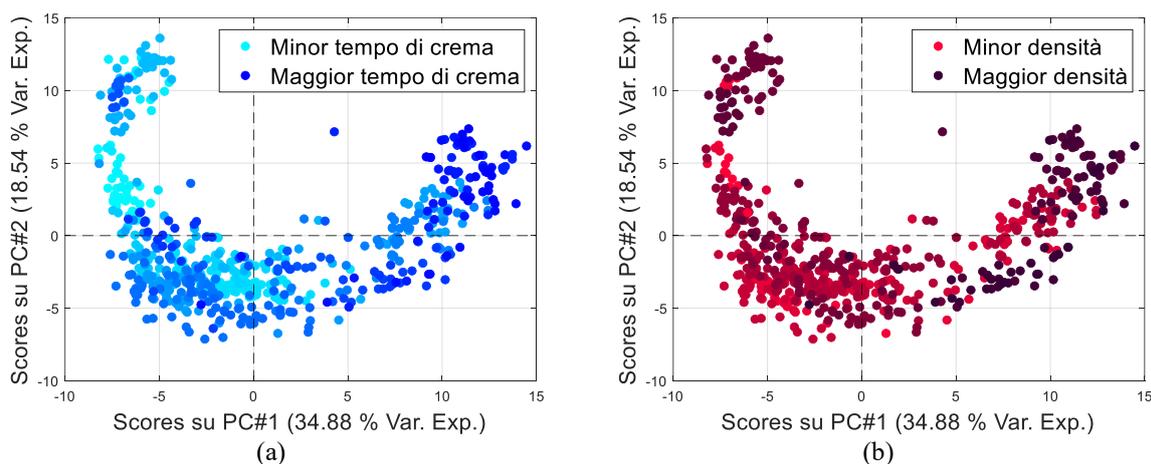


**Figura 3.20.** Diagramma degli score del modello PCA globale rappresentati in base a: (a) tempo di filo, (b) tempo di tocco.

Infatti, in Figura 3.20 (a) e 3.20 (b) è possibile vedere che il modello PCA globale è coerente con l'analisi dei tempi caratteristici: i formulati, colorati dal più veloce a più lento secondo i tempi di filo e di tocco, sono ordinati sulla PC1 come il tempo a cui il massimo del profilo di velocità viene raggiunto, come era stato rilevato nel Sotto-sottoparagrafo 3.2.2.2, quindi rappresenta la velocità generale del processo di espansione.

Invece, per quanto riguarda tempo di crema e densità, non si ottengono gli stessi riscontri perché il tempo di crema ha una misura meno accurata e ripetibile e non era atteso nemmeno che il modello generale costruito sui profili di velocità di crescita e gradiente medio di temperatura potesse in qualche modo essere collegato ai valori di densità, che sono determinati dal tipo e dalla quantità di espandente.

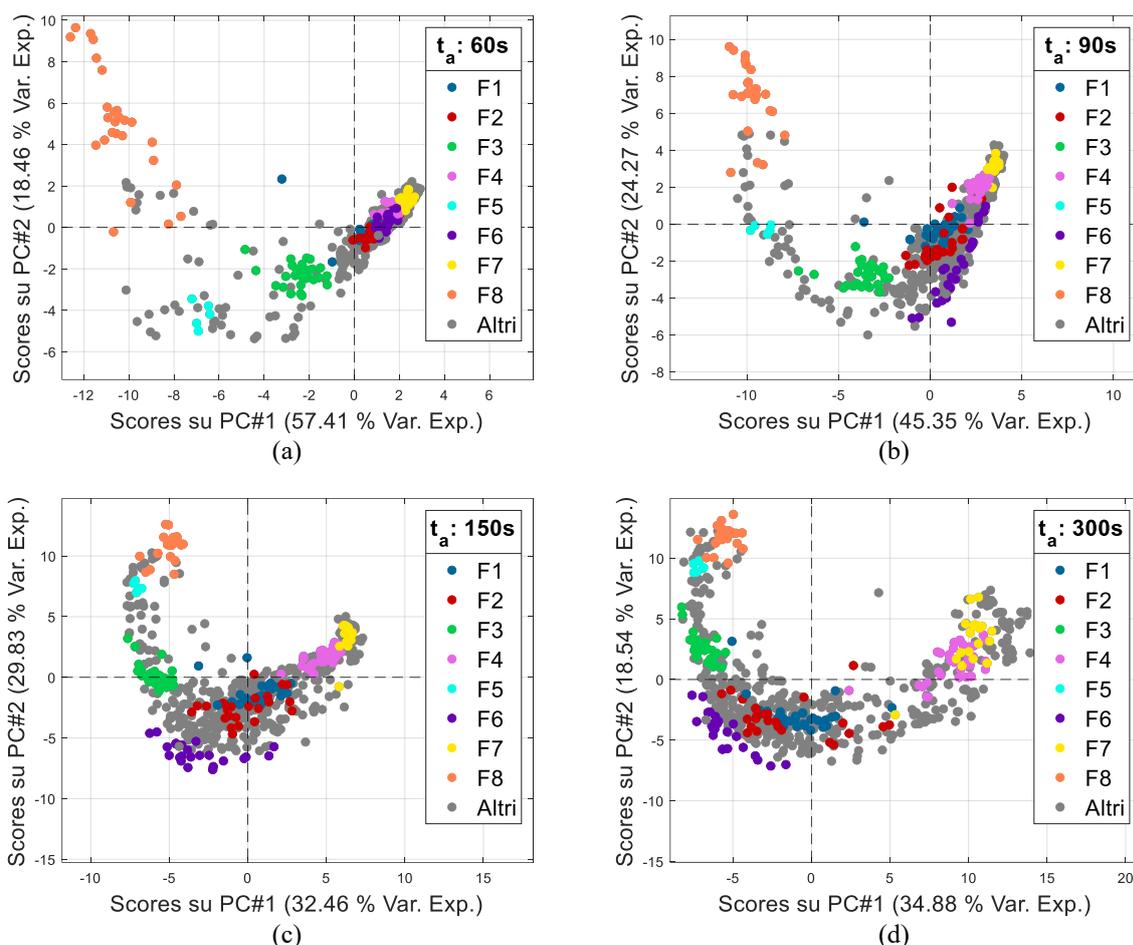
Inoltre, è possibile dire che l'effetto della densità è completamente rimosso dall'analisi dei profili in fase di normalizzazione dell'altezza. Non essendo la densità un parametro da controllare in modo stringente, e non essendo legato alla cinetica dell'espansione, non avrebbe senso attendersi una correlazione con i profili di temperatura e velocità. La dimostrazione è chiaramente visibile in Figura 3.21.



**Figura 3. 21.** Diagramma degli score per il modello PCA globale secondo:(a)tempo di crema; (b) densità.

### 3.3.1 Riduzione dei tempi di analisi

L'analisi dell'intera curva di crescita ha una durata complessiva di 615 s, circa 10 min . Come visto, l'analisi della velocità attraverso la nuova metodologia richiede solamente la metà del tempo, garantendo nonostante ciò una elevata accuratezza. È lecito quindi chiedersi se, la riduzione ulteriore del tempo di analisi, possa inficiare l'accuratezza del risultato. Per rispondere a questo quesito, sono stati costruiti diversi modelli generali, utilizzando i dati dei primi 60, 90, 150 e 300 s.



**Figura 3. 22.** Confronto fra i risultati dell'analisi PCA costruita considerando tempi di analisi crescenti (a) 60s, (b) 90s; (c) 150s; (d) 300s.

In Figura 3.22, sono riportati i diagrammi degli *score* al variare del tempo di analisi (cioè della lunghezza in termini di tempo dei profili utilizzati per la costruzione del modello PCA globale). Innanzitutto, occorre osservare che la forma dell'insieme degli *score* non viene del tutto distorta al variare del tempo di analisi, questo significa che il modello riconosce la differenza fra i formulati e li posiziona allo stesso modo. La correlazione fra le variabili e le osservazioni è quindi mantenuta anche con tempi di analisi più brevi. Analizzando però solamente i modelli calibrati su analisi di 60 s o 90 s, viene persa informazione riguardo la differenza fra i formulati F3 ed F5 lungo PC2. Lo stesso accade per i formulati F1, F2 ed F6. Allo stesso tempo però, i modelli costruiti su analisi brevi osservano una maggiore distinzione nella descrizione dei formulati veloci.

Quindi si conclude che l'analisi può essere ridotta ad 150 o 300 s ottenendo una diminuzione del tempo di analisi del 50-75%. Questo non solo renderebbe più veloce la formulazione di nuove miscele, ma permetterebbe al formulatore di ricavare più repliche di ogni profilo, rafforzando la qualità dell'analisi. Inoltre, permetterebbe un controllo dell'impianto di miscelazione dell'azienda molto più rapido ed efficiente.

### 3.4 Conclusioni e possibili sviluppi

Oltre ai vantaggi introdotti in termini di accuratezza di analisi, l'utilizzo di tecniche multivariate di analisi dei dati si è dimostrata efficace nell'approfondire lo studio della cinetica di espansione delle schiume poliuretatiche e ha permesso di proporre una sostanziale riduzione dei tempi di analisi. I risultati ottenuti sono riassunti di seguito:

- attraverso l'analisi esplorativa si ottiene una interpretazione accurata del fenomeno di espansione di un sistema poliuretatico ed è possibile analizzare, attraverso l'analisi degli scores, il parametro  $PV_j$  e i metodi di classificazione, sia formulati di riferimento che nuovi prodotti in modo completo ed oggettivo;
- l'analisi approfondita permette di concentrare lo studio dell'espansione su un singolo formulato e di determinare in modo molto preciso analogie e differenze tra i profili temporali di velocità di crescita media e gradiente medio di temperatura tra uno specifico formulato e nuove osservazioni;
- l'analisi comparativa, attraverso un singolo parametro,  $S_{i,j}$ , consente una veloce determinazione della somiglianza fra diverse tipologie di formulati; è possibile analizzare il profilo temporale di  $S_{i,j}$  e determinare a quale istante temporale i profili dei due formulati si differenziano;
- il nuovo sistema di analisi è coerente con le metodologie allo stato dell'arte;
- l'analisi attraverso la nuova metodologia permette una riduzione sensibile dei tempi di analisi, garantendo anche una maggiore tempestività di intervento in impianto.

Per migliorare il modello di analisi potrebbero essere utilizzati una serie di altri dati di, come spettri IR del prodotto, profili temporali di viscosità e del grado di polimerizzazione.

# Capitolo 4

## Sviluppo di un modello predittivo per la formulazione di sistemi poliuretanici rigidi

In questo Capitolo verranno illustrati i risultati delle campagne sperimentali per la costruzione di un modello di *machine learning* supervisionato che supporti il formulatore a determinare il dosaggio degli ingredienti che garantisca l'ottenimento di una cinetica di espansione desiderata ai prodotti poliuretanici.

### 4.1 Campagne sperimentali

Sono state organizzate due campagne sperimentali: la prima orientata allo studio degli effetti dei catalizzatori, la seconda allo studio degli effetti di tutti i possibili ingredienti della miscela.

Per ogni esperimento vengono raccolti i dati riportati in Tabella 4.1.

**Tabella 4. 1.** *Catalizzatori impiegati per la prima fase della campagna sperimentale.*

Variabile misurata	Unità di misura
Profilo di temperatura	[°C]
Profilo di distanza	[mm]
Dosaggio miscela B	[–]
Temperatura miscela B	[°C]
Temperatura miscela A	[°C]
Temperatura esterna	[°C]
Massa contenitore miscela B	[g]
Massa miscela B	[g]
Bicchiere Isocianato	[g]
Massa miscela A	[g]
Residuo miscela B	[g]
Rimasto Isocianato	[g]
Tempo di Filo	[s]
Tempo di Tocco	[s]
Densità del manufatto	[kg/m <sup>3</sup> ]

Gli esperimenti sono stati eseguiti ad una temperatura standardizzata compresa tra i 18 ed i 22°C, con temperature delle due miscele a circa 20°C. Questi vincoli sono molto importanti per l'analisi di una schiuma poliuretanic, in quanto l'espansione è fortemente condizionata dalla temperatura iniziale delle miscele e dalla temperatura ambiente. Per ogni manufatto è stata poi misurata la densità del cuore, per avere una conferma in merito alla buona riuscita dell'esperimento e per accertarsi di aver ottenuto un prodotto la cui qualità meccanica sia accettabile.

In Appendice A sono riportate tutte le misure sperimentali effettuate e i dati raccolti attraverso le campagne sperimentali. Inoltre, sono riportate tutte le combinazioni di fattori e livelli studiate.

I profili raccolti attraverso entrambe le campagne sperimentali sono stati pretrattati e manipolati come descritto al Capitolo 3.

#### 4.1.1 Campagna sperimentale per lo studio dell'effetto della catalisi

La prima fase sperimentale ha lo scopo di studiare gli effetti dei catalizzatori sulla cinetica di crescita di una schiuma, in quanto sono le sostanze con un impatto più sensibile sulla crescita del materiale; infatti, essi vanno a influire sulle reazioni chimiche che avvengono durante la produzione. Tuttavia, nella pratica industriale sono noti solamente gli effetti qualitativi di ogni catalizzatore, ovvero, si conosce in che modo ognuno intervenga sulla cinetica, è la specifica e singola attività sulla chimica del sistema è interpretata attraverso l'esperienza del formulatore.

Inoltre, i catalizzatori sono i primi componenti che si variano quando è necessario modulare i profili di velocità di crescita media e gradiente medio di temperatura. Idealmente, mantenendo la stessa base poliolicca, ma modulando la catalisi, un numero molto elevato di cinetiche diverse potrebbero essere ottenute, perciò, avere uno strumento che possa predire le quantità di catalizzatore da aggiungere ad una determinata miscela, potrebbe avere un impatto notevole in quanto si potrebbero ottenere diversi prodotti con una minima variazione di formulazione.

Sono stati scelti quattro catalizzatori su cui eseguire lo studio, in quanto essi sono i più largamente utilizzati e facilmente reperibili. Per motivi di riservatezza dei dati aziendali, le specifiche denominazioni dei catalizzatori saranno criptate. In Tabella 4.2 sono riportate le tipologie dei catalizzatori studiati.

**Tabella 4.2.** *Catalizzatori impiegati per la prima campagna sperimentale.*

Catalizzatore	Tipologia
C1	Catalizzatore generico 1
C2	Catalizzatore generico 2
C3	Catalizzatore ad effetto ritardato
C4	Catalizzatore di trimerizzazione

Dalla descrizione dei catalizzatori fornita dal produttore, è noto che C1 e C2 hanno un effetto generico andando ad accelerare tutte le reazioni in gioco in fase di espansione della schiuma, C3 è invece un catalizzatore ad effetto ritardato, ovvero che agisce successivamente allo sviluppo di calore a causa delle reazioni di *blow* e di poliaddizione. Infine, C4 agisce sulla *trimerizzazione*, ovvero catalizza la formazione di poliisocianurato.

Per questa campagna sperimentale, considerato il ridotto numero di fattori, si è utilizzato un piano degli esperimenti di tipo *Central Composite Design* circoscritto, nel quale si variano i fattori possono variare su cinque livelli, codificati in questo modo:  $-2, -1, 0, 1, 2$ . Quindi, sono state eseguite tutte le combinazioni dei livelli  $-1$  e  $1$ , tre repliche dei punti centrali, una replica per ogni vertice.

Si modificano quindi le quantità di catalizzatori, mantenendo costante la base costituita dalla miscela di polioli, di agenti espandenti fisici e chimici, di antifiama e di tensioattivi. La base della

componente B priva di catalizzatori è prodotta attraverso un miscelatore automatico, escludendo quindi differenze o errori in fase di produzione di tale miscela. Vengono inoltre mantenuti costanti l'indice di NCO e il rapporto A/B.

Tutti gli esperimenti sono stati condotti con 140 g di miscela A e 190 g di miscela B. Il rapporto B/A è stato mantenuto costante ad un valore di  $B/A = 1.36 [-]$ . Perciò, anche le quantità di catalizzatori corrispondono alle quantità inserite in quella massa di miscela B.

Prima di procedere con la campagna sperimentale, sono state eseguite una serie di prove preliminari di screening. I dettagli in merito a queste prove sono stati riportati anch'essi in appendice A.

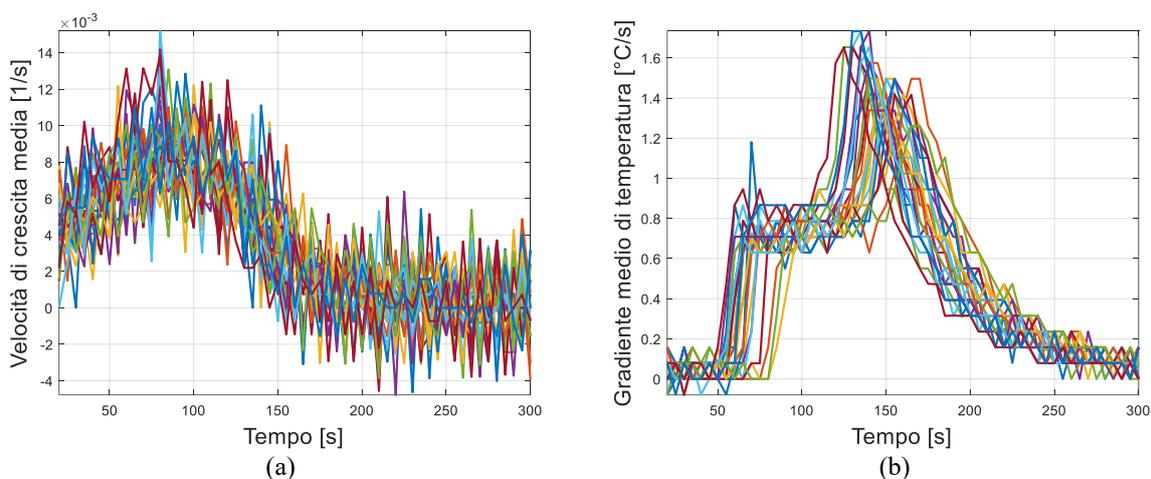
I range su cui variare le masse di catalizzatori sono stati scelti in base alle comuni quantità che vengono inserite per le miscele prodotte dall'azienda. Sulla base di questi range, il valore di massa di catalizzatore inserito per ogni livello è stato calcolato e riportato in Tabella 4.2:

**Tabella 4.3.** Massa di catalizzatore corrispondente ad ogni diverso livello per la prima fase della campagna sperimentale.

Catalizzatore	-2	-1	0	1	2
Massa di C1 [g]	0.58	0.69	0.81	0.92	1.05
Massa di C2 [g]	0.29	0.35	0.40	0.46	0.52
Massa di C3 [g]	0.12	0.16	0.20	0.24	0.28
Massa di C4 [g]	0.75	0.85	0.95	1.05	1.15

Gli esperimenti della campagna si compongono quindi di 16 combinazioni dei livelli -1 e 1, 3 punti centrali e 8 punti estremi; sono state anche effettuate 5 repliche, per un totale di 32 esperimenti. L'ordine con cui sono stati condotti gli esperimenti è casuale.

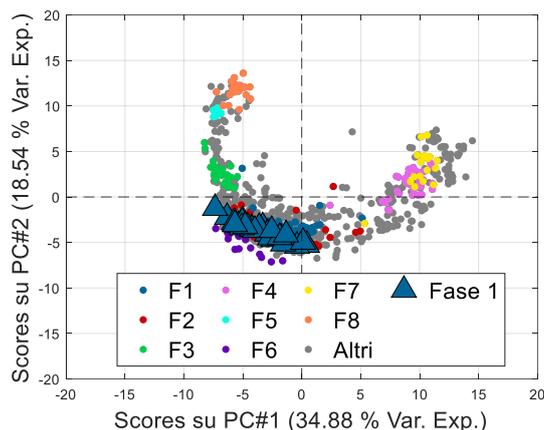
Una volta raccolti i dati, essi sono stati pretrattati e manipolati, i profili temporali di velocità di crescita media e gradiente medio di temperatura sono state riportate in Figura 4.1.



**Figura 4.1** Profili di: (a) velocità di crescita media e (b) gradiente medio di temperatura raccolti durante la prima campagna sperimentale.

I profili misurati risultano simili tra loro. Inoltre, il rumore sembra essere alto e paragonabile alla varianza fra i profili temporali di velocità media di crescita. Tuttavia, è necessario specificare che minime variazioni della curva di velocità hanno un impatto molto importante sul processo di produzione della schiuma. È stato deciso di eliminare una osservazione, non riportata in Figura 4.1, in quanto il suo profilo era stato raccolto in modo palesemente errato.

Proiettando i profili nel modello PCA generale discusso al Capitolo 3, è possibile osservare il posizionamento dei nuovi formulati ottenuti dalla prima campagna sperimentale (Figura 4.2).



**Figura 4.2.** Proiezione dei nuovi profili ottenuti attraverso la prima fase sperimentale nello score plot del modello generale.

La variabilità dei profili è elevata, inoltre essi si concentrano nella zona appartenente ai formulati standard, ovvero i composti prodotti in misura maggiore dall'azienda.

#### 4.1.2 Campagna sperimentale per lo studio dell'effetto congiunto di tutti i componenti

Lo scopo della seconda fase è lo studio degli effetti di tutti i componenti della miscela B sulla cinetica di espansione. Infatti, attraverso l'algoritmo *Kennard-Stone*, descritto al Capitolo 1, è stato possibile individuare un numero relativamente limitato di combinazioni di un numero altrimenti troppo elevato di combinazioni di fattori e livelli.

Le classi di componenti prese in considerazione sono: polioli, catalizzatori ed agenti espandenti. Ci si aspetta nuovamente che i catalizzatori siano i fattori più influenti sulla cinetica della schiuma. Tuttavia, anche i polioli e gli agenti espandenti hanno un impatto importante. Sono stati studiati sette polioli, tre dei quali sono comuni polieteri tra i maggiormente utilizzati. Si sono poi utilizzati un polietere bromurato, un poliestere e due polieteri meno generici. I catalizzatori sono gli stessi utilizzati anche durante la prima fase, con l'aggiunta di un quinto catalizzatore noto per avere un effetto particolare sulla reazione di *blow*. Tre agenti espandenti sono stati utilizzati, due dei quali sono agenti chimici, uno fisico. Si è deciso di mantenere la stessa tipologia di catalizzatori rispetto alla prima fase, poiché il loro effetto potrebbe variare in base alla tipologia di poliolo o di agente

espandente. Ancora una volta, le sigle dei componenti scelti per eseguire gli esperimenti sono state sostituite da nuove sigle, i dettagli riguardanti tutte le sostanze prese in considerazione sono riportati in Tabella 4.4:

**Tabella 4.4.** Componenti impiegati per la seconda fase della campagna sperimentale.

Componente	Classe di componente	Tipologia
P1	Poliolo	Polietero
P2	Poliolo	Polietero
P3	Poliolo	Polietero
P4	Poliolo	Poliestere
P5	Poliolo	Polietero bromurato
P6	Poliolo	Polietero
P7	Poliolo	Polietero
C1	Catalizzatore	Catalizzatore generico
C2	Catalizzatore	Catalizzatore generico
C3	Catalizzatore	Catalizzatore ad effetto ritardato
C4	Catalizzatore	Catalizzatore di trimerizzazione
C5	Catalizzatore	Catalizzatore di <i>blow</i>
E1	Agente espandente	Espandente chimico
E2	Agente espandente	Espandente chimico
E3	Agente espandente	Espandente fisico

Inoltre, si è cercato di capire se e come determinati componenti siano fra di loro legati e se interagiscano nella loro azione.

In questo caso, i componenti mantenuti costanti in massa e tipologia sono antinfiamma e tensioattivi. Considerando l'elevata variabilità di componenti soggetti a variazione nel piano sperimentale è stato necessario introdurre alcuni vincoli tra le combinazioni, per garantire l'ottenimento di un prodotto con proprietà meccaniche e di isolamento termico accettabili rispetto agli standard di qualità aziendali; in particolare, sono stati vincolati: l'*indice di NCO*, il *rapporto B/A*, e il rapporto fra le di espandente chimico e la massa totale: vincolare i primi due parametri garantisce l'ottenimento di un sistema poliuretanic rigido e meccanicamente stabile; vincolare l'espandente permette di ottenere densità all'interno del range di accettabilità descritto nel Capitolo 2. I vincoli imposti sono riportati in Tabella 4.5.

**Tabella 4.5.** Vincoli utilizzati per la seconda fase della campagna sperimentale.

Parametro	Vincolo Inferiore	Vincolo Superiore
<i>Indice di NCO</i> [-]	1.15	1.2
<i>Rapporto A/B</i> [-]	1.3	1.4
<i>mol<sub>esp</sub>/massa totale</i> [mol/g]	0.00086	0.0009

Inoltre, è stata mantenuta costante la massa totale di ogni classe di componente, in modo da variare solamente la percentuale di sostanza per ogni classe, ma di ottenere sempre una miscela coerente con gli standard aziendali. In Tabella 4.6 sono riportate le masse totali per ogni classe.

**Tabella 4.6.** Massa per ogni classe di sostanze utilizzate nella seconda fase della campagna sperimentale.

<b>Classe</b>	<b>Massa [g]</b>
<b>Polioli</b>	113.4
<b>Catalizzatori</b>	2.1
<b>Espandenti</b>	8.4
<b>Antifiamma</b>	14
<b>Silicone</b>	2.1
<b>Totale</b>	140

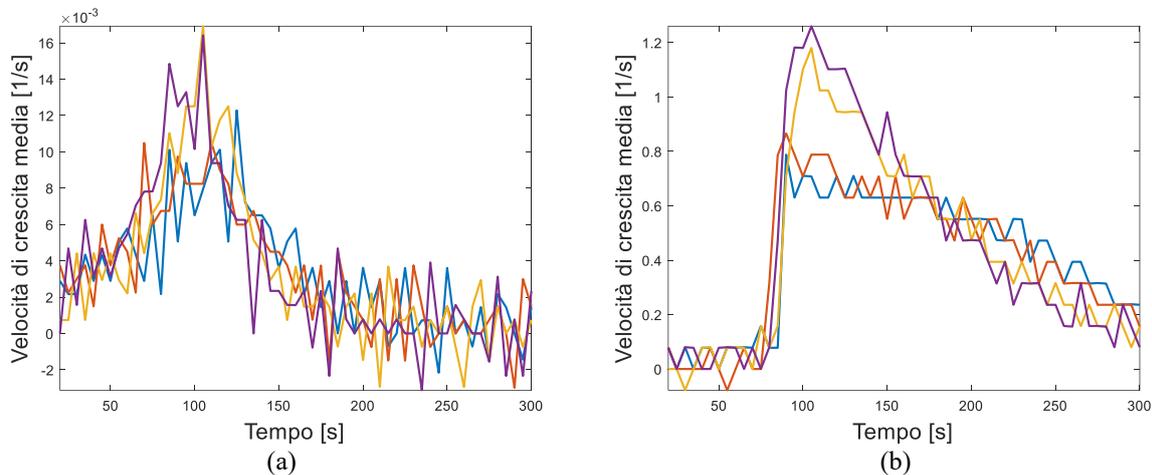
Con l'algoritmo di *Kennard-Stones* è stato possibile selezionare un numero limitato di combinazioni di fattori più rappresentative possibili, cioè che coprono nel modo più omogeneo possibile lo spazio del piano sperimentale. Perciò, è stato necessario costruire l'intero piano fattoriale di tutte le combinazioni ad ogni livello per ciascun componente. Per poter rispettare i vincoli introdotti, i livelli discreti di variazione di quantità di ogni fattore sono molto elevati: i polioli e gli agenti espandenti sono stati variati su undici livelli; per motivi computazionali, i catalizzatori invece sono stati variati su cinque livelli. Inoltre, i range di frazione ponderale per ogni componente della classe poliolica sono stati scelti in modo tale che ognuno non superi il 50% in massa della miscela poliolica. Il range è stato ristretto per quanto riguarda i polioli P6 e P7, in quanto essi non possono essere utilizzati in quantità elevate a causa di una ridotta reperibilità e di un costo elevato. È stato inoltre introdotto un vincolo per quanto riguarda la frazione di C5, in quanto, non è possibile introdurne più del 50% della massa totale di catalizzatori, poiché esso agisce principalmente sulle reazioni di *blow*. In Tabella 4.7 sono riportati i range e i livelli di variazione per ognuno dei fattori.

**Tabella 4.7.** Componenti impiegati per la seconda fase della campagna sperimentale, livelli e range di variazione.

<b>Componente</b>	<b>Livelli</b>	<b>Range di frazione ponderale [-]</b>
<b>P1</b>	11	0 – 0.5
<b>P2</b>	11	0 – 0.5
<b>P3</b>	11	0 – 0.5
<b>P4</b>	11	0 – 0.5
<b>P5</b>	11	0 – 0.5
<b>P6</b>	11	0 – 0.1
<b>P7</b>	11	0 – 0.1
<b>C1</b>	5	0 – 1
<b>C2</b>	5	0 – 1
<b>C3</b>	5	0 – 1
<b>C4</b>	5	0 – 1
<b>C5</b>	5	0 – 0.5
<b>E1</b>	11	0 – 1
<b>E2</b>	11	0 – 1
<b>E3</b>	11	0 – 1

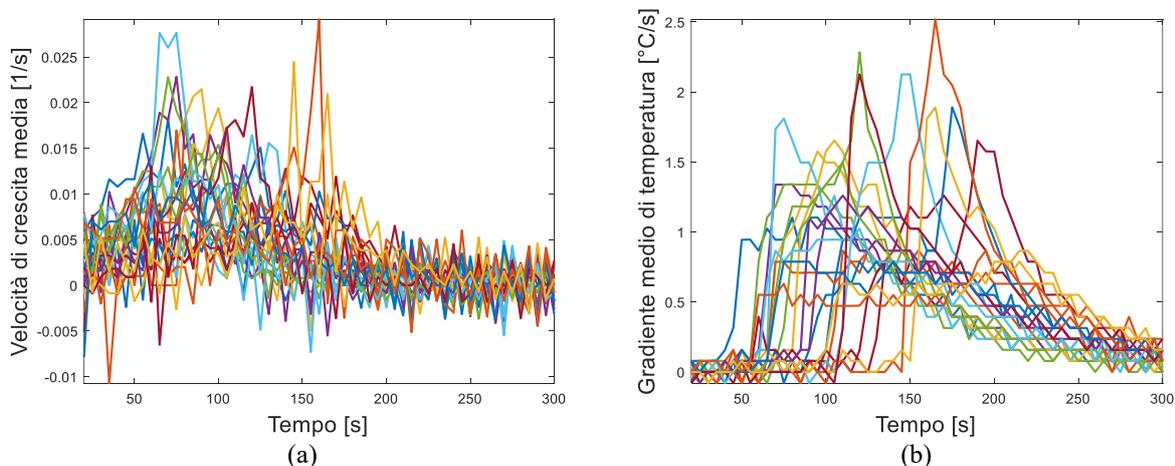
Una volta definite le combinazioni possibili, ci si è interrogati sulla effettiva influenza della componente poliolica sulla cinetica della schiuma poliuretana. Perciò, si è voluto procedere con una analisi di screening preliminare per accertarsi di ciò. Si è quindi deciso di costruire un nuovo dominio

nel quale catalizzatori ed espandenti sono stati tenuti a livello costante e si è variato solamente il dosaggio ponderale della parte poliolica. Sono state quindi estratte dal nuovo dominio preliminare cinque diverse combinazioni sperimentali attraverso l'algoritmo di KS. I profili di velocità di crescita media raccolti sono riportati in Figura 4.3.



**Figura 4.3.** Risultati dell'analisi preliminare sull'effetto dei polioli. (a) profili di velocità di crescita media; (b) profili di gradiente medio di temperatura.

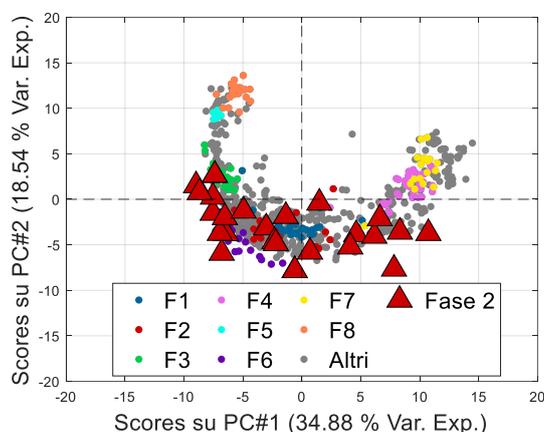
Come è possibile vedere dalla figura 4.3, i profili risultano sensibilmente differenti, perciò, si può dire che la variazione di miscela poliolica abbia un effetto rilevante sulla cinetica di crescita di una schiuma poliuretanic. L'analisi di una delle cinque combinazioni preliminari è stata tuttavia considerata inutilizzabile a causa di errori di misurazione. Una volta ottenuta la certezza dell'effetto del dosaggio della parte poliolica, sono state selezionate 25 formulazioni mediante l'algoritmo KS dall'intero dominio, che sarebbe costituito in totale da 41477 combinazioni (per cui la sperimentazione risulterebbe impossibile); quindi, sono state preparate tutte le miscele di componente A, e sono stati eseguiti gli esperimenti in ordine casuale. I risultati della seconda fase sperimentale sono stati riportati in Figura 4.4.



**Figura 4.4.** Risultati della seconda fase sperimentale. (a) profili di velocità di crescita media; (b) profili di gradiente medio di temperatura.

La variazione di tutti i possibili fattori genera una variabilità molto più elevata dei profili di crescita e di gradiente termico, rispetto alla prima campagna sperimentale.

Ancora una volta, è possibile proiettare i nuovi profili nel modello generale costruito ed analizzato al Capitolo 3. Lo score plot è riportato in Figura 4.5.



**Figura 4.5.** Proiezione dei nuovi profili ottenuti attraverso la seconda campagna sperimentale nello score plot del modello generale.

È in questo modo confermata l'elevata variabilità dei profili raccolti durante l'analisi della seconda fase sperimentale. Tuttavia, non si raggiungono mai i casi di cinetica estremamente veloce o lenta, questo perché sono stati rispettati i vincoli sull'*indice di NCO*, sul *rapporto A/B* e sulle moli di espandente utilizzate.

## 4.2 Costruzione del modello predittivo

Con i dati raccolti è stato possibile procedere con la costruzione del modello a superficie di risposta *PLS*. Questo modello non solo permette di generare un valore di formulazione a partire da un profilo di cinetica desiderato, ma permette anche di comprendere quali fattori influenzino maggiormente la cinetica e quali sono meno importanti. Si è deciso di non considerare i profili di gradiente medio di temperatura in fase di costruzione dei modelli, in quanto si vuole concentrare lo studio e la predizione sulla cinetica della schiuma.

### 4.2.1 Predizione e studio dell'effetto della catalisi

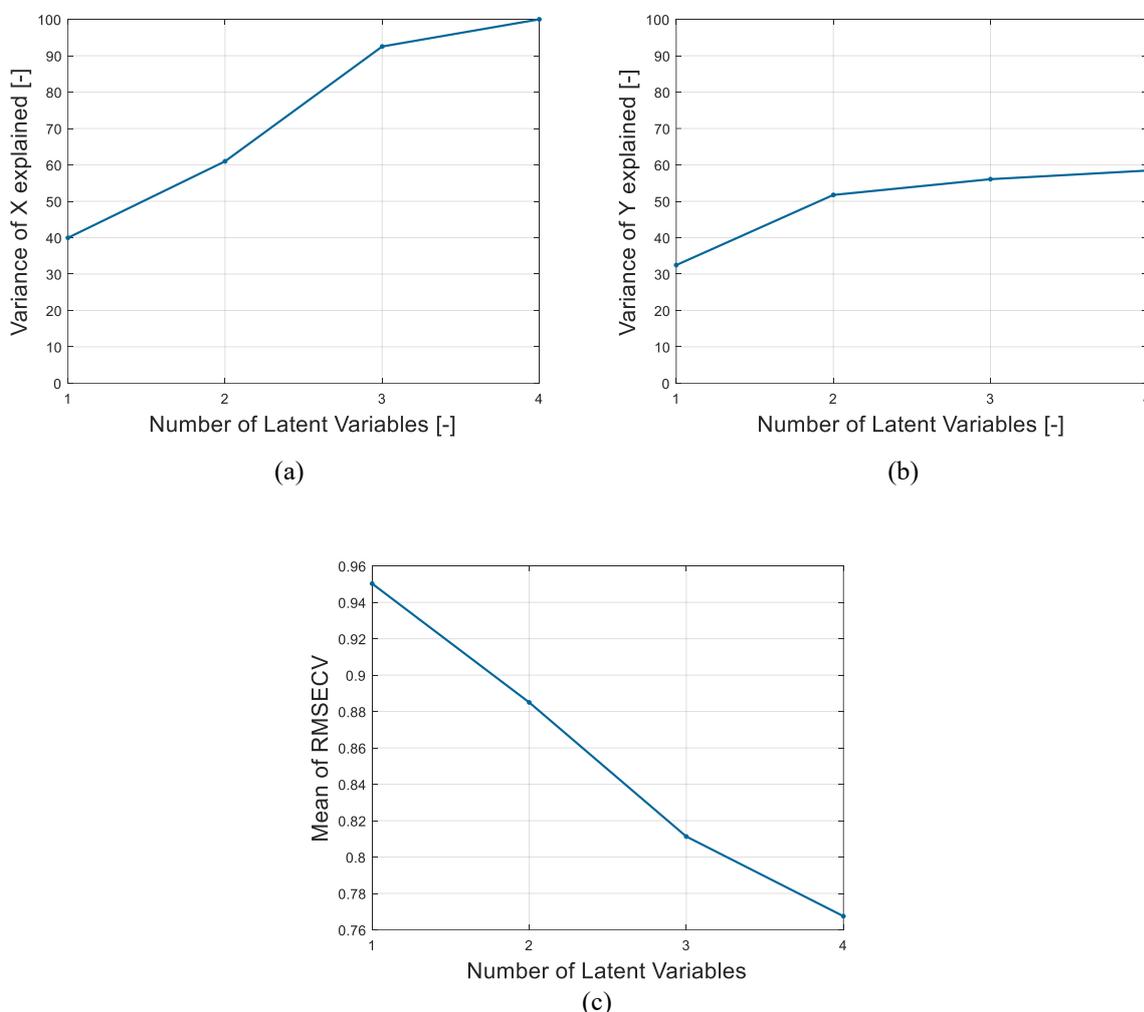
Il primo modello è costruito sui dati raccolti durante la prima campagna sperimentale; perciò, sarà in grado di cogliere tutti gli aspetti e gli impatti di ognuno dei quattro catalizzatori studiati. È possibile affermare che in questo caso non ci saranno confondimenti dell'effetto dei catalizzatori, in quanto i dati sperimentali provengono da un *Central Composite Design*, e tutte le combinazioni sono state eseguite sperimentalmente. Inoltre, tre esperimenti sono stati replicati.

Essendo i dati raccolti concentrati nella zona standard dello *score plot* del modello generale, ci si aspetta che la predizione sia molto accurati per quel tipo di profili e che l'accuratezza diminuisca man mano ci si allontana dal centro dell'arco degli *scores*.

I dati raccolti in fase di screening sono stati comunque introdotti nella fase di costruzione del modello ai fini di migliorarne la predizione, in quanto, la PLS non necessita l'utilizzo di dati esclusivamente derivanti da DoE. Il numero totale di osservazioni con le quali il modello è stato costruito ammonta a 50 prove.

#### 4.2.1.1 Costruzione del modello a superficie di risposta mediante PLS

Per la selezione del numero di LV più appropriato in un modello PLS che deve essere invertito è necessario spiegare una quota elevata di variabilità sia dalla matrice dei fattori **X** sia dalle variabili di qualità **Y**.



**Figura 4.6.** Test per la scelta del numero adatto di Latent Variables con il quale costruire il modello. Scree Test per (a) X e (b) Y; (c) RMSECV Test.

Scegliendo 4 LV si ottiene una riduzione dell'errore di predizione e si cattura tutta la varianza di  $\mathbf{X}$ . La varianza raccolta di  $\mathbf{Y}$  tuttavia non è molto elevata e sarebbero sufficienti due LV per spiegarne la maggior parte.

Sono stati quindi costruiti tipi diversi di modelli a seconda della variabile  $\mathbf{Y}$  usata nel modello. Questi sono descritti in Tabella 4.8. Questo modello predittivo nella forma diretta serve per stimare la risposta  $Y$  (cioè, i profili di espansione della schiuma) dalla formulazione dei catalizzatori. Nella forma inversa, dato un profilo temporale desiderato della cinetica di espansione il modello suggerisca quale sia la combinazione dei catalizzatori più appropriata per ottenere il prodotto desiderato.

**Tabella 4.8.** *Tipi di modelli PLS costruiti per la stima della variabile di qualità dalla formulazione dei catalizzatori e per la formulazione del prodotto.*

Tentativo	Descrizione
Modello #1	Modello che utilizza il solo profilo di velocità come variabile $\mathbf{Y}$ .
Modello #2	Considerando la non linearità della curva di velocità di crescita al variare della catalisi (la pendenza di entrambi i due rami aumenta al diminuire della scorrevolezza del formulato, e viceversa), si sono costruiti due modelli: uno per la prima parte, in cui si ha la crescita monotona dei profili e uno per la seconda, in cui si ha una diminuzione della velocità di crescita.
Modello #3	Modello che considera la derivata della velocità di crescita come variabile $\mathbf{Y}$
Modello #4	Modello che utilizza i profili di altezza normalizzata
Modello #5	La temperatura esterna, registrata ad ogni esperimento, è stata introdotta come un fattore determinante della cinetica. In fase di predizione, è stato vincolato al valore vero di temperatura esterna, in modo da ottenere comunque una predizione della catalisi.
Modello #6	Oltre alla velocità di crescita della schiuma, è stato introdotto il gradiente di temperatura come variabile risposta.
Modello #7	Modello che utilizza l'integrale dell'altezza normalizzata come variabile $\mathbf{Y}$ .
Modello #8	Modello che utilizza i profili di altezza normalizzati pretrattati mediante SNV come variabile $\mathbf{Y}$
Modello #9	Modello che utilizza i profili di altezza normalizzati e la velocità di crescita pretrattati mediante SNV come variabile $\mathbf{Y}$

Per ognuno dei tipi di modello illustrati in Tabella 4.8 è stata applicata la metodologia leave-one-out *LOO* ottenendo una stima dell'errore di predizione dei fattori e l'efficienza dell'inversione del modello. Quindi è stato possibile analizzare l'errore medio di previsione di ognuno dei fattori come:

$$err_x [g] = media(|x^{real} - x^{pred}|)$$

$$x^{real}[g] = \text{massa di catalizzatore realmente utilizzata} \quad (4.1)$$

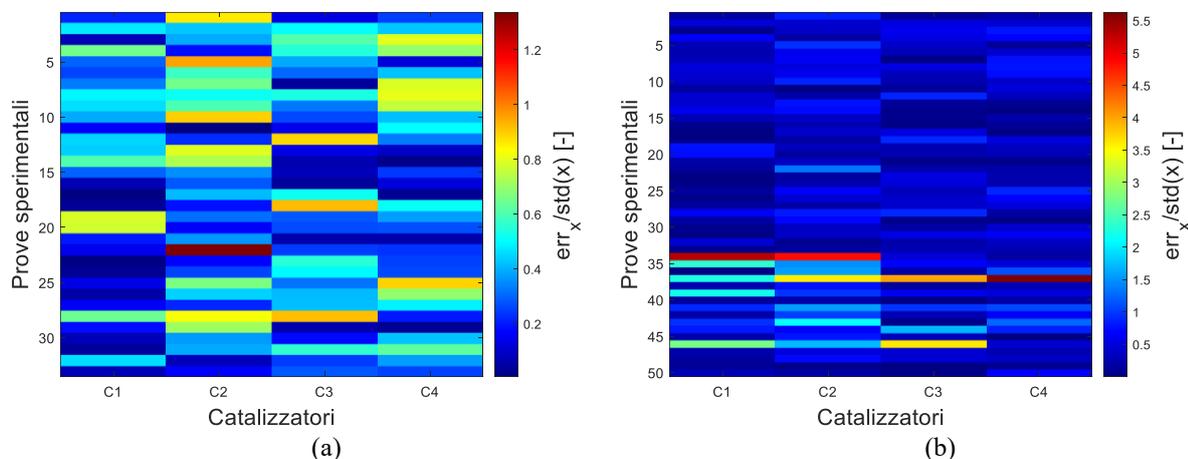
$$x^{pred}[g] = \text{massa di catalizzatore predetta dal modello}$$

**Tabella 4.9.** Risultato dell'applicazione della metodologia *Leave One Out* all'inversione del modello PLS per l'identificazione della ricetta ottimale per ottenere il profilo di espansione della schiuma desiderato.

Modelli	$err_{C1}$ [g]	$err_{C2}$ [g]	$err_{C3}$ [g]	$err_{C4}$ [g]	Media [g]
Modello #1	0.1130	0.0478	0.0623	0.0970	0.0801
Modello #2	0.2058	0.0628	0.0771	0.1283	0.1185
Modello #3	0.1506	0.1111	0.0651	0.1983	0.1313
Modello #4	0.1204	0.0528	0.0481	0.1237	0.0863
Modello #5	0.2236	0.058	0.0623	0.1358	0.1199
Modello #6	0.1374	0.0569	0.0634	0.1004	0.0895
Modello #7	0.2537	0.1244	0.1319	0.3122	0.2056
Modello #8	0.0916	0.0478	0.0495	0.0862	0.0688
Modello #9	0.0965	0.0468	0.0504	0.0844	0.0695

Il modello adottato è quello che utilizza come variabile risposta  $Y$  i profili di altezza normalizzata e velocità di crescita pretrattati con SNV (modello #9), in quanto presenta un minimo errore per quanto riguarda C2 e C4, che, come verrà discusso successivamente, sono i catalizzatori ad impatto maggiore. Modelli diversi non garantiscono una accuratezza migliore. In generale il pretrattamento con SNV aiuta a migliorare le prestazioni dei modelli, rendendo più importanti ai fini della predizione gli istanti iniziali.

In Figura 4.7 è possibile osservare l'errore standardizzato (cioè diviso per la variabilità di ogni fattore) per la stima del dosaggio di ogni catalizzatore ed ogni prova sperimentale calcolato con il LOO.



**Figura 4.7.** Errore di stima per ogni prova sperimentale: l'entità dell'errore aumenta secondo la barra, dal blu al rosso. (a) prove sperimentali e repliche pianificate (b) tutte le prove utilizzate per la costruzione del modello.

Come è possibile vedere in Figura 4.7 (a) l'errore si attesta attorno ad una deviazione standard di ogni fattore in tutte le prove sperimentali.

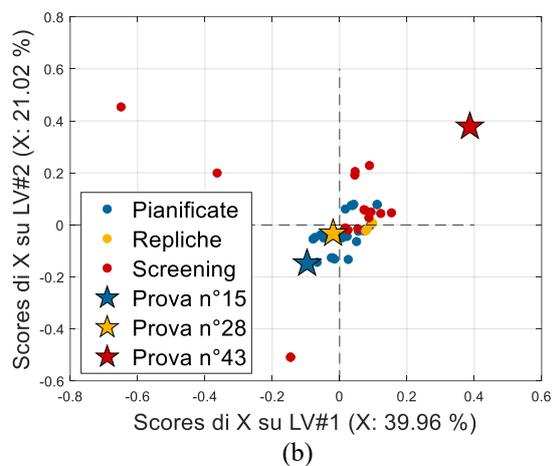
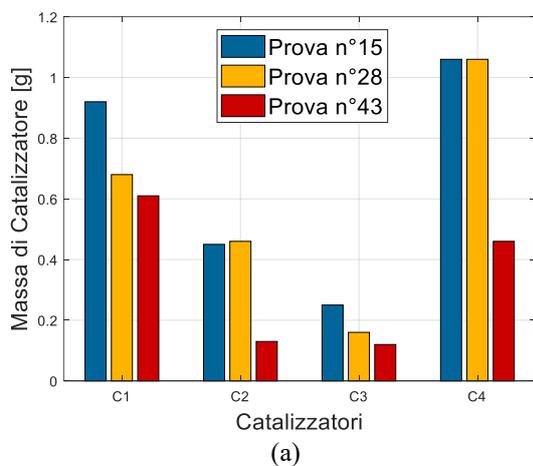
Si osserva che le prove sperimentali di screening che non derivano dal piano di *Central Composite Design* (prove 32-50) sono state effettuate a valori estremi di quantità di catalizzatore; perciò, l'errore di previsione è molto elevato, in quanto il modello deve lavorare in estrapolazione, come si vede da Figura 4.7 (b). Quindi, il modello funziona molto bene in fase di previsione di formulati standard o

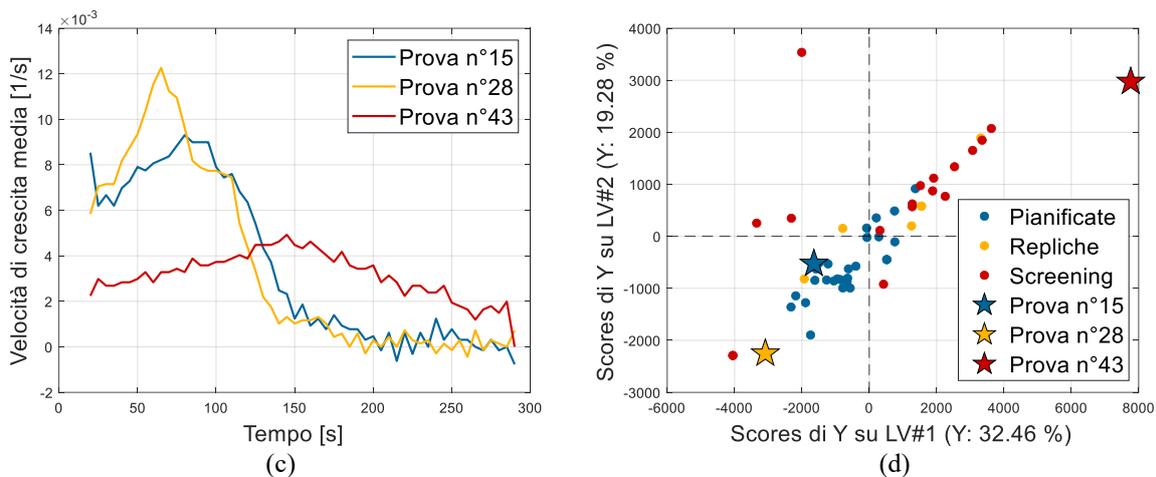
simili alle prove sperimentali eseguite, mentre in estrapolazione le prestazioni peggiorano in modo sensibile. Ciononostante, il risultato è comunque accettabile.

#### 4.2.1.2 Interpretazione del modello *PLS*

Per una comprensione approfondita della relazione tra il dosaggio dei catalizzatori e la cinetica di espansione dei sistemi poliuretanici si interpretano i parametri del modello PLS con le prestazioni di stima migliori. L'analisi verterà solamente sullo studio della prima e della seconda variabile latente, che rappresentano rispettivamente circa il 40% ed il 20% della varianza dei fattori ed il 30% e 20% della varianza dei profili di espansione.

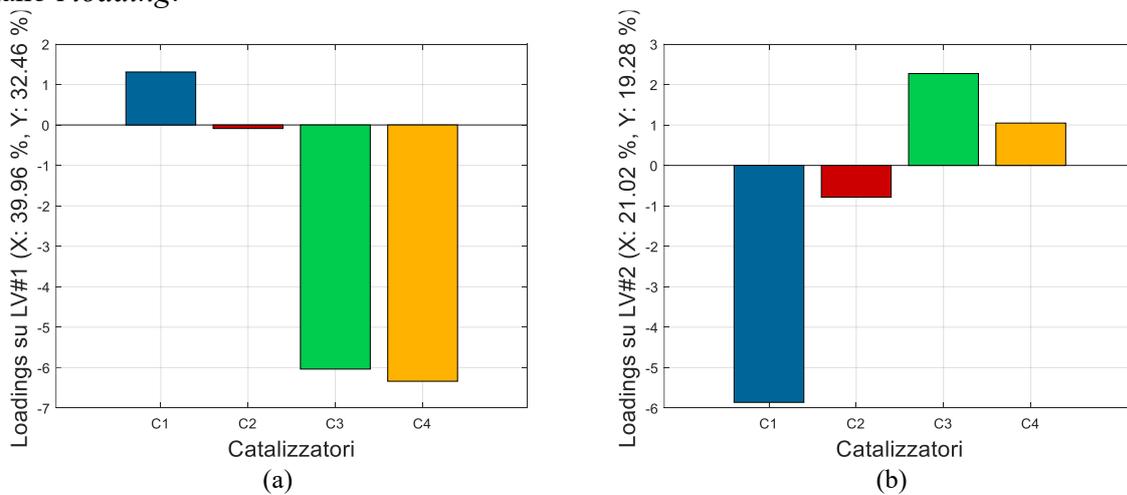
Innanzitutto, occorre analizzare lo *score plot* delle **X** e lo *score plot* delle **Y**; per quanto riguarda i fattori, il modello riesce ad interpretare correttamente i profili derivanti dall'analisi. Le osservazioni di screening non pianificate mediante DoE e che mostrano comportamenti estremi dei profili temporali delle variabili cinetiche, sono più distribuite nel grafico, hanno una maggiore variabilità ed talvolta escono dai limiti di confidenza. È possibile vedere da Figura 4.8 (a) e (b) che, spostandosi nello *score plot* verso valori negativi di entrambe le LV si ha un aumento della quantità di catalizzatori utilizzati. La distribuzione degli *score* delle **Y** mostra, Figura 4.8 (c) e (d) che maggiore è la velocità del formulato, più bassi sono gli *scores* sulle prime due LV.

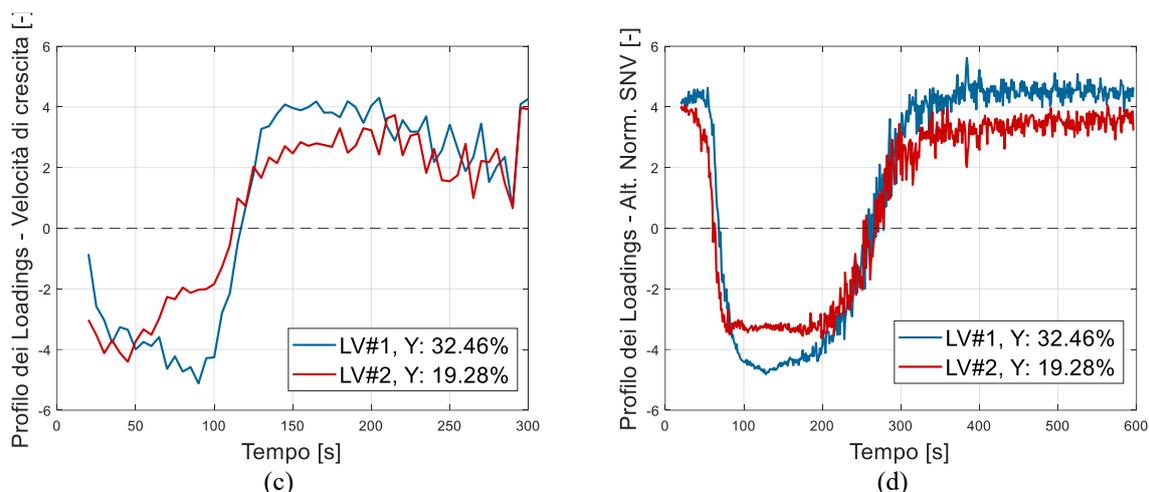




**Figura 4.8.** Score plot del modello PLS per la predizione della catalisi necessaria a partire da un profilo cinetico richiesto comparati con i valori effettivi di massa di catalizzatore e con i profili di velocità di crescita media; (a) masse di catalizzatore di tre prove sperimentali (b) score plot di X; (c) profili di velocità di crescita di tre prove sperimentali (d) score plot di Y.

Per capire e confermare la relazione tra i profili di altezza e velocità e il dosaggio dei catalizzatori si studiano i *loadings*.





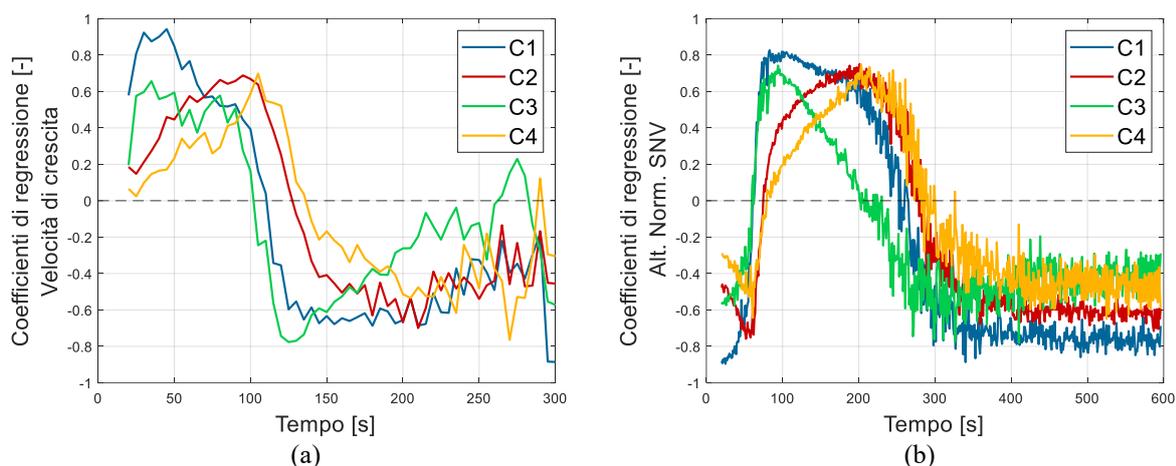
**Figura 4.9.** Loadings plots del modello per la predizione della catalisi necessaria a partire da un profilo cinetico richiesto (a) loading di  $\mathbf{X}$ , prima variabile latente; (b) loading di  $\mathbf{X}$ , seconda variabile latente (c) loading di  $\mathbf{Y}$  per la velocità di crescita; (d) loadings di  $\mathbf{Y}$  per l'altezza normalizzata.

In Figura 4.9, sono rappresentati i *loading* di entrambe le matrici sulle prime due componenti principali. Innanzitutto, è interessante notare che a valori negativi di LV1, corrispondono tutti i formulati il cui il dosaggio di C3 e C4 è più elevato. A valori negativi di LV2, si hanno formulati con molto C1. Perciò, nel primo quadrante dello *score plot* avremmo i formulati con poco catalizzatore. Per quanto riguarda i *loading* dei profili di altezza e velocità, il modello riconosce la correlazione tra velocità ed altezza e si nota come i profili delle due LV abbiano un andamento simile. I grafici confermano che più basso è lo *score* di una osservazione lungo entrambe le componenti principali, minore sarà la scala temporale su cui si sviluppa la schiuma.

Si osserva su LV1 che C3 e C4 agiscono nei primi 100 s, alzando la velocità, mentre dopo i 100 s la abbassano. Inoltre, su LV2 si nota che una elevata quantità di C1 agisce aumentando la velocità nei primi 50 s, accelerando la crescita tra i 50 e i 100 s, mentre abbassa la crescita dopo i 100s.

Questo è confermato dal fatto che C1 e C2 sono catalizzatori generici, C3 e C4, sono impiegati generalmente per la chiusura. Il risultato del modello PLS ha determinato alcune incoerenze con quanto definito da scheda tecnica dei catalizzatori C3 e C4: essi agiscono in fase di chiusura, ma non solo, come invece si pensava.

Infine, è possibile analizzare i coefficienti di regressione ottenuti dal modello PLS, che legano la matrice dei fattori con i profili  $\mathbf{Y}$ .



**Figura 4.10.** Profilo dei coefficienti di regressione: (a) velocità di crescita; (b) altezza normalizzata SNV.

Analizzando principalmente Figura 4.10 (a), in quanto le variabili rappresentano entrambe la cinetica del sistema, quindi si giungerebbe alla stessa conclusione, si nota come vengano confermate le ipotesi precedentemente riscontrate, ma è molto più semplice capirne i dettagli: C1 ha un effetto catalitico iniziale, ovvero agisce durante i primi istanti di crescita. C2, ha un effetto medio in tutto il tempo di analisi, tuttavia, fa aumentare la velocità di crescita successivamente a C1. C3 invece ha un effetto medio sull'aumento della cinetica della schiuma, ma attorno ai 100 s ha un effetto molto importante di diminuzione della velocità. Viene confermata la sua caratteristica di catalizzatore di chiusura ad effetto ritardato. C4 invece, sembra agire successivamente a tutti gli altri catalizzatori ed il suo effetto è limitato ad alzare le velocità attorno ai 200s. Questo conferma il fatto che non dovrebbe catalizzare la cinetica ma la *trimerizzazione*, in quanto agisce solamente una volta consumata l'acqua (che ha un effetto solvatante su di esso); tuttavia, seppur limitato ha comunque un effetto sulla velocità.

#### 4.2.1.3 Risultato della convalida del modello

La convalida del modello è stata fatta considerando tre diversi formulati standard, i cui profili di qualità sono l'obiettivo da ottenere. Mediante inversione del modello PLS si suggerisce quale sia il dosaggio più opportuno dei catalizzatori per ottenere la qualità desiderata.

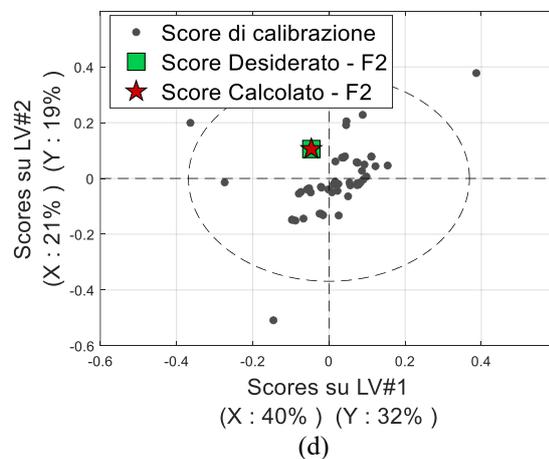
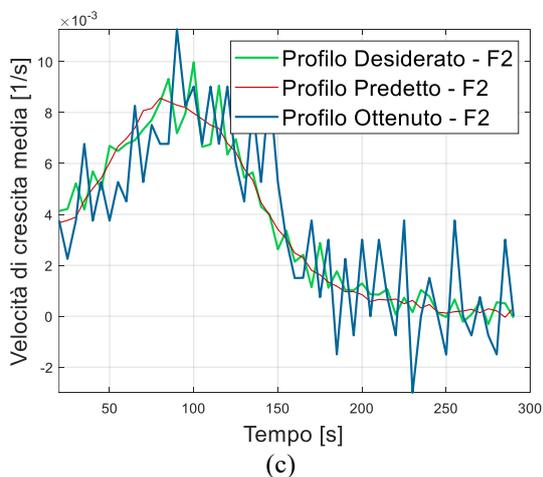
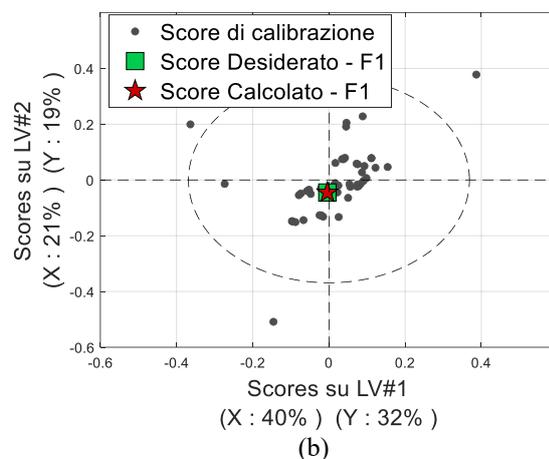
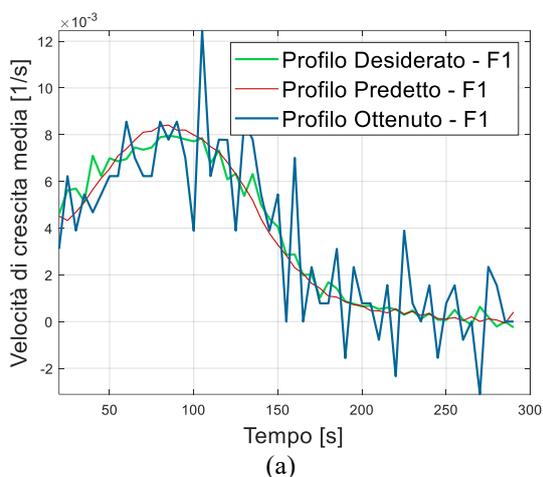
Questi tre formulati sono stati scelti per rappresentare scenari differenti: un formulato standard, un formulato di caratteristiche intermedie ed uno con caratteristiche di espansione particolarmente veloci. Questi corrispondono ai formulati F1, F2 ed F3, descritti al Capitolo 2.

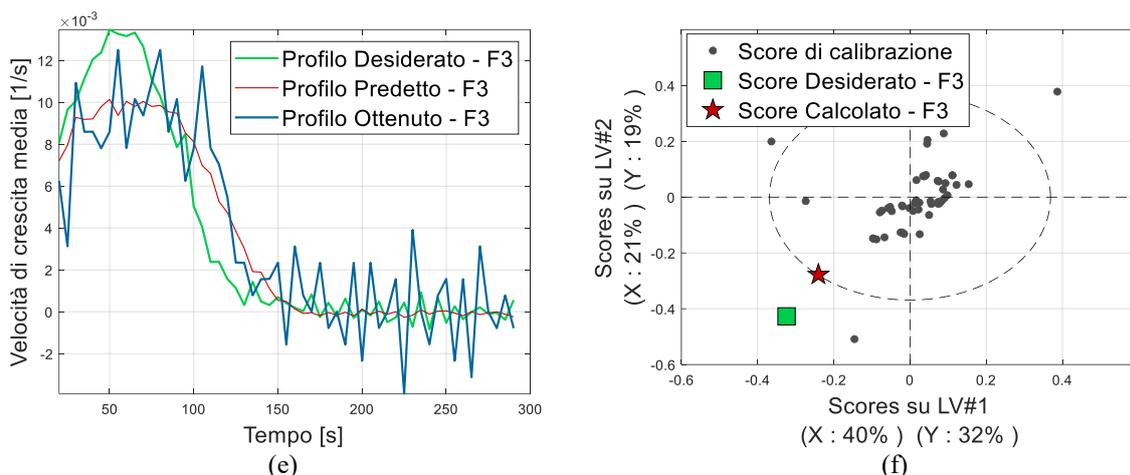
I risultati di previsione ottenuti sono riportati in Tabella 4.10.

**Tabella 4.10.** Risultato della previsione della catalisi per ottenere i profili di tre formulati richiesti.

Formulato	Massa di C1 [g]	Massa di C2 [g]	Massa di C3 [g]	Massa di C4 [g]
F1	0.66	0.50	0.17	0.99
F2	0.28	0.53	0.28	1.02
F3	1.50	0.25	0.41	1.20

Una volta ottenuta la previsione della catalisi, sono state eseguite le prove sperimentali con il dosaggio delle formulazioni suggerite. In tutti i casi, il risultato è molto soddisfacente. Ciò significa, che a partire da una base di miscela B, senza catalisi (relativamente semplice e poco costosa), è possibile imitare una grandissima quantità di profili diversi. Nonostante l'elevato rapporto rumore/segnale del profilo di velocità, il risultato di ottenere dei profili consistenti con i desiderati è stato raggiunto e il modello predittivo è molto accurato, come si può notare in Figura 4.11 (a) e (c) nei casi in cui lo *score* della formulazione si trova nei limiti del modello come nei casi di Figura 4.11 (b) e (d), anche se l'accuratezza diminuisce per i formulati con comportamenti più estremi.





**Figura 4. 11.** Risultati della previsione dei profili di tre differenti formulati. (a) score plot con la proiezione di F1; (b) confronto tra profilo desiderato ed ottenuto di F1. (c) score plot con la proiezione di F2; (d) confronto tra profilo desiderato ed ottenuto di F2. (e) score plot con la proiezione di F3; (f) confronto tra profilo desiderato ed ottenuto di F3.

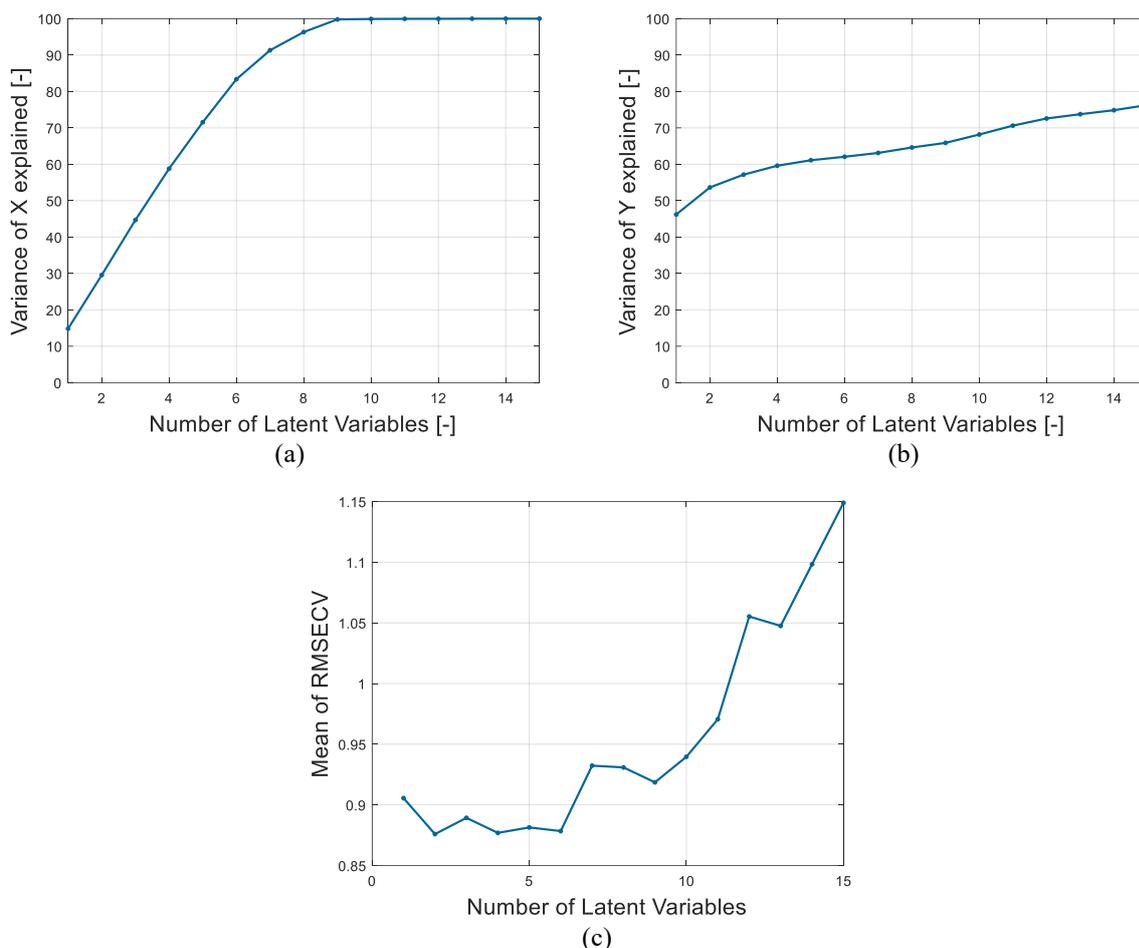
La distanza fra *score* calcolato e *score* desiderato nel terzo caso, Figura 4.11 (f) è dovuta al vincolo su  $T^2$  in fase di minimizzazione della funzione obiettivo, che impone al modello di imitare il profilo mantenendo una catalisi non eccessivamente distante da quelle utilizzate nelle formulazioni utilizzate per la calibrazione del modello. Si può tuttavia notare da Figura 4.11 (e) che il profilo ottenuto e il profilo desiderato sono molto simili, ad indicare che il modello predice molto bene l'effetto della catalisi, ovvero conosce molto bene la correlazione tra le variabili  $X$  ed  $Y$ .

#### 4.2.2 Predizione e studio dell'effetto di tutte le componenti

La costruzione del modello *PLS* con i dati sperimentali raccolti durante la seconda campagna è analoga al caso precedente. Tuttavia, occorre tener conto il numero molto elevato di fattori, in scala comparabile al numero di osservazioni. Non ci si aspetta dunque che questo modello sia accurato come il precedente.

##### 4.2.2.1 Costruzione del modello *PLS* e problema di aliasing

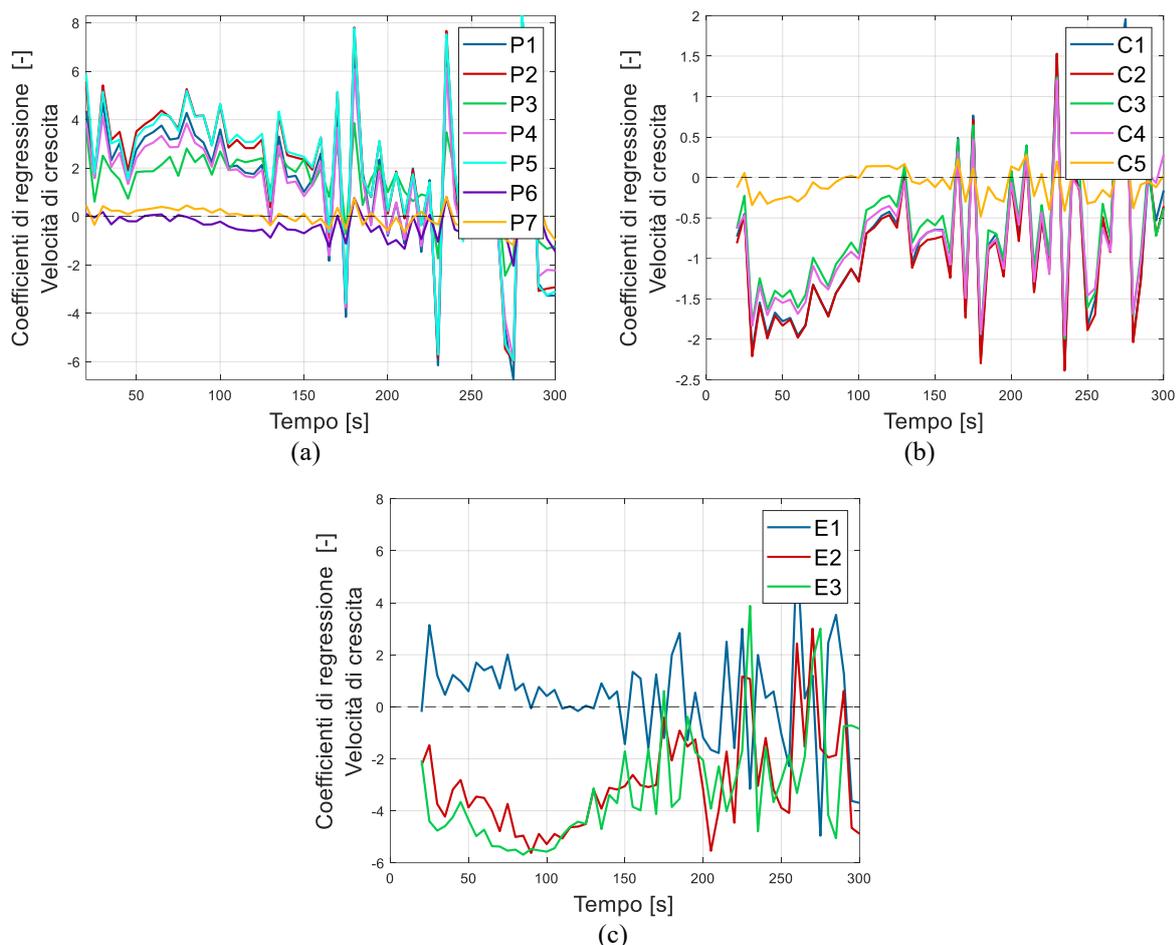
Il modello è stato costruito utilizzando le 25 prove sperimentali assieme alle 6 prove preliminari effettuate. In questo caso, il numero di fattori è molto elevato, perciò è possibile scegliere un numero altrettanto elevato di variabili latenti, i test eseguiti sono nuovamente lo *Scree Test* e l'*RMSECV Test*, i risultati sono presentati in Figura 4.12:



**Figura 4.12.** Test per la scelta del numero adatto di Latent Variables con il quale costruire il modello. Scree Test per (a)  $X$  e (b)  $Y$ ; (c) RMSECV Test.

È possibile notare che il 100% della varianza di  $X$  viene catturata utilizzando nove variabili latenti. Tuttavia, successivamente a sei LV, l'aumento della varianza di  $Y$  ha un cambio di pendenza e si ha un aumento progressivo di RMSECV, che è un indice di *overfitting*. Si è quindi deciso di utilizzare sei LV per costruire il modello. Il modello è stato costruito con le medesime variabili di qualità che nel caso precedente.

Una volta ottenuto il modello PLS è possibile osservare i coefficienti di regressione; questo passaggio è molto importante per quanto riguarda la seconda fase sperimentale, poiché si tratta di una pianificazione degli esperimenti non completa ma frazionata; perciò, è presente un confondimento di fattori. I profili dei coefficienti di regressione sono riportati in Figura 4.13.



**Figura 4. 13.** Profili dei coefficienti di regressione rispetto alla velocità di crescita media delle diverse classi di sostanze calcolati attraverso il modello PLS inverso che stima i profili di velocità di crescita media dal dosaggio dei componenti della miscela A; (a) polioli; (b) catalizzatori; (c) espandenti.

Osservando i profili dei coefficienti di regressione si nota la somiglianza tra i diversi profili: i primi cinque polioli vengono completamente confusi tra di loro e sembrano avere tutti lo stesso effetto; P6 e P7 sembrano non avere un effetto molto importante, allo stesso modo di C5 ed E1. I catalizzatori da C1 a C4 vengono anch'essi confusi, ma è noto dall'analisi precedente che il loro effetto è sostanzialmente diverso. E2 ed E3 sembrano comportarsi in un modo molto simile. Viste quindi le difficoltà del modello ad interpretare gli effetti di tutti i fattori, a causa del numero troppo ridotto di esperimenti rispetto al numero di fattori, si è deciso di non procedere ulteriormente con la costruzione di un modello predittivo basato su questa fase. Per ottenere risultati utili sarebbe necessario aumentare il numero di esperimenti, purtroppo, per questioni di tempo, non è stato possibile farlo.

### 4.3 Conclusioni e possibili sviluppi

Il dosaggio dei catalizzatori per garantire un profilo temporale desiderato della velocità e l'altezza di espansione di sistemi poliuretanici mediante tecniche di machine learning supervisionato fornisce

uno strumento efficace a supporto dei formulatori. In particolare, l'inversione di modelli di regressione a variabili latenti fornisce uno strumento predittivo accurato e sistematico.

È possibile affermare che lo strumento è molto accurato all'interno dei limiti di confidenza, ovvero quando si tratta di interpolazione, i profili ottenuti sperimentalmente con la catalisi suggerita rispecchiano perfettamente i profili desiderati. In caso di estrapolazione il modello non riesce a predire accuratamente un dosaggio di catalizzatori adeguato, ma questo risultato era atteso. Per migliorare tale modello, si suggerisce una ulteriore campagna sperimentale allargando i range di studio dei catalizzatori.

Come limite si è osservato che l'introduzione di un numero elevato di fattori richiederebbe una campagna sperimentale molto impegnativa in termini economici e di risorse richieste per la campagna sperimentale e la relativa analisi, mentre un'analisi mediante piani fattoriali frazionari è condizionata da *aliasing*.

# Conclusioni

In questa Tesi sono state sviluppate metodiche multivariate di machine learning per la caratterizzazione e la formulazione di nuovi sistemi poliuretanici.

Innanzitutto, è stata costruita una metodologia sistematica di analisi basata sui profili temporali di velocità di crescita media e gradiente medio di temperatura che è in grado di migliorare la comprensione dell'espansione di un sistema poliuretanico nel dettaglio delle diverse fasi cinetiche che la caratterizzano. Questa metodologia ha portato, inoltre, a proporre una riduzione del 50-75% dei tempi di analisi in quanto tutto il contenuto informativo dei profili temporali della schiuma è contenuto nella prima parte delle prove sperimentali di espansione.

È stato poi sviluppato un modello di inversione di PLS che suggerisce la formulazione più appropriata di un sistema poliuretanico al fine garantire la qualità del prodotto, ovvero un profilo di velocità di crescita media desiderato. È stata svolta una campagna sperimentale estensiva mirata a supportare la costruzione dei modelli matematici di machine learning supervisionato, attraverso tecniche di pianificazione degli esperimenti. È stato osservato che manipolando il dosaggio di catalizzatori utilizzati, con una metodologia parsimoniosa e semplice da produrre, si riesce a proporre una formulazione appropriata per ottenere la qualità desiderata. Questo strumento potrebbe essere notevolmente migliorato con ulteriori campagne sperimentali, ad esempio, allargando il range di variabilità dei livelli delle masse catalizzatori, in modo da raccogliere altri dati ed estendere il campo sperimentale.

Si può quindi concludere che le metodiche di machine learning multivariato sia supervisionato che non proposte in questa Tesi si sono dimostrate degli strumenti efficaci di analisi e formulazione di sistemi poliuretanici complessi.

Sviluppi futuri interessanti potrebbero derivare considerando nuove variabili di qualità, quali, ad esempio, spettri al vicino infrarosso e misure dei profili di viscosità e pressione.



# Appendice A

## Prove sperimentali ed ulteriori dati

In questa appendice sono riportati i dettagli in merito alle prove sperimentali eseguite ai fini della costruzione dei due modelli PLS, assieme agli ulteriori dati raccolti in fase di analisi.

In Tabella A.1, le masse di catalizzatore pianificate da introdurre, nei 140 g di miscela B priva di catalisi, in ogni prova della prima fase sperimentale.

**Tabella A.1.** *Quantità di catalizzatori inserite in ognuna delle miscele utilizzate per la prima fase sperimentale.*

Esp.	C1 [g]	C2 [g]	C3 [g]	C4 [g]
1	0.69	0.35	0.16	0.85
2	0.92	0.35	0.16	0.85
3	0.69	0.46	0.16	0.85
4	0.92	0.46	0.16	0.85
5	0.69	0.35	0.24	0.85
6	0.92	0.35	0.24	0.85
7	0.69	0.46	0.24	0.85
8	0.92	0.46	0.24	0.85
9	0.69	0.35	0.16	1.05
10	0.92	0.35	0.16	1.05
11	0.69	0.46	0.16	1.05
12	0.92	0.46	0.16	1.05
13	0.69	0.35	0.24	1.05
14	0.92	0.35	0.24	1.05
15	0.69	0.46	0.24	1.05
16	0.92	0.46	0.24	1.05
17	0.81	0.40	0.20	0.95
18	0.81	0.40	0.20	0.95
19	0.81	0.40	0.20	0.95
20	1.04	0.40	0.20	0.95
21	0.58	0.40	0.20	0.95
22	0.81	0.52	0.20	0.95
23	0.81	0.29	0.20	0.95
24	0.81	0.40	0.29	0.95
25	0.81	0.40	0.12	0.95
26	0.81	0.40	0.20	1.15
27	0.81	0.40	0.20	0.75
11b	0.69	0.46	0.16	1.05
16b	0.92	0.46	0.24	1.05
25b	0.81	0.40	0.11	0.92
27b	0.82	0.39	0.21	0.75
19b	0.81	0.39	0.10	0.92

In Tabella A.2 sono invece riportate le masse realmente inserite in miscela.

**Tabella A.2.** *Quantità di catalizzatori effettivamente utilizzate in ognuna delle prove sperimentali della prima fase.*

Esp.	C1 [g]	C2 [g]	C3 [g]	C4 [g]
1	0.73	0.34	0.16	0.84
2	0.93	0.38	0.16	0.89
3	0.71	0.46	0.16	0.87
4	0.91	0.47	0.17	0.86
5	0.68	0.36	0.25	0.89
6	0.91	0.37	0.25	0.85
7	0.69	0.46	0.24	0.87
8	0.92	0.45	0.24	0.89
9	0.69	0.35	0.16	1.07
10	0.91	0.34	0.15	1.07
11	0.69	0.45	0.16	1.07
12	0.91	0.45	0.17	1.10
13	0.69	0.36	0.25	1.07
14	0.91	0.35	0.24	1.05
15	0.69	0.46	0.24	1.04
16	0.92	0.45	0.25	1.06
17	0.83	0.41	0.20	0.94
18	0.81	0.39	0.21	0.94
19	0.81	0.39	0.10	0.94
20	1.03	0.39	0.21	0.94
21	0.59	0.40	0.20	0.94
22	0.80	0.51	0.19	0.95
23	0.81	0.28	0.23	0.96
24	0.82	0.39	0.30	0.97
25	0.81	0.40	0.11	0.92
26	0.81	0.39	0.20	1.16
27	0.82	0.39	0.21	0.75
11b	0.68	0.46	0.16	1.06
16b	0.91	0.46	0.23	1.06
25b	0.81	0.40	0.11	0.91
27b	0.82	0.38	0.21	0.74
19b	0.88	0.38	0.11	0.92

Vengono di seguito indicate le masse di catalizzatore presenti nelle miscele di screening utilizzate per la costruzione dei modelli, in Tabella A.3:

**Tabella A.3.** *Quantità di catalizzatori inserite in ognuna delle prove non pianificate aggiunte al modello.*

Esp.	C1 [g]	C2 [g]	C3 [g]	C4 [g]
1s	2.45	0.00	0.26	1.11
2s	0.11	0.51	0.21	0.86
3s	0.86	0.30	0.08	0.92
4s	0.00	0.11	0.82	2.20
5s	0.88	0.30	0.10	0.81
6s	0.13	0.51	0.20	0.97
7s	0.79	0.38	0.14	1.04
8s	0.57	0.28	0.25	0.98
9s	0.78	0.32	0.15	0.95
10s	0.61	0.13	0.12	0.46
11s	0.50	0.50	0.52	1.20
12s	0.88	0.35	0.17	0.91
13s	0.00	0.53	0.80	1.16
14s	0.76	0.34	0.18	0.89
15s	0.86	0.36	0.19	0.95
16s	0.74	0.36	0.14	0.89
17s	0.78	0.36	0.12	0.92
18s	0.91	0.47	0.26	1.04

In Tabella A.4, tutti i dati raccolti durante l'analisi sperimentale delle 50 prove della prima fase:

**Tabella A.4.** Dati raccolti manualmente durante l'analisi sperimentale dei formulati della prima fase sperimentale.

Esp.	$T_{mixA}$ [°C]	$T_{mixB}$ [°C]	$T_{amb}$ [°C]	$m_{bicch.B}$ [g]	$m_B$ [g]	$m_{bicch.A}$ [g]	$m_A$ [g]	$m_{resB}$ [g]	$m_{resA}$ [g]	$t_{filo}$ [s]	$t_{tocco}$ [s]	$\rho$ [ $\frac{kg}{m^3}$ ]
1	21.00	20.00	19.50	29.50	138.87	13.19	189.2	84.68	2.56	150.00	305.00	28.50
2	21.00	20.00	19.50	29.49	140.99	13.19	189.58	83.16	2.48	142.00	292.00	28.40
3	20.00	21.00	19.70	29.42	140.46	13.19	188.99	63.03	3.97	139.00	260.00	29.69
4	20.00	21.00	20.00	29.51	140.99	13.19	191.47	58.95	0.76	143.00	260.00	31.43
5	20.00	21.00	20.00	29.76	139.97	13.19	190.04	65.89	-2.15	142.00	280.00	31.40
6	21.00	21.00	20.50	32.11	140.25	13.19	189.72	81.80	0.29	140.00	290.00	29.90
7	20.00	21.00	20.80	30.82	139.74	13.19	189.51	82.01	0.57	132.00	290.00	29.45
8	21.00	21.00	20.30	29.97	140.70	13.19	189.95	97.11	-2.35	131.00	261.00	29.90
9	20.00	20.00	19.10	30.88	140.22	13.81	189.98	89.59	-2.44	170.00	310.00	31.00
10	20.00	20.00	19.30	29.56	140.51	13.81	189.55	77.10	-2.37	155.00	310.00	30.46
11	20.00	20.00	19.20	29.54	140.72	13.81	190.55	87.55	8.31	145.00	302.00	30.55
12	21.00	21.00	20.30	29.78	140.01	14.38	189.61	79.58	-0.91	133.00	250.00	28.26
13	21.00	21.00	20.50	30.52	139.35	13.19	189.85	85.31	1.21	140.00	290.00	30.85
14	21.00	21.00	20.90	30.32	140.31	13.81	189.92	86.84	1.06	130.00	253.00	30.69
15	21.00	21.00	20.60	30.40	139.50	14.38	190.74	78.50	0.57	130.00	250.00	28.15
16	21.00	21.00	20.70	29.83	140.13	13.81	189.67	84.16	1.21	130.00	250.00	29.54
17	21.00	21.00	21.30	30.44	139.80	13.81	189.84	72.79	-0.07	135.00	300.00	29.80
18	20.00	20.00	20.10	29.76	139.42	14.38	190.6	90.60	2.02	150.00	290.00	31.35
19	21.00	21.00	20.70	30.06	140.30	14.38	190.08	75.71	2.87	140.00	260.00	28.90
20	20.00	20.00	19.70	29.66	140.21	13.81	189.56	80.60	1.88	140.00	270.00	29.70
21	20.00	20.00	21.50	30.47	140.05	13.81	189.00	76.46	2.35	143.00	270.00	29.70
22	20.00	20.00	20.10	29.81	139.90	14.38	189.45	73.94	2.65	145.00	290.00	30.09
23	21.00	21.00	20.40	30.95	140.46	13.81	189.95	86.13	2.84	152.00	310.00	31.20
24	20.00	20.00	20.10	29.75	141.29	14.38	190.08	75.15	3.99	134.00	265.00	29.36
25	21.00	21.00	20.50	29.75	141.15	14.38	189.57	83.83	-0.26	155.00	300.00	29.58
26	20.00	21.00	20.00	30.68	140.02	13.81	189.75	78.50	-2.16	135.00	260.00	29.80
27	20.00	21.00	20.40	29.55	140.45	14.38	190.18	83.60	-0.48	160.00	320.00	30.25
11b	22.00	22.00	20.60	29.77	140.17	14.38	190.00	73.74	1.72	140.00	260.00	29.07
16b	22.00	22.00	20.50	30.32	140.30	14.38	189.53	74.37	5.40	130.00	250.00	28.18
25b	20.00	20.00	19.40	28.53	140.13	23.56	189.58	77.96	2.59	170.00	344.00	29.52
27b	20.00	20.00	19.50	29.57	140.57	24.50	190.06	74.14	1.47	160.00	360.00	30.09
19b	20.00	20.00	19.50	30.27	139.71	23.15	189.26	83.24	-0.50	175.00	378.00	28.78
1s	19.00	19.00	19.50	30.83	143.30	21.12	189.12	98.94	1.47	110.00	256.00	28.25
2s	19.00	19.00	19.50	70.06	139.96	17.17	189.55	125.81	4.43	181.00	340.00	31.01
3s	19.00	19.00	19.60	69.51	116.12	17.17	190.23	136.26	3.27	190.00	330.00	26.31
4s	19.00	19.00	19.70	69.83	138.87	17.17	189.75	130.84	5.36	140.00	260.00	28.06
5s	19.00	19.00	19.80	30.32	139.64	21.12	190.51	99.98	-2.36	160.00	290.00	29.27
6s	19.00	19.00	19.80	29.50	146.18	23.56	189.58	75.92	0.13	160.00	340.00	30.07
7s	20.00	20.00	19.40	28.69	143.71	24.50	190.06	73.89	-0.74	158.00	260.00	27.37
8s	20.00	20.00	19.40	28.40	140.41	23.50	189.73	70.79	-2.81	180.00	408.00	28.29
9s	20.00	20.00	19.40	29.60	140.00	18.80	190.31	78.52	1.54	160.00	320.00	27.97
10s	20.00	20.00	18.90	18.10	141.00	25.27	189.45	73.92	6.29	260.00	614.00	30.28
11s	20.00	20.00	18.90	29.14	140.27	23.35	188.22	75.75	-0.36	130.00	250.00	26.01
12s	20.00	20.00	18.90	30.45	143.30	23.01	189.73	71.82	-0.04	160.00	350.00	27.48
13s	20.50	20.50	20.00	29.76	141.61	25.19	191.75	80.76	1.07	125.00	230.00	25.82
14s	19.00	19.00	18.90	29.78	140.23	21.46	190.63	79.39	2.51	164.00	320.00	28.57
15s	20.00	20.00	19.20	29.55	140.31	19.86	189.54	82.64	1.78	164.00	340.00	28.15
16s	19.00	19.00	18.70	28.98	141.06	23.15	189.26	69.90	3.67	175.00	400.00	28.51
17s	19.00	19.00	18.90	30.48	139.49	7.44	190.34	78.77	-1.31	180.00	425.00	29.05
18s	20.00	20.00	18.90	30.04	140.45	19.86	189.54	87.85	1.68	140.00	315.00	27.46

Allo stesso modo di Tabella A.1, di seguito vengono riportate le masse di sostanze impiegate nella seconda fase sperimentale, per quanto riguarda sia le prove preliminari sia le prove pianificate in Tabella A.5 (a) e (b).

**Tabella A.5.** (a) *Quantità di polioli inserite in ognuna delle miscele utilizzate per la prima fase sperimentale.*

Esp.	P1 [g]	P2[g]	P3[g]	P4[g]	P5[g]	P6[g]	P7[g]
1s	0.00	51.03	51.03	0.00	0.00	11.34	0.00
2s	0.00	0.00	51.03	17.01	34.02	7.94	3.40
3s	42.53	8.51	51.03	0.00	0.00	1.13	10.21
4s	8.51	17.01	51.03	25.52	0.00	1.13	10.21
5s	0.00	42.53	17.01	0.00	42.53	0.00	11.34
6s	17.01	17.01	42.53	0.00	25.52	5.67	5.67
1	0.00	42.50	8.50	0.00	51.00	1.10	0.00
2	8.50	51.00	42.50	0.00	0.00	5.70	8.50
3	0.00	42.50	8.50	0.00	51.00	2.30	0.00
4	8.50	51.00	42.50	0.00	0.00	6.80	8.50
5	34.00	0.00	25.50	0.00	42.50	0.00	34.00
6	0.00	42.50	17.00	0.00	42.50	1.10	0.00
7	0.00	0.00	51.00	51.00	0.00	0.00	0.00
8	0.00	0.00	51.00	34.00	17.00	3.40	0.00
9	0.00	51.00	34.00	17.00	0.00	2.30	0.00
10	8.50	0.00	51.00	42.50	0.00	0.00	8.50
11	51.00	0.00	51.00	0.00	0.00	0.00	51.00
12	51.00	0.00	51.00	0.00	0.00	0.00	51.00
13	25.50	0.00	25.50	0.00	51.00	1.10	25.50
14	0.00	0.00	25.50	25.50	51.00	0.00	0.00
15	42.50	8.50	51.00	0.00	0.00	1.10	42.50
16	8.50	51.00	42.50	0.00	0.00	11.30	8.50
17	0.00	25.50	25.50	0.00	51.00	5.70	0.00
18	0.00	0.00	51.00	51.00	0.00	0.00	0.00
19	17.00	0.00	34.00	0.00	51.00	3.40	17.00
20	0.00	0.00	42.50	8.50	51.00	11.30	0.00
21	51.00	8.50	42.50	0.00	0.00	0.00	51.00
22	0.00	0.00	34.00	17.00	51.00	3.40	0.00
23	8.50	51.00	42.50	0.00	0.00	9.10	8.50
24	51.00	0.00	51.00	0.00	0.00	2.30	51.00
25	0.00	51.00	51.00	0.00	0.00	11.30	0.00

**Tabella A.5.** (b) *Quantità di catalizzatori ed agenti espandenti inserite in ognuna delle miscele utilizzate per la prima fase sperimentale.*

Esp.	C1[g]	C2[g]	C3[g]	C4[g]	C5[g]	E1[g]	E2[g]	E3[g]
1s	1.05	0.00	0.53	0.00	0.53	3.36	2.52	2.52
2s	1.05	0.00	0.53	0.00	0.53	3.36	2.52	2.52
3s	1.05	0.00	0.53	0.00	0.53	3.36	2.52	2.52
4s	1.05	0.00	0.53	0.00	0.53	3.36	2.52	2.52
5s	1.05	0.00	0.53	0.00	0.53	3.36	2.52	2.52
6s	1.05	0.00	0.53	0.00	0.53	3.36	2.52	2.52
1	10.20	1.60	0.00	0.50	0.00	0.00	0.80	6.70
2	5.70	1.60	0.00	0.50	0.00	0.00	3.40	2.50
3	9.10	0.00	1.60	0.50	0.00	0.00	1.70	5.00
4	4.50	0.00	1.10	1.10	0.00	0.00	4.20	0.80
5	11.30	0.50	0.00	1.60	0.00	0.00	1.70	5.00
6	10.20	0.50	0.00	1.60	0.00	0.00	4.20	0.80
7	11.30	0.00	0.50	1.60	0.00	0.00	1.70	5.00
8	7.90	1.10	0.00	0.50	0.00	0.50	1.70	5.00
9	9.10	1.10	0.50	0.00	0.50	0.00	0.80	6.70
10	11.30	0.00	1.60	0.00	0.50	0.00	0.80	6.70
11	11.30	0.00	1.60	0.00	0.50	0.00	4.20	0.80
12	11.30	0.50	0.00	1.10	0.50	0.00	4.20	0.80
13	10.20	0.00	1.10	0.00	1.10	0.00	0.80	6.70
14	11.30	0.50	0.00	0.50	1.10	0.00	1.70	5.00
15	10.20	0.50	0.00	0.00	1.60	0.00	3.40	2.50
16	0.00	0.00	0.50	0.00	1.60	0.00	0.80	6.70
17	5.70	0.00	0.50	0.00	1.60	0.00	4.20	0.80
18	11.30	1.60	0.00	0.00	0.50	0.00	4.20	0.80
19	7.90	1.10	0.00	0.50	0.00	0.50	3.40	2.50
20	0.00	0.50	0.50	0.50	0.00	0.50	0.80	6.70
21	11.30	0.00	1.10	0.50	0.00	0.50	1.70	5.00
22	7.90	1.10	0.00	0.50	0.00	0.50	4.20	0.80
23	2.30	0.50	0.00	1.10	0.00	0.50	0.80	6.70
24	9.10	1.10	0.00	0.00	0.50	0.50	0.80	6.70
25	0.00	0.50	0.50	0.00	0.50	0.50	3.40	2.50

Le quantità di sostanza effettivamente utilizzate sono leggermente diverse a causa di errori sperimentali in fase di preparazione dei formulati. In Tabella A.6 (a) e (b), le quantità di sostanza effettivamente impiegate.

**Tabella A.6.** (a) *Quantità di polioli, catalizzatori, agenti espandenti effettivamente utilizzate in ognuna delle prove sperimentali della prima fase.*

Esp.	P1 [g]	P2[g]	P3[g]	P4[g]	P5[g]	P6[g]	P7[g]
1s	0.00	51.05	51.34	0.00	0.00	11.14	0.00
2s	0.00	0.00	51.10	18.43	33.75	7.31	3.52
3s	42.28	8.94	51.03	0.00	0.00	1.02	10.34
4s	8.49	17.36	51.01	25.39	0.00	1.12	10.22
5s	0.00	42.42	16.94	0.00	48.38	0.00	11.88
6s	17.38	17.07	42.32	0.00	25.14	5.55	5.67
1	0.00	41.90	8.80	0.00	50.60	1.20	11.70
2	8.70	50.70	40.40	0.00	0.00	5.70	5.90
3	0.00	44.30	9.70	0.00	50.00	2.50	8.80
4	8.70	55.80	42.10	0.00	0.00	7.10	4.20
5	33.60	0.00	25.00	0.00	43.00	0.00	13.10
6	0.00	41.00	17.20	0.00	44.30	1.50	10.40
7	0.00	0.00	50.30	50.40	0.00	0.00	11.00
8	0.00	0.00	50.80	33.00	18.60	3.50	7.60
9	0.00	50.10	33.70	16.60	0.00	3.50	9.60
10	8.60	0.00	50.60	42.80	0.00	0.00	10.70
11	53.00	0.00	49.90	0.00	0.00	0.00	10.70
12	50.60	0.00	51.70	0.00	0.00	0.00	10.80
13	26.60	0.00	27.10	0.00	50.80	1.20	9.60
14	0.00	0.00	25.90	24.50	52.10	0.00	11.70
15	42.20	9.50	50.10	0.00	0.00	1.40	10.90
16	8.70	50.80	43.10	0.00	0.00	11.30	0.00
17	0.00	24.30	24.50	0.00	51.30	5.70	5.60
18	0.00	0.00	50.70	50.60	0.00	0.00	11.20
19	17.40	0.00	33.60	0.00	49.60	3.60	7.40
20	0.00	0.00	42.90	8.20	50.50	11.30	0.00
21	50.40	9.40	41.20	0.00	0.00	0.00	10.90
22	0.00	0.00	33.90	17.20	49.60	3.90	8.60
23	8.90	50.50	41.20	0.00	0.00	9.20	2.80
24	51.30	0.00	51.30	0.00	0.00	2.70	8.70
25	0.00	50.20	50.40	0.00	0.00	11.00	0.00

**Tabella A.6.** (b) *Quantità di catalizzatori, agenti espandenti effettivamente utilizzate in ognuna delle prove sperimentali della prima fase.*

Esp.	C1[g]	C2[g]	C3[g]	C4[g]	C5[g]	E1[g]	E2[g]	E3[g]
1s	1.02	0.00	0.53	0.00	0.52	3.44	2.52	2.56
2s	1.04	0.00	0.51	0.00	0.53	3.40	2.51	2.50
3s	1.07	0.00	0.53	0.00	0.53	3.36	2.53	2.50
4s	1.05	0.00	0.51	0.00	0.54	3.37	2.52	2.50
5s	1.03	0.00	0.53	0.00	0.54	3.37	2.50	2.50
6s	1.04	0.00	0.52	0.00	0.52	3.35	2.47	2.50
1	1.70	0.00	0.50	0.00	0.00	0.90	6.90	0.80
2	1.60	0.00	0.50	0.00	0.00	3.40	2.50	2.50
3	0.00	1.60	0.50	0.00	0.00	1.70	5.00	1.70
4	0.00	1.10	1.10	0.00	0.00	4.20	0.80	3.40
5	0.50	0.00	1.60	0.00	0.00	1.70	5.00	1.70
6	0.50	0.00	1.60	0.00	0.00	4.40	0.80	3.50
7	0.00	0.50	1.60	0.00	0.00	1.80	5.00	1.70
8	1.10	0.00	0.50	0.00	0.50	1.90	5.00	1.70
9	1.20	0.50	0.00	0.50	0.00	0.80	6.80	0.80
10	0.00	1.60	0.00	0.50	0.00	0.80	6.70	0.80
11	0.00	1.60	0.00	0.50	0.10	4.20	0.80	3.40
12	0.50	0.00	1.10	0.40	0.00	4.20	0.80	3.40
13	0.00	1.10	0.00	1.00	0.00	0.80	6.70	0.80
14	0.50	0.00	0.50	1.10	0.00	1.80	5.00	1.80
15	0.60	0.00	0.00	1.60	0.00	3.40	2.50	2.60
16	0.00	0.50	0.00	1.60	0.00	0.80	6.60	0.80
17	0.00	0.50	0.00	1.60	0.00	4.30	0.80	3.40
18	1.70	0.00	0.00	0.50	0.00	4.40	0.80	3.40
19	1.10	0.00	0.50	0.00	0.50	3.30	2.50	2.60
20	0.50	0.50	0.50	0.00	0.50	0.80	6.70	0.80
21	0.00	1.10	0.50	0.00	0.50	1.70	5.00	1.70
22	1.10	0.00	0.50	0.00	0.50	4.20	0.80	3.40
23	0.50	0.00	1.10	0.00	0.50	0.80	6.70	0.80
24	1.20	0.00	0.00	0.70	0.50	0.80	6.70	0.80
25	0.50	0.50	0.00	0.60	0.50	3.40	2.70	2.50

Di seguito in Tabella A.7 vengono riportati i dati raccolti manualmente durante la seconda prova sperimentale.

**Tabella A.7.** Dati raccolti manualmente durante l'analisi sperimentale dei formulati della seconda fase sperimentale.

Esp.	$T_{mixA}$ [°C]	$T_{mixB}$ [°C]	$T_{amb}$ [°C]	$m_{bicch.B}$ [g]	$m_B$ [g]	$m_{bicch.A}$ [g]	$m_A$ [g]	$m_{resB}$ [g]	$m_{resA}$ [g]	$t_{filo}$ [s]	$t_{tocco}$ [s]	$\rho$ [ $\frac{kg}{m^3}$ ]
1s	21.00	19.00	20.00	29.54	139.74	17.09	189.65	68.70	10.20	170.00	400.00	30.98
2s	21.00	20.00	19.70	29.50	141.18	16.40	189.60	70.54	5.60	150.00	260.00	33.00
3s	20.00	19.00	19.70	29.55	140.04	16.40	190.53	874.79	-0.88	164.00	350.00	32.14
4s	20.50	19.00	19.90	29.50	139.67	16.40	190.69	72.92	-1.18	148.00	280.00	33.00
5s	20.00	19.00	19.90	29.53	146.90	16.40	190.02	73.89	1.29	140.00	275.00	34.50
6s	20.50	19.00	19.80	29.53	169.83	21.04	189.65	76.43	-0.89	143.00	270.00	33.60
1	18.00	19.00	19.70	70.02	138.46	16.89	190.02	122.60	2.44	200.00	350.00	35.10
2	19.00	19.00	19.40	72.72	135.60	16.40	190.85	124.78	4.74	185.00	289.00	39.85
3	19.00	19.00	19.80	71.33	138.90	21.12	189.97	122.91	2.13	155.00	265.00	35.22
4	19.00	19.00	19.80	71.80	141.20	21.12	189.50	127.97	-0.36	151.00	241.00	35.32
5	19.00	19.00	20.20	72.02	138.80	21.12	189.17	127.61	1.08	120.00	254.00	26.53
6	19.00	19.00	20.20	73.17	135.40	21.12	190.08	132.34	-0.22	144.00	260.00	35.16
7	20.00	20.00	20.70	71.47	141.50	20.70	189.49	132.40	4.88	103.00	213.00	28.49
8	20.00	20.00	20.70	71.73	141.50	20.70	190.05	131.38	-1.17	110.00	180.00	31.34
9	20.00	20.00	20.70	73.33	139.20	20.70	189.90	127.97	-3.18	95.00	160.00	34.08
10	19.50	20.00	20.70	72.31	134.50	20.70	189.70	126.30	0.27	180.00	250.00	41.71
11	19.50	20.00	20.70	72.80	142.40	20.70	190.42	121.19	1.27	200.00	500.00	33.25
12	20.00	20.00	20.70	71.40	138.90	20.70	191.99	122.16	2.50	144.00	420.00	29.47
13	20.00	20.00	20.70	76.68	139.40	20.70	191.39	124.16	-2.21	104.00	200.00	31.17
14	20.00	20.00	20.90	75.22	136.20	23.06	191.58	121.59	-1.11	120.00	210.00	32.51
15	20.00	20.00	20.90	73.66	135.70	23.06	189.14	129.00	-0.82	152.00	315.00	36.35
16	20.00	20.50	21.00	71.14	140.80	23.06	189.66	144.26	0.12	154.00	530.00	30.13
17	20.00	20.50	21.00	79.31	140.20	23.06	191.67	128.80	0.39	122.00	217.00	31.35
18	20.00	20.50	21.00	72.33	136.10	23.06	189.92	125.91	2.09	144.00	250.00	35.22
19	20.00	20.00	20.80	70.86	139.60	22.89	189.93	124.65	-0.20	120.00	220.00	32.17
20	20.00	20.00	20.90	72.80	135.10	22.89	195.50	124.58	-0.55	180.00	460.00	34.25
21	20.00	20.00	20.80	73.17	141.30	22.89	190.18	124.35	3.19	108.00	220.00	32.65
22	20.00	20.00	20.90	73.47	149.20	22.89	190.02	130.48	-0.91	190.00	423.00	32.64
23	20.00	20.00	21.00	71.41	139.10	22.89	190.14	118.47	1.12	180.00	360.00	29.00
24	20.50	20.50	21.00	71.66	136.60	22.89	192.05	128.25	-3.52	135.00	280.00	31.63
25	21.00	21.00	21.00	71.48	142.70	22.89	191.08	141.15	2.22	166.00	395.00	29.80



# Nomenclatura

$A$	=	numero di componenti principali o variabili latenti [-]
$a$	=	generica componente principale o variabile latente [-]
$a_f$	=	parametro utilizzato per la filtrazione dei dati sperimentali [-]
$\mathbf{B}$	=	coefficienti di regressione complessivi di un modello pls [-]
$\mathbf{B}_r$	=	coefficienti di regressione per metodi di regressione comuni [-]
rapporto $B/A$	=	rapporto fra le masse di componente $b$ e di componente $a$ [-]
$b_a$	=	coefficienti di regressione interni di un modello pls
$b_f$	=	parametro utilizzato per la filtrazione dei dati sperimentali [-]
$c$	=	numero di componenti di una classe di osservazioni [-]
$c_f$	=	parametro utilizzato per la filtrazione dei dati sperimentali [mm]
$D_{t_K}$	=	distanza misurata dal sensore di distanza al tempo $t_k$
$d_{i,j}$	=	distanza fra due combinazioni in un algoritmo kennard-stone [-]
$d_f$	=	parametro utilizzato per la filtrazione dei dati sperimentali [mm]
$\mathbf{E}$	=	errore residuo dall'approssimazione di $\mathbf{X}$ [-]
$\mathbf{e}$	=	errore residuo dall'approssimazione di $\mathbf{x}$ [-]
eq. di $i$	=	numero di equivalenti di $i$ [eq]
$err_x$	=	errore di predizione di $\mathbf{x}$ [g]
$err_{ass}$	=	errore di assegnazione [%]
$\mathbf{F}$	=	errore residuo dall'approssimazione di $\mathbf{Y}$ [-]
$F_{A,N-A,1-0.95}$	=	valore corrispondente ad una probabilità del 95% di una distribuzione di $F$ con $A$ e $N - A$ gradi di libertà [-]
$G$	=	numero di fattori in una pianificazione degli esperimenti [-]
Indice di $NCO$	=	rapporto fra gli equivalenti di gruppi isocianati effettivi e stechiometrici [-]
$H_{t_K}$	=	altezza della schiuma al tempo $t_k$
$HN_{t_K}$	=	altezza normalizzata della schiuma al tempo $t_k$
$h_0$	=	parametro utilizzato per il calcolo del limite di confidenza di $Q$ [-]
$K$	=	numero di istanti temporali [-]
$L$	=	numero di livelli di variazione di un fattore [-]
$M$	=	numero di variabili in un dataset $\mathbf{Y}$ [-]
$MW_i$	=	massa molare di $i$ [g/mol]
$m$	=	variabile di un dataset $\mathbf{Y}$ [-]
$m_i$	=	massa di $i$ [kg]
$mv_{HN_{t_j}}$	=	velocità di crescita media [s <sup>-1</sup> ]
$mv_{T_{t_j}}$	=	gradiente medio di temperatura [°C s <sup>-1</sup> ]
$N$	=	numero di osservazioni in un dataset [-]
$N_c$	=	numero di combinazioni di possibili di fattori [-]
$N_e$	=	numero di esperimenti [-]
$N_m$	=	numero totale di osservazioni di una classe
$N_t$	=	numero totale di osservazioni per una <i>linear discriminant analysis</i>
$N^\circ OH_i$	=	numero di idrossido di un poliuretano [mgkoh/g poliolo]
$n$	=	osservazione generica di un dataset [-]
$n_c$	=	ennesima iterazione in un algoritmo kennard-stone [-]

<b>P</b>	= matrice dei loadings di <b>X</b> [-]
<b>P<sup>w</sup></b>	= matrice dei loadings di <b>X</b> , pesati secondo la varianza spiegata di ogni pc
<b>p<sub>a</sub></b>	= <i>loading</i> di <b>X</b> lungo la componente principale o variabile latente <i>a</i> [-]
<b>p<sub>a</sub><sup>w</sup></b>	= <i>loading</i> di <b>X</b> lungo la componente principale <i>a</i> [-] pesata secondo la varianza spiegata dalla rispettiva pc
<b>PE<sub>i</sub></b>	= peso equivalente di <i>i</i> [g/eq]
<b>PV<sub>j</sub></b>	= parametro di velocità del formulato <i>j</i> [-]
<b>Q</b>	= matrice dei loadings di <b>Y</b> [-]
<b>Q</b>	= squared prediction error di una osservazione [-]
<b>Q<sub>95%CL</sub></b>	= limite di confidenza del 95% di square prediction error <i>Q</i> [-]
<b>q<sub>a</sub></b>	= loading di <b>Y</b> lungo la componente principale o variabile latente <i>a</i> [-]
<b>S</b>	= matrice di covarianza fra <b>X</b> e <b>Y</b>
<b>S<sub>i,j</sub></b>	= parametro di similarità fra due modelli [-]
<b>T</b>	= matrice degli score di <b>X</b> [-]
<b>T<sub>t<sub>K</sub></sub></b>	= temperatura al tempo <i>t<sub>K</sub></i> [°c]
<b>T<sub>mixA</sub></b>	= temperatura miscela a [°c]
<b>T<sub>mixB</sub></b>	= temperatura miscela b [°c]
<b>T<sub>amb</sub></b>	= temperatura ambientale [°c]
<b>T<sup>2</sup></b>	= hotelling <i>T<sup>2</sup></i> di una osservazione [-]
<b>T<sup>2</sup><sub>95%CL</sub></b>	= limite di confidenza del 95% di hotelling <i>T<sup>2</sup></i> [-]
<b>T<sup>2</sup><sub>score calcolato</sub></b>	= valore di <i>T<sup>2</sup></i> per lo score calcolato [-]
<b>t<sub>a</sub></b>	= score di <b>X</b> lungo la componente principale o variabile latente <i>a</i> [-]
<b>t<sub>des</sub></b>	= score desiderato da una inversione di un modello [-]
<b>t<sub>calc</sub></b>	= score calcolato attraverso ottimizzazione [-]
<b>t<sub>K</sub></b>	= istante di tempo <i>k</i> [s]
<b>t<sub>filo</sub></b>	= tempo di filo [s]
<b>t<sub>tocco</sub></b>	= tempo di tocco [s]
<b>t<sub>a</sub></b>	= tempo di analisi [s]
<b>U</b>	= matrice degli score di <b>Y</b> [-]
<b>u<sub>a</sub></b>	= score di <b>Y</b> lungo la componente principale o variabile latente <i>a</i> [-]
<b>V</b>	= numero di variabili in un dataset <b>X</b> [-]
<b>Vol<sub>PU</sub></b>	= volume totale di materiale espanso [m <sup>3</sup> ]
<b>v</b>	= variabile di un dataset <b>X</b> [-]
<b>v<sub>HN<sub>t<sub>j</sub></sub></sub></b>	= velocità di crescita [s <sup>-1</sup> ]
<b>v<sub>T<sub>t<sub>j</sub></sub></sub></b>	= gradiente di temperatura [°c s <sup>-1</sup> ]
<b>v<sub>a</sub></b>	= loading normalizzato di <b>X</b> lungo la variabile latente <i>a</i> [-]
<b>X</b>	= dataset di <i>N</i> osservazioni e <i>V</i> variabili [-]
<b>W</b>	= matrice dei weights [-]
<b>w<sub>a</sub></b>	= weight lungo la componente principale o variabile latente <i>a</i> [-]
<b>x</b>	= osservazione del dataset <b>X</b> [u.d.m.]
<b>x̂</b>	= approssimazione di <b>x</b> [u. d. m.]
<b>x<sup>real</sup></b>	= massa di catalizzatore realmente utilizzata nella prima campagna sperimentale [g]
<b>x<sup>pred</sup></b>	= massa di catalizzatore predetta dal modello [g]
<b>x<sub>autoscalata</sub></b>	= valore di <i>x</i> sottoposto ad autoscaling [-]
<b>x<sub>i</sub></b>	= misura di una osservazione [u.d.m.]
<b>x</b>	= prima variabile generica [u.d.m.]

$\mathbf{Y}$	= dataset di $N$ osservazioni e $M$ variabili [-]
$y$	= seconda variabile generica [u.d.m.]
$\mathbf{y}$	= osservazione del dataset $\mathbf{Y}$ [u.d.m]
$\hat{\mathbf{y}}$	= approssimazione di $\mathbf{y}$ [u.d.m]

Lettere greche:

$\alpha_j$	= angolo dello <i>score</i> di un formulato $j$ [°]
$\delta_c$	= funzione discriminante [-]
$\Delta H_r^0$	= entalpia di reazione in condizioni standard [-]
$\varepsilon$	= rumore [u.d.m.]
$\eta$	= parte sistematica dell'informazione [u.d.m.]
$\theta$	= angolo fra il piano originale e la pc [°]
$\theta_i$	= parametri utilizzati per il calcolo del limite di confidenza di $Q$ [°]
$\mathbf{\Lambda}$	= matrice degli autovalori [-]
$\lambda$	= autovalore [-]
$\mu$	= media [u.d.m.]
$\mu_m$	= media delle osservazioni di una classe
$\pi_m$	= rapporto fra numero di osservazioni di una classe e il numero di osservazioni totali
$\rho_{bulk}$	= densità di bulk di un poliuretano espanso [kg/m <sup>3</sup> ]
$\Sigma$	= matrice di covarianza fra le osservazioni di una stessa classe [-]
$\sigma_{xy}$	= covarianza [u.d.m.]
$\chi^2_{1-0.95}$	= valore corrispondente ad una probabilità del 5% di una distribuzione $\chi^2$ [-]

Acronimi

CL	=	Confidence Limit
CCD	=	Central Composite Design
C	=	Catalizzatore
DoE	=	Design of Experiments
E	=	Agente Espandente
FFD	=	Fractional Factorial Design
F	=	Formulato
KS	=	Kennard - Stones
LV	=	Latent Variable
MDI	=	Difenil-Metano Di-Isocianato
NF	=	Nuovo Formulato
PCA	=	Principal Components Analysis
PLS	=	Projection to Latent Structure o Partial Least Squares
PC	=	Principal Components
P	=	Poliolo
RMSECV	=	Root of Minimum Squared Error in Cross Validation
SNV	=	Standard Normal Variate
TDI	=	Toluene Di Isocianato



# Riferimenti bibliografici

- Al-Moameri H., Zhao Y., Ghoreishi R., Suppes G.R. (2016). Simulation Blowing Agent Performance, Cell Morphology, and Cell Pressure in Rigid Polyurethane Foams, *Industrial & Engineering Chemistry Research* 2016 55 (8), 2336-2344; DOI: 10.1021/acs.iecr.5b04711
- Camacho J., Picò J., Ferrer A. (2008). Bilinear modelling of batch processes. Part I: theoretical discussion. *Journal of Chemometrics*, Volume 22, Issue 5, p.299-308. <https://doi.org/10.1002/cem.1113>
- De Jong S. (1993). SIMPLS: an alternative approach to partial least squares regression. *Chemometrics and Intelligent Laboratory Systems*, 18: p.251-263.
- Fisher R.A. (1936). The Use of Multiple Measurements in Taxonomic Problems. *Annals of Eugenics*, Vol. 7, p. 179–188. <https://digital.library.adelaide.edu.au/dspace/handle/2440/15227>.
- Gomes C., Claeys-Bruno M., Sergent. M. (2018). Space-filling designs for mixtures. *Chemometrics and Intelligent Laboratory Systems* 111-127.
- Hart H., Hadad C.M., Craine L.E., Hart D.J. (1985). *Chimica Organica*, Seconda Edizione.
- Hartmann. K., Krois J., Waske B. (2018). E-Learning Project SOGA: Statistics and Geospatial Data Analysis. Department of Earth Sciences, Freie Universitaet Berlin.
- Ionescu M. (2016). *Chemistry and Technology of Polyols for Polyurethanes*, 2nd Edition, Volume I.
- Lee S.T., Ramesh N.S. (2004). *Polymeric Foams: Mechanisms and Materials*.
- Ottavian M., Facco P., Fasolato L., Barolo M. (2012). Multispectral data classification using similarity factors. *Chemometrics and Intelligent Laboratory System*, 118: p. 12-23.
- Ryan T.P. (2007). *Modern Experimental Design*. Wiley. ISBN: 978-0-471-21077-1
- Sonnenschein M.F. (2015). *Polyurethanes Science, Technology, Markets, and Trends*.
- Szycher M. (2013). *Szycher's Handbook of Polyurethanes*, 2nd Edition.
- Skleničková K., Abbrent S., Halecký M., Kočí V., Beneš H. (2020). Biodegradability and ecotoxicity of polyurethane foams: A review, *Critical Reviews in Environmental Science and Technology*, DOI: 10.1080/10643389.2020.1818496
- Tomba E., Barolo M., Garcia-Muñoz S. (2016). General Framework for Latent Variable Model Inversion for the Design and Manufacturing of New Products. Computer-Aided Process Engineering Laboratory (CAPE-Lab), Dipartimento di Ingegneria Industriale, Università di Padova, via Marzolo 9, 35131 Padova PD, Italy; Pfizer Worldwide R&D, 445 Eastern Point Road, Groton, Connecticut 06340, United States
- Wise B.M., Gallagher N. B. (1996). The process chemometrics approach to process monitoring and fault detection. *Journal of Process Control*, Volume 6, Issue 6: p.329-348, ISSN 0959-1524, [https://doi.org/10.1016/0959-1524\(96\)00009-1](https://doi.org/10.1016/0959-1524(96)00009-1).
- Woods G. (1987). *The ICI Polyurethanes Book*. 1st Edition. Wiley.
- Xu T., Ting Y. (2004). Optimisation on bioleaching of incinerator of fly ash by aspergillus niger – use of central composite design. Department of Chemical and Biomolecular Engineering, National

University of Singapore, Kent Ridge Crescent, Singapore 119260, Singapore. *Enzyme and Microbial Technology*, 35: 444–454.

# Ringraziamenti

Il lavoro di Tesi è dedicato a tutti coloro che in questi mesi ed anni mi sono stati accanto e mi hanno aiutato a raggiungere questo traguardo, senza i quali tutto ciò non sarebbe stato possibile.

Innanzitutto, sarò sempre grato i miei genitori, Flavia e Daniele, che con molta pazienza e tanti sacrifici mi hanno sostenuto e permesso di percorrere questa strada arrivando fin qui, passo dopo passo. Ringrazio anche Luca, che mi ha stimolato a dare sempre il massimo.

Ad Alice, per essermi stata a fianco in tutti i momenti, in particolare durante la realizzazione di questo lavoro, per avermi dato una spinta e una motivazione a puntare sempre in alto e avermi insegnato che c'è sempre da migliorare.

Ad Alex, Nicolò, Leonardo, Riccardo e Leonardo, per aver portato pazienza in questi anni in cui ho dato priorità all'università invece che a loro; A Giulia e Ottavio per aver affievolito tutti i momenti di stress e difficoltà nel percorso universitario, e per aver reso l'ambiente di studio sereno e piacevole, riuscendo sempre a strapparmi un sorriso; A Linda, Alberto e Serena, per tutti i bei momenti durante i lavori di gruppo affrontati nel corso della laurea magistrale, e per averci messo sempre il massimo impegno mirato a produrre dei risultati ottimali. A Nicola, Elena, Francesco, Martina e Fabrizio per tutte le sere passate a ricordare gli anni delle scuole superiori, dove mi sono appassionato per la prima volta alla Chimica e alla sua parte ingegneristica. Per questo, dedico questo lavoro anche alle professoresse G. Casati e M. Zanato, che mi hanno trasmesso la passione verso questo campo di studi ed avermi trasmesso le basi che mi sono state poi utili in tutti gli anni universitari.

Ringrazio il professor P. Facco, relatore di questa Tesi, che mi ha proposto il progetto e mi ha fornito indicazioni, correzioni e consigli fondamentali alla sua realizzazione e alla sua buona riuscita; ringrazio anche E.A. Feffin, per il sostanziale aiuto nello sviluppo dei metodi basati sull'inversione dei modelli PLS e tutta la squadra di Europoliuretani s.r.l. per avermi permesso di lavorare a questo progetto ed avermi accolto in un contesto gradevole ed ospitale, in particolare ringrazio l'Ingegnere N. Mela, che mi ha aiutato alla realizzazione di questo lavoro e aver assecondato con impegno ed entusiasmo tutte le mie idee e proposte. Infine, ringrazio Fabio e i colleghi di MPM s.r.l., che mi hanno fatto sentire a casa e mi hanno permesso di lavorare ed autosostenermi negli ultimi mesi.