POLITECNICO DI TORINO

Corso di Laurea Magistrale in Ingegneria Gestionale

Tesi di Laurea Magistrale

Studio degli effetti della Complessità di prodotto nei processi di Assemblaggio e Disassemblaggio sui tempi di processo, sulla generazione dei difetti e sull'apprendimento degli operatori

Relatori:

Prof. Gianfranco Genta

Dott.ssa Elisa Verna

Candidato:

Giammarco Arcieri

Anno Accademico 2021/2022

INDICE

INTRODUZIONE

Il seguente elaborato è frutto di un progetto di sperimentazione di assemblati molecolari, nell'ambito dell'ingegneria della qualità, svoltosi tra i mesi di settembre 2021 e dicembre 2021 condotto presso il Politecnico di Torino. Lo scopo della campagna sperimentale è di studiare gli effetti della complessità di prodotto nei processi di assemblaggio e disassemblaggio sui tempi di processo, sulla generazione dei difetti e sull'apprendimento degli operatori. Allo scopo di condurre l'esperimento, sono stati selezionati alcuni studenti del corso di *Ingegneria della Qualità* del Prof. Maurizio Galetto i quali hanno eseguito l'assemblaggio e il disassemblaggio di dodici strutture molecolari preselezionate, effettuato il controllo di qualità durante l'esecuzione delle due fasi e riportato i risultati ottenuti in apposite tabelle.

I giorni di sperimentazione sono stati suddivisi in giornate lavorative in cui sono state eseguite le *prove Randomiche* e giornate lavorative in cui sono state eseguite le *prove Ripetute*.

Nelle giornate lavorative in cui sono state eseguite le *prove Randomiche* gli studenti avevano il compito di assemblare e disassemblare tutte e dodici le strutture molecolari secondo un ordine randomico precedentemente stabilito. Gli studenti, massimo 14 studenti a giornata per rispettare le norme Anti-Covid, sono stati divisi coppie e ad ogni studente della coppia è stata affidata una mansione da svolgere.

Durante le fasi di assemblaggio e di disassemblaggio è stato effettuato il controllo della qualità della fase, ovvero uno degli studenti ha controllato che le due fasi fossero svolte correttamente e segnalato la presenza di eventuali difetti riscontarti.

Nelle giornate lavorative in cui sono state eseguite le *prove Ripetute*, invece, gli studenti avevano il compito di assemblare e disassemblare solo alcune delle dodici strutture molecolari definite ma l'assemblaggio e il disassemblaggio di ogni struttura è stato ripetuto un numero dichiarato di volte, per poter valutare l'effetto di apprendimento. Anche in queste giornate, durante le fasi di assemblaggio e di disassemblaggio, è stato effettuato il controllo della qualità della fase.

Il seguente elaborato di tesi si struttura in quattro capitoli.

Il primo capitolo, *Complessità di Assemblaggio e di Disassemblaggio,* è un'introduzione ai concetti principali utilizzati nell'elaborato. In esso si forniscono la storia dell'evoluzione delle definizioni della complessità oggettiva e della complessità soggettiva.

Il secondo, *Pianificazione del Caso di Studio*, presenta il caso di studio e i suoi obiettivi descrivendone lo scopo, le molecole analizzate, i dati raccolti, l'organizzazione e i criteri di valutazione della complessità soggettiva e i metodi di calcolo della complessità oggettiva.

Il terzo capitolo, *Metodologie di Analisi*, descrive la metodologia utilizzata per analizzare i dati, spiegando come ricavare le informazioni principali e come scegliere il modello adatto ad interpolare i dati osservati.

Il quarto ed ultimo capitolo, *Risultati Sperimentali e Analisi dei Risultati,* presenta i dati raccolti in fase di sperimentazione e vengono presentati i risultati ottenuti dalle varie analisi svolte.

CAPITOLO 1: Complessità di Assemblaggio e di Disassemblaggio

La crescente richiesta da parte del mercato di prodotti vari e con componenti sempre più articolati ha generato un aumento della complessità dei prodotti e degli errori di qualità originati nei sistemi di produzione.
L'aumento della complessità innesca una serie di effetti come l'aumento dei costi totali, la nascita di problemi operativi, il deterioramento della qualità di produzione e il deterioramento dell'efficienza di progettazione, di funzionamento, di gestione e di manutenzione del sistema (Schuh, 2015). Non solo, i sistemi tendono a diventare meno reattivi al cambiamento, più difficili da gestire e i tempi di produzione e di consegna dei prodotti tendono a diventare più lunghi.

In fase di assemblaggio e disassemblaggio, lavorare con prodotti complessi può comportare un abbassamento della qualità e alla nascita di un maggior numero di difetti dovuti alle operazioni svolte dagli operatori. Quest'ultimi, infatti, dovendo lavorare con prodotti più complessi, riceveranno maggiori istruzioni e quindi, avranno la possibilità di eseguire le operazioni con più opzioni di manovra. Secondo Zhu (2008), più opzioni di assemblaggio sono disponibili per l'operatore, più è probabile che si verifichino degli errori legati alla fase di assemblaggio.

Gli operatori delle fasi di processo ricoprono un ruolo fondamentale per l'abbassamento dei difetti di prodotto. Tramite le ricerche condotte da Vineyard (1999) e, successivamente riprese da Shibata (2002), è stato dimostrato che circa il 40% dei difetti totali deriva da un errore commesso dall'operatore, il quale è sottoposto a numerosi sforzi cognitivi e fisici che aumentano con il grado di complessità totale del sistema.

In quest'ottica, progettare correttamente le mansioni che saranno affidate agli operatori risulta essere fondamentale in quanto, se progettate male si possono raggiungere i limiti cogniti e fisici dell'uomo che influiscono negativamente sull'incertezza creata dalla varietà di prodotto (Alkan, 2016).

Per Samy e ElMaraghy (2012) la gestione della complessità può aiutare a ridurre i costi e i tempi di produzione aumentando così la produttività, la redditività e la qualità dei prodotti. Sulla base degli studi condotti da Falck e Rosenqvist (2019), i tempi, i costi di processo e la complessità di prodotto sono strettamente correlati tra loro. Per ridurre al minimo i tempi e i costi e, contemporaneamente, per aumentare l'efficienza delle operazioni e la qualità stessa dei prodotti finiti è necessario che strategie complesse siano eliminate o comunque ridotte al minimo (Alkan, 2019).

Lo studio e la comprensione della complessità risultano imprescindibili, oltre che alla riduzione dei problemi sopracitati, per la determinazione e la risoluzione dei punti di stress dei sistemi di produzione e per la scelta di una strategia più adeguata al ciclo di vita dei prodotti. Riuscire a studiare la

complessità di un sistema, riuscire ad identificarla e a ridurla risulta un passo necessario ed obbligatorio per poter rispondere adeguatamente alle richieste del mercato. Questo studio è consigliabile effettuarlo durante le prime fasi di progettazione in quanto il costo per la correzione dei difetti aumenta esponenzialmente tanto più tardi essi sono individuati.

In tale ottica, un aumento della complessità nei sistemi di produzione è accettabile solo se si riescono a migliorare le capacità, l'usabilità e le prestazioni del sistema, altrimenti risulta essere un problema (Samy e ElMaraghy, 2012).

La complessità è un concetto astratto la cui natura e definizione precisa, universale e ampiamente accettata non è stata ancora raggiunta nonostante le varie ricerche in diverse discipline scientifiche, come la biologia o la fisica.

Generalmente, la complessità è definita come la caratteristica di un sistema il cui comportamento globale non è immediatamente riconducibile a quello delle singole parti, ma dipende dal modo i cui esse agiscono. Un sistema è concepito come un insieme organico di parti interagenti tra di loro. In letteratura, la complessità è stata studiata ampiamente in diverse aree di ricerca ottenendo così diverse visioni e definizioni in base agli obiettivi posti.

Secondo Peng Liu e Zhizhong (2012), la complessità può essere raggruppata in tre gruppi:

- Complessità strutturalista;
- Complessità del fabbisogno di risorse o di impiego;
- Complessità dell'interazione.

La complessità strutturalista è associata al numero di elementi totali presenti nel processo e come questi sono in relazione tra di loro. Secondo questa definizione la complessità è una funzione tra la complessità dei singoli elementi del sistema e il modo in cui essi entrano in relazione. Un processo, quindi, risulta tanto più complesso quanti più elementi sono presenti e interconnessi tra loro.

Robert Wood (1986) ha rielaborato questa visione di complessità introducendo e mettendo in relazione degli elementi statici, come la progettazione del processo, con degli elementi dinamici. Egli ha ipotizzato che la complessità strutturalista potesse essere scomposta in complessità dei componenti, che fa riferimenti agli elementi del processo, complessità coordinativa, che fa riferimento alle informazioni affidata ad inizio processo e agli elementi facenti parte di quest'ultimo, e la complessità dinamica, che misura la stabilità della relazione tra informazione ed elementi.

Dong Han Ham, Jinkyun Park e Wondea Jung (2012) sulla base delle ricerche precedentemente

condotte sulla complessità strutturalista, hanno definito un processo come un sistema astratto dove il livello di complessità del sistema descrive la complessità del processo. Secondo questa teoria, un processo è composto da tre aspetti fondamentali: un aspetto funzionale, basato sulla ricerca degli obiettivi, uno aspetto comportamentale, basato sulla ricerca delle informazioni necessarie, e un aspetto strutturale, basato sulla ricerca delle forme strutturali.

I risultati ottenuti, usando modelli che si basano sulla definizione della complessità strutturalista, analizzano i sistemi studiando le caratteristiche dei componenti, come la dimensione, l'architettura del sistema e le informazioni associate.

La complessità del fabbisogno di risorse fa riferimento alle richieste fisiche e mentali impiegate dell'operatore e ai requisiti necessari per acquisire ed elaborare le informazioni.

Il termine "risorse" si riferisce alle risorse intrinseche dell'operatore come l'acquisizione e l'elaborazione delle informazioni mediante risorse visive, uditive o la sua abilità manuale. Secondo questa visione, un processo aumenta la propria complessità tante più richieste vengono affidate all'operatore poiché l'elaborazione delle istruzioni comportano un consumo maggiore di risorse. La complessità del fabbisogno delle risorse non può non considerare la complessità intrinseca del processo in quanto, aumentando gli elementi costituenti e la loro varietà, lo sforzo mentale e fisico richiesto all'operatore aumenta.

L'ultimo gruppo, la complessità dell'interazione, considera la complessità come il prodotto dell'interazione tra caratteristiche del processo e le caratteristiche dell'esecutore.

Riallacciandosi a questa visione, Bystrom e Barfield (1999) hanno ipotizzato che la percezione della complessità dell'operatore deve essere necessariamente presa in considerazione quando si modella la complessità del processo in quanto un obiettivo dello stesso processo può essere interpretato, raggiunto e percepito in maniera differente a seconda dell'operatore. Questo comporta che uno stesso sistema, con una certa complessità oggettiva, può essere percepito differentemente da due operatori andando ad aumentare, o diminuire, la complessità del sistema.

I modelli che adottano quest'ottica di complessità classificano i processi in cinque gruppi a seconda di come sono elaborate le informazioni, le procedure e i risultati. Nello specifico:

- Processi con automatica elaborazione delle informazioni, nei quali le informazioni e il risultato sono completamente determinabili;
- Processi con normale elaborazione delle informazioni, nei quali le informazioni sono quasi determinabili ma che richiedono qualche decisione specifica per ottenere dei risultati;

- Processi con decisioni normali, nei quali le decisioni specifiche risultano essere fondamentali per il conseguimento del risultato;
- Processi con decisioni conosciute e genuine, nei quali il risultato e le informazioni sono determinabili ma si ha una procedura coerente;
- Procedure sconosciute e genuine, nei quali il risultato, la procedura e le informazioni sono non strutturati e sconosciuti.

Frizelle e Woodcock (1995) introducono un nuovo concetto di complessità distinguendola in complessità statica e complessità dinamica.

La complessità statica è il tipo di complessità indipendente dal tempo che fa riferimento alla struttura del sistema, nello specifico ai tipi di sottoinsiemi, alla varietà degli elementi costituenti e alle forze delle interazioni tra loro (Rodriguez e Toro, 2004). Si distingue in complessità statica combinatoria e periodica in base a come cresce il livello di incertezza, ovvero se aumenta in un certo tempo indeterminato o se si ferma occasionalmente in un determinato processo e poi torna ai livelli iniziali.

È definita complessità dinamica, o operativa, l'imprevedibilità influenzata dal tempo che fa riferimento al comportamento e alle caratteristiche operative del sistema. Essa si manifesta a causa della mancata soddisfazione dei requisiti funzionali, dovuta alla non comprensione delle informazioni, alla scarsa conoscenza del sistema, all'incapacità nel gestire una grande quantità di componenti e di interazioni o, ancora, all'imprevedibilità del sistema stesso (ElMaraghy, 2012).

La complessità dinamica, a sua volta, si può distinguere, a seconda della causa della mancata soddisfazione dei requisiti funzionali, in reale ed immaginaria. Infatti, è definita come complessità reale dinamica la complessità generata dalla scarsa conoscenza del contenuto informativo e come complessità immaginaria dinamica quella causata dall'imprevedibilità del sistema.

Partendo dai risultati ottenuti dalle ricerche di Frizelle e Woodcock (1995), Rodriguez e Toro (2004) hanno classificato la complessità in: complessità dei componenti, legata alla geometria dei componenti costituenti il sistema, e complessità dell'assemblaggio, dipendente dalla scomposizione strutturale del prodotto e dal numero di operazione necessaria per effettuare l'assemblaggio del prodotto finito.

In generale, la complessità di un qualsiasi sistema produttivo può essere scomposta ed analizzata secondo due punti di vista: la prospettiva soggettiva e la prospettiva oggettiva.

La prospettiva oggettiva fa riferimento direttamente alle caratteristiche proprie del processo mentre la prospettiva soggettiva considera la complessità come il prodotto tra le caratteristiche proprie del

processo e le caratteristiche intrinseche dell'operatore. Quest'ultima può essere descritta come la complessità percepita dall'operatore e dipende da numerosi fattori, come la strategia progettata ed adoperata dall'operatore per eseguire le operazioni del processo, la difficoltà oggettiva delle mansioni che deve effettuare, la capacità di comprensione e di gestione del sistema con cui interagisce, la complessità oggettiva del sistema, il livello di esperienza e di formazione dell'operatore ma anche dal suo grado di coinvolgimento e dal contesto di lavoro (Liu e Li, 2012 e Alkan, 2019).

1.1 Complessità Soggettiva

Per poter valutare la complessità di un processo non è sufficiente analizzare solamente la complessità intrinseca del processo, ma è necessario determinare e considerare la complessità percepita dall'operatore. La complessità rilevata da un operatore umano fa emergere una natura soggettiva che la rende dipendente dalle personali interpretazioni e dal processo stesso (Mattson, 2013).

Con la seguente visione di complessità può accade, quindi, che un operatore, non dotato di un'adeguata conoscenza o di consoni strumenti tecnici, percepisca un processo più complicato di quanto lo sia oggettivamente o potrebbe accadere che due operatori differenti, dotati della stessa conoscenza e di consoni strumenti tecnici, percepiscano un diverso grado di complessità di uno stesso sistema.

Secondo Zaeh (2009), il coinvolgimento degli operatori dovrebbe essere basato su tre fattori connessi tra di loro: fattori temporali, cognitivi e basati sulla conoscenza. Il fattore temporale considera il tempo necessario per effettuare la mansione assegnata all'operatore, il fattore cognitivo valuta il tempo di elaborazione delle informazioni necessarie per effettuare correttamente le mansioni e il fattore basato sulla conoscenza osserva il grado di familiarità dei componenti e del sistema di produzione.

Per poter catturare il grado di complessità di un sistema percepito da un operatore, nella letteratura, sono stati proposti metodi di indagine basati su questionari strutturati e su interviste. In particolare, il questionario permette di rilevare dati e ottenere informazioni di natura quantitativa, facilmente generalizzabili, che possono essere analizzati, permettendo al soggetto di esprimere una preferenza personale, eliminando la componente oggettiva. Per ottenere dei risultati affidabili è necessario stabilire correttamente i criteri del questionario in maniera da non creare confusione nei soggetti e rischiando di ricevere risposte imprecise. Il questionario deve essere costruito in modo da risultare efficace e di facile comprensione al fine di ottenere dei risultati significativi. La costruzione del questionario è il passaggio più importante poiché, dalle domande che si pongono e dalla loro chiarezza, dipendono la qualità e la quantità di informazioni ottenute.

Per la determinazione della complessità soggettiva, Alkan (2019) ha proposto l'adozione di 16 criteri, definiti criteri di bassa complessità dell'assemblaggio manuale, che includono sia fattori fisici che cognitivi. I criteri si basano sugli studi congiunti di Falck, Örtengren e Rosenqvist (2012) effettuati su alcune aziende manifatturiere svedesi che si pongono l'obiettivo di aumentare l'efficienza generale, la qualità dell'assemblaggio e la produttività, mirando a ridurre gli errori generati durante la fase di assemblaggio ed i costi necessari per individuarli e correggerli.

Per poter valutare i criteri sottoposti ai soggetti tramite questionario è necessario definire ed impiegare delle scale di misura. La tipologia di scala di misura che meglio si adatta all'impiego è una scala ti tipo ordinale. Le scale di misura si possono distinguere in quattro tipologie, a seconda delle proprietà formali delle quali godono e a seconda di quali operazioni permettono l'impiego.

La scala nominale è il livello più basso di misurazione ed è un livello puramente qualitativo. La si può adoperare quando i risultati osservati si possono raggruppare in categorie qualitative, ovvero in gruppi in cui è espressa la qualità senza l'utilizzo di valori numerici. La scala nominale è caratterizzata dalla proprietà di esclusività: gli elementi appartenenti ad una stessa categoria qualitativa sono tra loro equivalenti, mentre tutti gli elementi appartenenti a categorie differenti, sono tra loro differenti. Quindi, un valore attribuito ad una categoria della scala rappresenterà solamente gli elementi appartenenti a quella categoria. L'unica operazione definita in questa tipologia di scala di misura è il conteggio degli elementi facenti parte alle categorie qualitative definite. La scala ordinale contiene una quantità di informazioni superiore rispetto a quella nominale poiché utilizza modalità sequenziali. Questa tipologia di scala di misura, oltre a godere della proprietà dell'esclusività, gode della proprietà dell'ordinamento, che permette di distinguere e ordinare le diverse categorie. Non è tuttavia possibile esprimere il concetto di distanza tra le categorie. Con i valori della scala ordinale non è possibile effettuare nessuna operazione aritmetiche poiché si assegnano alle categorie dei valori arbitrari. Un livello superiore di informazioni è recepito mediante l'utilizzo della scala intervallo, la quale si basa sulla definizione di intervalli costanti ed uniformi, in quanto aggiunge la proprietà di poter quantificare la distanza e la proprietà di esprimere la differenza tra due coppie di valori. In essa, quindi, è possibile stabilire se due valori sono equivalenti tra di loro, determinare una relazione d'ordine ed esprimere i valori con un'unità di misura. Il poter assegnare un'unità di misura ai valori permette l'impiego delle operazioni aritmetiche sulle differenze dei valori ma non tra i valori stessi poiché il punto di origine (lo zero) e l'unità di misura sono arbitrarie. L'ultima tipologia di scala di misura è la scala di rapporti. Essa presenta le stesse caratteristiche di una scala intervallo, con il vantaggio di avere un'origine reale, ovvero uno zero assoluto che rappresenta la quantità nulla. Questa ultima caratteristica permette l'impiego delle operazioni aritmetiche su tutti i valori e non solo per le differenze.

Come detto precedentemente, l'impiego di un questionario rende necessaria la definizione di una scala di misura. In letteratura la tipologia di scala maggiormente utilizzata è la scala ordinale.

1.2 Complessità Oggettiva

Secondo la prospettiva oggettiva, la complessità è definita come una proprietà quantificabile di un processo, la quale valuta le sue caratteristiche, senza considerare le caratteristiche degli operatori (Alkan, 2019). Il modello *Sinha-de Weck-Alkan* permette di determinare la complessità oggettiva, detta anche strutturale, come una funzione della complessità di manipolazione dei singoli componenti, della complessità di interazione tra i componenti e della complessità topologica, ovvero dell'architettura risultante (Alkan & Harrison, 2019).

La formula per il calcolo della complessità oggettiva di una generica struttura molecolare è la seguente:

$$C = C_1 + C_2 \cdot C_3 \tag{1}$$

dove:

C rappresenta la complessità strutturale;

C_1 la complessità di manipolazione (handling complexity);

C_2 la complessità delle connessioni (complexity of connections);

C_3 la complessità topologica (topological complexity).

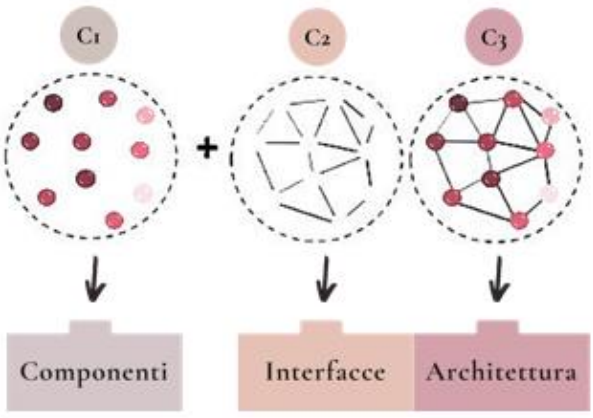


Figura 1 - Complessità di Manipolazione, Complessità delle Connessioni e Complessità Topologica

La Complessità di Manipolazione (C_1) è definita come la somma delle complessità dei singoli componenti ed è descritta mediante la formula:

$$C_1 = \sum_{p=1}^{N} \gamma_p \tag{2}$$

dove:

N rappresenta il numero totale di componenti;

γ_p rappresenta la complessità di manipolazione del singolo componente p.

La complessità di manipolazione del componente γ_p fa riferimento alla difficoltà tecnica nel gestire il singolo componente, ovvero è lo sforzo richiesto all'operatore per programmare, usare o controllare il componente, senza tenere in considerazione le difficoltà riferite alle interfacce del componente e delle informazioni architetturali del sistema (Alkan & Harrison, 2019).

Il termine γ_p è calcolato come segue:

$$\gamma_p = 5(1 - e^{-(\sum_{i=1}^{m} k_i n_i)}) \tag{3}$$

dove:

m rappresenta il numero totale di tipologie di elementi che costituiscono il componente;

n_i è il numero totale di elementi appartenenti alla specifica tipologia i;

k_i, il cui valore è compreso tra 0 e 1, rappresenta il parametro della funzione esponenziale corrispondente all'elemento appartenete alla tipologia i.

Per calcolare la complessità del componente si è deciso di adottare una funzione esponenziale per due precisi motivi: il primo è per definire un punteggio di complessità compreso tra 0 e 5 in modo da poter utilizzare un range globale per tutti i componenti e il secondo motivo è per limitare il punteggio massimo di complessità soprattutto per quei componenti per i quali lo sviluppo e la gestione vanno oltre la comprensione umana. La complessità percepita da un individuo, infatti, non può eccedere i limiti di comprensione dell'osservatore (Alkan & Harrison, 2019). Per facilità di utilizzo, nonostante la validità della formula precedente, Alkan (2019) ha stimato il termine γ_p in funzione del tempo di interazione con il componente generico durante la fase di assemblaggio, includendo il tempo per la localizzazione del contenitore giusto, il tempo necessario per recuperare il componente dal contenitore e il tempo necessario per ritornare alla postazione di lavoro.

La Complessità delle Connessioni (C_2) indica lo sforzo associato allo sviluppo e alla gestione di ogni interazione a coppie. È definita come la somma delle complessità delle connessioni tra coppie del sistema ed è calcolata come segue:

$$C_2 = \sum_{p=1}^{N-1} \sum_{r=p+1}^{N} \varphi_{pr} \cdot A_{pr} \quad (4)$$

dove il termine A_{pr} rappresenta la matrice di adiacenza binaria e permette di visualizzare la possibile interazione tra due specifici componenti. A_{pr} assume valore 1, se esiste la connessione tra i componenti p ed r, e valore pari a 0, se non esiste alcuna connessione tra loro.

Il termine φ_{pr} rappresenta la complessità nel realizzare la connessione tra i componenti p ed r e può essere espresso mediante la relazione che esiste tra gli elementi e la natura della connessione. φ_{pr} è calcolato come:

$$\varphi_{pr} = (\sum_{k=1}^{l} C_k)^{\frac{(\alpha_i + \alpha_j)}{2}} \quad (5)$$

dove:

C_k è il coefficiente di interfaccia, il quale definisce la natura della connettività;
I è il numero di connessioni tra gli elementi i e j.

Alkan (2019) ha proposto di stimare φ_{pr} come il tempo necessario per connettere due elementi in condizioni isolate. Questo tempo comprende, oltre al tempo necessario per collegare i due componenti al legame, il tempo impiegato per individuare gli elementi, il tempo necessario per effettuare il posizionamento degli elementi e il tempo impiegato per effettuare il controllo finale.

L'ultima complessità, la Complessità Topologica (C_3), rappresenta la complessità della dipendenza tra le entità del sistema, ovvero cattura gli effetti dell'architettura del sistema e aumenta passando da strutture con architetture centralizzate a strutture con architetture più distribuite. La Complessità Topologica è una misura globale che considera la disposizione intrinseca delle connessioni ed è definita mediante la formula:

$$C_3 = \frac{E_A}{N} = \frac{\sum_{q=1}^{N} \delta_q}{N} \quad (6)$$

dove:

E_A rappresenta l'energia associata alla matrice di adiacenza ed è calcolata tramite la somma dei valori singolari δ_q

N rappresenta il numero totale dei valori singolari della matrice di adiacenza;

δ_q rappresenta i valori singolari della matrice di adiacenza.

I valori singolari di una generica matrice A sono calcolati sfruttando le radici quadrate degli autovalori non negativi della matrice A^TA.

A seconda del tipo di architettura che possiede il sistema, è possibile definire complessità topologica in tre regioni:

- Se l'architettura della struttura è centralizzata, C_3 assumerà valori inferiori a 1 e il sistema si definisce ipoenergetico o energetico;
- Se l'architettura della struttura è gerarchica o stratificata, C_3 assume valori compresi tra 1, incluso, e 2 e il sistema si definisce transitorio;
- Se l'architettura della struttura è distribuita, C_3 assume valori maggiori o uguale a 2 e il sistema si definisce iperenergetico.

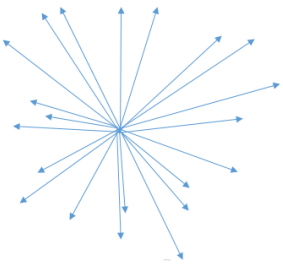


Figura 2 - Architettura Centralizzata: C3 < 1

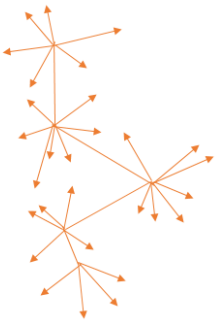


Figura 3 - Architettura Gerarchica: 1 =< C3 < 2

Figura 4 – Architettura Distribuita: C3 >=2

Inoltre, il termine C_3 descrive un effetto globale che può essere percepito durante l'assemblaggio del sistema e quindi, nell'equazione della complessità oggettiva, il prodotto tra C_2 e C_3 può essere considerato come un indicatore generale dello sforzo di integrazione del sistema (Alkan, 2019).

CAPITOLO 2: Pianificazione del Caso di Studio

Nel seguente capitolo verrà presento il caso di studio oggetto del presente lavoro di tesi e i suoi obiettivi, descrivendone lo scopo, le molecole scelte per condurre le analisi, la metodologia di raccolta dei dati e l'organizzazione dell'esperimento. Verranno, inoltre, descritti i criteri di valutazione della complessità soggettiva.

2.1 Obiettivi della Ricerca

Nell'ambito dei sistemi di produzione, la complessità di prodotto comporta un aumento dei costi, problemi operativi e tempi di realizzazione del prodotto più lunghi (Alkan, 2019). Le attività di assemblaggio dei prodotti rappresentano circa il 50% del tempo totale di produzione e circa il 40% del costo totale quindi un'analisi della complessità, e degli effetti che essa comporta sui tempi di processo e sul comportamento degli operatori, può portare ad un miglioramento del processo produttivo ed un sostanzioso risparmio sul costo di produzione.

L'esperimento condotto si ispira all'esperimento di Alkan (2019), in riferimento ai processi di assemblaggio e di disassemblaggio di diverse strutture molecolari, e ha come obiettivo la determinazione delle relazioni tra:

- complessità oggettiva e tempi di assemblaggio e disassemblaggio;
- complessità oggettiva e difetti introdotti dagli operatori;
- complessità oggettiva e apprendimento degli operatori;
- complessità oggettiva e complessità soggettiva.

Il prodotto oggetto della ricerca sono delle strutture molecolari, composte da atomi e legami, le quali sono state scelte per essere facilmente assemblabili e disassemblabili. Gli studenti aderenti alla sperimentazione hanno avuto il compito di riprodurre le strutture molecolari il più fedelmente e nel minor tempo possibile, riportando il numero di difetti commessi e i tempi impiegati a realizzare l'assemblaggio e il disassemblaggio delle strutture in apposite tabelle. Sono state selezionate dodici molecole con differente complessità oggettiva per poter determinare la relazione che esiste tra complessità oggettiva e tempi di processo e la relazione tra la complessità oggettiva e i difetti introdotti dagli operatori.

2.2 Kit Ball and Stick e Strutture Molecolari

Le dodici strutture molecolari scelte come oggetto dell'esperimento sono state realizzate tramite l'utilizzo di un kit *Orbit Molecular Building System* (Orbit™ by 3B Scientific®), un kit basato sul modello ball and stick nel quale delle palline fungono da atomi e dei tubicini da legami, e seguendo delle riproduzioni in 2D e in 3D delle molecole.

Sono stati utilizzati sette kit, uno per coppia di studenti, ognuno dei quali includeva 145 atomi di Carbonio, ball di colore nero, di cui 100 di forma tetraedrica, 20 di forma lineare e 25 di forma trigonale planare, 85 atomi di Ossigeno, ball di colore rosso di cui 10 di forma tetraedrica, 25 di forma bivalente e 50 di monovalente, 100 di Idrogeno, ball di colore bianco e di forma monovalente, 45 atomi di Azoto, ball di colore blu di cui 10 di forma tetraedrica, 10 di forma trigonale planare, 5 di forma lineare e 20 di forma monovalente, 50 atomi di Zolfo, ball di colore giallo di cui 5 di forma ottaedrica, 10 di forma tetraedrica, 25 di forma bivalente e 10 di forma monovalente, 20 atomi di fosforo, ball di colore viola di cui 15 di forma tetraedrica e 5 di forma trigonale bipiramidale, 30 atomi di Cloro, ball di colore verde scuro di forma monovalente, 20 atomi di Fluoro, ball di colore verde chiaro di forma monovalente, e 10 atomi di Metalli, ball di colore grigio di cui 5 di forma ottaedrica e 5 di forma tetraedrica. I kit includevano, inoltre, 100 tubi grigi di lunghezza pari a 3 millimetri utilizzati per rappresentare i legami singoli e 20 tubi trasparenti di lunghezza pari a 5 millimetri che simulavano i legami doppi.

Gli studenti partecipanti, divisi in coppie, trovavano le postazioni di assemblaggio e disassemblaggio perfettamente organizzate. Prima di ogni giornata lavorativa gli atomi venivano distinti per colore e forma e venivano posizionati negli appositi recipienti della postazione di assemblaggio. Anche i legami venivano divisi per tipologia e posizionati negli appositi recipienti. Gli studenti che eseguivano l'assemblaggio prelevavano le entità dai contenitori e procedevano a costruire la molecola basandosi sulle strutture 2D e 3D fornite.

Ogni postazione di disassemblaggio era formata da tanti recipienti quanti atomi di forma e colore diversa e da due recipienti per le tipologie di legami. Inizialmente questi recipienti erano vuoti e durante la fase di disassemblaggio gli operatori smontavano le strutture molecolari realizzate dal compagno di lavoro e ponevano ogni entità nell'apposito contenitore. Una volta terminata la fase di disassemblaggio ed effettuato il controllo di qualità, le entità presenti nei recipienti venivano trasferite nei recipienti della fase di assemblaggio. Per facilitare gli operatori, ogni recipiente è stato contrassegnato da un'etichetta in cui è stato riportato il nome dell'atomo e la forma.

Le molecole scelte, dalla meno complessa alla più complessa sono:

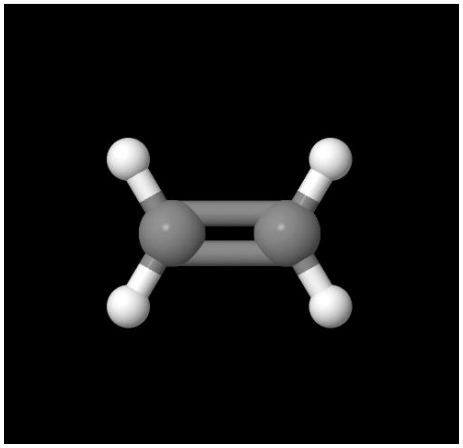


Figura 5 - Molecola 1: Ethylene

Formula bruta: C_2H_4

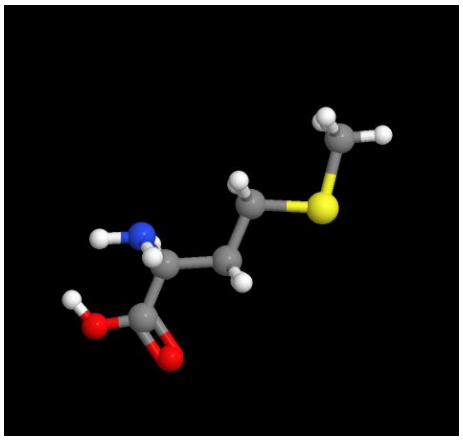


Figura 6 - Molecola 2: DL-Methionine

Formula bruta: $C_5H_{11}NO_2S$

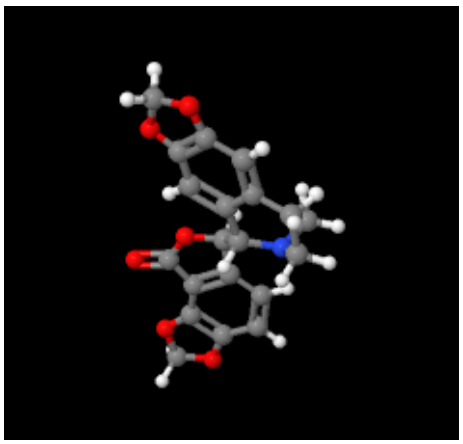


Figura 7 - Molecola 3: (+)-Bicuculline

Formula bruta: $C_{20}H_{17}NO_6$

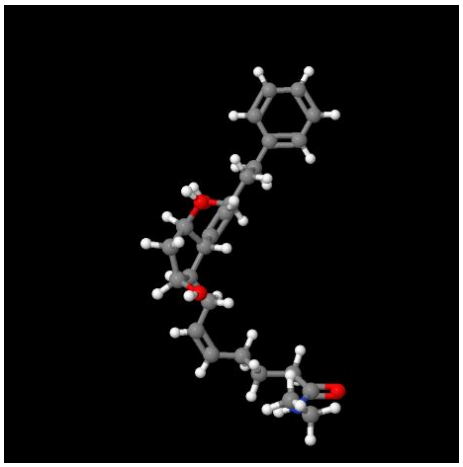


Figura 8 - Molecola 4: Bimatoprost

Formula bruta: $C_{25}H_{37}NO_4$

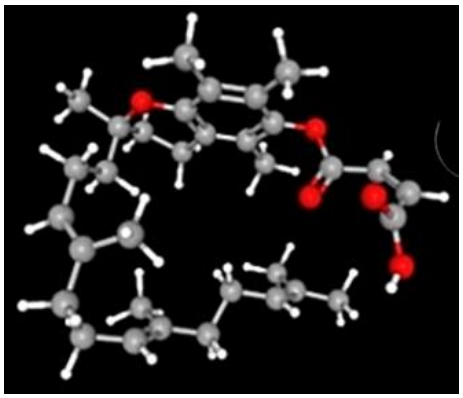


Figura 9 - Molecola 5: Alpha-Tocotrienol maleate

Formula bruta: $C_{33}H_{46}O_5$

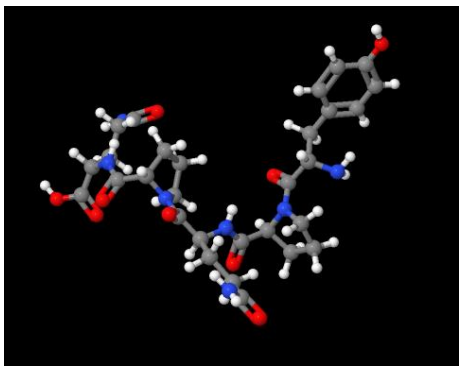


Figura 10 - Molecola 6: L-Tyrosyl-L-prolyl-L-glutaminyl-L-prolyl-L-glutamine

Formula bruta: $C_{29}H_{41}N_7O_9$

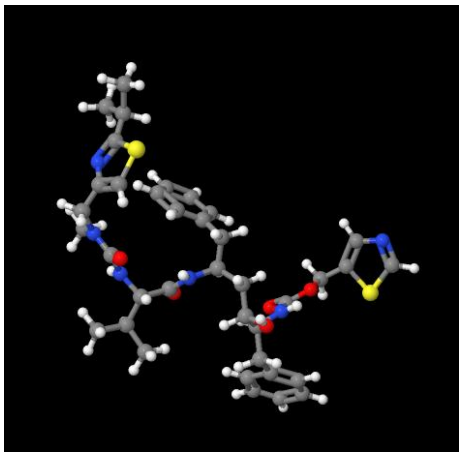


Figura 11 - Molecola 7: Ritonavir

Formula bruta: $C_{37}H_{48}N_6O_5S_2$

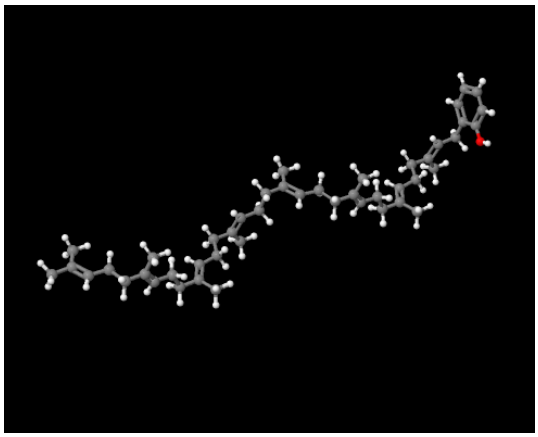


Figura 12 - Molecola 8: 2-Octaprenylphenol

Formula bruta: $C_{46}H_{70}O$

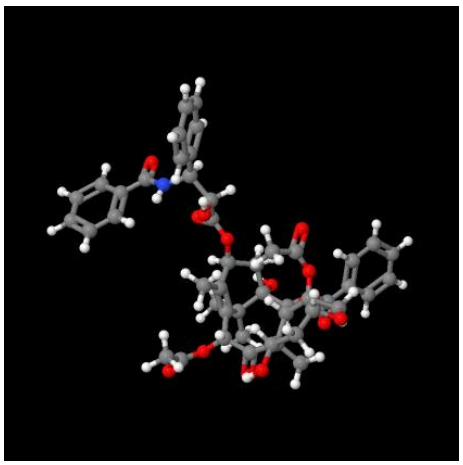


Figura 13 - Molecola 9: Paclitaxel

Formula bruta: $C_{47}H_{51}NO_{14}$

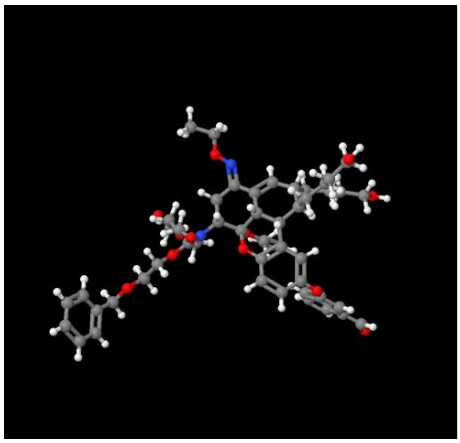


Figura 14 - Molecola 10: 2-(Benzyloxy)ethyl [6a-(allyloxy)-4-(ethoxyimino)-10-(3-formylphenoxy)-1,2-bis(4-hydroxybutyl)-1,2,4,5,6,6a,11b,11c-octahydrobenzo[kl]xanthen-6-yl][2-(2-hydroxyethoxy)ethyl]carbamate

Formula bruta: $C_{50}H_{64}N_2O_{12}$

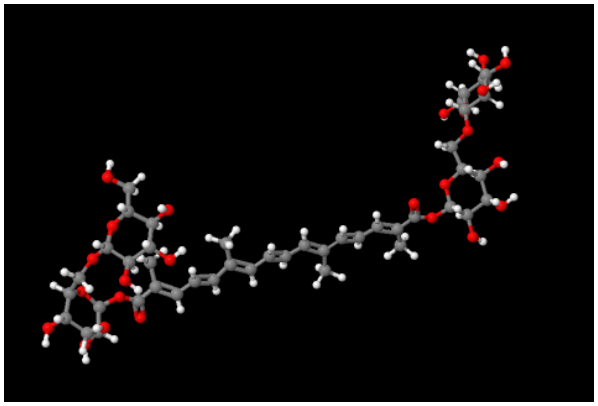


Figura 15 - Molecola 11: Crocin

Formula bruta: $C_{44}H_{64}O_{24}$

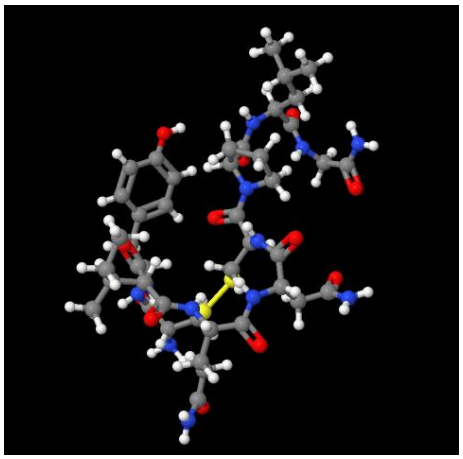


Figura 16 - Molecola 12: Oxytocin

Formula bruta: $C_{43}H_{66}N_{12}O_{12}S_2$

Nella tabella sottostante, *Tabella 1*, sono riporte le complessità oggettiva delle 12 molecole oggetto di indagine nel presente elaborato:

Molecola	ID1	ID2	ID3	ID4	ID5	ID6	ID7	ID8	ID9	ID10	ID11	ID12
Complessità Oggettiva (C)	6,4	24,1	66,9	90,47	108,32	117,44	134,81	147,73	159,88	177,07	181,07	181,04

Tabella 1 - Complessità Oggettiva (C = C1 + C2 · C3) delle 12 strutture molecolari selezionate

Per poter stimare correttamente il termine γ_p, proposto da Alkan (2019), fondamentale per il calcolo della Complessità di Manipolazione (C1), è stata fornita ad ogni coppia di studenti partecipante la Tabella sotto riportata, *Tabella 2*. In essa sono stati indicati i tempi necessari per individuare e prendere un atomo dal suo contenitore, individuare, prendere dal proprio contenitore e collegare due atomi tramite un legame doppio e, infine, il tempo necessario per individuare e prendere un atomo dal suo contenitore, individuare, prendere dal proprio contenitore e collegare due atomi tramite un legame singolo. Ogni operazione è stata eseguita solamente da uno studente della coppia, prima di iniziare l'assemblaggio delle molecole, ed è stata eseguita per tre volte consecutive. Infine, per le operazioni, sono stati considerati solamente i tempi medi.

Entità	T1	T2	T3	Tempo Medio [s]
Atomo				
Legame Singolo				
Legame Doppio				

Tabella 2 - Tabella fornita agli studenti per il calcolo del termine γ_p

Di seguito (*Tabella 3* e *Tabella 4*) sono riportate le Matrici di Adiacenza delle molecole ID1 e ID2, necessarie per il calcolo della Complessità delle Connessioni (C_2):

ID1	C1	C2	H1	H2	H3	H4
C1	0	1	1	0	1	0
C2	1	0	0	1	0	1
H1	1	0	0	0	0	0
H2	0	1	0	0	0	0
H3	1	0	0	0	0	0
H4	0	1	0	0	0	0

Tabella 3 - Matrice di Adiacenza molecola ID1

ID2	C1	S	C2	C3	C4	N	C5	O1	O2	H1	H2	H3	H4	H5	H6	H7	H8	H9	H10	H11
C1	0	1	0	0	0	0	0	0	0	1	1	1	0	0	0	0	0	0	0	0
S	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C2	0	1	0	1	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0
C3	0	0	1	0	1	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0
C4	0	0	0	1	0	1	1	0	0	0	0	0	0	0	0	0	1	0	0	0
N	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0
C5	0	0	0	0	1	0	0	1	1	0	0	0	0	0	0	0	0	0	0	1
O1	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
O2	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
H1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
H2	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
H3	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
H4	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
H5	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
H6	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
H7	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
H8	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
H9	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
H10	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
H11	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0

Tabella 4 - Matrice di Adiacenza molecola ID2

Nel caso di studio, per poter determinare e valutare la complessità soggettiva, ovvero la complessità percepita dall'operatore, è stato deciso di adottare un metodo basato sull'indagine mediante l'utilizzo di un questionario. Sono stati rielaborati i 16 criteri di bassa complessità proposti da Alkan (2019) e definite due scale ordinali linguistiche. Gli operatori, dopo aver ricevuto e assimilato le informazioni necessarie ed eseguito le operazioni di assemblaggio e disassemblaggio delle molecole, hanno attribuito un valore a ciascun criterio, in base a quanto si trovano d'accordo con quest'ultimo e in base a quanto esso è stato considerato importante al fine di un assemblaggio non complesso.

Le due scale di misura adottate, una per esprimere il grado di accordo con il criterio e l'altra per esprimere l'importanza per un assemblaggio non complesso, fanno parte della famiglia delle scale Likert le quali sono una via di mezzo tra una scala ordinale e una quantitativa.

La scala Likert è una scala di valutazione che, sfruttando un questionario composto da risposte chiuse e precompilate, identifica le opinioni dei soggetti tramite delle opzioni numeriche e/o verbali. Ad ogni domanda del questionario, l'intervistato deve assegnare un punteggio sulla base della sua persona esperienza. A tale scopo sono stati definiti 5 livelli di risposta, che spaziano dall'ipotesi peggiore (trascurabile\totalmente in disaccordo) a quella migliore (indispensabile\totalmente d'accordo), con i quali il soggetto può esprimere il suo grado di importanza e di accordo a ciascun criterio presente nel questionario.

L'utilizzo di una scala Likert è giustificato dalla semplicità di applicazione di quest'ultima ma anche perché comporta una serie di vantaggi come la possibilità di distinguere velocemente le risposte e di poterle velocemente quantificare, tramite l'impiego dei valori numerici.

L'utilizzo di una scala ti tipo ordinale per la raccolta dell'importanza e delle valutazioni attribuite ai criteri fa sorgere alcuni problemi. In essa è definita la proprietà dell'ordinamento, ma non può essere applicato il concetto di distanza. Inoltre, il suo impiego permette di esprimere una preferenza soggettiva che però risulta essere difficile da giustificare. Ogni operatore attribuisce un valore al criterio in base alle sue sensazioni e sulla base della sua esperienza personale, senza fornire una motivazione. Il principale problema derivante dall'utilizzo di una scala di tipo ordinale linguistica è il significato recepito da ogni operatore e la relativa codifica delle informazioni. Ad esempio, una stessa classe di appartenenza potrebbe essere interpretata differentemente da soggetti diversi e definire un limite univoco in modo concreto tra due classi risulta impossibile.

2.3 Pianificazione delle Prove

Prima di partire con la sperimentazione vera e propria è stato necessario pianificare le prove, basandosi su dati preliminari, per poter organizzare le giornate lavorative in maniera ottimale.

Gli studenti partecipanti sono stati divisi in due gruppi con istruzioni e obiettivi diversi: il gruppo A, sul quale si è deciso di non osservare gli effetti di apprendimento, e il gruppo B sul quale, invece, si è deciso di osservare gli effetti di apprendimento degli operatori. Al gruppo A è stato chiesto di assemblare e poi disassemblare tutte e dodici le strutture molecolari seguendo un ordine randomico precedentemente stabilito. Al gruppo B, invece, è stato chiesto di assemblare e disassemblare per sette volte consecutive una struttura con bassa complessità oggettiva e per sette volte consecutive una struttura con alta complessità oggettiva. Per l'assemblaggio e il disassemblaggio del gruppo B sono state selezionate sei strutture molecolari dalle dodici totali: 1,3,5,8,10 e 12 e ad ogni coppia di studenti è stata affidata la riproduzione di due strutture molecolari. L'analisi preliminare condotta sull'assemblaggio e il disassemblaggio delle strutture molecolari, dalla molecola con complessità molecolare minore alla molecola con complessità molecolare maggiore, hanno permesso di ottenere dei tempi indicativi di processo che sono stati utilizzati nella scelta delle tre coppie di molecole affidate agli studenti appartenenti al gruppo B. Le sei molecole sono, infatti, state scelte in base ai loro tempi di processo in modo che, ripetendo l'assemblaggio e il disassemblaggio per sette volte consecutive, non si sforasse l'orario prestabilito e cercando di affidare ad ogni coppia lo stesso carico di lavoro. Inoltre, ogni coppia di molecole doveva necessariamente essere composta da una molecola con bassa complessità oggettiva e una molecola con alta complessità oggettiva. Le coppie formatosi sono state le seguenti: la molecola 1 con la molecola 12, la molecola 3 con la molecola 10 e la molecola 5 con la molecola 8.

Di seguito sono riportati i piani operativi contenente le date e il numero di studenti partecipanti e i tempi di assemblaggio e disassemblaggio richiesti.

Programmazione Giornate Lavorative			
Esecuzioni Randomiche	Numero Studenti	Bumero Assemblatori	Numero Disassemblatori
04/10/2021	10	5	5
11/10/2021	10	5	5
22/10/2021	14	7	7
05/11/2021	14	7	7
12/11/2021	14	7	7
22/11/2021	14	7	7
26/11/2021	14	7	7
10/12/2021	14	7	7
Totale	104	52	52

Tabella 5 - Studenti, divisi in assemblatori e disassemblatori, partecipanti nelle giornate in cui si sono effettuate le prove Randomiche

Programmazione Giornate Lavorative	
Esecuzioni Ripetute	Numero Studenti
08/10/2021	8
15/10/2021	8
18/10/2021	8
29/10/2021	12
08/11/2021	12
15/11/2021	12
19/11/2021	12
03/12/2021	12
Totale	84

Tabella 6 - Studenti partecipanti nelle giornate in cui si sono effettuate le prove Ripetute

In *Figura 17* e in *Figura 18* è riportata la schedulazione della giornata lavorativa e il tempo previsto per eseguire l'assemblaggio e il disassemblaggio delle 12 molecole, nel caso di *prove Randomiche*, e il tempo previsto per la ripetizione dell'assemblaggio e del disassemblaggio delle molecole ID1, ID3, ID5, ID8, ID10 e ID12, nel caso di *prove Ripetute*.

PROVE RANDOMICHE

										Giornata 1 GRUPPO A ASSEMBLAGGIO					
Lunedì 04/10	inizio	pausa	ripresa	pausa		ripresa	pausa	ripresa	fine	n° studenti totali	n° operatori	n° controllori	n° molecole	tempo richiesto [min]	
	08:30	11:00	11:15	13:00		14:00	16:30	16:45	19:00	14	7	7	12	289	
											Giornata 1 GRUPPO A DISASSEMBLAGGIO				
Lunedì 4/10	inizio	pausa				ripresa	pausa	ripresa	fine						
	08:30	11:00	11:15	13:00		14:00	16:30	16:45	19:00	n° operatori	n° controllori	n° molecole	tempo richiesto [min]		
										7	7	12	100		

Figura 17 - Pianificazione delle prove per le giornate con prove Randomiche

PROVE RIPETUTE

	Giornata 2 GRUPPO B ASSEMBLAGGIO										
	n° studenti totali	n° molecole	tempo richiesto ID 1 [min]	tempo richiesto ID 3 [min]	tempo richiesto ID 5 [min]	tempo richiesto ID 8 [min]	tempo richiesto ID 10 [min]	tempo richiesto ID 12 [min]	tempo richiesto ID 1 + ID 12 [min]	tempo richiesto ID 3 + ID 10 [min]	tempo richiesto ID 5 + ID 8 [min]
	12	6	7	49	105	196	238	259	266	287	301
	n° coppie di studenti	n° studenti	n° molecole	tempo richiesto [min]							
ID 1 + ID 12	2	2	2	266							
ID 3 + ID 10	2	2	2	287							
ID 5 + ID 8	2	2	2	301							

Figura 18 - Pianificazione delle prove per le giornate con prove Ripetute

La buona riuscita del progetto è stata possibile grazie alla partecipazione degli studenti del prof. Maurizio Galetto del corso di Ingegneria della Qualità.

Gli studenti interessati a partecipare potevano aderire alla sperimentazione registrandosi al seguente link *https://politecnico-di-torino5.reservio.com* inserendo il proprio nome, cognome, matricola e scegliendo la preferenza del giorno tra una delle sedici date disponibili.

In *Figura 19* è mostrata l'anteprima del sito per le registrazioni.

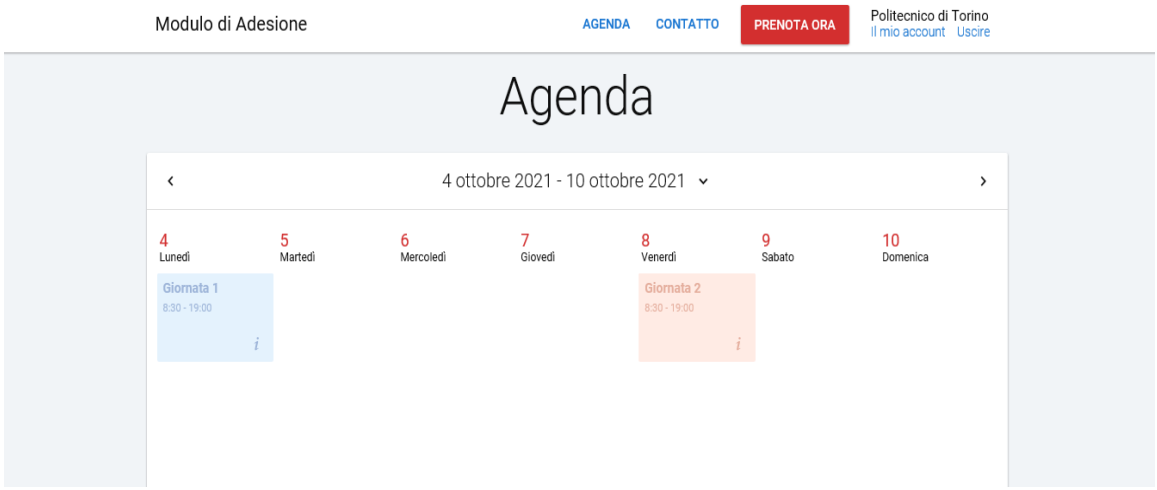


Figura 19 - Metodo di adesione al progetto

I giorni proposti sono stati organizzati con il seguente orario:

- Dalle ore 08.30 alle ore 09.00 breve spiegazione delle regole della campagna sperimentale e degli obiettivi da conseguire;
- Dalle ore 09.00 alle ore 13.00 assemblaggio e disassemblaggio delle strutture molecolari;
- Dalle ore 13.00 alle ore 14.00 pausa pranzo;
- Dalle ore 14.00 alle ore 19.00 assemblaggio e disassemblaggio delle strutture molecolari.

Per riuscire a raggiungere tutti gli obiettivi posti, si è deciso di suddividere le giornate lavorative in due tipologie: giornate lavorative in cui sono state effettuate *prove Randomiche*, a cui ha partecipato il gruppo A e giornate lavorative in cui sono state effettuate *prove Ripetute*, a cui ha partecipato il gruppo B. Entrambi i gruppi si sono occupati dell'assemblaggio, del controllo delle fasi e del disassemblaggio delle molecole.

Per entrambe le tipologie di giornata lavorativa, gli studenti sono stati divisi in coppie e ad ogni studente è stata affidata una doppia mansione:

- Uno studente si occupava della fase di assemblaggio, ovvero effettuava la riproduzione delle molecole assegnategli, e fungeva da controllore della qualità durante la fase di disassemblaggio, ovvero supervisionava e controllava la presenza di eventuali difetti;
- L'altro studente si occupava della fase di disassemblaggio, ovvero effettuava il disassemblaggio della struttura realizzata riponendo le entità nei corretti recipienti, e fungeva da controllore della qualità durante l'assemblaggio, ovvero supervisionava e controllava la presenza di eventuali difetti.

Le mansioni venivano affidate all'inizio di ogni giornata lavorativa e dovevano rimanere invariate per tutto l'arco della giornata al fine di non falsificare i risultati ottenuti. Venivano, inoltre, fornite le istruzioni e i modelli rappresentanti le strutture molecolari in 3D e in 2D come nell'esempio sotto riportato in figura.

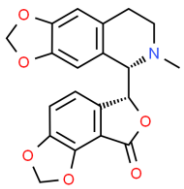


Figura 20 - Esempio di rappresentazione 2D di una struttura molecolare

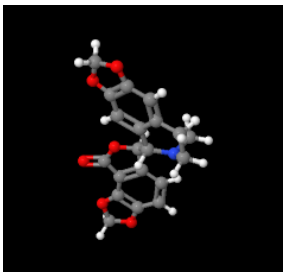


Figura 21 - Esempio di rappresentazione 3D di una struttura molecolare

Agli studenti, prima di eseguire l'assemblaggio delle strutture, indistintamente se si dovevano effettuare le *prove Ripetute* o le *prove Randomiche*, è stato chiesto di effettuare delle operazioni e riportare il tempo necessario per effettuarle. I tempi necessari ad effettuare queste operazioni risultano fondamentali per determinare la complessità oggettiva. Uno studente per coppia doveva riportare il tempo impiegato per prendere un atomo da un qualsiasi recipiente e posizionarlo sul banco da lavoro, il tempo impiegato per prendere due atomi differenti da qualsiasi recipiente, un legame singolo dall'apposito recipiente e assemblare i due atomi con il legame singolo e infine il tempo impiegato per prendere due atomi differenti da un qualsiasi recipiente, un legame doppio dall'apposito recipiente e assemblare i due atomi con il legame doppio. Ogni operazione doveva essere effettuata per tre volte consecutive e gli studenti dovevano riportare in apposite tabelle i singoli tempi impiegati a compiere le operazioni e i loro tempi medi. L'immagine sottostante riporta l'esempio di una tabella in cui venivano inseriti i tempi delle operazioni.

Entità	T1	T2	T3	Tempo Medio [s]
Atomo				
Legame Singolo				
Legame Doppio				

Tabella 7 - Tabella fornita agli studenti per il calcolo del termine γ_p

Nelle giornate lavorative in cui sono state effettuate le *prove Randomiche* gli studenti hanno assemblato e poi disassemblato tutte le dodici strutture molecolari secondo un ordine casuale precedentemente stabilito. In particolare, per ogni giornata lavorativa, uno studente si è occupato dell'assemblaggio della molecola e l'altro del relativo disassemblaggio. Terminate le fasi di assemblaggio e di disassemblaggio è stato effettuato il controllo della qualità della fase, ovvero gli studenti a fine assemblaggio dovevano verificare che la struttura fosse stata realizzata in maniera corretta e segnalare la presenza di eventuali difetti così come a fine disassemblaggio dovevano verificare che ogni componente fosse stato inserito nel recipiente corretto e nel caso segnalare la presenza di eventuali difetti.

Nelle giornate lavorative in cui sono state effettuate le *prove Ripetute*, invece, ad ogni coppia sono state affidate due strutture molecolari, tra 1-12, 3-10 e 5-8, e uno studente della coppia si è occupato della riproduzione delle strutture molecolari per sette volte e l'altro studente si è occupato di eseguire il relativo disassemblaggio. In particolare, per ogni giornata lavorativa, ad ogni coppia di studenti sono state affidate due strutture molecolari. Uno studente si è occupato dell'assemblaggio della prima struttura molecolare e l'altro del relativo disassemblaggio. È stato deciso di effettuare l'assemblaggio e il disassemblaggio di ogni struttura per sette volte consecutive per poter analizzare l'effetto

apprendimento degli operatori e per poter valutare come, lasciando immutata la complessità oggettiva della struttura, i tempi di assemblaggio e disasemblaggio e il numero di difetti commessi dagli operatori, diminuiva all'aumentare delle prove effettuate. Terminata la settima riproduzione, gli studenti hanno realizzato la seconda molecola nella quale i ruoli sono stati scambiati, Durante le fasi di assemblaggio e di disassemblaggio è stato effettuato il controllo della qualità della fase.

In *Appendice* sono riportate alcune foto che mostrano le molecole assemblate dagli studenti tramite l'utilizzo del kit *Ball and Stick* e alcune foto che mostrano come sono state strutturate le postazioni di assemblaggio e di disassemblaggio.

2.4 Tipologie di Difetti

Sono state introdotte due tipologie di difetti che potevano essere commessi da parte degli studenti: difetti riscontrati in fase di assemblaggio e difetti riscontrati in fase di disasemblaggio. Per la prima tipologia sono stati considerati come difetti tutte le entità che differiscono dalla struttura originaria. Per entità è inteso sia la tipologia di legame, quindi legame singolo o doppio, sia la forma la forma o il colore dell'atomo. Per fare un'analisi più accurata è stato deciso di fare un'ulteriore distinzione dei difetti che si potevano riscontrare durante la fase di assemblaggio: difetti di processo e difetti di prodotto.

Sono definiti difetti di processo tutti i difetti riscontrati durante la fase di assemblaggio della struttura dall'operatore che esegue questa fase. Nel momento in cui lo studente si accorgeva di aver commesso un errore, il controllore della qualità riportava il difetto nella tabella. Questa tipologia di difetti poteva essere riscontrata esclusivamente dallo studente che eseguiva l'assemblaggio e doveva essere considerato come difetto qualsiasi legame o atomo che mancava o differiva dalla struttura originaria.

Sono definiti difetti di prodotto, invece, tutti i difetti riscontrati durante la fase di controllo da parte dell'operatore che funge da controllore della qualità. Questa tipologia di difetti raccoglie tutti gli errori commessi dagli operatori che eseguono l'assemblaggio senza rendersene conto. Anche in questo caso, sono stati considerati come difetti tutte le entità che differivano dalla struttura originaria.

L'ultima tipologia di difetti considerata sono quelli riscontrati in fase di disassemblaggio. Rientrano in questa tipologia tutte le entità che venivano posizionate nei recipienti sbagliati. In questo caso l'errore poteva essere rilevato indistintamente sia dallo studente che si occupava della fase di disassemblaggio che dall'operatore che funge da controllore della qualità.

2.5 Dati da Rilevare

La ricerca si basa sull'analisi dei tempi e dei difetti riscontrati dagli studenti. Per poter analizzare le relazioni tra la complessità oggettiva e i tempi di assemblaggio e disassemblaggio, tra la complessità oggettiva e difetti introdotti dagli operatori e tra la complessità oggettiva e l'apprendimento degli operatori è stato chiesto loro di rilevare il tempo necessario ad effettuare l'assemblaggio delle strutture, il tempo necessario ad effettuare il disassemblaggio delle strutture, il numero di difetti di processo riscontrati in fase di assemblaggio, il numero di difetti di prodotto riscontrati in fase di assemblaggio, il numero di difetti totale riscontrati in fase di assemblaggio, ovvero la somma dei difetti di processo e di prodotto, e il numero di difetti riscontrati in fase di disassemblaggio.

Tutti i dati rilevati dovevano essere riportati nella seguente tabella:

ID	Nome molecola	Struttura 2D e 3D da utilizzare come istruzione per l'assemblaggio	Tempo di assemblaggio [min]	Tempo di disassemblaggio [min]	Numero difetti di *processo* riscontrati in assemblaggio	Numero difetti di *prodotto* riscontrati in assemblaggio	Numero difetti totali riscontrati in assemblaggio	Numero difetti riscontrati in disassemblaggio
2	DL-Methionine	http://www.chemspider.com/Chemical-Structure.853.html?rid=f1d67131-ea51-47af-8b45-6473b7e3f65f&pags_num=0						
6	L-Tyrosyl-L-prolyl-L-glutaminyl-L-prolyl-L-glutamine	http://www.chemspider.com/Chemical-Structure.169226.html?rid=521918a4-f5cc-4f6c-920f-6bb331337488&pags_num=0						
4	Bimatoprost	http://www.chemspider.com/Chemical-Structure.4470565.html?rid=54da8d60-8329-4834-9a6c-4ac5c40f437d						
5	alpha-Tocotrienol maleate	https://pubchem.ncbi.nlm.nih.gov/compound/43870345						
8	2-Octaprenylphenol	http://www.chemspider.com/Chemical-Structure.4444381.html?rid=85a4a34a-b51d-4393-a36a-7c6d4409b1b6						
7	Ritonavir	http://www.chemspider.com/Chemical-Structure.347980.html?rid=f045cc7c-8133-4ccb-8f00-01711c2dc04c						
1	Ethylene	http://www.chemspider.com/Chemical-Structure.6085.html?rid=372ddc03-fe16-4cce-b343-8324ae35bd0d						
10	2-(Benzyloxy)ethyl [6a-(allyloxy)-4-(ethoxyimino)-10-(3-formylphenoxy)-1,2-bis(4-hydroxybutyl)-	http://www.chemspider.com/Chemical-Structure.353947.1.html?rid=cf16a7d2-a2ba-49ff-8ddb-c7e1dc7e8d25&pags_num=0						
3	(+)-Bicuculline	http://www.chemspider.com/Chemical-Structure.3820.html?rid=ad1b3225-3def-4c2d-b377-a028b261e39f&pags_num=0						
12	Oxytocin	http://www.chemspider.com/Chemical-Structure.388434.html?rid=02f3ca8c-e389-484e-b50f-2aa637a42324						
9	Paclitaxel	https://www.biotopics.co.uk/jsmol/taxol.html						
11	Crocin	http://www.chemspider.com/Chemical-Structure.4444645.html?rid=73b5470b-0baa-4072-880a-067359aef6a3&pags_num=0						

Tabella 8 - Tabella fornita agli studenti per la raccolta dei tempi e dei difetti di assemblaggio e di disassemblaggio nelle prove Randomiche

Nel caso di giornate lavorative in cui si sono effettuate le *prove Ripetute*, è stato chiesto di riportare i tempi e i difetti di ogni ripetizione.

Prova	Tempo di assemblaggio [min]	Tempo di disassemblaggio [min]	Numero difetti di *processo* riscontrati in assemblaggio	Numero difetti di *prodotto* riscontrati in assemblaggio	Numero difetti totali riscontrati in assemblaggio	Numero difetti riscontrati in disassemblaggio
1						
2						
3						
4						
5						
6						
7						

Tabella 9 - Tabella fornita agli studenti per la raccolta dei tempi e dei difetti di assemblaggio e di disassemblaggio nelle prove Ripetute

Per poter valutare la relazione che sussiste tra la complessità oggettiva e la complessità soggettiva è stato stilato un questionario basato su sedici criteri di bassa complessità soggettiva. La complessità soggettiva è definita come la complessità percepita dagli studenti durante l'esecuzione della fase di assemblaggio e che può risultare differente dalla difficoltà reale di prodotto in quanto è influenzata da alcuni aspetti personali dell'operatore, come l'esperienza pregressa, la creatività o il grado di coinvolgimento, dalle interpretazioni soggettive e dall'ipotizzare delle strategie e dalla capacità di attuarle.

Il questionario è rivolto agli studenti che si occupavano dell'assemblaggio delle strutture. Gli studenti che effettuavano le *prove Randomiche*, dopo aver assemblato una struttura molecolare, esprimevano una valutazione sul grado di accordo, ovvero assegnavano un valore compreso tra 1 e 5 in base a quanto si trovava d'accordo con il criterio in relazione alla struttura molecolare appena realizzata. A fine giornata lavorativa, dopo aver assemblato tutte le strutture molecolari, dovevano esprimere una valutazione sull'importanza dei criteri, ovvero assegnavano un valore compreso tra 1 e 5 in base a quanto percepivano importante quel determinato criterio affinché avvenga un assemblaggio semplice.

Gli studenti che effettuavano le *prove Ripetute*, invece, esprimevano una valutazione sul grado di accordo e una valutazione sull'importanza dei criteri al termine del settimo assemblaggio della struttura.

Di seguito sono riportati i sedici criteri di bassa complessità:

1. Vi è un'unica (o poche) strategie per eseguire l'assemblaggio (NON ci sono molte strategie diverse per completare l'assemblaggio);
2. Pochi pezzi da montare; è possibile lavorare per blocchi preassemblati, cioè in maniera modulare;
3. L'assemblaggio dei componenti è facile e veloce (non ci sono operazioni lunghe);
4. La posizione di assemblaggio delle parti/componenti è chiara (capisco subito dove posizionare i pezzi all'interno della struttura);
5. L'accessibilità alla struttura è buona durante il montaggio;
6. Operazioni completamente visibili (le operazioni non richiedono di orientare l'assemblaggio per una migliore visibilità);
7. Maneggiare la struttura è facile da un punto di vista ergonomico;
8. Non è richiesto sforzo cognitivo e strategia durante l'assemblaggio da parte dell'operatore;

9. L’assemblaggio può essere fatto senza seguire un ordine preciso (indipendenza dell’ordine di montaggio);
10. Non sono necessari ricontrolli intermedi durante l’assemblaggio per valutare la qualità e correttezza della struttura (non ci si perde durante l’assemblaggio della struttura);
11. Le operazioni non richiedono precisione, accuratezza e attenzione;
12. Non è necessario effettuare aggiustamenti (a causa di errori o imprecisioni);
13. Componenti che rimangono stabili, che non è necessario spostare (aggiustare) per inserire nuovi pezzi;
14. Non è necessario avere istruzioni dettagliate, cioè l'operatore può procedere in maniera intuitiva (non è necessario ruotare più volte l’immagine 2D o 3D di istruzione durante l’assemblaggio);
15. La struttura è resistente e compatta, quindi non cambia dimensione e non si deforma se manipolata e spostata;
16. C'è un immediato feedback di corretto assemblaggio.

Le legende delle importanze e dei livelli di accordo utilizzate sono le seguenti:

Legenda importanze	
Definizione scala importanze	
1	trascurabile
2	preferibile
3	importante
4	molto importante
5	indispensabile

Legenda livelli di accordo	
Definizione scala livelli di concordanza	
1	totalmente in disaccordo
2	in disaccordo
3	relativamente d'accordo
4	d'accordo
5	totalmente d'accordo

Tabella 10 - Legenda delle importanze dei criteri e legenda dei livelli di accordo con i criteri

CAPITOLO 3: Metodologie di Analisi

Nel seguente capitolo sono presentate le metodologie di analisi utilizzate nell'ambito della campagna sperimentale. In particolare, sono descritti i metodi statistici dell'ANOVA, tecnica che permette lo studio della varianza dei dati, e della Regressione, che ci permette di trovare l'equazione che meglio spiega la relazione tra una variabile indipendente e una dipendente.

3.1 Analisi della Varianza (ANOVA)

L'analisi della Varianza, detta comunemente ANOVA, è un metodo statistico nel quale si pone come obiettivo quello di valutare gli effetti di uno o più fattori su una variabile di interesse. Nello specifico è utilizzato per confrontare le varianze e le medie di gruppi diversi e rivelare se queste sono statisticamente significative.

Prima di poter effettuare l'ANOVA è necessario che i dati raccolti soddisfino le seguenti ipotesi:

- Indipendenza: il valore di un'osservazione della variabile dipendente deve essere indipendente dal valore di qualsiasi altra osservazione;
- Normalità: i valori delle osservazioni della variabile dipendente devono seguire una distribuzione Normale;
- Continuità: la variabile dipendente deve essere continua e quindi le osservazioni devono poter essere misurate attraverso l'utilizzo di una apposita scala.

Esistono vari tipi di ANOVA che differiscono in base al numero di variabili dipendenti e indipendenti che modello oggetto di studio impiega.

Se si impiega una sola variabile indipendenti è consigliabile eseguire una One-Way ANOVA; se il modello prevede l'utilizzo di due o più variabili indipendente si adopererà una Two-Way ANOVA; se il modello prevede una sola variabile dipendente allora verrà eseguita una ANOVA Univariata e se il modello prevede due o più variabili dipendenti si effettuerà una MANOVA.

Poiché il caso studio in esame prevede l'utilizzo due variabili indipendenti e di una variabile dipendente, si ricade nel caso in cui è necessario condurre dei test Two-Way ANOVA. In particolare, nelle *prove Randomiche* le variabili indipendenti impiegate sono la *Complessità Oggettiva* e l'*Operatore* e nelle *prove Ripetute* le variabili sono il *Numero di Prova* (o *Esecuzione*) e l'*Operatore*. Per entrambe le prove le variabili dipendenti analizzate sono i *Tempi*, e poi i *Difetti di Assemblaggio* e *di Disassemblaggio*.

Nello specifico, per le entrambe le tipologie di prova, sono stati condotti test Two-Way ANOVA su sei variabili dipendenti: *Tempo di Assemblaggio*, *Tempo di Disassemblaggio*, *Numero di Difetti di Processo di Assemblaggio*, *Numero di Difetti di Prodotto di Assemblaggio*, *Numero di Difetti Totali di assemblaggio* e *Numero di Difetti di Disassemblaggio*.

Per poter valutare il test statistico e valutare se accettare o rifiutare l'ipotesi nulla è necessario confrontare il P-value del termine con il livello di significatività. L'ipotesi nulla testata è che le medie dei gruppi siano uguali tra loro; quindi, si è voluto verificare se le variabili indipendenti influiscono in modo significativo sulla variabile dipendente di osservazione. Il livello di significatività indica la soglia entro la quale un risultato è considerato significativo e, generalmente, è posto al 5%. Il valore è determinato a priori e indica, in questo caso, un rischio del 5% di concludere che esiste una correlazione tra la variabile indipendente e quella indipendente quando, in realtà, non esista alcuna associazione effettiva.

Qualora dall'analisi dell'ANOVA si è determinato un valore del P-value del termine minore, o pari, al livello di significatività, si può concludere che esiste un'associazione statisticamente significativa tra le due variabili e quindi si rifiuta l'ipotesi nulla. Invece, se si determina un valore del P-value maggiore della soglia, non si rifiuta l'ipotesi nulla e quindi esiste un'associazione tra le variabili dipendente e quelle indipendenti.

L'ANOVA risulta essere molto utile per l'analisi di dati, ma presenta anche dei limiti. Uno dei limiti più influenti è che l'analisi effettuata può indicare esclusivamente se è presente, o meno, una differenza significativa tra le medie dei gruppi e presuppone che il campione di dati utilizzato sia uniformemente distribuito e che le variazioni standard dei gruppi siano simili tra loro. Proprio a causa di questi limiti, l'ANOVA è utilizzata in combinazione ad altri metodi statistici, come la regressione.

3.2 Regressione

La regressione è un metodo di analisi statistica adottato per analizzare un set di dati. L'obiettivo è di stimare una possibile relazione funzionale tra la variabile dipendete e una, o più, variabili indipendenti attraverso un'equazione che preveda i valori della variabile dipendente.

Dopo aver effettuato l'analisi della varianza, per ognuna delle sei variabili dipendenti individuate sono state condotte delle analisi statistiche di regressione, lineare e non lineare, in modo da scegliere il miglior modello che interpolasse i dati osservati. Il miglior modello è scelto basandosi sull'analisi della curva di adattamento, dei valori restituiti dagli indicatori S e R^2, del valore restituito dal P-value e in base all'analisi condotte sui residui.

Tramite l'utilizzo di un grafico di dispersione, lo Scatterplot, è possibile evidenziare la possibile relazione esistente tra due variabili. Il set di dati da analizzare viene disposto in uno spazio cartesiano, dove lungo l'asse delle ascisse è posta la variabile indipendente e lungo l'asse delle ordinate è posta quella indipendente, e vengono individuati mediante l'utilizzo di coordinate. Tramite il grafico di dispersione si può ipotizzare quale tipo di regressione, se lineare o no, adottare.

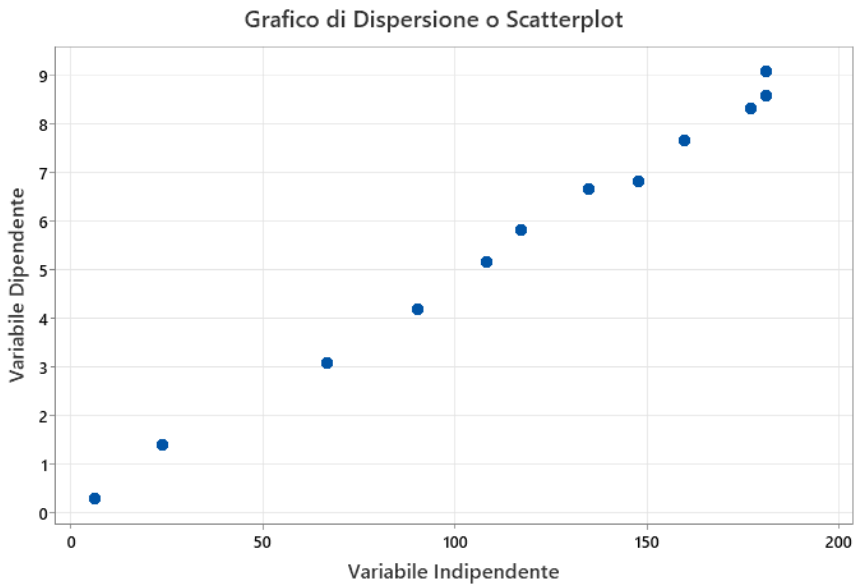


Figura 22 - Esempio di Grafico di Dispersione, o Scatterplot

Dall'analisi della regressione ci vengono forniti dei valori di alcuni indici. L'indicatore S, detto errore standard di regressione o errore standard della stima, fornisce una stima di quanto il modello si adatta al set di dati. L'indicatore rappresenta la distanza media, espressa nella stessa unità della variabile dipendente, dei valori osservati dalla curva che descrive il modello. Più il valore restituito dall'indicatore è piccolo, più i valori reali si avvicinano alla curva di regressione e più il modello è adatto al set di dati.

L'indicatore R^2 restituisce un valore statistico che permette di capire quanto il modello adottato fitta bene i dati. Il valore restituito stima in che percentuale la variazione della variabile indipendente influisce sulla variabile dipendente, ovvero restituisce una valutazione di quello che è il rapporto tra le due variabili. Quindi, più è alto il valore di R^2 più il modello è adatto ai dati osservati e più è prossimo allo zero più il modello scelto è inadatto.

La differenza tra i due indicatori è che l'indicatore S può essere stimato indipendentemente dal tipo di regressione adottata mentre l'indicatore R^2 ha significato solamente all'interno della regressione di tipo lineare.

L'analisi dell'indicatore S, o R^2 in ambito di regressione lineare, è necessaria ma non sufficiente per determinare la bontà del modello.

Dopo aver valutato i valori degli indicatori è necessario analizzare i residui associati alla curva. I residui sono definiti come la differenza tra il punto empirico e i punti osservati dalla curva calcolata dal modello, ovvero:

$$e_i = y_i - \bar{y}_i \tag{7}$$

dove:

y_i è il valore osservato dalla raccolta dei dati della prova i;
$\bar{y}_i$ è il valore osservato mediante la curva del modello della prova i.

Affinché il modello adottato descriva correttamente le osservazioni è necessario che i residui rispettino quattro ipotesi di base. Essi devono essere indipendenti tra loro, devono essere distribuiti in modo casuale e secondo una distribuzione normale e devono avere varianza costante.

L'analisi dei residui è condotta attraverso lo studio di quattro grafici: *Normal Probability Plot, Versus Fits, Histogram* e *Versus Order* (*Figura 23*).

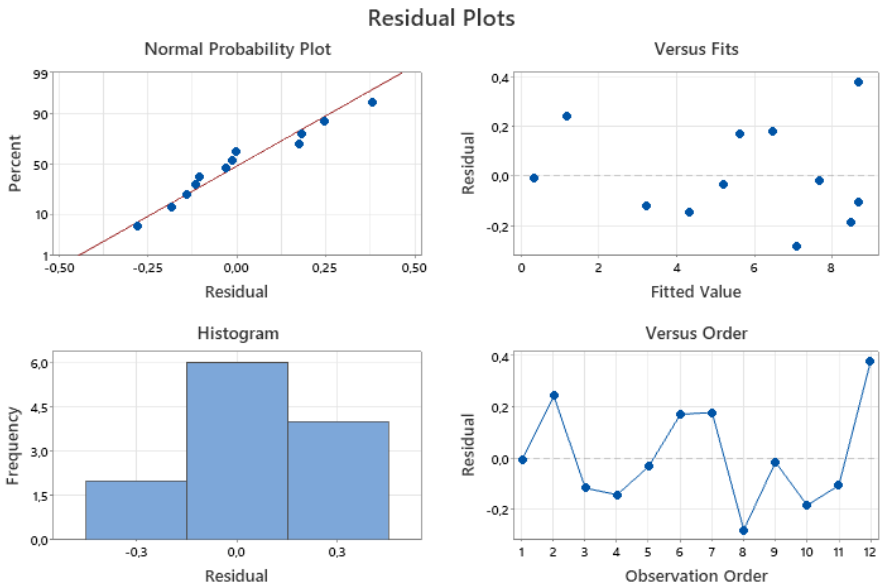


Figura 23 - Esempio di output grafico dell'analisi dei residui: Normal Probability Plot, Versus Fits, Histogram, Versus Order

Il *Normal Probability Plot*, ossia il Grafico di Probabilità Normale, è un grafico a due dimensioni in cui sono riportati i residui sull'asse dell'ascisse e sull'asse delle ordinate il quantile relativo di una distribuzione normale standard.

Nel grafico è riportata, in rosso, anche una retta inclinata positivamente. Attraverso il suo studio è possibile verificare l'ipotesi per cui i residui sono distribuiti secondo una distribuzione normale. Infatti, se i punti individuati dal grafico si posizionano approssimativamente lungo la linea retta si può affermare che i residui sono distribuiti secondo una distribuzione normale.

Mediante l'analisi dei grafici *Versus Fits* e *Versus Order* è possibile verificare le ipotesi per cui i residui si devono distribuire in maniera casuale, al di sotto e al di sopra della linea dello 0.0, e che devono avere una varianza costante. Infatti, affinché il modello scelto interpoli correttamente il set di dati osservati è necessario che i punti dei due grafici siano disposti in modo casuale su entrambi i lati dello zero, che non siano presenti agglomerati di punti e che non sia possibile osservare particolari tendenze.

L'Histogram è un altro strumento il cui studio ci permette di valutare l'ipotesi circa la distribuzione normale dei residui. Per poter affermare che i residui seguano una distribuzione normale è necessario che le classi di osservazioni, contenute nel grafico, seguano una curva di tipo gaussiano, tipica di un andamento con distribuzione normale.

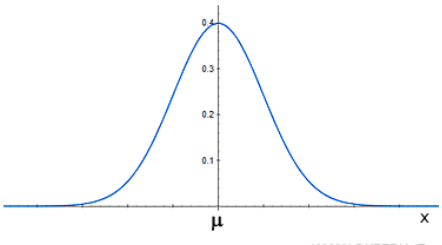


Figura 24 - Esempio di curva di tipo Gaussiano

Per avere un'ulteriore conferma circa la distribuzione normale seguita dai residui, è possibile condurre il test di Anderson-Darling, il quale ci restituisce un valore del P-value da confrontare con il livello di significatività.

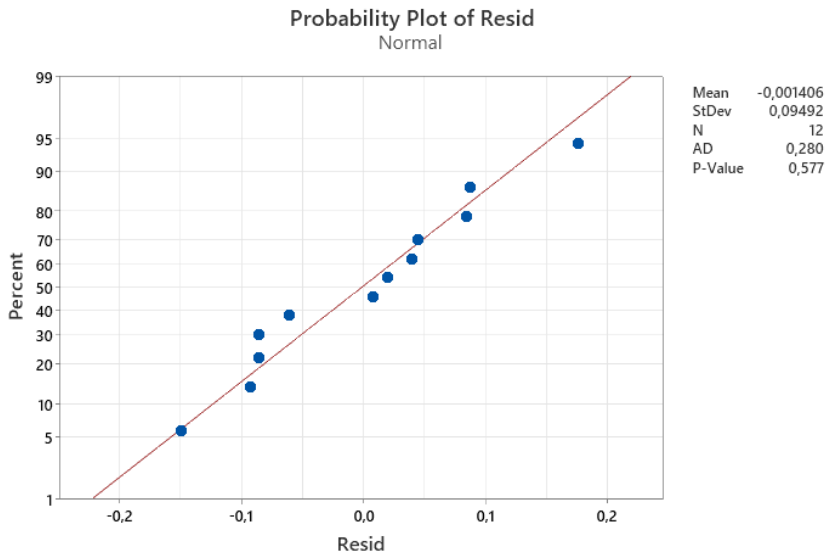


Figura 25 - Esempio di Grafico di Probabilità Normale con test di Anderson–Darling

Questo è uno dei test statistici più usati per stabilire se un determinato campione di dati segue una distribuzione normale o meno. L'ipotesi nulla che si va a testare è che il set di dati segua una distribuzione normale e il suo rifiuto, o meno, è determinato con il confronto il livello di significatività.

Un livello di significatività pari al 5% indica che c'è un rischio del 5% di affermare che i dati non seguano una distribuzione normale quando, in realtà, essi la seguono.

Qualora il valore del P-value sia uguale o minore al valore del livello di significatività stabilito, allora si rifiuta l'ipotesi nulla e quindi si può concludere che i residui non seguono una distribuzione normale. Se, invece, il valore restituito dal P-value è maggiore della soglia imposta allora non si accetta l'ipotesi nulla e si conclude che i residui seguono una distribuzione normale.

Qualora tutte le assunzioni sugli indicatori e sui residui restituiti dalla regressione non fossero rispettate il modello statistico scelto deve essere rigettato poiché non interpola correttamente i dati osservati ed è necessario cambiare il modello.

Le analisi condotte sui tempi e i difetti di assemblaggio e disassemblaggio nelle *prove Randomiche* e nelle *prove Ripetute* sono state effettuate utilizzando una regressione lineare o una regressione con andamento a potenza.

Per poter analizzare la relazione esistente tra la complessità oggettiva e la complessità delle dodici molecole percepita dagli operatori durante le operazioni di assemblaggio e di disassemblaggio si è adoperata una regressione logistica di tipo ordinale.

La regressione logistica, detta anche modello logit, fa parte dei modelli di regressioni non lineari e si adopera quando la variabile dipendente è di tipo binario ovvero quando è una variabile che può assumere solamente due valori: 0 in caso di assenza e 1 in caso di presenza.

L'obiettivo della regressione logistica è quello di stimare la probabilità per la quale si verifica un evento piuttosto che un altro in relazione alle variabili considerate. Esistono varie tipologie di regressioni logistiche che si differenziano a seconda di come è espressa la variabile dipendente. Se questa è espressa mediante l'utilizzo di soli due valori si parla di regressione logistica binaria; se la variabile è espressa mediante due o più valori nominali si utilizza una regressione logistica nominale. I valori delle variabili nominali non hanno nessun ordine specifico. Un esempio di valori nominale è: *acqua*, *terra* e *fuoco*. L'ultima tipologia di regressione logistica, che sarà di fatto quella applicata al caso studio in esame, è la regressione logistica ordinale nella quale la variabile dipendente è espressa da due o più valori. Questi valori, a differenza di quelli nominali, hanno un ordine definito. Un esempio è: *1,2, 3* e *4*.

Per le variabili indipendenti non è posto alcun vincolo e possono assumere uno o più valori, con o senza un ordine definito.

La regressione logistica ordinale stima un coefficiente per ogni variabile indipendente impiegata dal modello e una serie di coefficienti costanti per ogni categoria di esito, meno una. Combinando i valori è possibile strutturare delle equazioni binarie tramite le quali stimare la probabilità si verifichi un evento. In particolare, la prima equazione stima la probabilità per la quale si verifica il primo evento, la seconda equazione stima la probabilità per la quale si verifica il primo o il secondo evento, e così via per i restanti valori dei coefficienti costanti. Quest'ultimi sono etichettati da Minitab® come *Const (1), Const (2)* e così via in base a quante sono le categorie di esito.

Logistic Regression Table

						95% CI	
Predictor	Coef	SE Coef	Z	P	Odds Ratio	Lower	Upper
Const(1)	-1,31163	0,811978	-1,62	0,106			
Const(2)	-0,797106	0,768218	-1,04	0,299			
Const(3)	-0,209999	0,747603	-0,28	0,779			
Const(4)	0,292147	0,748430	0,39	0,696			
Const(5)	0,452648	0,751620	0,60	0,547			
Const(6)	0,948210	0,769913	1,23	0,218			
Const(7)	1,75733	0,829993	2,12	0,034			
Const(8)	2,80743	0,998196	2,81	0,005			
Variabile Indipendente	-0,0265413	0,0486778	-0,55	0,586	0,97	0,89	1,07

Figura 26 - Esempio della stima dei Coefficienti di una Regressione Logistica Ordinale

Il solo valore restituito dai coefficienti ci permette di comprendere come varia la probabilità che si verifichi un determinato evento in seguito ad una variazione della variabile predittiva. In generale, si

può affermare che un valore positivo del coefficiente costante rende più probabile il verificarsi del primo evento, e degli eventi ad esso più vicino, all'aumentare della variabile predittiva. Contrariamente, valori negativi del coefficiente rende più probabile il verificarsi dell'ultimo evento, e degli eventi ad esso più vicino, man mano che a variabile predittiva aumenta. Un valore del coefficiente molto prossimo allo zero, indica che l'impatto del predittore non è influente.

Nel caso studio in esame, volendo valutare la relazione tra la complessità oggettiva e la complessità soggettiva, un valore positivo del coefficiente indica che all'aumentare della complessità oggettiva, gli operatori hanno una maggiore probabilità di scegliere l'opzione "Totalmente d'accordo", mentre un valore negativo indica che all'aumentare della complessità oggettiva, gli operatori hanno una maggiore probabilità di selezionare l'opzione "Totalmente in disaccordo".

Minitab®, inoltre, ci restituisce per ogni predittore un valore del P-value, il quale ci permette di determinare se la relazione esistente tra ogni termine implementato nel modello e la risposta è statisticamente significativa.

L'ipotesi nulla testata è che il coefficiente del termine in esame sia uguale a zero, ovvero che non esista nessuna associazione tra la risposta e il termine in oggetto.

Dopo aver stabilito il livello di significatività pari al 5%, lo si confronta con il valore del P-value. Se il P-value è minore o pari alla soglia imposta, non si accetta l'ipotesi nulla e si conclude che esiste un'associazione statisticamente significativa tra il termine e la variabile risposta. Contrariamente, se il valore del P-value è maggiore della soglia si accetta l'ipotesi nulla e si conclude che non esiste un'associazione statisticamente significativa tra il termine e la risposta.

Un ulteriore analisi è condotta sul comportamento dell'odd ratio.

Questo indicatore ci permette di confrontare la probabilità di due eventi. È calcolato come il rapporto tra la probabilità che si verifichi l'evento e la probabilità che non si verifichi. L'interpretazione del valore restituito cambia a seconda della tipologia del predittore esaminato, se categorico o continuo.

Infatti, se il predittore esaminato è di tipo continuo, un valore dell'odd ratio maggiore di 1 indica che il primo evento, e gli eventi più vicini ad esso, sono più probabili man mano che aumenta il valore del predittore mentre un valore dell'odd ratio minore di 1, indica che all'aumentare del valore del predittore, è più probabile il verificarsi dell'ultimo evento, e degli eventi ad esso vicino.

Se il predittore è categorico, l'odd ratio confronta le probabilità che l'evento si verifichi a due livelli diversi del predittore, quello contenuto nella "Logistic Regression Table" e quello di riferimento, il quale non è presente.

Se è restituito un valore minore di 1, allora l'ultimo evento, e quelli più vicino ad esso, sono più probabili al livello del predittore contenuto nella "Logistic Regression Table" rispetto al livello del predittore. Viceversa, se restituisce un valore maggiore di 1, il primo evento, e quelli più vicino ad esso, sono più probabili al livello del predittore contenuto nella "Logistic Regression Table" rispetto al livello di riferimento del predittore.

Per poter sfruttare le valutazioni basate sull'indicatore odd ratio è necessario che le categorie di risposta siano in ordine.

Un altro strumento che ci permette di valutare la bontà del modello in relazione al set di dati utilizzati è il log-likelihood. Il valore restituito è sempre negativo e valori prossimi allo zero, indicano un migliore adattamento del modello ai dati.

Log-Likelihood = -53,281

Figura 27 - Esempio di Log-Likelihood di una Regressione Logistica Ordinale

Minitab® ci restituisce una tabella contenente i valori di alcuni degli indici principali. Nella *Figura* sottostante sono riportati i valori relativi agli indici di Somers, Goodman-Kruskal e Kendall la cui analisi ci permetto di avere maggiori informazioni sul modello adottato.

Measures of Association:

(Between the Response Variable and Predicted Probabilities)

Pairs	Number	Percent	Summary Measures	Value
Concordant	149	54,4	Somers' D	0,09
Discordant	125	45,6	Goodman-Kruskal Gamma	0,09
Ties	0	0,0	Kendall's Tau-a	0,08
Total	274	100,0		

Figura 28 - Valori degli indici di Somers, Goodman-Kruskal e Kendall restituiti da una Regressione Logistica Ordinale

L'indice di Somers ci permette di analizzare le prestazioni predittive del modello e calcolato come la differenza percentuale tra le coppie concordanti e quelle discordanti, comprendendo i pareggi. I valori spaziano tra 0 e 1 e valori prossimi a 1 indicano una migliore capacità predittiva. L' indice di Goodman-Kruskal, anch'esso definito in un range tra 0 e 1, è statisticamente uguale all'indice di Somers qualora nel modello non sono presenti parità. Infatti, quest'ultimo sintetizza le capacità predittive di un modello considerando solamente la differenza percentuale tra le coppie concordanti e quelle discordanti.

L'ultimo indice riportato è l'indice di Kendall. È calcolato come la differenza percentuale di coppie concordanti e discordanti su tutte le coppie permesse dal modello, comprese quelle con stesso valore

di risposta. Anch'esso ci permette di analizzare e confrontare le capacità predittive del modello adottato e può assume valori compresi tra -2/3 e 2/3.

In generale, più elevati sono i valori restituiti dai tre indici migliore è la capacità predittive del modello.

Come ultimo output, Minitab® ci restituisce la tabella "*Test of All Slopes Equal to Zero*". Grazie ad essa è possibile fare delle considerazioni generali analizzando tutti i coefficienti dei predittori del modello. In esso viene testata l'ipotesi nulla per cui non esiste una relazione statisticamente significativa tra i predittori e gli eventi risposta, ovvero che tutti i coefficienti dei predittori siano uguali a zero.

Ancora una volta è necessario prefissare la soglia di significatività (5%) e confrontarla con il valore restituito dal P-value. Se il valore del P-value è minore o uguale al livello di significatività si può concludere che esiste un'associazione statisticamente significativa tra l'evento risposta ed almeno un predittore. Qualora il valore del P-value sia maggiore della soglia imposta, non si può concludere che esista una relazione significativa tra la risposta e uno tra i predittori.

CAPITOLO 4: Risultati Sperimentali e Analisi dei Risultati

In questo capitolo sono riportati i risultati ottenuti dall'analisi dell'ANOVA e dall'analisi delle regressioni, lineari e non lineari, ottenute tramite il software Minitab®.

4.1 ANOVA

Per poter valutare la relazione tra i difetti o i tempi di lavorazione delle molecole e la complessità oggettiva, è stata condotta un'analisi Two-Way ANOVA al fine di valutare gli effetti che due variabili indipendenti potrebbero avere su una variabile dipendente.

I test sono stati condotti sia per le *prove Ripetute* sia per le *prove Randomiche*. Per entrambe le prove si è scelto di valutare gli effetti che si potrebbero generare sui tempi e sui difetti nelle operazioni di assemblaggio e di disassemblaggio. Per le analisi che hanno come oggetto le osservazioni rilevate nelle *prove Randomiche*, si è deciso di valutare gli effetti che la *Complessità Oggettiva* e l'*Operatore* potrebbero avere sulla variabile dipendente. Per le analisi delle *prove Ripetute*, invece, si è deciso di valutare gli effetti che il *Numero di Prova* e l'*Operatore* potrebbero avere sui tempi e i difetti riscontrati nelle singole molecole.

Mentre per le *prove Randomiche* la variabile dipendente comprende tutti i tempi, o difetti, delle molecole, nelle *prove Ripetute* si è deciso di effettuare un'analisi sui tempi e difetti delle singole molecole.

Per entrambe le prove, le variabili dipendenti analizzate sono: *Difetti di Processo di Assemblaggio*, *Difetti di Prodotto di Assemblaggio*, *Difetti Totali di Assemblaggio*, *Tempi di Assemblaggio*, *Difetti di Disassemblaggio* e *Tempi di Disassemblaggio*.

Come esempio, di seguito sono riportate le analisi del test Two-Way ANOVA condotte sui tempi (*Figura 29*) e sui difetti totali (*Figura 30*) riscontrati in fase di assemblaggio delle *prove Randomiche* e le analisi condotte sui tempi (*Figura 31*) e sui difetti di assemblaggio della molecola ID8 (*Figura 32*) delle *prove Ripetute*.

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Complessità	11	32597	2963,4	25,46	0,000
Operatore	51	15678	307,4	2,64	0,000
Error	561	65301	116,4		
Total	623	113577			

Figura 29 - Analisi Two-Way ANOVA Tempi di Assemblaggio vs Complessità Oggettiva vs Operatore nelle prove Randomiche

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Complessità	11	123803	11254,8	154,23	0,000
Operatore	51	9230	181,0	2,48	0,000
Error	561	40938	73,0		
Total	623	173971			

Figura 30 - Analisi Two-Way ANOVA Difetti Totali di Assemblaggio vs Complessità Oggettiva vs Operatore nelle prove Randomiche

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
ESECUZIONE	6	2165	360,88	4,81	0,000
OPERATORE	13	1415	108,87	1,45	0,156
Error	78	5851	75,01		
Total	97	9431			

Figura 31 - Analisi Two-Way ANOVA Tempi di Assemblaggio vs Complessità Oggettiva vs Operatore della molecola ID8 nelle prove Ripetute

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
ESECUZIONE	6	1503,3	250,543	29,03	0,000
OPERATORE	13	1036,5	79,729	9,24	0,000
Error	78	673,2	8,630		
Total	97	3212,9			

Figura 32 - Analisi Two-Way ANOVA Difetti Totali di Assemblaggio vs Complessità Oggettiva vs Operatore della molecola ID8 nelle prove Ripetute

Come si può notare, il valore del P-value restituito dalle analisi è maggiore del livello di significatività, 5%, solo per la variabile *Operatore* nel caso in cui si analizza la variabile dipendente *Tempo di Assemblaggio* della molecola ID8 durante lo svolgimento delle *prove Ripetute*. Quindi possiamo rifiutare l'ipotesi nulla e si può affermare che la variabile indipendente *Operatore* non è significativa.

In tutti gli altri casi in esame, è restituito un valore del P-value minore della soglia e quindi si accetta l'ipotesi nulla e si conclude che la variabile indipendente in esame è significativa.

4.2 Regressione prove Randomiche

Dopo aver condotto le analisi sull'ANOVA e aver stabilito quali sono le variabili significative e quali no, si è effettuata un'analisi di regressione per trovare l'equazione che meglio spiegasse la relazione che esiste tra la variabile dipendente e la variabile indipendente. Le analisi di regressione condotte di seguito sono state effettuare considerando i tempi e dei difetti medi riscontrati nell'assemblaggio e disassemblaggio delle molecole. Non si è analizzato l'effetto che la variabile Operatore potesse avere sulle variabili dipendenti. Si è deciso di condurre le analisi in questo modo per poter eliminare la molta dispersione contenuta nei dati grezzi osservati.

Per le analisi di regressione condotte sui valori osservati nelle *prove Randomiche*, si è stabilito di utilizzare come variabile dipendente la media dei tempi, o dei difetti, osservati nelle dodici molecole oggetto di studio. Mentre, per le analisi di regressione condotte sui valori osservati nelle *prove Ripetute*, si è valutata la relazione tra la media, per ogni esecuzione, dei tempi, o dei difetti, e l'esecuzione stessa.

Per ogni regressione è stato effettuato un confronto tra gli esiti di una regressione lineare e gli esiti di una regressione non lineare al fine di valutare quale tra le due interpolasse meglio i dati.

Come esempio, è riportata l'analisi di regressione effettuata considerando come variabile dipendente i difetti medi totali di assemblaggio riscontrati nelle *prove Randomiche* e il confronto tra una regressione di tipo lineare e una non lineare.

L'output restituito da Minitab® di un modello di regressione di tipo lineare è riportato di seguito:

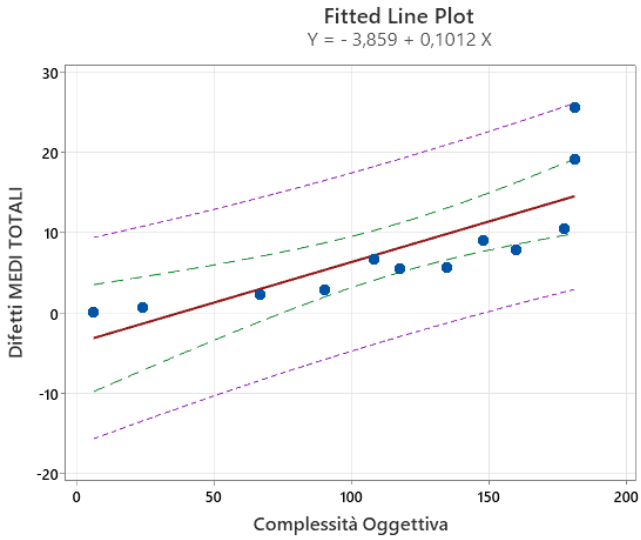


Figura 33 - Curva di Regressione di una Regressione Lineare Difetti Medi Totali di assemblaggio vs Complessità Oggettiva nelle prove Randomiche

Model Summary

S	R-sq	R-sq(adj)
4,76635	63,76%	60,14%

Figura 34 - Principali statistiche riassuntive di una Regressione Lineare Difetti Medi Totali di assemblaggio vs Complessità Oggettiva nelle prove Randomiche

Il valore dell'indicatore R^2 pari a 63,76% indica la percentuale della variabilità spiegata dal modello lineare. È riportato anche il valore dell'indicatore R^2 (adj), il quale svolge la stessa funzione dell'indicatore R^2 ma nel suo calcolo è considerato anche il numero di elementi analizzati. In questo caso indica che il 60,14% della variabilità è spiegata dal modello adottato.

Per poter effettuare un confronto con una regressione non lineare, è necessario valutare il valore restituito dall'indicatore S poiché l'indicatore R^2 ha significato solamente in ambito di regressione lineare.

Il modello che fitterà meglio i dati sarà quello che presenta un valore minore dell'indicatore S.

Al fine di valutare quale modello di regressione non lineare adottare per l'analisi statistica è necessario inserire i dati osservati all'interno di uno Scatterplot, in cui l'asse delle ordinate rappresenta la variabile indipendente, ovvero la complessità oggettiva, e l'asse delle ascisse rappresenta i difetti medi totali raccolti durante l'esecuzione di assemblaggio delle molecole.

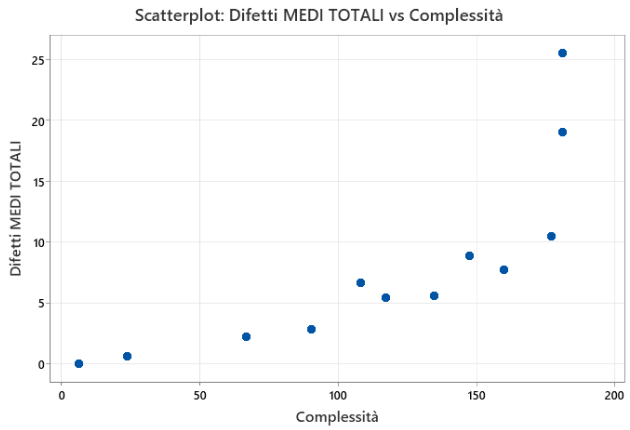


Figura 35 - Grafico di Dispersione o Scatterplot Difetti Medi Totali di Assemblaggio vs Complessità Oggettiva nelle prove Randomiche

Come si può notare in *Figura 35*, l'andamento che sembrerebbe meglio descrivere i dati osservati è l'andamento a potenza.

L'output restituito da Minitab® di un modello di regressione con andamento a potenza è il seguente:

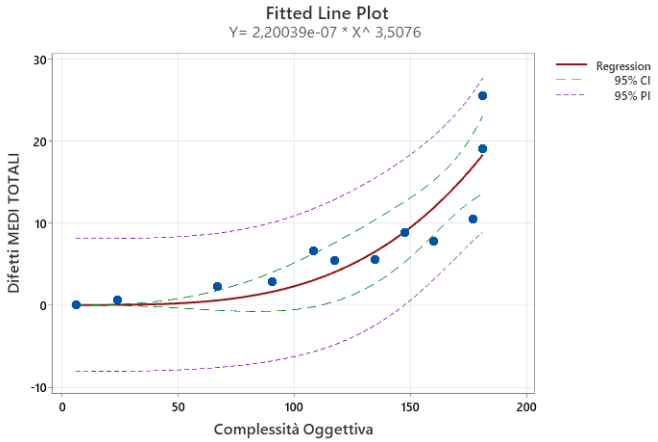


Figura 36 - Curva di Regressione di una Regressione con andamento a Potenza Difetti Medi Totali di Assemblaggio vs Complessità Oggettiva nelle prove Randomiche

Summary

Iterations	57
Final SSE	131,859
DFE	10
MSE	13,1859
S	3,63124

Figura 37 - Principali statistiche riassuntive di una Regressione con andamento a Potenza Difetti Medi Totali di Assemblaggio vs Complessità Oggettiva nelle prove Randomiche

Per stabilire il miglior modello per l'interpolazione dei dati è necessario confrontare il valore restituito dall'indicatore S in quanto più piccolo è il valore, più i valori reali si avvicinano alla curva di regressione e quindi migliore è il modello.

Nel caso di utilizzo di un modello lineare, il valore dell'indicatore S restituito è pari a 4,76635 mentre nel caso di utilizzo di un modello a potenza è pari a 3,63124: il modello migliore risulta essere quello descritto da un'equazione a potenza.

Prima di poter concludere l'analisi affermando che la relazione tra le due variabili può essere espressa con l'adozione di un modello con andamento a potenza è necessario condurre l'analisi sui residui restituiti.

In questo caso, il confronto tra l'analisi condotta sui residui del modello lineare e sui residui del modello con andamento a potenza conferma quanto detto sopra: il modello a potenza fitta meglio i dati osservati.

L'output dei residui del modello lineare evidenzia il mancato soddisfacimento delle ipotesi sui residui.

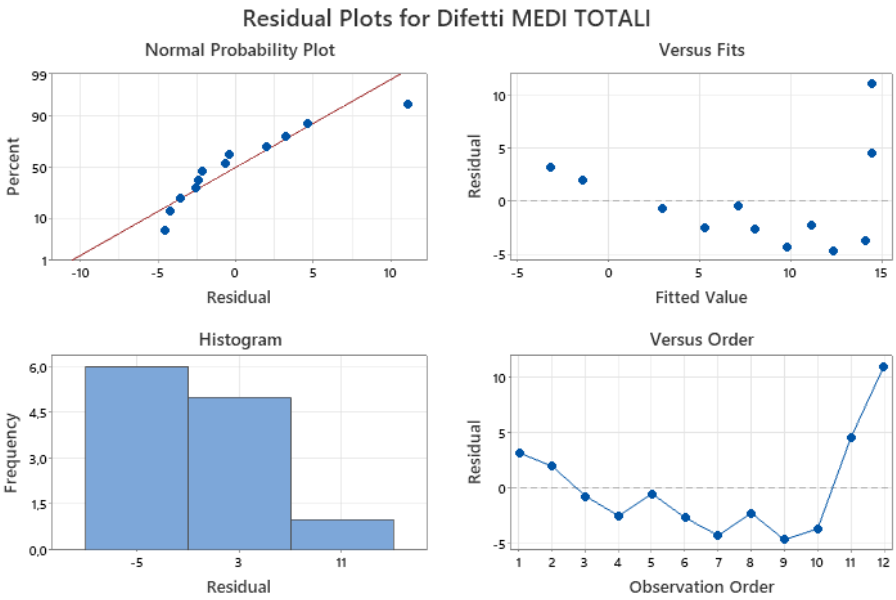


Figura 38 - Output dei Grafici dei Residui restituiti dal modello Lineare

Il *Normal Probability Plot* e l'*Histogram* non confermano l'ipotesi forte per cui i residui devono seguire una distribuzione normale. Nel primo, infatti, si nota che i residui non sono disposti lungo la linea retta inclinata e nel secondo, le classi di residui non seguono l'andamento della curva gaussiana, tipico di una distribuzione normale. Ulteriori conferme sulla non bontà del modello sono fornire dall'analisi dei grafici *Versus Fits* e *Versur Order.* In essi si nota che i punti non si dispongono in modo omogeneo al di sopra e al di sotto della linea centrale dello 0, ma tendono a posizionarsi al di sotto.

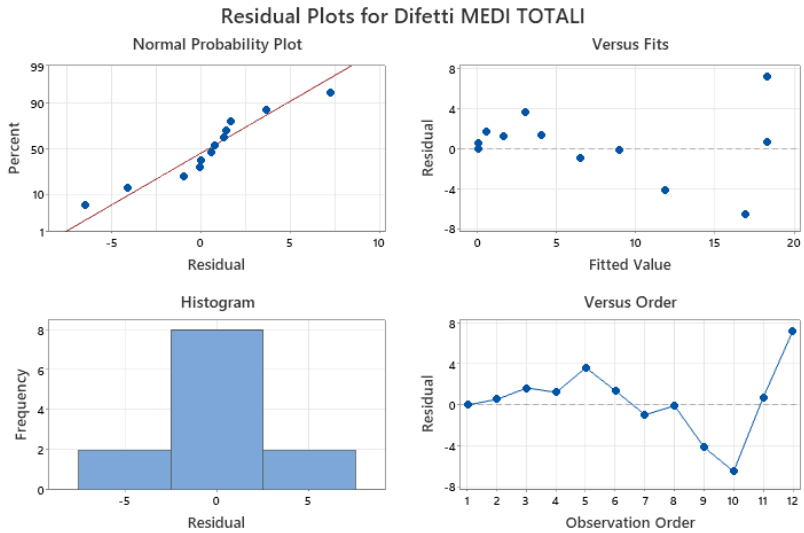


Figura 39 - Output dei Grafici dei Residui restituiti dal modello con andamento a potenza

Dall'analisi dei residui della regressione con andamento a potenza si nota, invece, che tutte le ipotesi sui di essi sono confermate. Sia il *Normal Probability* Plot sia l'*Histogram* confermano che i residui

seguono una distribuzione normale. Nel primo grafico i punti sono collocati tutti in prossimità della retta inclinata e nel secondo si evince l'andamento a curva di campana seguito dalle classi di residui.

Anche il *Versus Fits* e *il Versus Order* confermano la bontà del modello in quanto i punti sono disposti in modo omogeno al di sopra e al di sotto della linea centrale dello zero e non si individuano né nuvole di punti né particolari trend.

Un'analisi quantitativa, circa la distribuzione normale dei residui, è ottenuta eseguendo il test di normalità di Anderson-Darling riportato di seguito:

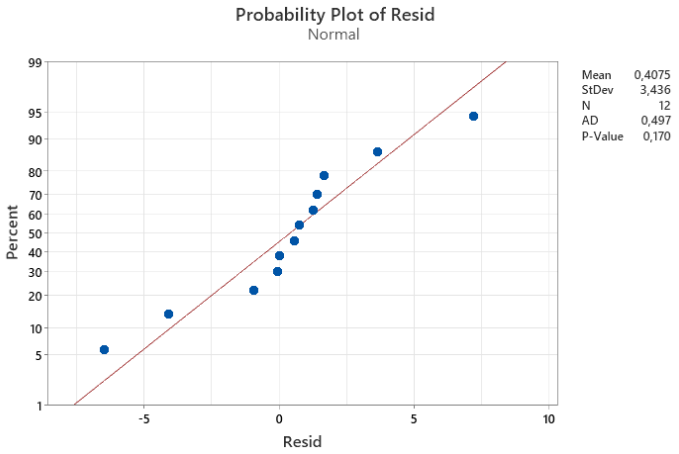


Figura 40 - Risultato test di Anderson-Darling condotto sui Residui ottenuti da un modello con Potenza Difetti Medi Totali di Assemblaggio vs Complessità Oggettiva nelle prove Randomiche

Essendo il valore restituito dal P-value maggiore del livello di significatività stabilito, 5%, si può concludere che si accetta l'ipotesi nulla e quindi i residui seguono una distribuzione normale.

Per quanto analizzato, si arriva a conclusione che una legge a potenza, del tipo $Y = \theta 1 X^{\theta 2}$, è l'equazione che riesce meglio a descrivere la relazione esistente tra i difetti medi totali e la complessità oggettiva. In particolare, $\theta 1$ e $\theta 2$ rappresentano i coefficienti della curva e sono rispettivamente pari a $2{,}20 \cdot 10^{-07}$ e $3{,}51$.

L'equazione della curva, pertanto, è la seguente:

$$\textbf{Difetti Medi Totali} = \mathbf{2,20 \cdot 10^{-07} \cdot C^{3,51}}$$

Il valore positivo dell'esponente della variabile indipendente *Complessità Oggettiva* descrive la correlazione positiva con la variabile dipendente *Difetti* Medi Totali e quindi si può affermare che all'aumentare della complessità oggettiva delle molecole da assemblare maggiore è il numero di difetti totali commessi dagli operatori durante la fase di assemblaggio.

Anche il confronto tra una regressione lineare ed una regressione con andamento a potenza considerando i tempi medi di assemblaggio osservati durante le *prove Randomiche* e la complessità oggettiva delle molecole, mostra che un'equazione con andamento a potenza è quella che interpola meglio i dati.

In *Figura 41* e *Figura 42* e in *Figura 43* e *Figura 44* sono riportate rispettivamente i risultati ottenuti dalla regressione lineare e dalla regressione con andamento a potenza.

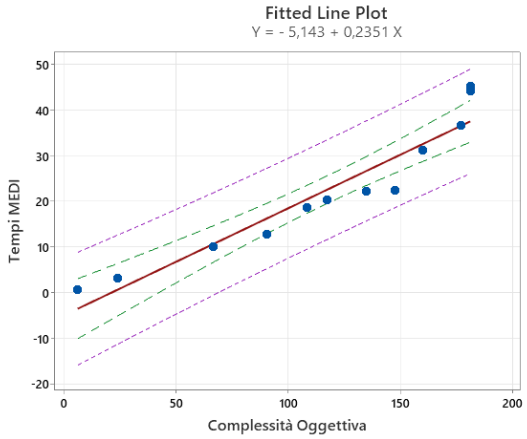


Figura 41 - Curva di Regressione di una Regressione Lineare Tempi Medi di Assemblaggio vs Complessità Oggettiva nelle prove Randomiche

Model Summary

S	R-sq	R-sq(adj)
4,70578	90,70%	89,77%

Figura 42 - Principali statistiche riassuntive di una Regressione lineare Tempi Medi di Assemblaggio vs Complessità Oggettiva nelle prove Randomiche

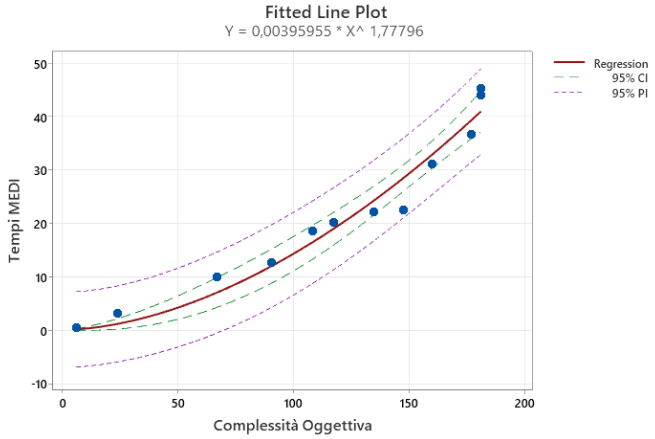


Figura 43 - Curva di Regressione di una Regressione con andamento a Potenza Tempi di Assemblaggio vs Complessità Oggettiva nelle prove Randomiche

Summary

Iterations	20
Final SSE	100,142
DFE	10
MSE	10,0142
S	3,16452

Figura 44 - Principali statistiche riassuntive di una Regressione con andamento a potenza Tempi Medi di Assemblaggio vs Complessità Oggettiva nelle prove Randomiche

Come si nota da una prima analisi, i valori restituiti dalla regressione lineare degli indicatori R^2 e R^2 (adj) sono abbastanza buoni e quindi si può affermare che circa il 90% della variabilità è spiegata dal modello. Confrontando il valore dell'indicatore S restituito della regressione lineare e quello restituito dalla regressione con andamento a potenza, si può concludere che, nonostante un buon valore dell'indicatore R^2, il modello con andamento a potenza la relazione tra le variabili.

Infatti, il valore restituito da S dall'adozione di un modello a potenza è pari a 3,16 il quale risulta minore del valore restituito dal modello lineare (4,71) e giustifica la scelta del modello. L'analisi effettuata sui residui conferma la scelta del modello a potenza in quanto sono rispettate le assunzioni di una distribuzione normale, del collocamento casualmente e omogeneamente attorno alla linea centrale dello zero e non sono presenti evidenze di particolari trend.

Nella figura sottostante sono riportati i grafici utilizzati per condurre l'analisi dei residui in relazione ad una regressione lineare e ad una regressione con andamento a potenza.

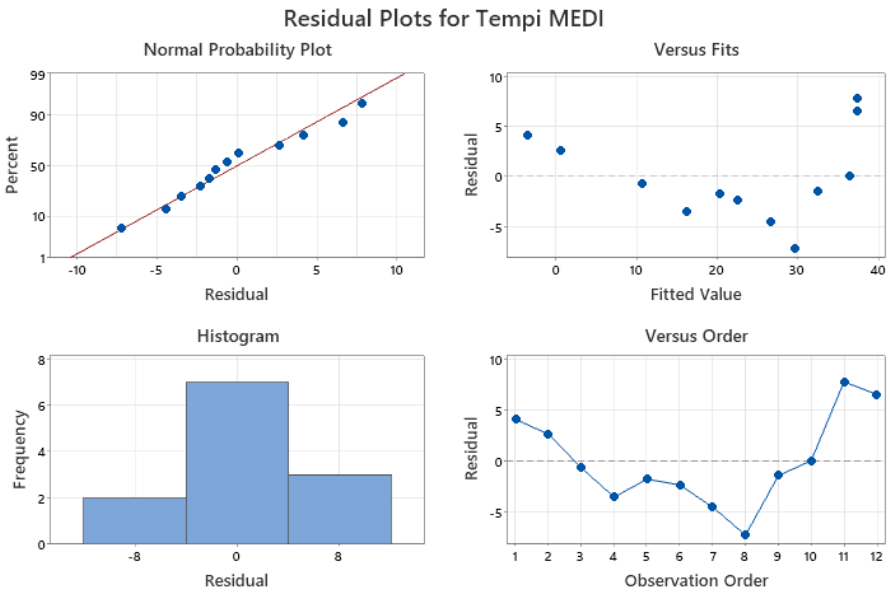


Figura 45 - Output dei Grafici dei Residui restituiti dal modello lineare

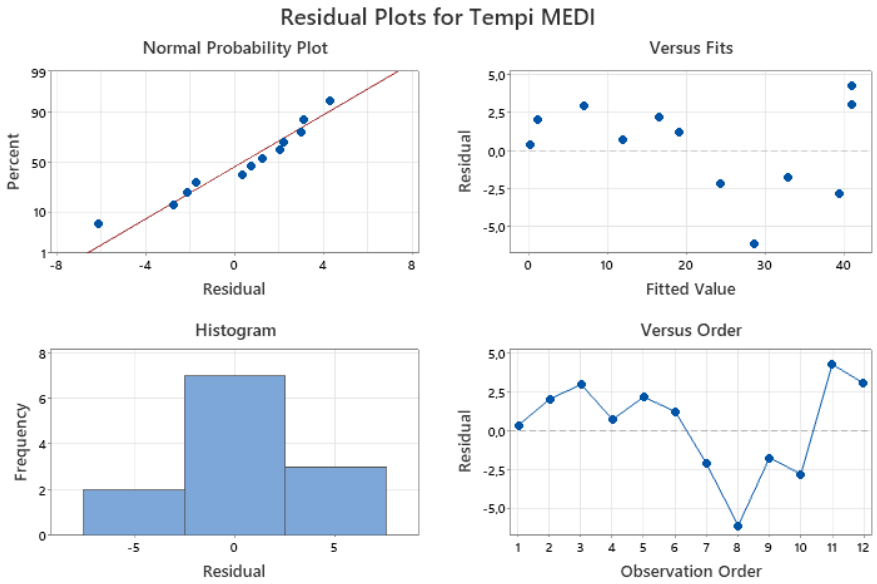


Figura 46 - Output dei Grafici dei Residui restituiti dal modello con andamento a potenza

È stato effettuato inoltre il test di Anderson-Darling sui residui del modello con andamento a potenza (*Figura 49)*. Si ricava un valore di P-value pari a 0,429, superiore alla soglia del 5% prefissata, e quindi non si può affermare che i residui non seguano una distribuzione normale.

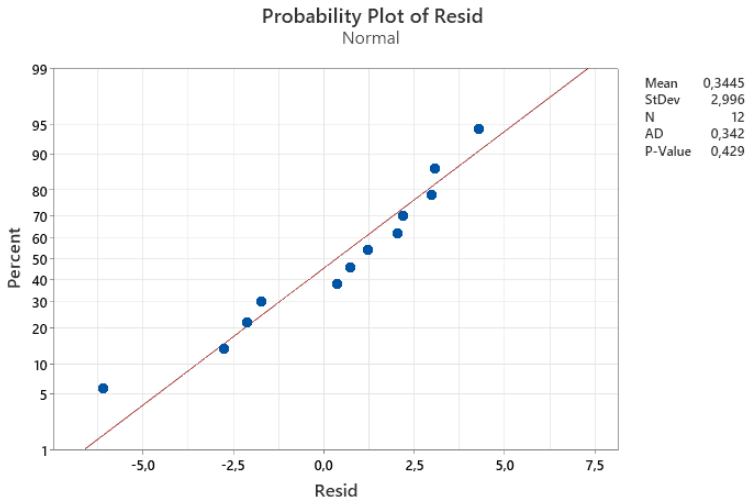


Figura 47 - Risultato test di Anderson-Darling condotto sui Residui ottenuti da un modello con andamento a Potenza Tempi Medi di Assemblaggio vs Complessità Oggettiva nelle prove Randomiche

Alla luce dei delle evidenze riscontrate, l'equazione che meglio descrive la relazione tra i tempi medi di assemblaggio e la complessità oggettiva nelle *prove Randomiche* è la seguente:

$$\textbf{Tempi Medi} = 0,004 \cdot \mathbf{C}^{1,78}$$

4.3 Regressione prove Ripetute

Per le *prove Ripetute* è stata riportata come esempio la relazione esistente tra i difetti medi riscontrati in fase di disassemblaggio della molecola ID10 e le esecuzioni effettuate e la relazione esistente tra il tempo di disassemblaggio della molecola ID10 e le esecuzioni effettuate. Per esecuzioni effettuate si intende il numero di volte che è stata ripetuta l'operazione di disassemblaggio per il singolo studente nell'arco di una giornata lavorativa.

Come per le analisi sopra riportate, per determinare il modello da utilizzare, è stato effettuato un confronto tra una regressione lineare e una regressione non lineare e la scelta del modello è stata presa valutando i valori restituiti dai principali indicatori e verificando il rispetto delle ipotesi sui residui delle due regressioni.

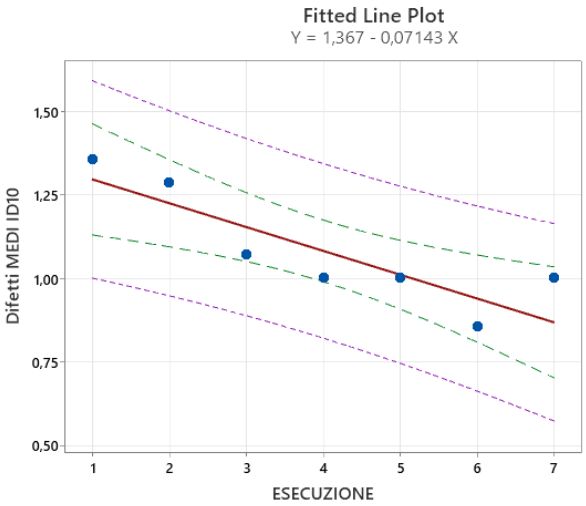


Figura 48 - Curva di Regressione di una Regressione Lineare Difetti Medi Totali di Disassemblaggio vs Esecuzione della Molecola ID10 nelle prove Ripetute

Model Summary

S	R-sq	R-sq(adj)
0,0950679	75,97%	71,16%

Figura 49 - Principali statistiche riassuntive di una Regressione Lineare Difetti Medi Totali di Disassemblaggio vs Esecuzione della Molecola ID10 nelle prove Ripetute

Come si nota in *Figura 48*, la regressione lineare restituisce dei valori degli indicatori prossimi a 1. Il modello lineare spiega circa il 76% della variabilità; infatti, l'indicatore R^2 è pari a 75,97%. Il valore dell'indicatore S, pari a 0,095, prossimo allo zero suggerisce un buon adattamento del modello al set

di dati osservati. Tuttavia, per effettuare la scelta del modello da utilizzare, è necessario confrontare questo valore con quello restituito dalla regressione non lineare.

Riportando i valori osservati all'interno di uno *Scatterplot*, in cui le ascisse sono rappresentanti il numero di esecuzioni effettuate e le ordinate rappresentano la media dei difetti osservati nelle singole esecuzioni, si nota che è possibile interpolare i dati mediante un andamento a potenza. In figura sottostante è riportato lo *Scatterplot* ottenuto tramite Minitab ®:

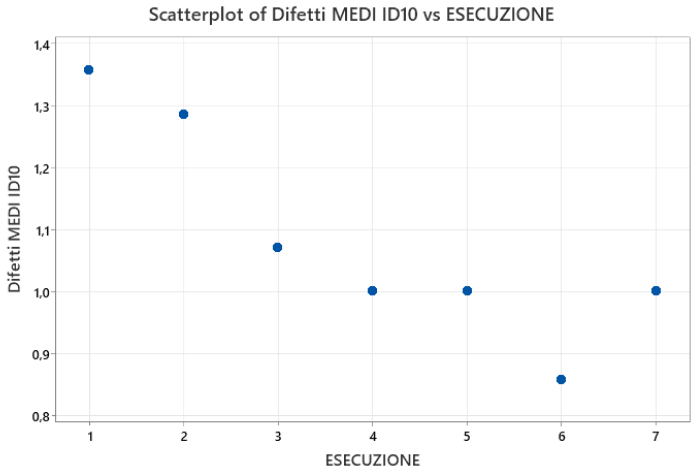


Figura 50 - Grafico di Dispersione o Scatterplot Difetti Medi Totali di Disassemblaggio vs Esecuzione della Molecola ID10 nelle prove Ripetute

Il confronto tra i valori dell'indicatore S restituiti dalla regressione lineare e dalla regressione con andamento a potenza, suggerisce che quest'ultimo andamento si adatta meglio ai dati osservati.

Infatti, il valore di S restituito, pari a 0,072, è minore del valore restituito dalla regressione lineare (0,095).

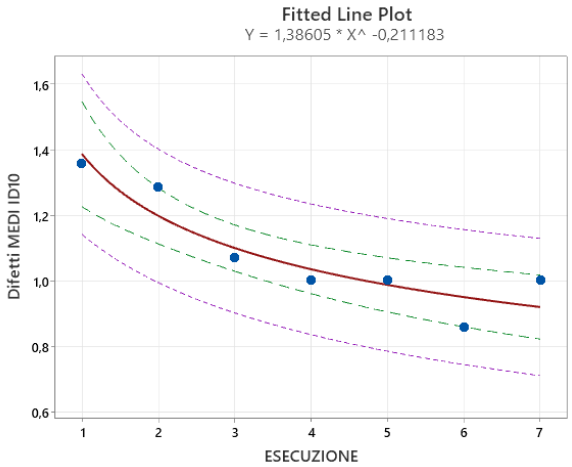


Figura 51 - Curva di Regressione di una Regressione con andamento a Potenza Difetti Medi Totali di Disassemblaggio vs Esecuzione della Molecola ID10 nelle prove Ripetute

Summary

Iterations	5
Final SSE	0,0258402
DFE	5
MSE	0,0051680
S	0,0718891

Figura 52 - Principali statistiche riassuntive di una Regressione con andamento a Potenza Difetti Medi Totali di Disassemblaggio vs Esecuzione della Molecola ID10 nelle prove Ripetute

La bontà del modello è confermata anche dall'analisi eseguita sui residui, la quale conferma l'ipotesi di normalità e l'ipotesi per la quale essi si dispongono in maniera casuale, senza evidenziare trend e nuvole di punti, attorno la linea dello zero.

Infatti, il *Normal Probability Plot* e l'*Histogram* confermano la prima ipotesi sui residui. Dal primo grafico si osserva che i punti si dispongono lungo una linea retta e nel secondo che le classi di punti seguono un andamento a curva di campana, tipica di una regressione normale. Osservando il *Versus Fits* e il *Versus Order* si nota che i punti si dispongono casualmente e omogeneamente lungo la retta dello zero e non sono riconoscibili né particolari trend né nuvole di punti.

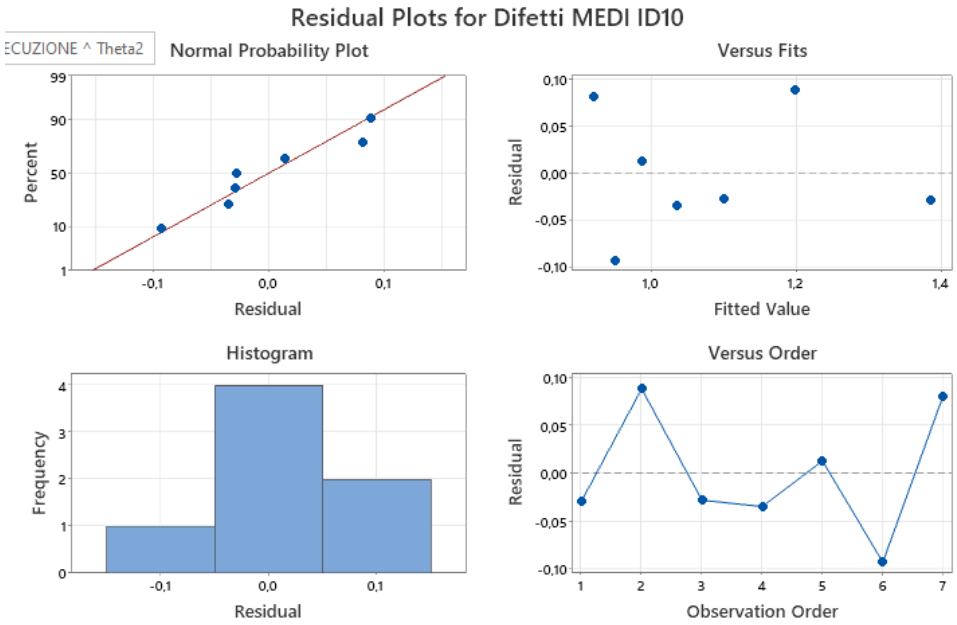


Figura 53 - Output dei Grafici dei Residui restituiti dal modello con andamento a potenza

Anche in questo caso è stato effettuato il test di normalità di Anderson-Darling per dimostrare l'ipotesi di normalità dei difetti attraverso il valore restituito dal P-value.

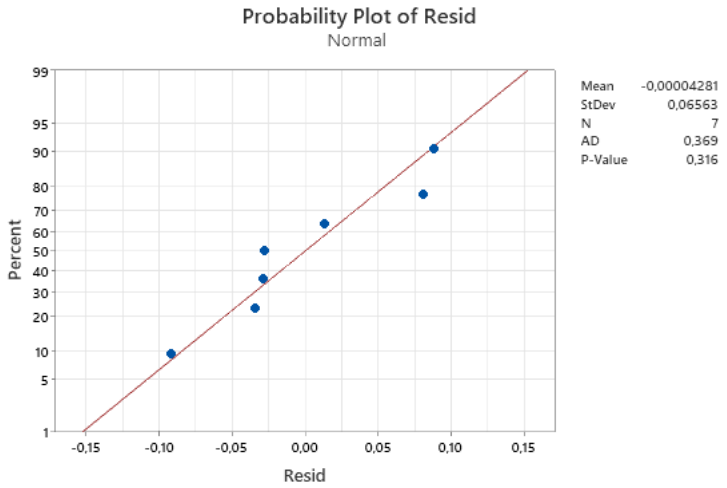


Figura 54 - Risultato test di Anderson-Darling condotto sui Residui ottenuti da un modello con andamento a Potenza Tempi Medi di Disssemblaggio vs Esecuzione della Molecola ID10 nelle prove Ripetute

Come si osserva dall'output del test (*Figura 54*), il P-value restituito è pari a 0,316 e quindi maggiore del livello di significatività scelto. Questo è sufficiente per poter concludere che i dati seguono una distribuzione normale.

Alla luce di quanto detto, l'equazione che descrive meglio il modello è la seguente:

$$\textbf{Difetti Medi ID10} = 1,39 \cdot \textbf{Esecuzione}^{-0,21}$$

Il valore negativo dell'esponente della variabile indipendente mostra una relazione inversamente proporzionale con la variabile dipendente. Al crescere del numero di prove di disassemblaggio eseguita dagli operatori, i difetti totali riscontrati nelle singole prove sono minori.

Anche per lo studio della relazione tra i tempi di disassemblaggio della molecola ID10 e il numero di esecuzioni effettuate è stato realizzato un confronto tra i principali indicatori ottenuti dalla regressione lineare ed una relazione non lineare al fine di scegliere il modello migliore che descrivesse tale relazione.

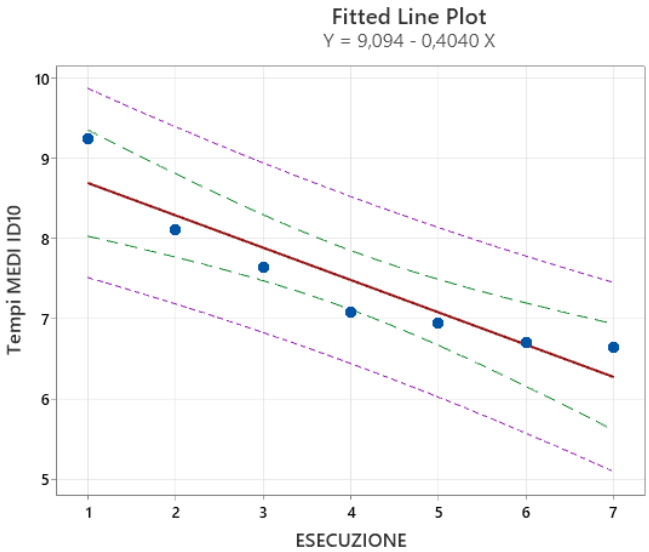


Figura 55 - Curva di Regressione di una Regressione Lineare Tempi Medi di Disassemblaggio vs Esecuzione della Molecola ID10 nelle prove Ripetute

Model Summary

S	R-sq	R-sq(adj)
0,379226	86,41%	83,69%

Figura 56 - Principali statistiche riassuntive di una Regressione Lineare Tempi Medi di Disassemblaggio vs Esecuzione della Molecola ID10 nelle prove Ripetute

I valori restituiti dalla regressione lineare degli indicatori S e R^2 potrebbero suggerire che il modello lineare descrive correttamente i dati. Infatti, circa l'87% della variabilità è spiegata dal modello e un valore di S prossimo allo zero indica un buon adattamento al set di dati osservati.

Tuttavia, analizzando i residui ottenuti mediando i dati attraverso la regressione lineare, si evince che l'ipotesi per cui si devono disporre in maniera casuale attorno alla linea dello zero, senza che si notino determinati trend, non è rispettata.

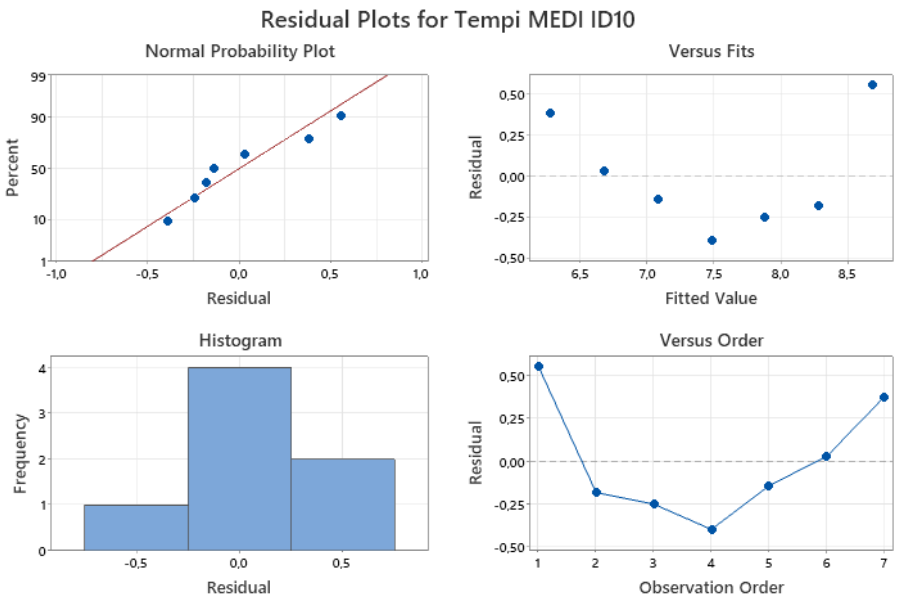


Figura 57 - Output dei Grafici dei Residui restituiti dal modello lineare

Se l'ipotesi di normalità sembra essere confermata dall' *Histogram* e dal *Normal Probability Plot*, il *Versus Fits* e *il Versus Order* risaltano un trend a parabola con concavità verso l'alto, sufficiente per rigettare il modello lineare.

L'andamento evidenziato dallo *Scatterplot* sottostante suggerisce, ancora una volta, l'impiego di una regressione con andamento a potenza.

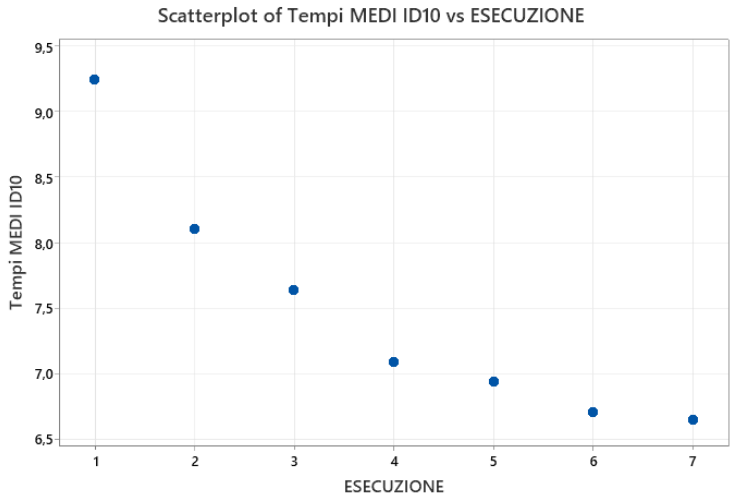


Figura 58 - Grafico di Dispersione o Scatterplot Tempi Medi di Disassemblaggio vs Esecuzione della Molecola ID10 nelle prove Ripetute

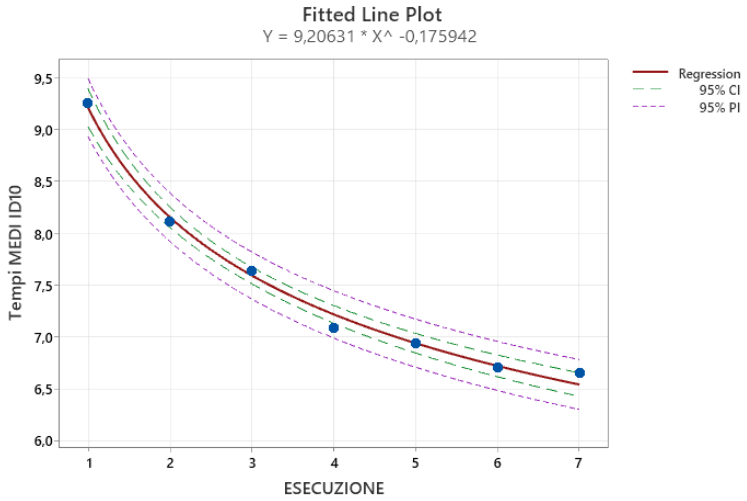


Figura 59 - Curva di Regressione di una Regressione con andamento a Potenza Tempi Medi di Disassemblaggio vs Esecuzione della Molecola ID10 nelle prove Ripetute

Summary

Iterations	8
Final SSE	0,0339362
DFE	5
MSE	0,0067872
S	0,0823847

Figura 60 - Principali statistiche riassuntive di una Regressione con andamento a Potenza Tempi Medi di Disassemblaggio vs Esecuzione della Molecola ID10 nelle prove Ripetute

Il set di dati raccolti è descritto bene dal modello a potenza in quanto il valore dell'indicatore S è prossimo allo zero. Come si è visto, la bontà del modello non può essere confermata prendendo in considerazione solamente quest'ultimo valore. È necessario studiare il comportamento dei residui generati.

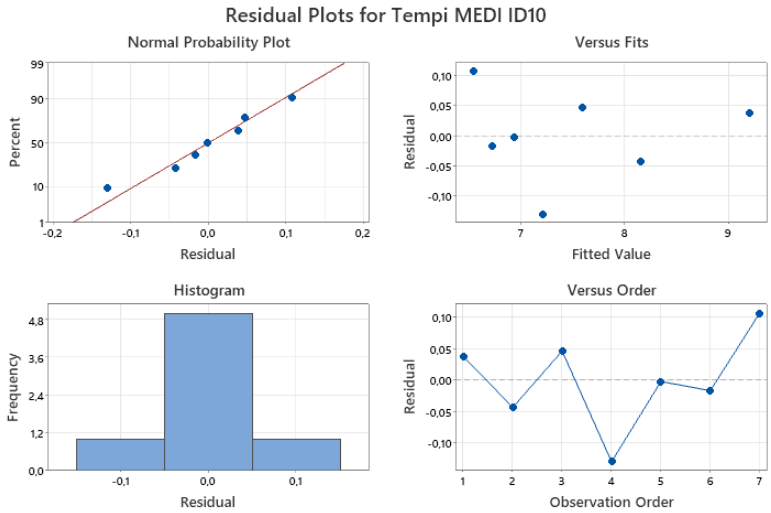


Figura 61 - Output dei Grafici dei Residui restituiti dal modello con andamento a potenza

L'ipotesi di normalità dei residui è confermata dal *Normal Probability Plot*, in quanto i punti si adagiano lungo la linea retta inclinata, e dall' *Histogram*, in quanto le classi di punti seguono un andamento di una curva a campana.

Nei restanti grafici si nota che i residui si dispongono in maniera casuale, senza far emergere particolari tendenze o nuvole di punti, lungo la linea centrale dello 0.

Come ultima verifica è stato effettuato un test di normalità di Anderson-Darling, il quale restituendo un valore del P-value di 0,857 maggiore del livello di significatività del 5%, permette di concludere l'analisi, accettando l'ipotesi nulla per cui i residui seguono una distribuzione normale.

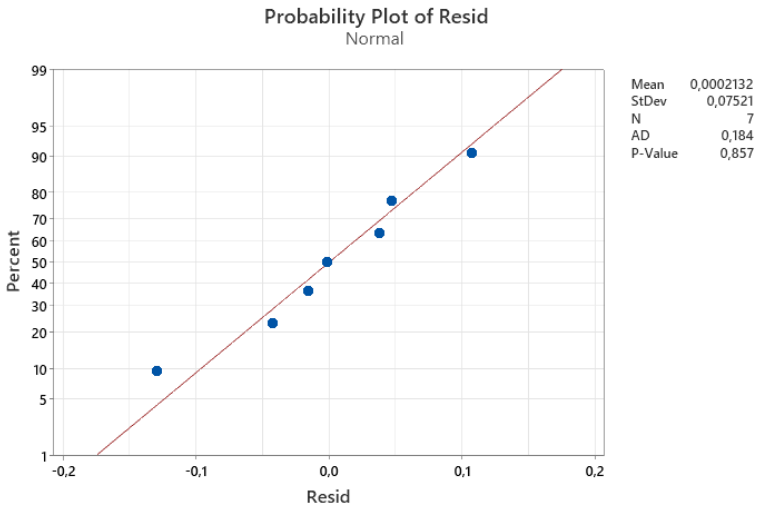


Figura 62 - Risultato test di Anderson-Darling condotto sui Residui ottenuti da un modello con andamento a Potenza Tempi Medi di Disassemblaggio Esecuzione della Molecola ID10 nelle prove Ripetute

Alla luce di quanto analizzato, la migliore equazione che interpola il set di dati raccolti e per descrivere la relazione esistente tra i tempi medi di disassemblaggio della molecola ID10 e il numero di volte che l'operatore ha dovuto ripetere l'operazione nelle *prove Ripetute* è la seguente:

$$\textbf{Tempi ID10} = 9,21 \cdot \textbf{Esecuzione}^{-0,18}$$

L'esponente negativo, e il conseguente andamento decrescente della curva, indica che i tempi di disassemblaggio della molecola tendono a diminuire maggiori sono le volte che si esegue l'operazione di disassemblaggio.

In *Appendice* sono riportati i valori dell'indicatore S, dell'indicatore R2, in caso di regressione lineare, e l'equazione che meglio spiega la relazione esistente tra la variabile dipendente e quella indipendente.

4.4 Regressione Complessità Soggettiva vs Complessità Oggettiva

Come abbiamo visto finora, per analizzare le relazioni tra la complessità oggettiva e tra i difetti e i tempi di assemblaggio e disassemblaggio nelle *prove Randomiche* e nelle *prove Ripetute* si è utilizzata la regressione con andamento lineare e con andamento a potenza.

Per analizzare i dati relativi alla complessità soggettiva, ovvero la complessità della molecola percepita dall'operatore che esegue l'operazione di assembaggio o di disassemblaggio, si è sfruttata la regressione logistica ordinale poiché permette di non lavorare con il concetto di "distanza", il quale con l'adozione di scale ordinali non dovrebbe essere introdotto. Quest'ultima tipologia di regressione è un'estensione della regressione logistica e si applica quando la variabile dipendente può assumere più di due valori. In questo caso, infatti, la complessità soggettiva può assumere solo categorie di risposte prestabilite, ovvvero; *Totalmente in disaccordo; In disaccordo; Relativamente d'accordo*; *D'accordo* e *Totalmente d'accordo*. Ogni categoria, successivamente, è stata convertita in una scala numerica da 1 a 5, dove 1 rappresenta la categoria di risposta *Totalmente in disaccordo* e 5 la categoria di risposta *Totalmente d'accordo*.

Affinchè il modello di regressione logistica ordinale abbia senso è necessario che la variabile dipendente soddisfi due assunzioni di base. È necessario che essa sia categorica, cioè che possa assumere più di due valori compresi categorie di risposta, e che sia ordinata, cioè che esista un ordine naturale nella classificazione delle categorie.

Si sono analizzate le complessità oggettive delle singole molecole e le valutazioni soggettive delle molecole per operatore, pesate sulla base del grado di importanze affidato al criterio.

Inoltre, sfruttando l'indicatore OWA (Ordered Weighted Averanging), il quale ci permette di adoperare gli operatori di aggregazione, come la media, in un contesto ordinale (Franceschini, 2005) dove il concetto di distanza non trova una definizione, sono stati studiati i valori medi dei criteri pesati sulla base del loro grado di importanza.

I dati raccolti, ottenuti sottoponendo gli studenti ad un questionario e permettendogli di valutare i criteri e la loro importanza attraverso l'utilizzo di scale di tipo ordinali linguistiche, fanno parte di una tipologia di insiemi definiti fuzzy.

Questo è un concetto che si differenzia dal classico concetto di insieme. È definito come inisieme, un qualsiasi raggruppamento di oggetti o dati, in cui è sempre possibile definire con assoluta certezza se un qualsiasi oggetto o dato appartiene al raggruppamento o meno. Come si evince dalla definizione, il concetto di insieme è strettamente collegato al concetto di appartenenza ad un insieme.

In una logica classica, definito il dominio entro il quale si vuole adoperare, per poter stabilire se un determinato elemento fa parte di un insieme o meno è necessario stabilire la funzione di appartenenza dell'insieme $F_{A,}$ definita come segue:

- $F_A(W)$= 1 se 0 <= w <= Th
- $F_a(W)$ = 0 altrimenti

Dove *Th* raprensenta i valore limite oltre il quale ogni elemento del dominio W non appartiene all'insieme A.

Ad esempio, si vogliono determinare gli elementi facenti parte dell insieme delle persone che indossano delle scarpe di taglia piccola. Dopo aver stabilito di voler lavorare all'interno del dominio W del numero di scarpe indossato dalle persone, si stabilisce la funzione di appertenenza F_A dell'insieme A e il limite soglia Th. Impostando la soglia *Th* uguale a 37 allora, fanno parte dell'insieme A tutte le persone che indossano un numero di scarpe pari o minore di 37 mentre tutte quelle che indossano un numero superiore sono escluse dall'insieme A.

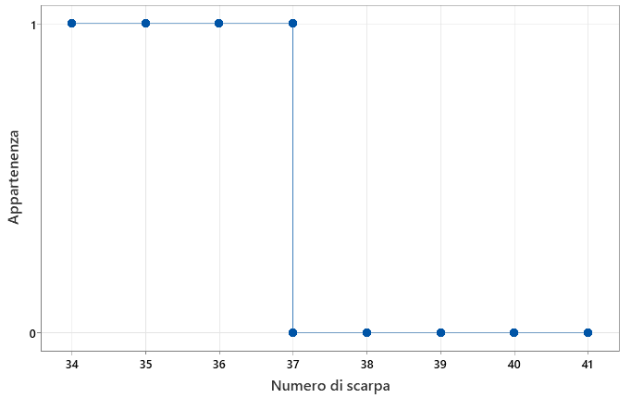


Figura 63 - Esempio di rappresentazione della funzione di appartenenza

In quest'ottica, la funzione di appertanenza F_A è una funzione booleane che puo' assumero solamente valore 1, se l'elemento appartiene all'insieme A, e 0, se non appartiene.

Il concetto di insieme così definito risulta essere molto stringente e limitato se si vuole considerare la conoscenza del pensiero umano. Volendo identificare l'insieme delle personi giovani, si fissa soglia sull'età di 40 anni. Allora, in base alla definizione data, una persona avente 39 anni appartiene all'insieme delle persone giovani ma una persona avente 40 anni e 1 minuto non appartiene all'insieme.

Questa definizione di insieme risulta inappropriata soprattutto quando si vogliono interpretare concetti vaghi come " le persono abbastanza alte" o "le temperature molto fredde".

Gli insiemi di tipo fuzzy, ideati nel 1965 da L. Zadeh, rilassano il concetto del valore delal soglia per poter considerare la logica del pensiero umano e tutte quelle situzioni limite. In questo modo è possibile descrivere e operare con definizioni vaghe tipiche dell'ambiemte reale.

Il concetto classico di insieme è stato così rimodellato introducendo il concetto del grado di appartenenza. Il valore che può assumere la funzione di appartenza non è più solo 0 o 1, ma può assumere anche tutti i valori intermedi, in modo da rappresentare tutte le situazioni vaghe che si presentano nel mondo reale, come "una persona è abbastanza bassa".

Allora, il valore assunto dalla funzione di appartenenza F_A rappresenta il grado di appartenenza dell'elemento all'insieme A e:

- Se $F_A(w)= 1$ allora l'elemento w appartiene all'insieme A;
- Se $0<F_A(w)<1$ allora l'elemento w appartiene parzialmente all'insieme A;
- Se Se $F_A(w)= 0$ allora l'elemento w non appartiene all'insieme A.

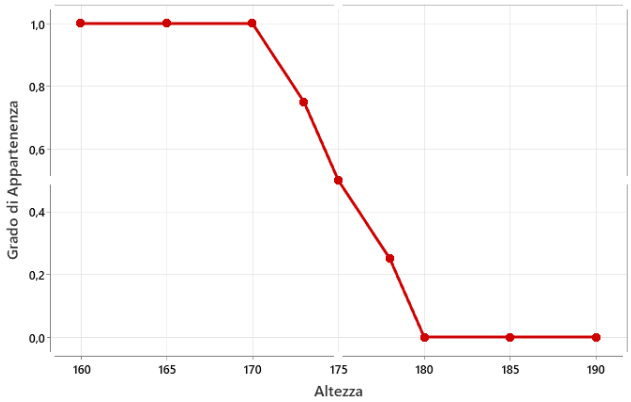


Figura 64 - Esempio di rappresentazione del grado di appartenenza

In conclusione, possiamo definire un insieme fuzzy come un insieme nel quale oggetti e dati sono raggruppati senza la definizione di un confine ben delimitato e preciso sull'appartenenza o meno all'insieme.

Prima di poter effettuare la regressione logistica ordinale è necessario definire le variabili con cui si vuole operare. È stato necessario lavorare i dati da interpolare in modo che ogni valutazione espressa in considerazione al criterio considerasse anche il grado di importanza che l'operatore ha attribuito al criterio stesso.

Seguendo tale logica, la complessità soggettiva percepita dal singolo operatore per una singola molecola in è definita come segue:

$$CS_{n,j} = min_i \{max[Neg(a_{n,i}); V(a_{n,i,j})]\} \quad (8)$$

dove:

n = 1,…,52 rappresenta gli operatori che hanno partecipato alla compilazione del questionario;

i = 1,…,16 rappresenta il criterio;

j = 1,…,12 rappresenta la molecola;

$V(a_{i,j})$ = 1,…,5 rappresenta la valutazione che l'operatore assegna in base al grado di accordo con il criterio in riferimento ad una determinata molecola;

Il valore di $Neg(a_i)$ fa riferimento al grado di importanza attribuito dall'operatore al criterio ed è indipendente dalla molecola che si fa a considerare. È definito come segue:

$$Neg(a_{n,i}) = a_{n,(l-i+1)} = 1, \dots, 5 \quad (9)$$

dove:

$a_{n,l}$ rappresenta il valore massimo di importanza imponibile ad un criterio. In questo caso si ha $a_l = 5$ poichè la scala delle importanze dei criteri adottata permette la scelta tra cinque livelli;

$a_{n,i}$ rappresenta il valore di importanza assegnato al criterio i dall'operatore.

Il valore finale della complessità della molecola percepita dall'operatore è stato ricavato effettuando il minimo tra i due valori.

Dopo aver ricavato i valori delle complessità soggettive degli operatori riferite ad una molecola, si è deciso di invertire i valori della complessità in modo da ottenere dei punteggi crescenti con la complessità oggettiva.

Come esempio di seguito, sono riportate le importanze assegnate ai 16 criteri, il valore assegnato in base al grado di accordo dell'operatore per le molecole ID1, ID6 e ID12 e il valore di complessità della molecola percepita dall'operatore:

Importanza assegnata ai criteri	
Criteri	Operatori 1
1	4
2	4
3	2
4	5
5	4
6	5
7	5
8	2
9	3
10	4
11	4
12	4
13	5
14	4
15	3
16	4

Tabella 11 - Importanza assegnata ai criteri

Valori dei gradi di Accordo			
Criteri	ID1	ID6	ID12
1	5	3	3
2	4	5	5
3	5	4	3
4	5	3	3
5	5	4	4
6	5	5	4
7	5	5	5
8	5	3	3
9	5	4	3
10	5	4	4
11	5	4	4
12	5	3	4
13	5	3	3
14	4	4	3
15	5	3	3
16	5	4	2

Tabella 12 - Valori del grado di accordo con i criteri delle Molecole ID1, ID6 e ID12

Complessità Soggettiva			
Operatore	CS ID1	CS ID6	CS ID12
1	4	3	2

Complessità Soggettiva Invertita			
Operatore	CS ID1	CS ID6	CS ID12
1	2	3	4

Tabella 13 - Valori finali della Complessità Soggettiva e Valori finali della Complessità Soggettiva Invertita

Nei sottostanti diagrammi cartesiani, in cui le ordinate rappresentano le molecole e le ascisse il range che può assumere la complessità soggettiva, sono riportati i valori della complessità percepita dall'operatore 1 e dall'operatore 26:

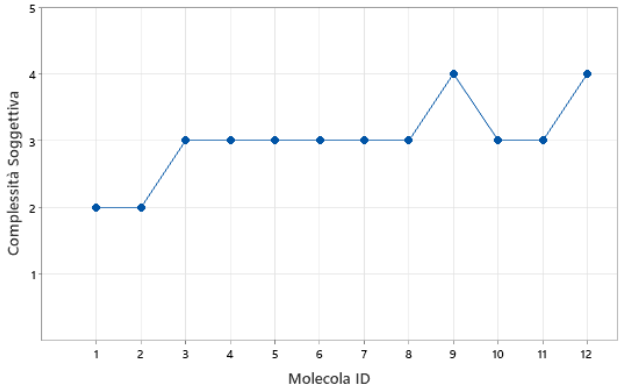


Figura 65 - Scatterplot della Complessità Soggettiva vs Complessità Oggettiva dell'operatore 1

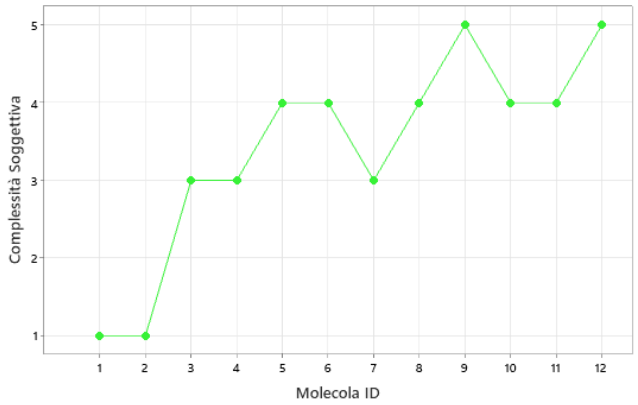


Figura 66 - Scatterplot della Complessità Soggettiva vs Complessità Oggettiva dell'operatore 26

Al fine di ottenere delle valutazioni medie degli operatori, pesate per il grado di importanza affidato ai criteri, si è scelto di utilizzare l'indicatore OWA. Questo indicatore è stato introdotto dal ricercatore americano R. Yager (1988) ed è definito come segue:

$$OWA = max_{k=1}^{n} [min\{Q(k), b_k\}] \quad (10)$$

dove:

n = 1,…,52 rappresenta gli operatori che hanno partecipato alla compilazione del questionario;

Q(k) = Sf_k = Int{1+[k*((t-1)/n)]} rappresenta il peso dell'indicatore;

la funzione Int(a) permette di considerare solamente l'intero del numero, senza considerare i numeri dopo la virgola;

b_k rappresenta la valutazione soggettiva del singolo partecipante per ogni molecola studiata.

$$Fk = int(1 + (k \cdot ((t - 1) / n)))$$

Operatore	Fk
1	1
2	1
3	2
4	2
5	3
6	3
7	3
8	4
9	4
10	5

t	5
k	n operatore
n	10

Tabella 14 - Esempio di calcolo di FK con 10 operatori utilizzando una scala con 5 livelli di Risposta

Di seguito è mostrato l'andamento della complessità soggettiva ottenuto riportando nel *Grafico* i dati mediati tramite l'operatore OWA delle *prove Randomiche*. Nel diagramma cartesiano l'asse delle ascisse rappresenta la complessità oggettiva delle 12 molecole esaminate e l'asse delle ordinate il valore mediato tramite l'operatore.

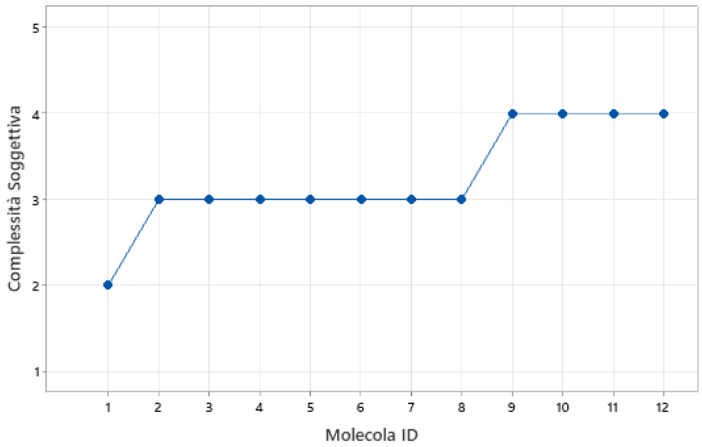


Figura 67 - Scatterplot Complessità Soggettiva vs Complessità Oggettiva delle 12 Molecole con i risultati della Complessità Soggettiva mediati tramite l'operatore OWA

Grazie al software per analisi Minitab®, è stato possibile effettuare una regressione di tipo logistica ordinale che evidenziasse la possibile relazione tra la *Complessità Oggettiva* (variabile indipendente) e la *Complessità Soggettiva* (variabile dipendente) delle dodici molecole.

Di seguito è riportato l'output ottenuto:

Ordinal Logistic Regression: Complessità Soggettiva versus Complessità Oggettiva

Link Function: Logit

Response Information

Variable	Value	Count
Complessità Soggettiva	1	27
	2	43
	3	176
	4	259
	5	119
	Total	624

Figura 68 - Output restituito dalla Regressione Logistica Ordinale Complessità Soggettiva vs Complessità Oggettiva delle 12 Molecole

Sono state processate un totale di 624 valutazioni della complessità percepita da 52 operatori affidate alle 12 molecole. Di queste:

- 27 volte è stata selezionata la risposta "Totalmente d'accordo";
- 43 volte è stata selezionata la risposta "D'accordo";
- 176 volte è stata selezionata la risposta "Relativamente d'accordo";
- 259 volte è stata selezionata la risposta "In disaccordo";
- 119 volte è stata selezionata la risposta "Totalmente in disaccordo".

Logistic Regression Table

						95% CI	
Predictor	Coef	SE Coef	Z	P	Odds Ratio	Lower	Upper
Const(1)	-1,24487	0,225475	-5,52	0,000			
Const(2)	0,0104822	0,178069	0,06	0,953			
Const(3)	2,18486	0,203028	10,76	0,000			
Const(4)	4,51967	0,253430	17,83	0,000			
Complessità Oggettiva	-0,0226111	0,0015999	-14,13	0,000	0,98	0,97	0,98

Figura 69 - Output restituito dalla Regressione Logistica Ordinale Complessità Soggettiva vs Complessità Oggettiva delle 12 Molecole

La *Logistic Regression Table* ci permette di valutare i coefficienti costanti per le categorie degli esiti, meno una, e il coefficiente relativo all'unica variabile predittiva del modello: *Complessità Oggettiva*. Essendo i valori dei coefficienti di *Const(3)* e *Const(4)* positivi, allora si può affermare che all'aumentare della complessità oggettiva è più probabile il verificarsi del primo evento e degli eventi più vicini ad esso. Il valore del coefficiente di *Const(1)* essendo negativo implica che all' aumentare

della complessità oggettiva è più probabile il verificarsi l'ultimo evento e degli eventi più vicini ad esso. Il coefficiente di *Const(2)* è positivo ma prossimo allo zero quindi l'impatto del predittore non è influente.

Nella riga del predittore *Complessità Oggettiva* è necessario valutare il valore restituito dall'odd ratio. Utilizzando come predittore una variabile continua, il valore restituito minore di 1 (pari a 0,98) indica che al crescere della complessità oggettiva delle molecole è più probabile aspettarsi come risposta l'ultimo evento e gli eventi ad esso più vicino. Quest'ultima affermazione viene confermata anche dal valore negativo restituito dal coefficiente.

I valori restituiti dal P-value dei predittori confermano i risultati sopra ricavati. Il P-value di *Const(1), Const(3), Const(4)* e di *Complessità Oggettiva* è uguale a 0,0005, minore del livello di significatività del 5%, e quini si può rifiutare l'ipotesi nulla e poter concludere che esiste una relazione statisticamente significativa tra il termine e la variabile risposta. Il valore *di Const(2)* essendo pari a 0,953, maggiore della soglia limite, permette di accettare l'ipotesi nulla e concludere che non esiste una relazione tra la risposta e il termine.

Log-Likelihood = -733,245

Test of All Slopes Equal to Zero

DF	G	P-Value
1	228,509	0,000

Figura 70 - Output restituito dalla Regressione Logistica Ordinale Complessità Soggettiva vs Complessità Oggettiva delle 12 Molecole

Il valore del P-value restituito dal *Test of All Slopes Equal to Zero* è pari a 0,005, minore della soglia del 5% e quindi si può affermare che esiste un'associazione statisticamente significativa tra la risposta ed almeno uno dei predittori.

Goodness-of-Fit Tests

Method	Chi-Square	DF	P
Pearson	10,5168	14	0,724
Deviance	12,9615	14	0,530

Figura 71 - Output restituito dalla Regressione Logistica Ordinale Complessità Soggettiva vs Complessità Oggettiva delle 12 Molecole

Measures of Association:

(Between the Response Variable and Predicted Probabilities)

Pairs	Number	Percent	Summary Measures	Value
Concordant	96562	70,3	Somers' D	0,49
Discordant	29007	21,1	Goodman-Kruskal Gamma	0,54
Ties	11721	8,5	Kendall's Tau-a	0,35
Total	137290	100,0		

Figura 72 - Output restituito dalla Regressione Logistica Ordinale Complessità Soggettiva vs Complessità Oggettiva delle 12 Molecole

I valori restituiti dagli indicatori Somers, Goodman-Kruskal e Kendall sono rispettivamente 0,49, 0,54 e 0,35.

In base ai risultati ottenuti dalla regressione di tipo logistica ordinale, è possibile affermare che esiste una relazione tra la complessità oggettiva e la complessità soggettiva. Più specificamente, si può concludere che la complessità soggettiva è positivamente correlata alla complessità oggettiva e che al crescere di quest'ultima, la probabilità che gli operatori percepiscano le strutture molecolari con un'alta complessità è maggiore.

Volendo sfruttare e analizzare i valori medi di complessità soggettiva percepita dagli operatori in relazione alle 12 molecole esaminate, ottenute mediante l'impiego dell'indicatore OWA, è possibile determinare la relazione esiste tra le due variabili oggetto di studio.

Per poter applicare e confrontare la regressione lineare e non lineare è necessario introdurre il concetto di distanza.

Così come accaduto in precedenza, per poter scegliere il miglior modello che interpola il set di dati osservato, è necessario confrontare il valore restituito dell'indicatore S dalla regressione lineare (*Figura 73*) e il valore restituito dalla regressione non lineare (*Figura 74*).

Model Summary

S	R-sq	R-sq(adj)
0,349720	71,22%	68,34%

Figura 73 - Principali statistiche riassuntive restituiti da una Regressione con andamento Lineare Complessità Soggettiva vs Complessità Oggettiva

Summary

Iterations	7
Final SSE	1,55868
DFE	10
MSE	0,155868
S	0,394801

Figura 74 - Principali statistiche riassuntive restituiti da una Regressione con andamento a Potenza Complessità Soggettiva vs Complessità Oggettiva

Dal confronto si evince che il modello con andamento lineare è quello che descrive meglio il set di dati poiché restituire un minore valore dell'indicatore S.

Il valore di R^2 pari a 71,22% indica che circa il 72% della variabilità è spiegata dal modello lineare.

L'analisi dei residui (*Figura 75*) e il test di Anderson-Darling eseguito sui residui (*Figura 76*), restituendo un valore P-value pari a 0,528 e quindi maggiore della soglia imposta del 5%, confermano la bontà del modello con andamento lineare.

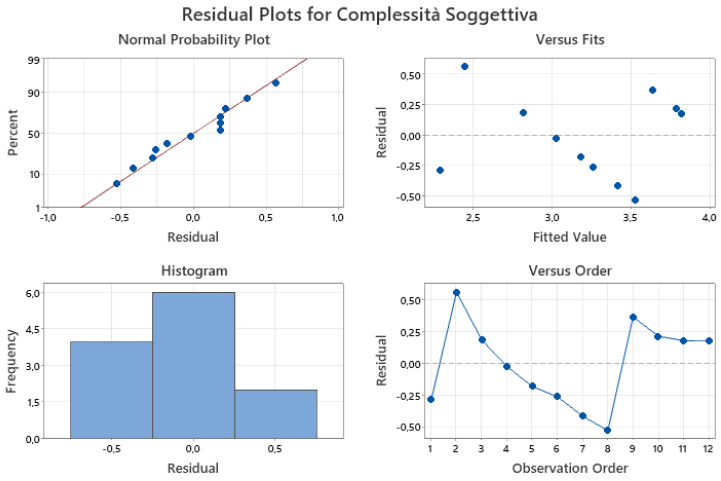


Figura 75 - Output dei Grafici dei Residui restituiti dal modello con andamento Lineare

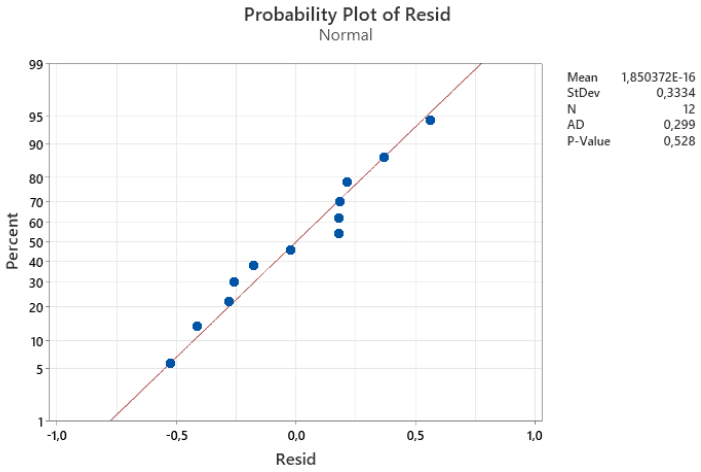


Figura 76 - Risultato test di Anderson-Darling condotto sui Residui ottenuti da un modello con andamento Lineare Complessità Soggettiva vs Complessità Oggettiva

Alla luce di quanto detto, la relazione tra la complessità oggettiva e la complessità soggettiva può essere espressa dalla seguente equazione:

$$\textbf{Complessità Soggettiva} = 2,23 + 0,009 \cdot \textbf{C}$$

Il valore positivo del coefficiente che determina la pendenza della retta (0,009) suggerisce che all'aumentare della complessità oggettiva delle molecole, gli operatori percepiscono un grado di complessità più alto.

4.5 Equazioni e Principali statistiche riassuntive di Regressione

Nella seguente sezione sono riportate le tabelle riassuntive delle analisi svolte. Le tabelle presentano in colonna la variabile dipendente, il modello, ovvero l'equazione che meglio spiega la relazione tra le variabile dipendete e quella indipendente, e il valore restituito dagli indicatori S e R^2.

4.5.1 Difetti Medi vs Complessità Oggettiva prove Randomiche

Variabile Dipendente	Modello	S
Difetti Medi di Processo Assemblaggio	Difetti = $9{,}05e^{-05} \cdot C^{2{,}14}$	1,54
Difetti Medi di Prodotto Assemblaggio	Difetti = $1{,}14e^{-09} \cdot C^{4{,}44}$	2,16
Difetti Medi Totali Assemblaggio	Difetti = $2{,}20e^{-07} \cdot C^{3{,}51}$	3,63
Difetti Medi Totali Disassemblaggio	Difetti = $2{,}16e^{-05} \cdot C^{2{,}15}$	0,13

Tabella 15 - Equazione Difetti Medi vs Complessità Oggettiva

4.5.2 Tempi Medi vs Complessità Oggettiva prove Randomiche

Variabile Dipendente	Modello	S	R^2
Tempi Medi Assemblaggio	Tempi = $0{,}004 \cdot C^{1{,}78}$	3,16	
Tempi Medi Disassemblaggio	Tempi = $0{,}04 + 0{,}05 \cdot C$	0,20	99,53 %

Tabella 16 - Equazione Tempi Medi vs Complessità Oggettiva

4.5.3 Difetti Medi di Processo di Assemblaggio vs Esecuzioni prove Ripetute

Variabile Dipendente	Modello	S	R^2
Difetti Medi ID1 di Processo Assemblaggio	Difetti ID1 = $0{,}28 \cdot \text{Esecuzione}^{-1{,}24}$	0,06	
Difetti Medi ID3 di Processo Assemblaggio	Difetti ID3 = $2{,}51 \cdot \text{Esecuzione}^{-0{,}66}$	0,26	
Difetti Medi ID5 di Processo Assemblaggio	Difetti ID5 = $4{,}85 \cdot \text{Esecuzione}^{-0{,}94}$	0,51	
Difetti Medi ID8 di Processo Assemblaggio	Difetti ID8 = $3{,}06 - 0{,}39 \cdot$ Esecuzione	0,49	77,50 %
Difetti Medi ID10 di Processo Assemblaggio	Difetti ID10 = $6{,}22 \cdot \text{Esecuzione}^{-0{,}63}$	0,30	
Difetti Medi ID12 di Processo Assemblaggio	Difetti ID12 = $7{,}62 \cdot \text{Esecuzione}^{-0{,}72}$	1,25	

Tabella 17 - Equazione Difetti Medi di Processo di Assemblaggio vs Esecuzione

4.5.4 Difetti Medi di Prodotto di Assemblaggio vs Esecuzioni prove Ripetute

Variabile Dipendente	Modello	S	R^2
Difetti Medi ID1 di Prodotto Assemblaggio	Difetti ID1 = 0,26 - 0,05 · Esecuzione	0,14	38,60%
Difetti Medi ID3 di Prodotto Assemblaggio	Difetti ID3 = 0,91 · Esecuzione $^{-0,48}$	0,25	
Difetti Medi ID5 di Prodotto Assemblaggio	Difetti ID5 = 2,98 - 0,39 · Esecuzione	1,02	45,30 %
Difetti Medi ID8 di Prodotto Assemblaggio	Difetti ID8 = 12,76 · Esecuzione $^{-1,82}$	0,54	
Difetti Medi ID10 di Prodotto Assemblaggio	Difetti ID10 = 7,13 · Esecuzione $^{-0,92}$	0,91	
Difetti Medi ID12 di Prodotto Assemblaggio	Difetti ID12 = 9,12 · Esecuzione $^{-0,58}$	1,70	

Tabella 18 - Equazione Difetti Medi di Prodotto di Assemblaggio vs Esecuzione

4.5.5 Difetti Medi Totali vs Esecuzioni prove Ripetute

Variabile Dipendente	Modello	S	R^2
Difetti Medi Totali ID1 Assemblaggio	Difetti ID1 = 0,71 · Esecuzione $^{-2,10}$	0,08	
Difetti Medi Totali ID3 Assemblaggio	Difetti ID3 = 3,40 · Esecuzione $^{-0,60}$	0,48	
Difetti Medi Totali ID5 Assemblaggio	Difetti ID5 = 8,72 · Esecuzione $^{-0,99}$	0,90	
Difetti Medi Totali ID8 Assemblaggio	Difetti ID8 = 15,55 · Esecuzione $^{-1,35}$	0,37	
Difetti Medi Totali ID10 Assemblaggio	Difetti ID10 = 13,32 · Esecuzione $^{-0,77}$	1,11	
Difetti Medi Totali ID12 Assemblaggio	Difetti ID12 = 13,88 - 1,347 · Esecuzione	3,81	41,20 %

Tabella 19 - Equazione Difetti Medi Totali di Assemblaggio vs Esecuzione

4.5.6 Tempi Medi di Assemblaggio vs Esecuzioni prove Ripetute

Variabile Dipendente	Modello	S
Tempi Medi ID1 Assemblaggio	Tempi ID1 = 0,71 · Esecuzione $^{-0,31}$	0,04
Tempi Medi ID3 Assemblaggio	Tempi ID3 = 13,46 · Esecuzione $^{-0,43}$	0,59
Tempi Medi ID5 Assemblaggio	Tempi ID5 = 22,29 · Esecuzione $^{-0,35}$	0,43
Tempi Medi ID8 Assemblaggio	Tempi ID8 = 24,22 · Esecuzione $^{-0,35}$	0,49
Tempi Medi ID10 Assemblaggio	Tempi ID10 = 40,66 · Esecuzione $^{-0,37}$	0,43
Tempi Medi ID12 Assemblaggio	Tempi ID12 = 51,74 · Esecuzione $^{-0,40}$	1,36

Tabella 20 - Equazione Tempi Medi di Assemblaggio vs Esecuzione

4.5.7 Difetti Medi di Disassemblaggio vs Esecuzioni prove Ripetute

Variabile Dipendente	Modello	S	R^2
Difetti Medi ID1 di Disassemblaggio	Difetti ID1 = 0,21 · Esecuzione $^{-1,62}$	0,04	
Difetti Medi ID3 di Disassemblaggio	Difetti ID3 = 0,43 · Esecuzione ^{0,03}	0,04	
Difetti Medi ID5 di Disassemblaggio	Difetti ID5 = 0,65 - 0,089 · Esecuzione	0,07	89,50 %
Difetti Medi ID8 di Processo Disassemblaggio	Difetti ID8 = 1,11 · Esecuzione $^{-0,54}$	0,04	
Difetti Medi ID10 di Disassemblaggio	Difetti ID10 = 1,39 · Esecuzione $^{-0,21}$	0,07	
Difetti Medi ID12 di Disassemblaggio	Difetti ID12 = 1,47 · Esecuzione $^{-0,18}$	0,10	

Tabella 21 - Equazione Difetti Medi di Disassemblaggio vs Esecuzione

4.5.8 Tempi Medi di Disassemblaggio vs Esecuzioni prove Ripetute

Variabile Dipendente	Modello	S
Tempi MEDI ID1 Disassemblaggio	Tempi ID1 = 0,29 · Esecuzione $^{-0,12}$	0,009
Tempi MEDI ID3 Disassemblaggio	Tempi ID3 = 3,43 · Esecuzione $^{-0,19}$	0,11
Tempi MEDI ID5 Disassemblaggio	Tempi ID5 = 5,20 · Esecuzione $^{-0,15}$	0,06
Tempi MEDI ID8 Disassemblaggio	Tempi ID8 = 7,06 · Esecuzione $^{-0,13}$	0,05
Tempi MEDI ID10 Disassemblaggio	Tempi ID10 = 9,21 · Esecuzione $^{-0,18}$	0,08
Tempi MEDI ID12 Disassemblaggio	Tempi ID12 = 9,90 · Esecuzione $^{-0,17}$	0,25

Tabella 22 - Equazione Tempi Medi di Disassemblaggio vs Esecuzione

CONCLUSIONI

Il presente lavoro di tesi si pone l'obiettivo di analizzare e spiegare gli effetti che la complessità di prodotto produce nei processi di assemblaggio e disassemblaggio manuali sui tempi di processo, sulla generazione dei difetti e sull'apprendimento degli operatori.

La sperimentazione è stata divisa in due fasi principali: la progettazione delle giornate lavorative e la raccolta dei dati osservati in esse e l'analisi per determinare la relazione esistente tra le varie variabili esaminate.

Nella prima fase, le giornate lavorative sono state suddivise in due tipologie: *prove Randomiche* e *prove Ripetute,* le quali si differenziano in base alla tipologia di assemblaggio e di disassamblaggio eseguita per le molecole. Durante le prove sono stati raccolti i dati relativi ai difetti e ai tempi registrati in fase di assemblaggio e di disassemblaggio. Sono stati utilizzati dodici modelli molecolari ball and stick di complessità oggettiva differente al fine di valutare come quest'ultima influenzasse i tempi di esecuzione, i difetti commessi e l'apprendimento degli operatori. Inoltre, sono state raccolte le valutazioni e il grado di importanza relativi ai 16 criteri di bassa complessità al fine di valutare come la complessità oggettiva condizionasse la complessità soggettiva.

Nella seconda fase sono stati analizzati i dati raccolti e sono state ricavate le equazioni che spiegano le relazioni esistenti tra i tempi e la complessità oggettiva, i difetti la complessità oggettiva, tra l'apprendimento degli operatori e la complessità oggettiva e tra la complessità soggettiva e la complessità oggettiva.

In seguito alle analisi condotte, si può concludere che le relazioni tra le variabili dipendenti e indipendenti, oggetto di studio, confermano i risultati attesi in base alla sperimentazione condotta da Alkan (2019). Nello specifico, analizzando i dati riscontrati nelle *prove Randomiche*, si può dedurre che le relazioni esistenti tra tempi di assemblaggio e disassemblaggio con la complessità oggettiva e tra i difetti di assemblaggio e disassemblaggio con la complessità oggettiva seguono un andamento a potenza. Fa eccezione la relazione con i tempi medi riscontrati in fase di disassemblaggio, la quale segue un andamento di tipo lineare. In generale, sia i difetti sia i tempi tendono ad aumentare al crescere della complessità oggettiva delle molecole.

Le analisi condotte sui tempi e sui difetti di assemblaggio e di disassemblaggio delle *prove Ripetute* evidenziano l'effetto di apprendimento degli operatori. Nello specifico, si può affermare che le relazioni tra i tempi di assemblaggio e di disassemblaggio delle molecole ripetute con il numero di volte in cui si è ripetuta l'operazione e tra i difetti di assemblaggio e di disassemblaggio delle molecole ripetute con il numero di volte in cui si è ripetuta l'operazione, seguono un andamento a potenza.

Fanno eccezione le relazioni con i difetti medi di processo riscontrati in fase di assemblaggio della molecola ID8, con i difetti medi di prodotto riscontrati in fase di assemblaggio della molecola ID1, con i difetti medi di prodotto riscontrati in fase di assemblaggio della molecola ID5, con i difetti medi totali riscontrati in fase di assemblaggio della molecola ID12 e con i difetti medi di disassemblaggio della molecola ID5, che risultano essere spiegate da un andamento di tipo lineare. In generale, si può osservare l'effetto di apprendimento dell'operatore, in quanto sia i tempi sia i difetti tendono a diminuire all'aumentare del numero di ripetizione delle esecuzioni. Tale effetto è giustificato dal fatto che si effettuano ripetutamente e consecutivamente le stesse operazioni manuali e per ognuna di queste l'operatore riesce a memorizzare gli errori commessi andando a migliorare, in questo modo, il tempo di esecuzione.

I risultati ottenuti dal confronto tra la complessità soggettiva e la complessità oggettiva tendono, invece, ad allontanarsi dall'applicazione di Alkan (2019). L'analisi effettuata mediante la regressione logistica ordinale della complessità percepita dai singoli operatori ci permette di stabilire l'esistenza di una relazione statisticamente positiva con la complessità oggettiva e di affermare che, al crescere di quest'ultima, è maggiore la probabilità che gli operatori percepiscano le strutture molecolari con un'alta complessità.

Volendo confrontare i risultati dell'analisi della complessità soggettiva ottenuti da Alkan (2019), è stato necessario lavorare con i dati mediati dall'indicatore OWA e suppore che possa essere applicato, in tale contesto, il concetto di distanza. Elaborando e confrontando i dati ottenuti mediante una regressione di tipo lineare e una regressione con andamento a potenza, si può concludere che esista una relazione lineare tra la complessità soggettiva e quella oggettiva. Infatti, al crescere della complessità oggettiva, gli operatori tendono a percepire le strutture molecolari con un grado più alto di complessità.

In conclusione, solo modellizzando correttamente la complessità di prodotto e distinguendola nella componente oggettiva e nella componente soggettiva è possibile prevedere e gestire i suoi effetti. In un contesto industriale, una corretta gestione della complessità permette una riduzione dei costi totali, dei tempi di produzione e dei tempi di consegna del prodotto finale. In aggiunta, è necessario monitorare e gestire la complessità percepita dagli operatori, poiché essa influisce, come dimostrato, sulla complessità di prodotto. Per poter migliorare le prestazioni offerte dagli operatori è necessario fornir loro un numero maggiore di informazioni comprensibili e dettagliate e dotarli dei migliori strumenti possibili. Inoltre, per poter trarre beneficio dell'effetto di apprendimento degli operatori, è opportuno allenare gli operatori poiché più volte ripetono le operazioni manuali, meno difetti sono riscontrati.

APPENDICE A

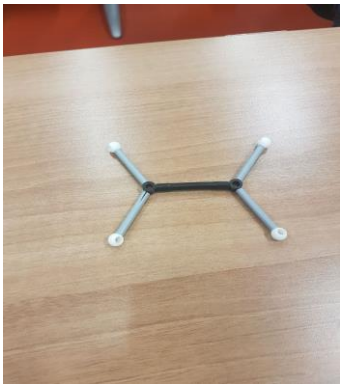


Figura 77 - Assemblaggio completo Molecola ID1

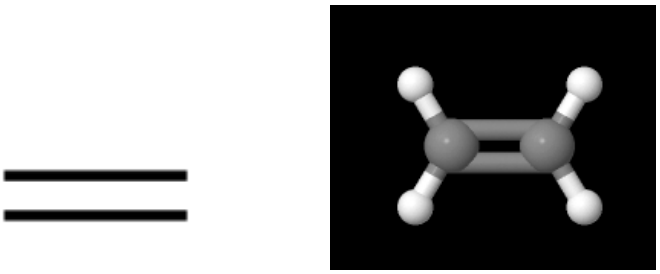


Figura 78 - Rappresentazione 2D e 3D della struttura molecolare ID1

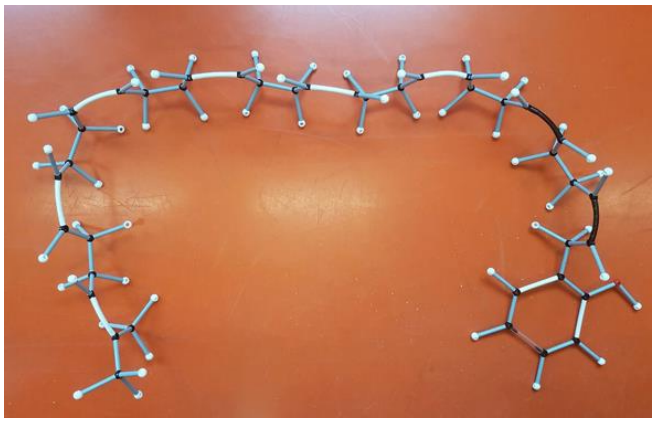
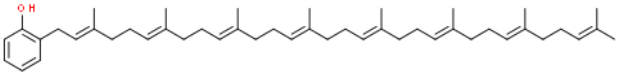


Figura 79 - Assemblaggio Completo Molecola ID8



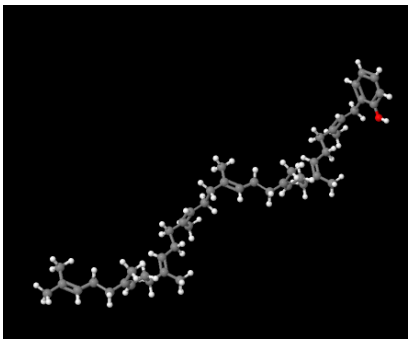


Figura 80 - Rappresentazione 2D e 3D della struttura molecolare ID8

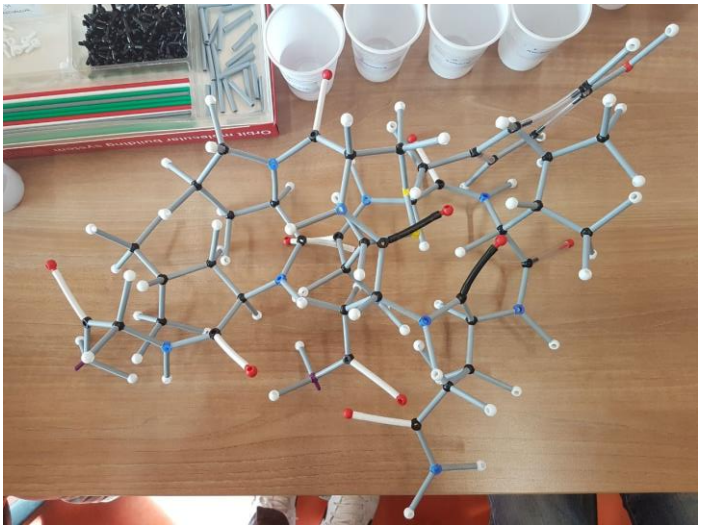
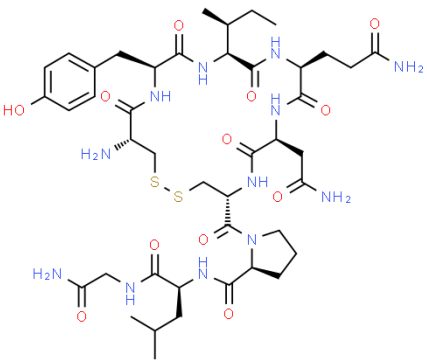


Figura 81 - Assemblaggio completo Molecola ID12



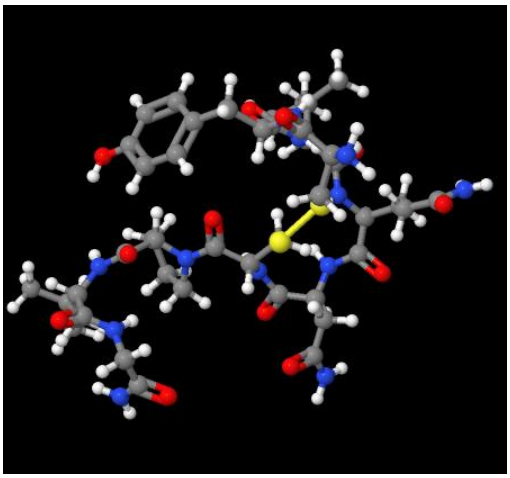


Figura 82 - Rappresentazione 2D e 3D della struttura molecolare ID12

Figura 83 - Kit Orbit Molecular Building System, basato sul modello Ball and Stick, dato in dotazione ad ogni coppia di studenti

Figura 84 - Postazione di Assemblaggio (a sinistra) e di Disassemblaggio (a destra) per l'esecuzione delle prove Randomiche e delle prove Ripetute

Figura 85 - Postazione di Assemblaggio per l'esecuzione delle prove Randomiche e delle prove Ripetute

Figura 86 - Postazione di Disassemblaggio per l'esecuzione delle prove Randomiche e delle prove Ripetute

APPENDICE B

B.1 Analisi Two-Way ANOVA delle prove Randomiche

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Complessità	11	3874	352,21	13,37	0,000
Operatore	51	3475	68,13	2,59	0,000
Error	561	14781	26,35		
Total	623	22130			

Figura 87 - Analisi Two-Way ANOVA Difetti di Processo di Assemblaggio vs Complessità Oggettiva vs Operatore nelle prove Randomiche

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Complessità	11	14459	1314,44	15,17	0,000
Operatore	51	9772	191,60	2,21	0,000
Error	561	48607	86,64		
Total	623	72837			

Figura 88 - Analisi Two-Way ANOVA Difetti di Prodotto di Assemblaggio vs Complessità Oggettiva vs Operatore nelle prove Randomiche

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Complessità	11	194,5	17,685	10,45	0,000
Operatore	51	312,5	6,127	3,62	0,000
Error	561	949,1	1,692		
Total	623	1456,1			

Figura 89 - Analisi Two-Way ANOVA Difetti di Disassemblaggio vs Complessità Oggettiva vs Operatore nelle prove Randomiche

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Complessità	11	4634,9	421,355	243,48	0,000
Operatore	51	527,2	10,338	5,97	0,000
Error	561	970,8	1,731		
Total	623	6133,0			

Figura 90 - Analisi Two-Way ANOVA Tempi di Disassemblaggio vs Complessità Oggettiva vs Operatore nelle prove Randomiche

B.2 Analisi Two-Way ANOVA delle prove Ripetute

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
ESECUZIONE	6	0,9592	0,15986	1,93	0,087
OPERATORE	13	2,7449	0,21115	2,55	0,006
Error	78	6,4694	0,08294		
Total	97	10,1735			

Figura 91 - Analisi Two-Way ANOVA Difetti di Processo di Assemblaggio vs Complessità Oggettiva vs Operatore della molecola ID1 nelle prove Ripetute

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
ESECUZIONE	6	37,69	6,282	5,13	0,000
OPERATORE	13	74,98	5,768	4,71	0,000
Error	78	95,45	1,224		
Total	97	208,12			

Figura 92 - Analisi Two-Way ANOVA Difetti di Processo di Assemblaggio vs Complessità Oggettiva vs Operatore della molecola ID3 nelle prove Ripetute

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
ESECUZIONE	6	178,1	29,677	8,34	0,000
OPERATORE	13	111,6	8,587	2,41	0,009
Error	78	277,7	3,560		
Total	97	567,3			

Figura 93 - Analisi Two-Way ANOVA Difetti di Processo di Assemblaggio vs Complessità Oggettiva vs Operatore della molecola ID5 nelle prove Ripetute

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
ESECUZIONE	6	76,06	12,677	4,21	0,001
OPERATORE	13	115,35	8,873	2,94	0,002
Error	78	235,08	3,014		
Total	97	426,49			

Figura 94 - Analisi Two-Way ANOVA Difetti di Processo di Assemblaggio vs Complessità Oggettiva vs Operatore della molecola ID8 nelle prove Ripetute

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
ESECUZIONE	6	210,5	35,082	4,97	0,000
OPERATORE	13	326,1	25,087	3,55	0,000
Error	78	550,7	7,060		
Total	97	1087,3			

Figura 95 - Analisi Two-Way ANOVA Difetti di Processo di Assemblaggio vs Complessità Oggettiva vs Operatore della molecola ID10 nelle prove Ripetute

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
ESECUZIONE	6	400,2	66,71	3,92	0,002
OPERATORE	13	481,3	37,02	2,18	0,018
Error	78	1327,5	17,02		
Total	97	2209,0			

Figura 96 - Analisi Two-Way ANOVA Difetti di Processo di Assemblaggio vs Complessità Oggettiva vs Operatore della molecola ID12 nelle prove Ripetute

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
ESECUZIONE	6	2,143	0,3571	1,76	0,119
OPERATORE	13	2,500	0,1923	0,95	0,511
Error	78	15,857	0,2033		
Total	97	20,500			

Figura 97 - Analisi Two-Way ANOVA Difetti di Prodotto di Assemblaggio vs Complessità Oggettiva vs Operatore della molecola ID1 nelle prove Ripetute

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
ESECUZIONE	6	7,551	1,259	1,04	0,405
OPERATORE	13	102,694	7,900	6,54	0,000
Error	78	94,163	1,207		
Total	97	204,408			

Figura 98 - Analisi Two-Way ANOVA Difetti di Prodotto di Assemblaggio vs Complessità Oggettiva vs Operatore della molecola ID3 nelle prove Ripetute

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
ESECUZIONE	6	133,5	22,255	2,67	0,021
OPERATORE	13	192,8	14,832	1,78	0,061
Error	78	649,3	8,325		
Total	97	975,7			

Figura 99 - Analisi Two-Way ANOVA Difetti di Prodotto di Assemblaggio vs Complessità Oggettiva vs Operatore della molecola ID5 nelle prove Ripetute

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
ESECUZIONE	6	1611	268,47	3,71	0,003
OPERATORE	13	1130	86,93	1,20	0,294
Error	78	5638	72,28		
Total	97	8379			

Figura 100 - Analisi Two-Way ANOVA Difetti di Prodotto di Assemblaggio vs Complessità Oggettiva vs Operatore della molecola ID8 nelle prove Ripetute

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
ESECUZIONE	6	428,1	71,35	2,07	0,066
OPERATORE	13	882,1	67,86	1,97	0,035
Error	78	2688,4	34,47		
Total	97	3998,7			

Figura 101 - Analisi Two-Way ANOVA Difetti di Prodotto di Assemblaggio vs Complessità Oggettiva vs Operatore della molecola ID10 nelle prove Ripetute

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
ESECUZIONE	6	558,6	93,11	1,52	0,181
OPERATORE	13	1628,4	125,26	2,05	0,027
Error	78	4762,2	61,05		
Total	97	6949,3			

Figura 102 - Analisi Two-Way ANOVA Difetti di Prodotto di Assemblaggio vs Complessità Oggettiva vs Operatore della molecola ID12 nelle prove Ripetute

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
ESECUZIONE	6	5,388	0,8980	3,60	0,003
OPERATORE	13	4,531	0,3485	1,40	0,180
Error	78	19,469	0,2496		
Total	97	29,388			

Figura 103 - Analisi Two-Way ANOVA Difetti Totali di Assemblaggio vs Complessità Oggettiva vs Operatore della molecola ID1 nelle prove Ripetute

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
ESECUZIONE	6	72,78	12,129	4,62	0,000
OPERATORE	13	289,35	22,257	8,47	0,000
Error	78	204,94	2,627		
Total	97	567,06			

Figura 104 - Analisi Two-Way ANOVA Difetti Totali di Assemblaggio vs Complessità Oggettiva vs Operatore della molecola ID3 nelle prove Ripetute

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
ESECUZIONE	6	584,3	97,38	8,44	0,000
OPERATORE	13	287,8	22,14	1,92	0,040
Error	78	899,4	11,53		
Total	97	1771,6			

Figura 105 - Analisi Two-Way ANOVA Difetti Totali di Assemblaggio vs Complessità Oggettiva vs Operatore della molecola ID5 nelle prove Ripetute

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
ESECUZIONE	6	1207	201,23	4,17	0,001
OPERATORE	13	1440	110,78	2,29	0,013
Error	78	3767	48,30		
Total	97	6415			

Figura 106 - Analisi Two-Way ANOVA Difetti Totali di Assemblaggio vs Complessità Oggettiva vs Operatore della molecola ID10 nelle prove Ripetute

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
ESECUZIONE	6	1727	287,8	2,83	0,015
OPERATORE	13	2424	186,5	1,83	0,052
Error	78	7940	101,8		
Total	97	12090			

Figura 107 - Analisi Two-Way ANOVA Difetti Totali di Assemblaggio vs Complessità Oggettiva vs Operatore della molecola ID12 nelle prove Ripetute

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
ESECUZIONE	6	1,121	0,18680	13,56	0,000
OPERATORE	13	1,284	0,09874	7,17	0,000
Error	78	1,075	0,01378		
Total	97	3,479			

Figura 108 - Analisi Two-Way ANOVA Tempi di Assemblaggio vs Complessità Oggettiva vs Operatore della molecola ID1 nelle prove Ripetute

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
ESECUZIONE	6	598,5	99,748	15,48	0,000
OPERATORE	13	482,4	37,105	5,76	0,000
Error	78	502,6	6,443		
Total	97	1583,4			

Figura 109 - Analisi Two-Way ANOVA Tempi di Assemblaggio vs Complessità Oggettiva vs Operatore della molecola ID3 nelle prove Ripetute

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
ESECUZIONE	6	1251,3	208,550	37,87	0,000
OPERATORE	13	2155,5	165,810	30,11	0,000
Error	78	429,5	5,506		
Total	97	3836,3			

Figura 110 - Analisi Two-Way ANOVA Tempi di Assemblaggio vs Complessità Oggettiva vs Operatore della molecola ID5 nelle prove Ripetute

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
ESECUZIONE	6	4518	753,01	57,81	0,000
OPERATORE	13	1797	138,26	10,61	0,000
Error	78	1016	13,03		
Total	97	7332			

Figura 111 - Analisi Two-Way ANOVA Tempi di Assemblaggio vs Complessità Oggettiva vs Operatore della molecola ID10 nelle prove Ripetute

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
ESECUZIONE	6	8281	1380,15	29,09	0,000
OPERATORE	13	4654	358,00	7,54	0,000
Error	78	3701	47,45		
Total	97	16636			

Figura 112 - Analisi Two-Way ANOVA Tempi di Assemblaggio vs Complessità Oggettiva vs Operatore della molecola ID12 nelle prove Ripetute

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
ESECUZIONE	6	0,4898	0,08163	1,46	0,204
OPERATORE	13	0,7755	0,05965	1,07	0,401
Error	78	4,3673	0,05599		
Total	97	5,6327			

Figura 113 - Analisi Two-Way ANOVA Difetti di Disassemblaggio vs Complessità Oggettiva vs Operatore della molecola ID1 nelle prove Ripetute

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
ESECUZIONE	6	0,1020	0,01701	0,05	0,999
OPERATORE	13	13,3878	1,02983	3,00	0,001
Error	78	26,7551	0,34301		
Total	97	40,2449			

Figura 114 - Analisi Two-Way ANOVA Difetti di Disassemblaggio vs Complessità Oggettiva vs Operatore della molecola ID3 nelle prove Ripetute

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
ESECUZIONE	6	3,490	0,5816	1,94	0,084
OPERATORE	13	3,561	0,2739	0,91	0,542
Error	78	23,367	0,2996		
Total	97	30,418			

Figura 115 - Analisi Two-Way ANOVA Difetti di Disassemblaggio vs Complessità Oggettiva vs Operatore della molecola ID5 nelle prove Ripetute

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
ESECUZIONE	6	5,408	0,9014	0,99	0,439
OPERATORE	13	12,694	0,9765	1,07	0,397
Error	78	71,163	0,9123		
Total	97	89,265			

Figura 116 - Analisi Two-Way ANOVA Difetti di Disassemblaggio vs Complessità Oggettiva vs Operatore della molecola ID8 nelle prove Ripetute

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
ESECUZIONE	6	2,633	0,4388	0,44	0,851
OPERATORE	13	86,776	6,6750	6,68	0,000
Error	78	77,939	0,9992		
Total	97	167,347			

Figura 117 - Analisi Two-Way ANOVA Difetti di Disassemblaggio vs Complessità Oggettiva vs Operatore della molecola ID10 nelle prove Ripetute

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
ESECUZIONE	6	2,531	0,4218	0,26	0,952
OPERATORE	13	35,888	2,7606	1,72	0,072
Error	78	124,898	1,6013		
Total	97	163,316			

Figura 118 - Analisi Two-Way ANOVA Difetti di Disassemblaggio vs Complessità Oggettiva vs Operatore della molecola ID12 nelle prove Ripetute

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
ESECUZIONE	6	0,04755	0,007926	4,75	0,000
OPERATORE	13	0,49226	0,037866	22,67	0,000
Error	78	0,13026	0,001670		
Total	97	0,67007			

Figura 119 - Analisi Two-Way ANOVA Tempi di Disassemblaggio vs Complessità Oggettiva vs Operatore della molecola ID1 nelle prove Ripetute

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
ESECUZIONE	6	12,729	2,1215	10,11	0,000
OPERATORE	13	7,935	0,6104	2,91	0,002
Error	78	16,366	0,2098		
Total	97	37,031			

Figura 120 - Analisi Two-Way ANOVA Tempi di Disassemblaggio vs Complessità Oggettiva vs Operatore della molecola ID3 nelle prove Ripetute

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
ESECUZIONE	6	19,185	3,1974	26,41	0,000
OPERATORE	13	72,020	5,5400	45,76	0,000
Error	77	9,322	0,1211		
Total	96	100,526			

Figura 121 - Analisi Two-Way ANOVA Tempi di Disassemblaggio vs Complessità Oggettiva vs Operatore della molecola ID5 nelle prove Ripetute

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
ESECUZIONE	6	27,09	4,5152	11,29	0,000
OPERATORE	13	59,78	4,5984	11,50	0,000
Error	78	31,20	0,4000		
Total	97	118,07			

Figura 122 - Analisi Two-Way ANOVA Tempi di Disassemblaggio vs Complessità Oggettiva vs Operatore della molecola ID8 nelle prove Ripetute

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
ESECUZIONE	6	74,05	12,3415	15,77	0,000
OPERATORE	13	107,12	8,2403	10,53	0,000
Error	78	61,03	0,7824		
Total	97	242,20			

Figura 123 - Analisi Two-Way ANOVA Tempi di Disassemblaggio vs Complessità Oggettiva vs Operatore della molecola ID10 nelle prove Ripetute

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
ESECUZIONE	6	82,30	13,7165	16,71	0,000
OPERATORE	13	212,76	16,3661	19,94	0,000
Error	78	64,01	0,8206		
Total	97	359,07			

Figura 124 - Analisi Two-Way ANOVA Tempi di Disassemblaggio vs Complessità Oggettiva vs Operatore della molecola ID12 nelle prove Ripetute

APPENDICE C

C.1 Regressione delle prove Randomiche

C.1.1 Difetti Medi di Processo di Assemblaggio

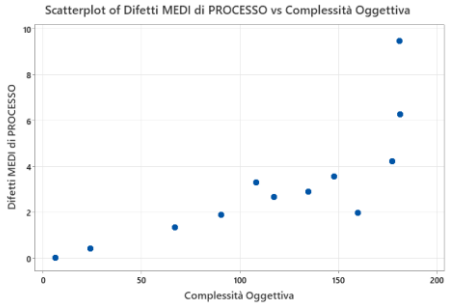


Figura 125 - Grafico di Dispersione o Scatterplot Difetti Medi di Processo di Assemblaggio vs Complessità Oggettiva nelle prove Randomiche

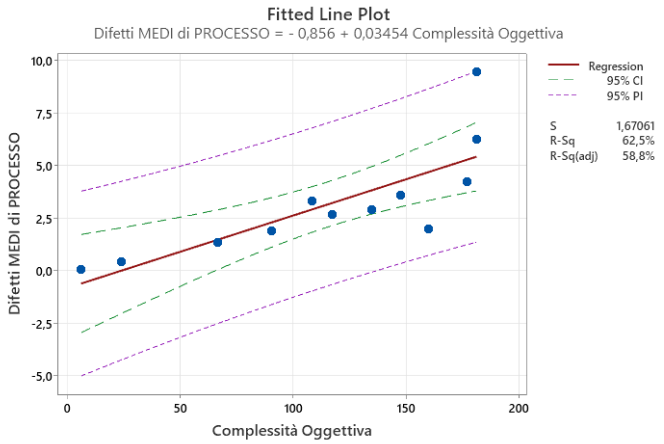


Figura 126 - Curva di Regressione e principali statistiche riassuntive di una Regressione Lineare Difetti Medi di Processo di Assemblaggio vs Complessità Oggettiva nelle prove Randomiche

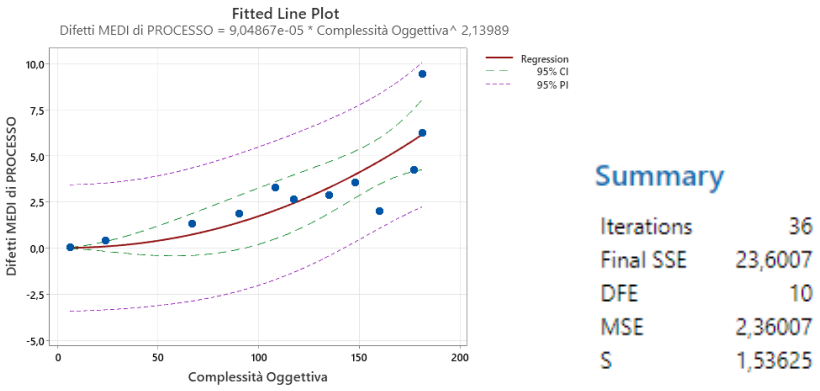


Figura 127 - Curva di Regressione principali statistiche riassuntive di una Regressione con andamento a Potenza Difetti Medi di Processo di Assemblaggio vs Complessità Oggettiva nelle prove Randomiche

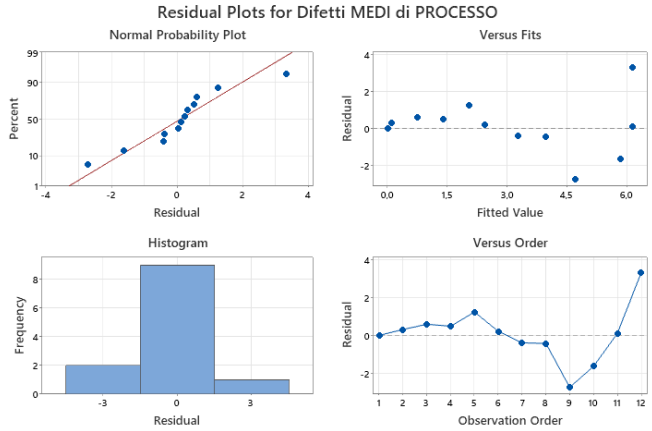


Figura 128 - Output dei Grafici dei Residui restituiti dal modello con andamento a Potenza

C.1.2 Difetti Medi di Prodotto di Assemblaggio

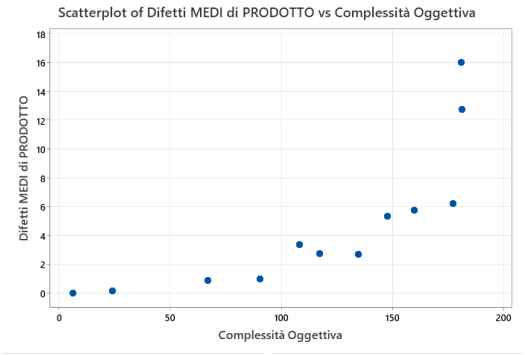


Figura 129 - Grafico di Dispersione o Scatterplot Difetti Medi di Processo di Assemblaggio vs Complessità Oggettiva nelle prove Randomiche

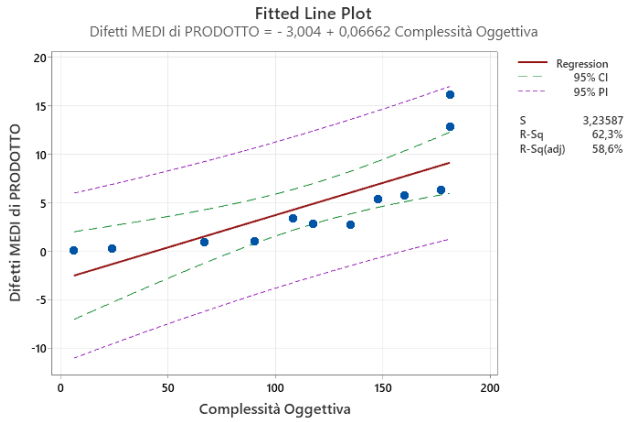


Figura 130 - Curva di Regressione e principali statistiche riassuntive di una Regressione Lineare Difetti Medi di Prodotto di Assemblaggio vs Complessità Oggettiva nelle prove Randomiche

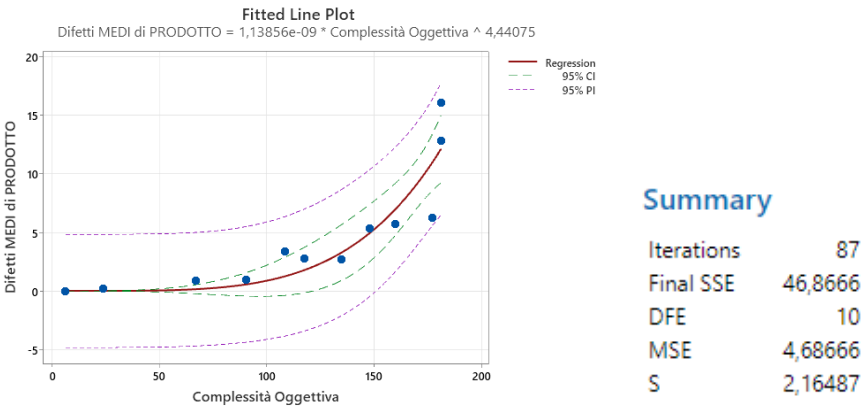


Figura 131 - Curva di Regressione e principali statistiche riassuntive di una Regressione con andamento a Potenza Difetti Medi di Prodotto di Assemblaggio vs Complessità Oggettiva nelle prove Randomiche

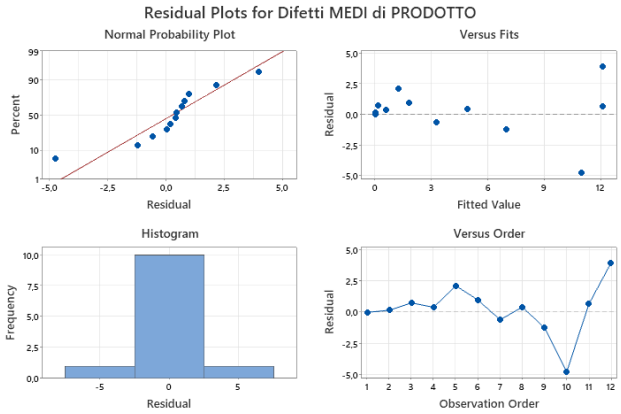


Figura 132 - Output dei Grafici dei Residui restituiti dal modello con andamento a Potenza

C.1.3 Difetti Medi di Disassemblaggio

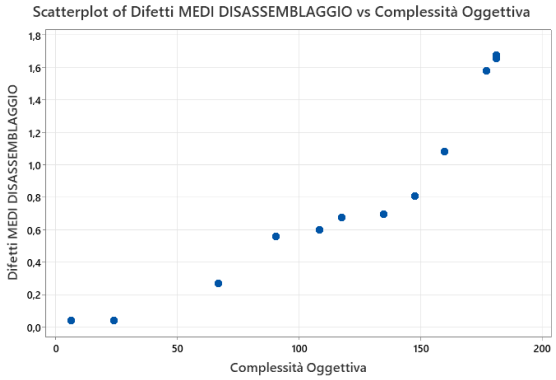


Figura 133 - Grafico di Dispersione o Scatterplot Difetti Medi di Disassemblaggio vs Complessità Oggettiva nelle prove Randomiche

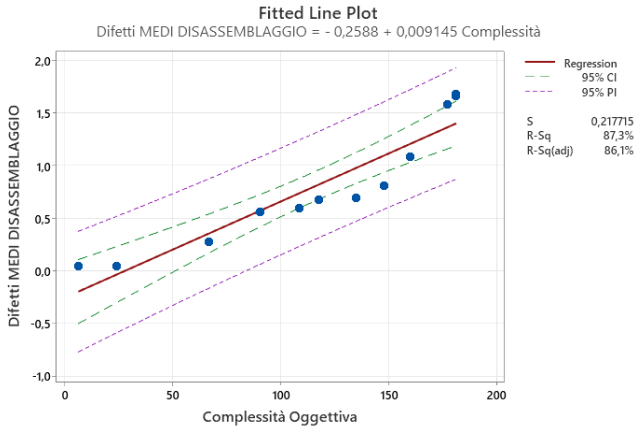


Figura 134 - Curva di Regressione e principali statistiche riassuntive di una Regressione Lineare Difetti Medi di Disassemblaggio vs Complessità Oggettiva nelle prove Randomiche

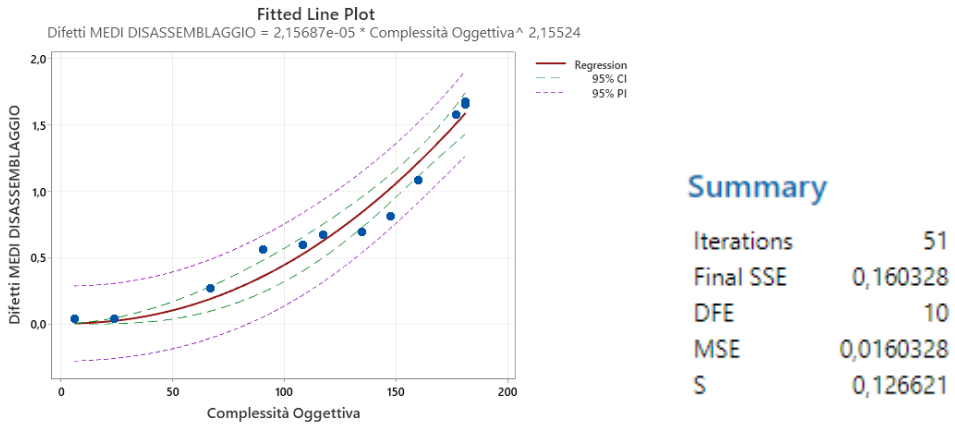


Figura 135 - Curva di Regressione e principali statistiche riassuntive di una Regressione con andamento a Potenza Difetti Medi Disassemblaggio vs Complessità Oggettiva nelle prove Randomiche

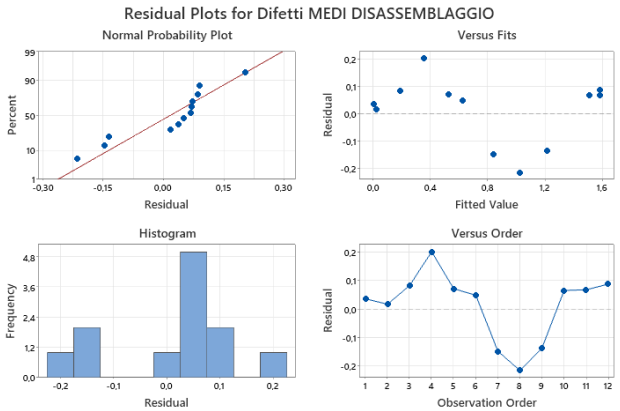


Figura 136 - Output dei Grafici dei Residui restituiti dal modello con andamento a Potenza

C.1.4 Tempi Medi di Disassemblaggio

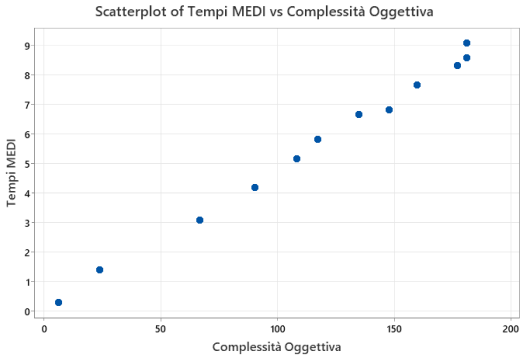


Figura 137 - Grafico di Dispersione o Scatterplot Tempi Medi di Disassemblaggio vs Complessità Oggettiva nelle prove Randomiche

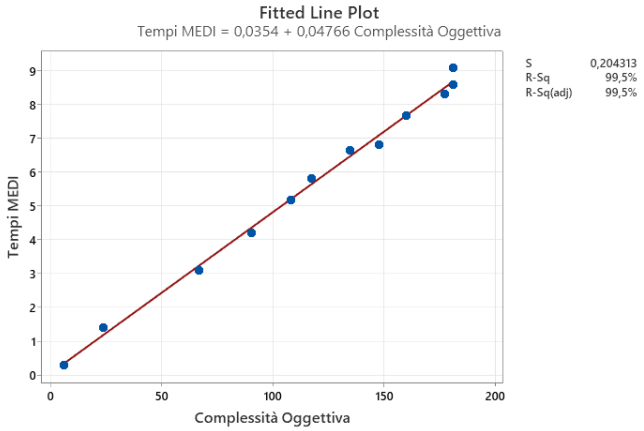


Figura 138 - Curva di Regressione e principali statistiche riassuntive di una Regressione Lineare Tempi Medi di Disassemblaggio vs Complessità Oggettiva nelle prove Randomiche

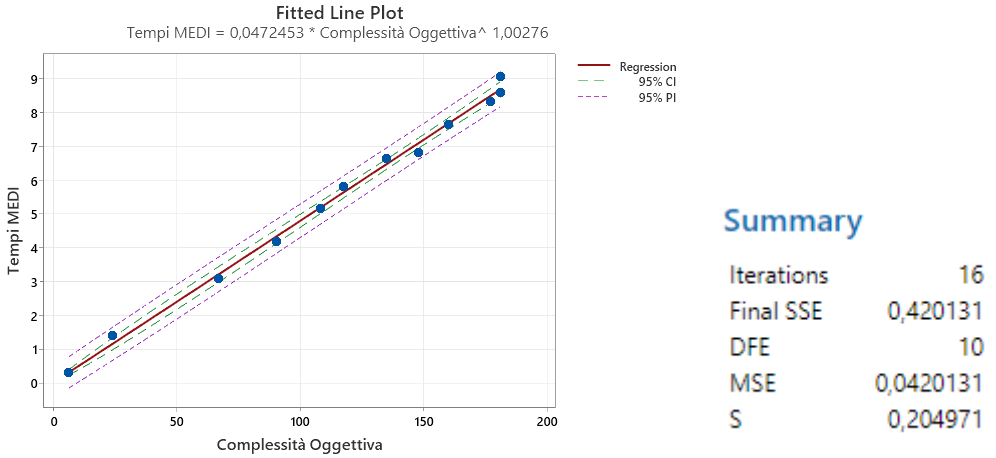


Figura 139 - Curva di Regressione e principali statistiche riassuntive di una Regressione con andamento a Potenza Tempi Medi Disassemblaggio vs Complessità Oggettiva nelle prove Randomiche

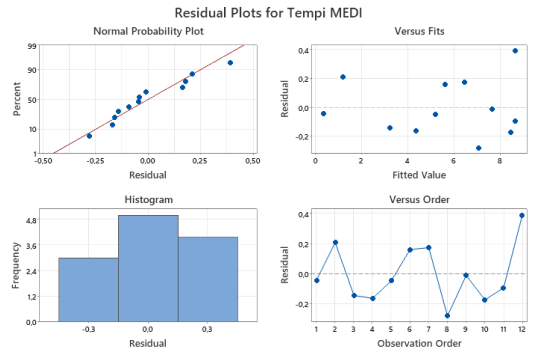


Figura 140 - Output dei Grafici dei Residui restituiti dal modello con andamento Lineare

C.2 Regressione delle prove Ripetute

C.2.1 Difetti Medi di Processo di Assemblaggio

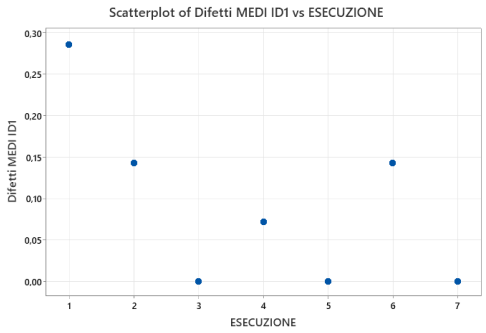


Figura 141 - Grafico di Dispersione o Scatterplot Difetti Medi di Processo di Assemblaggio vs Esecuzione della molecola ID1 nelle prove Ripetute

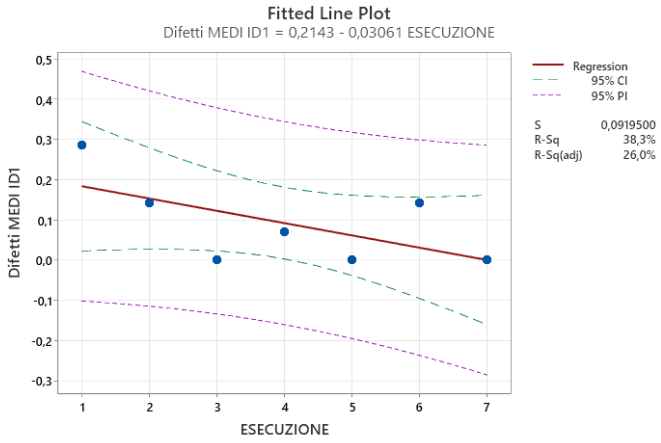


Figura 142 - Curva di Regressione e principali statistiche riassuntive di una Regressione Lineare Difetti Medi di Processo di Assemblaggio vs Esecuzione della molecola ID1 nelle prove Ripetute

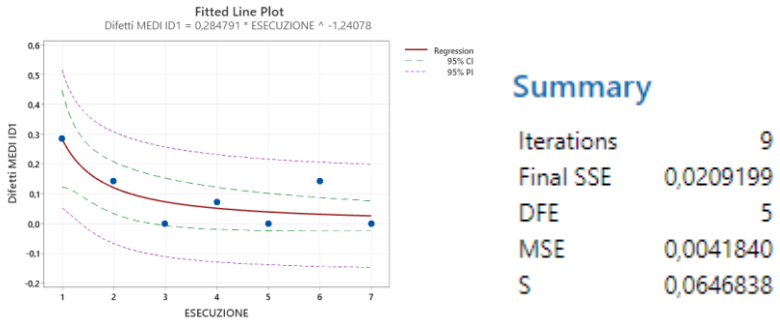


Figura 143 - Curva di Regressione e principali statistiche riassuntive di una Regressione con andamento a Potenza Difetti Medi di Processo Assemblaggio vs Esecuzione della molecola ID1 nelle prove Ripetute

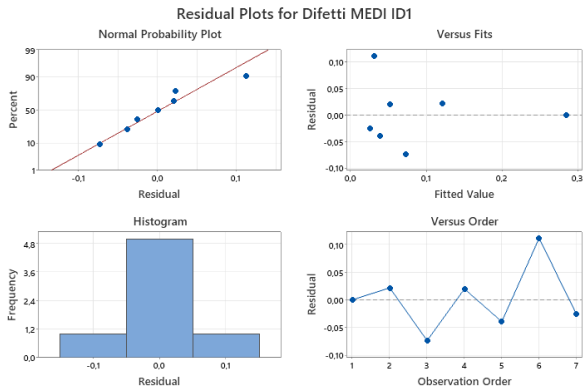


Figura 144 - Output dei Grafici dei Residui restituiti dal modello con andamento a Potenza

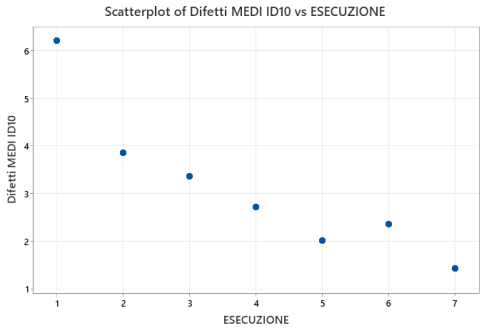


Figura 145 - Grafico di Dispersione o Scatterplot Difetti Medi di Processo di Assemblaggio vs Esecuzione della molecola ID10 nelle prove Ripetute

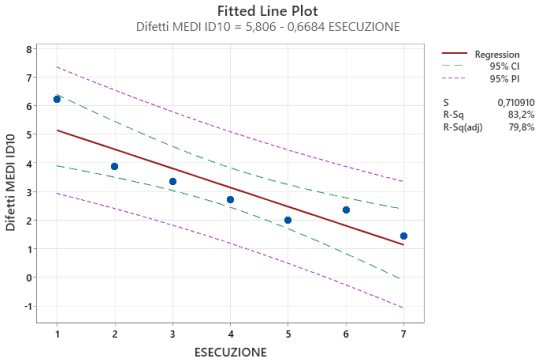


Figura 146 - Curva di Regressione e principali statistiche riassuntive di una Regressione Lineare Difetti Medi di Processo di Assemblaggio vs Esecuzione della molecola ID10 nelle prove Ripetute

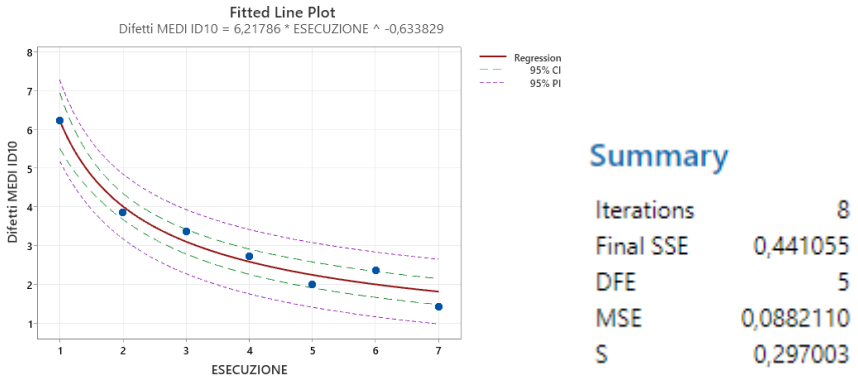


Figura 147 - Curva di Regressione e principali statistiche riassuntive di una Regressione con andamento a Potenza Difetti Medi di Processo Assemblaggio vs Esecuzione della molecola ID10 nelle prove Ripetute

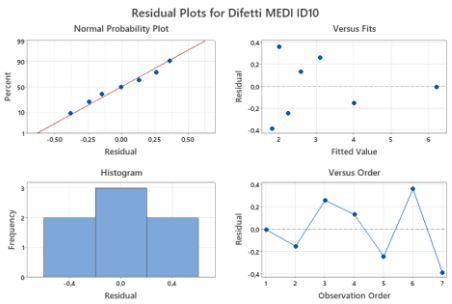


Figura 148 - Output dei Grafici dei Residui restituiti dal modello con andamento a Potenza

C.2.2 Difetti Medi di Prodotto di Assemblaggio

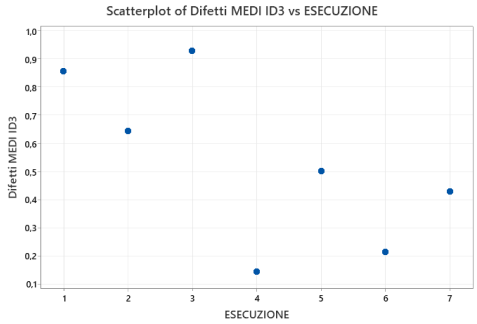


Figura 149 - Grafico di Dispersione o Scatterplot Difetti Medi di Prodotto di Assemblaggio vs Esecuzione della molecola ID3 nelle prove Ripetute

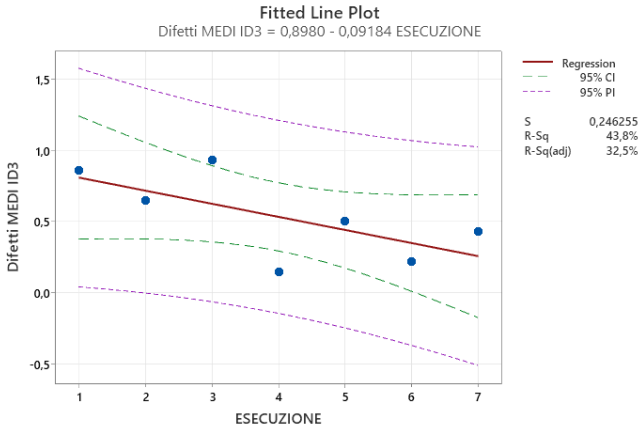


Figura 150 - Curva di Regressione e principali statistiche riassuntive di una Regressione Lineare Difetti Medi di Prodotto di Assemblaggio vs Esecuzione della molecola ID3 nelle prove Ripetute

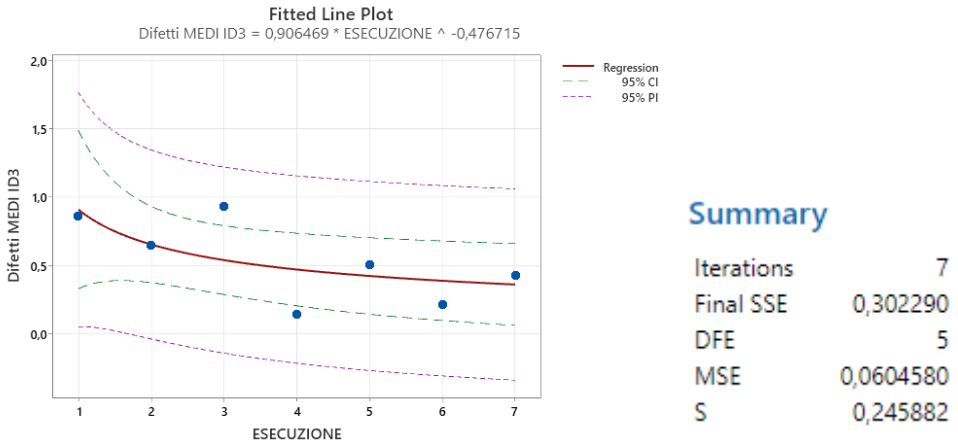


Figura 151 - Curva di Regressione e principali statistiche riassuntive di una Regressione con andamento a Potenza Difetti Medi di Processo Assemblaggio vs Esecuzione della molecola ID3 nelle prove Ripetute

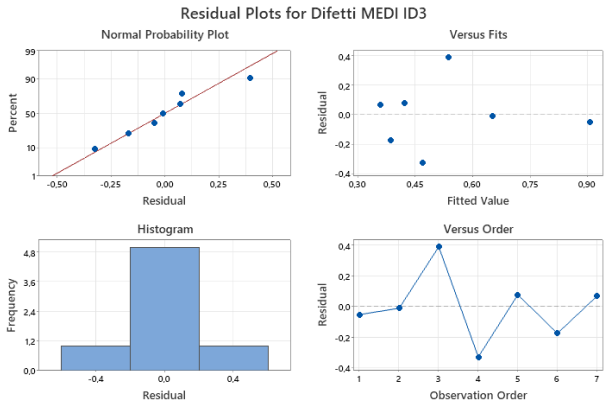


Figura 152 - Output dei Grafici dei Residui restituiti dal modello con andamento a Potenza

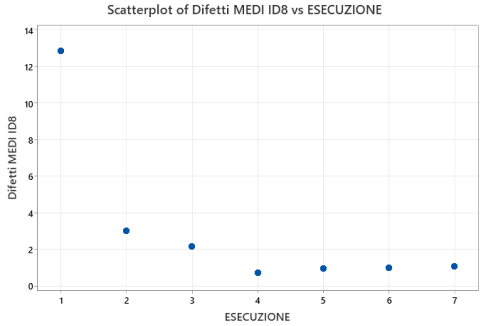


Figura 153 - Grafico di Dispersione o Scatterplot Difetti Medi di Prodotto di Assemblaggio vs Esecuzione della molecola ID8 nelle prove Ripetute

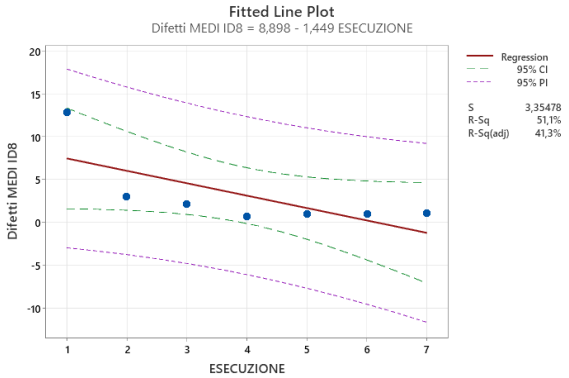


Figura 154 - Curva di Regressione e principali statistiche riassuntive di una Regressione Lineare Difetti Medi di Prodotto di Assemblaggio vs Esecuzione della molecola ID8 nelle prove Ripetute

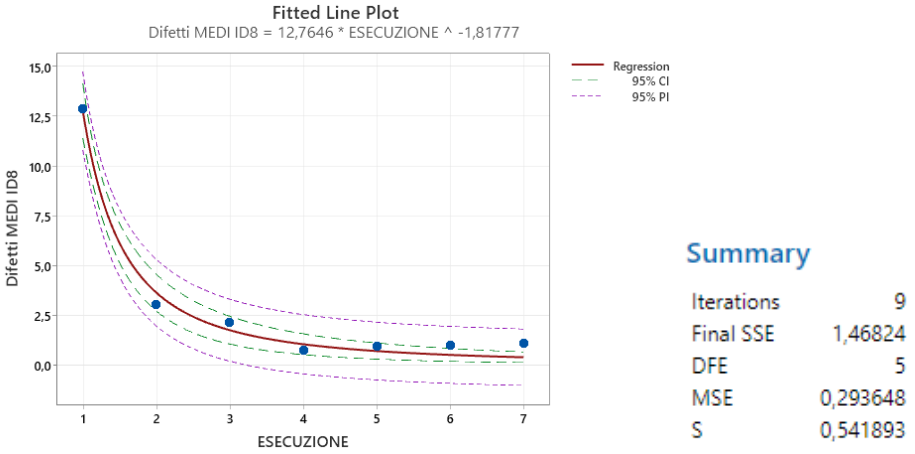


Figura 155 - Curva di Regressione e principali statistiche riassuntive di una Regressione con andamento a Potenza Difetti Medi di Processo Assemblaggio vs Esecuzione della molecola ID8 nelle prove Ripetute

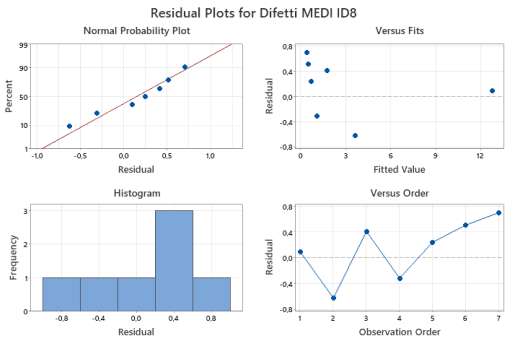


Figura 156 - Output dei Grafici dei Residui restituiti dal modello con andamento a Potenza

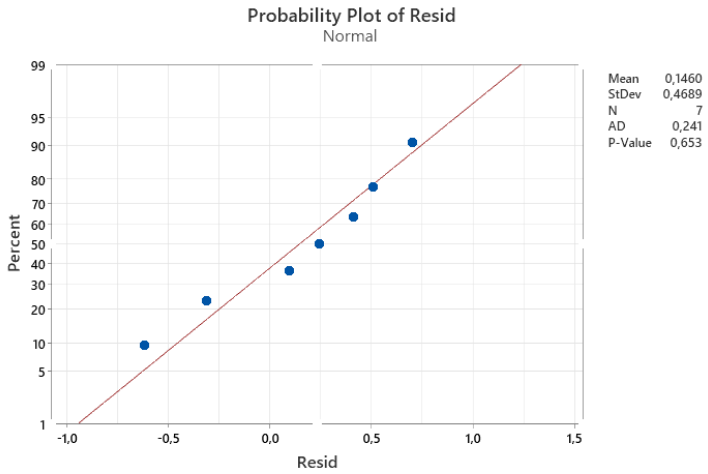


Figura 157 - Risultato test di Anderson-Darling condotto sui Residui ottenuti da un modello con andamento a Potenza Difetti Medi di Prodotto Assemblaggio vs Esecuzione della Molecola ID8 nelle prove Ripetute. Essendo il P-value > 0,05

C.2.3 Difetti Medi Totali di Assemblaggio

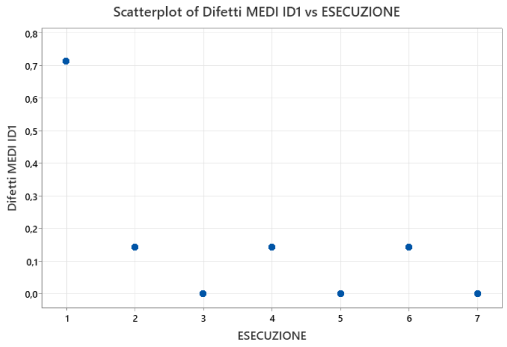


Figura 158 - Grafico di Dispersione o Scatterplot Difetti Medi Totali di Assemblaggio vs Esecuzione della molecola ID1 nelle prove Ripetute

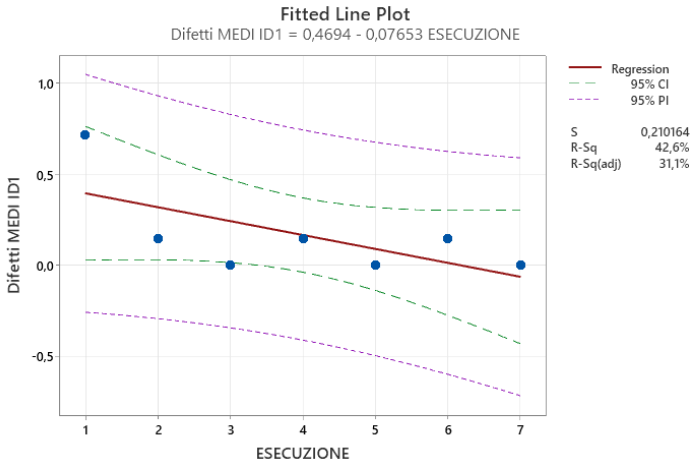


Figura 159 - Curva di Regressione e principali statistiche riassuntive di una Regressione Lineare Difetti Medi Totali di Assemblaggio vs Esecuzione della molecola ID1 nelle prove Ripetute

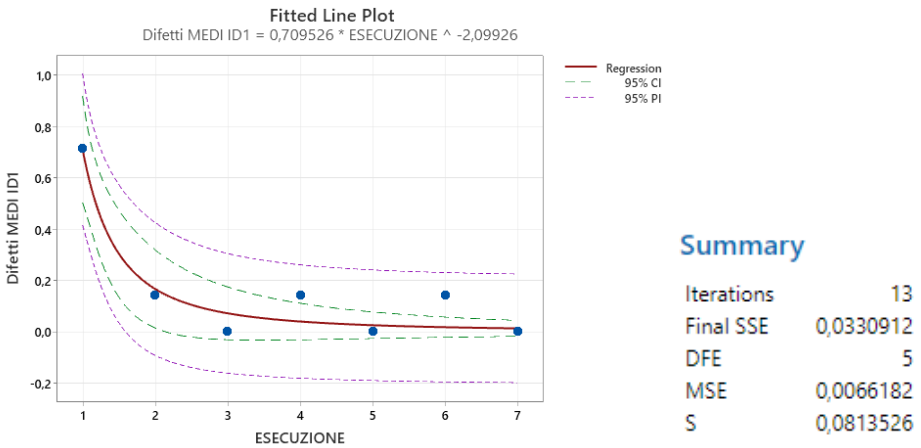


Figura 160 - Curva di Regressione e principali statistiche riassuntive di una Regressione con andamento a Potenza Difetti Medi Totali di Assemblaggio vs Esecuzione della molecola ID1 nelle prove Ripetute

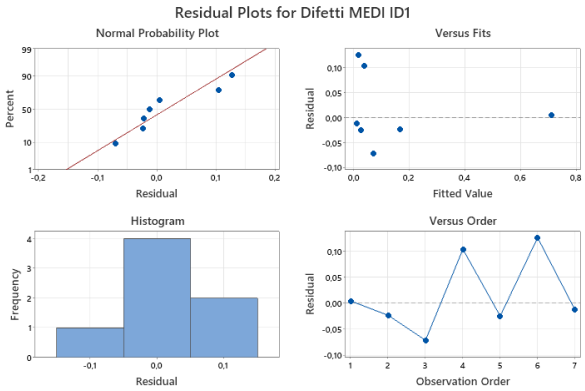


Figura 161 - Output dei Grafici dei Residui restituiti dal modello con andamento a Potenza

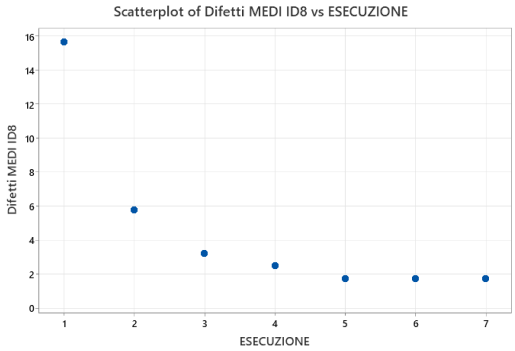


Figura 162 - Grafico di Dispersione o Scatterplot Difetti Medi Totali di Assemblaggio vs Esecuzione della molecola ID8 nelle prove Ripetute

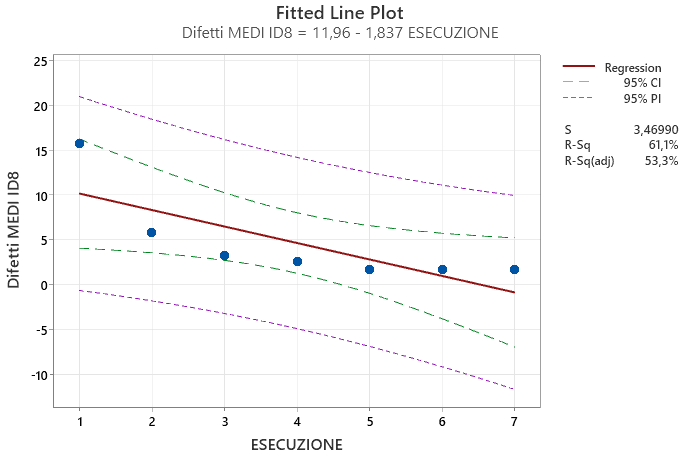


Figura 163 - Curva di Regressione e principali statistiche riassuntive di una Regressione Lineare Difetti Medi Totali di Assemblaggio vs Esecuzione della molecola ID8 nelle prove Ripetute

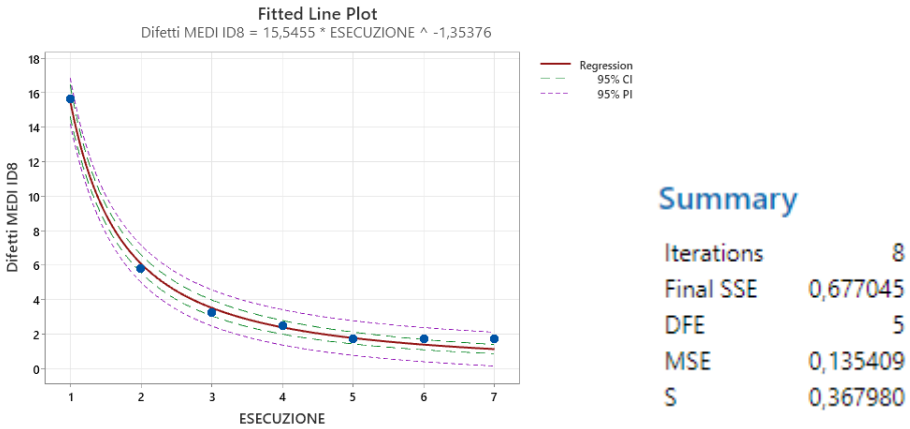


Figura 164 - Curva di Regressione e principali statistiche riassuntive di una Regressione con andamento a Potenza Difetti Medi Totali di Assemblaggio vs Esecuzione della molecola ID8 nelle prove Ripetute

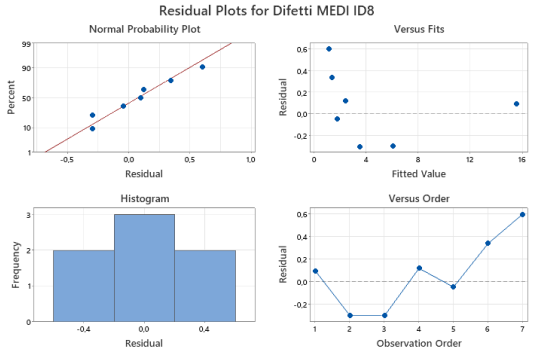


Figura 165 - Output dei Grafici dei Residui restituiti dal modello con andamento a Potenza

C.2.4 Tempi Medi Totali di Assemblaggio

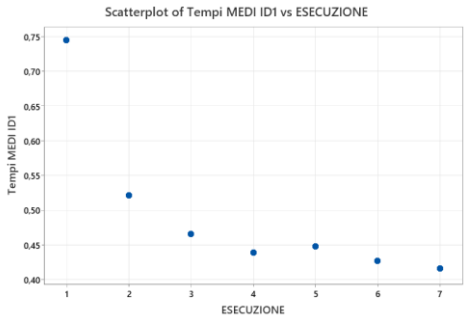


Figura 166 - Grafico di Dispersione o Scatterplot Tempi Medi di Assemblaggio vs Esecuzione della molecola ID1 nelle prove Ripetute

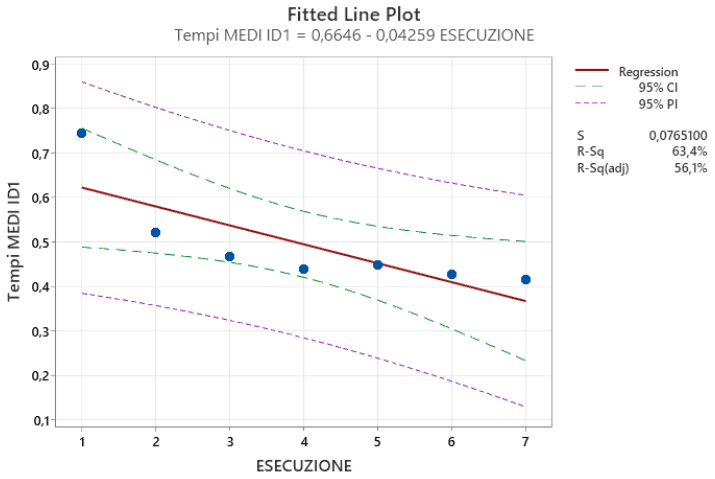


Figura 167 - Curva di Regressione e principali statistiche riassuntive di una Regressione Lineare Tempi Medi di Assemblaggio vs Esecuzione della molecola ID1 nelle prove Ripetute

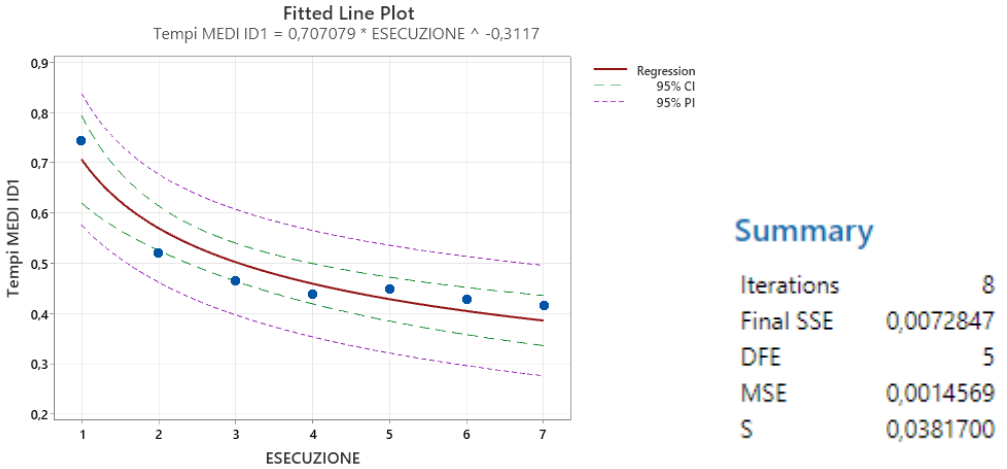


Figura 168 - Curva di Regressione e principali statistiche riassuntive di una Regressione con andamento a Potenza Tempi Medi di Assemblaggio vs Esecuzione della molecola ID1 nelle prove Ripetute

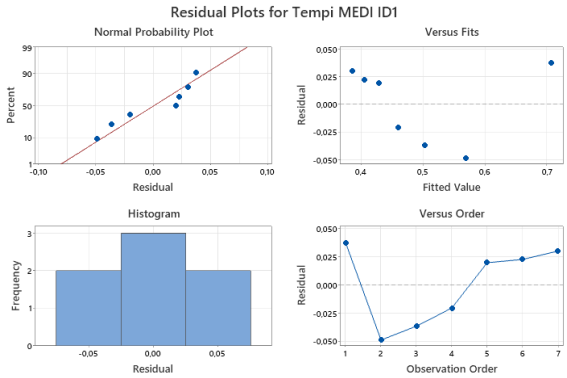


Figura 169 - Output dei Grafici dei Residui restituiti dal modello con andamento a Potenza

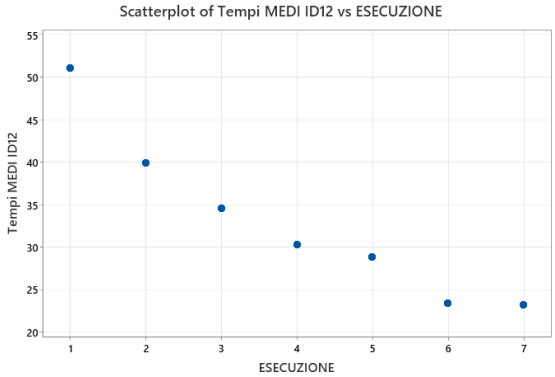


Figura 170 - Grafico di Dispersione o Scatterplot Tempi Medi di Assemblaggio vs Esecuzione della molecola ID12 nelle prove

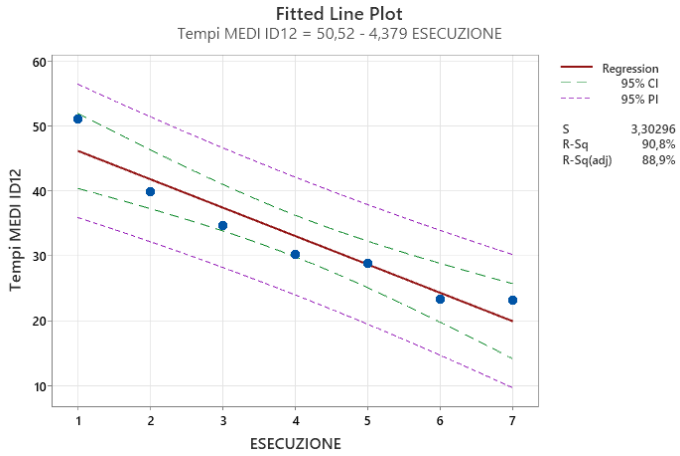


Figura 171 - Curva di Regressione e principali statistiche riassuntive di una Regressione Lineare Tempi Medi di Assemblaggio vs Esecuzione della molecola ID12 nelle prove Ripetute

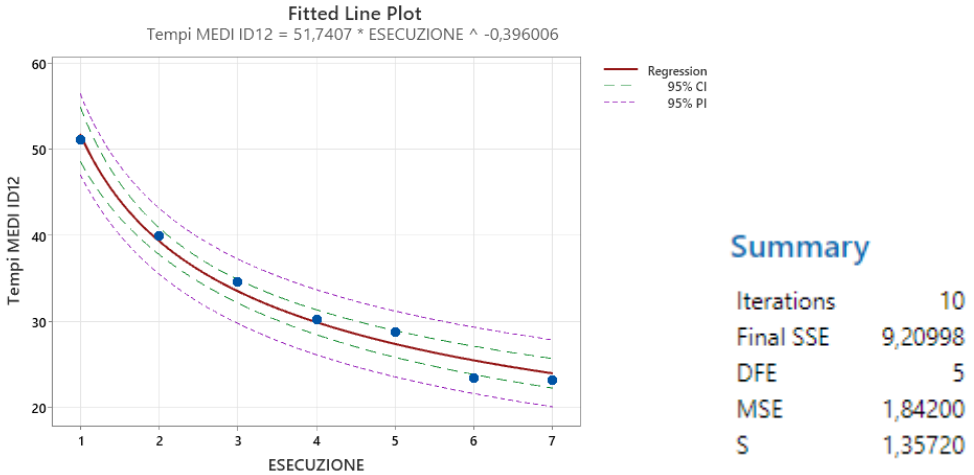


Figura 172 - Curva di Regressione e principali statistiche riassuntive di una Regressione con andamento a Potenza Tempi Medi di Assemblaggio vs Esecuzione della molecola ID12 nelle prove Ripetute

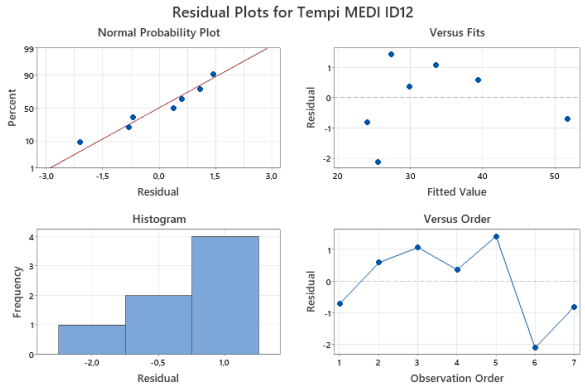


Figura 173 - Output dei Grafici dei Residui restituiti dal modello con andamento a Potenza

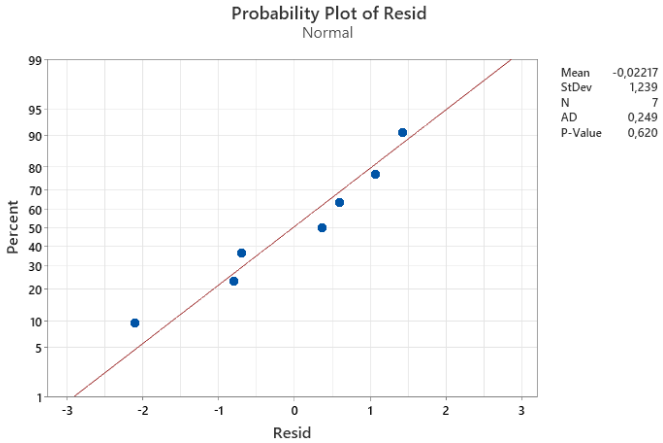


Figura 174 - Risultato test di Anderson-Darling condotto sui Residui ottenuti da un modello con andamento Lineare Tempi Medi di Assemblaggio vs Esecuzione della Molecola ID12 nelle prove Ripetute. Essendo il P-value > 0,05 allora si può confermare che i Residui seguono una distribuzione Normale e quindi la bontà del modello

C.2.5 Difetti Medi Totali di Disassemblaggio

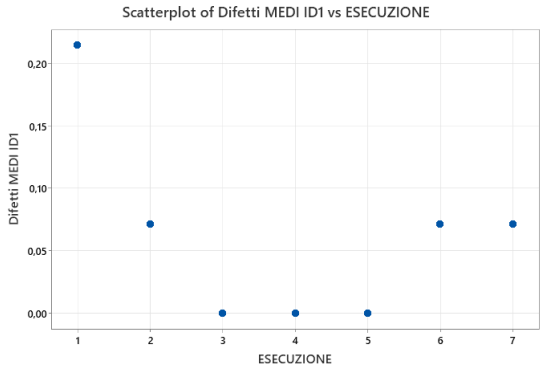


Figura 175 - Grafico di Dispersione o Scatterplot Difetti Medi di Disassemblaggio vs Esecuzione della molecola ID1 nelle prove

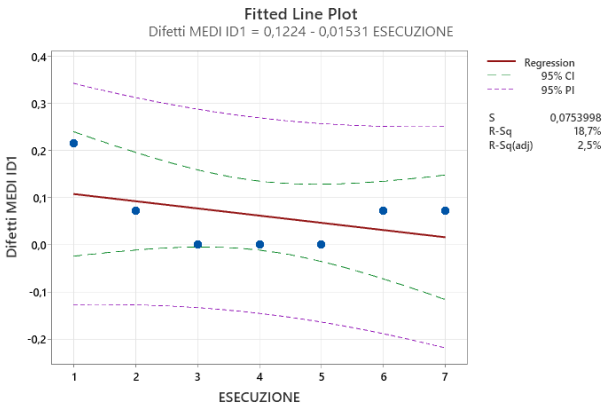


Figura 176 - Curva di Regressione e principali statistiche riassuntive di una Regressione Lineare Difetti Medi di Disassemblaggio vs Esecuzione della molecola ID1 nelle prove Ripetute

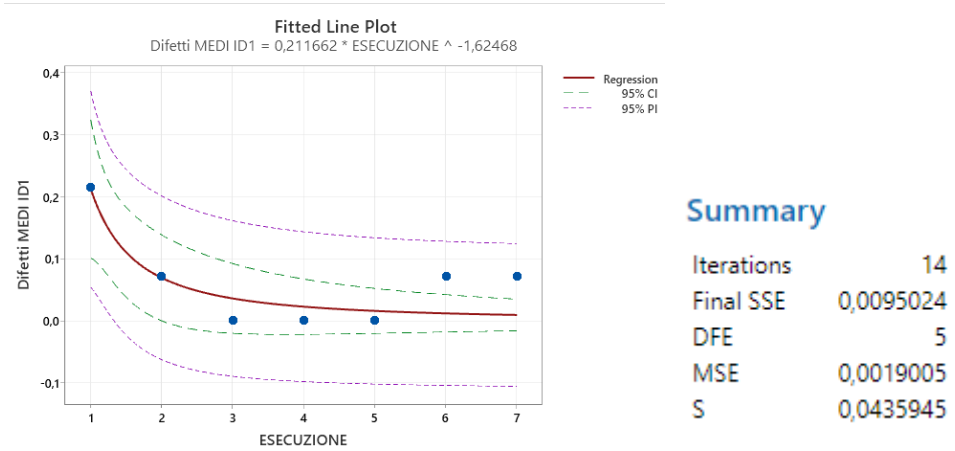


Figura 177 - Curva di Regressione e principali statistiche riassuntive di una Regressione con andamento a Potenza Difetti Medi di Disassemblaggio vs Esecuzione della molecola ID1 nelle prove Ripetute

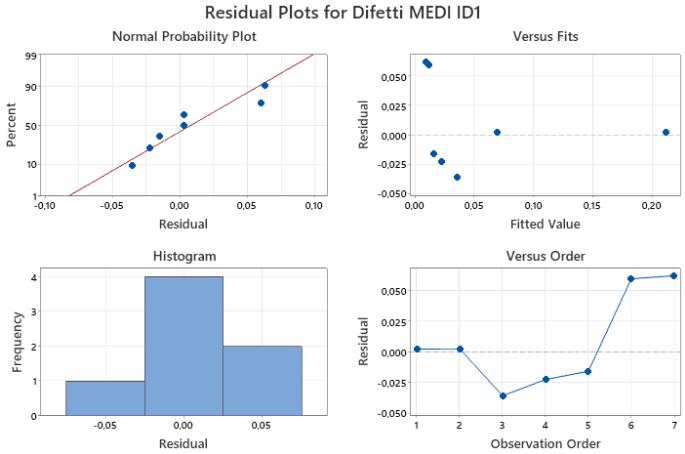


Figura 178 - Output dei Grafici dei Residui restituiti dal modello con andamento a Potenza

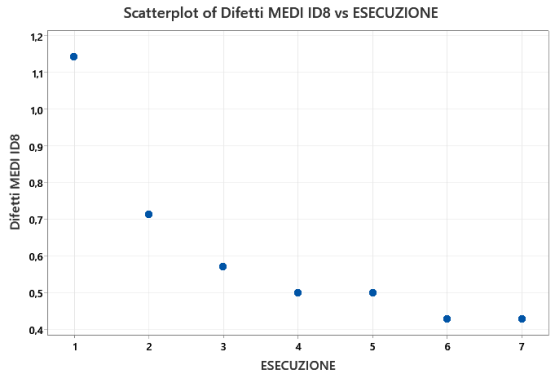


Figura 179 - Grafico di Dispersione o Scatterplot Difetti Medi di Disassemblaggio vs Esecuzione della molecola ID8 nelle prove Ripetute

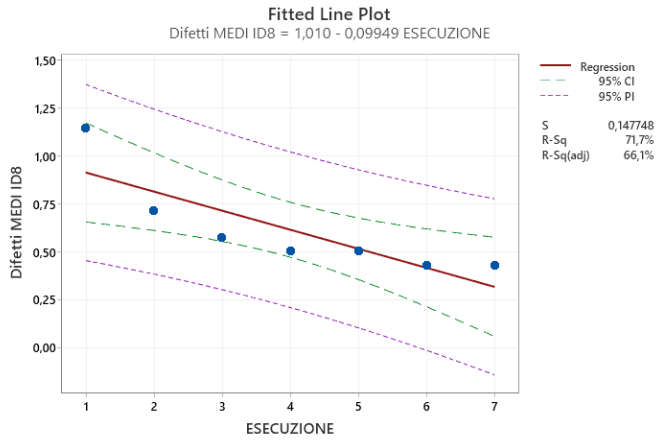


Figura 180 - Curva di Regressione e principali statistiche riassuntive di una Regressione Lineare Difetti Medi di Disassemblaggio vs Esecuzione della molecola ID8 nelle prove Ripetute

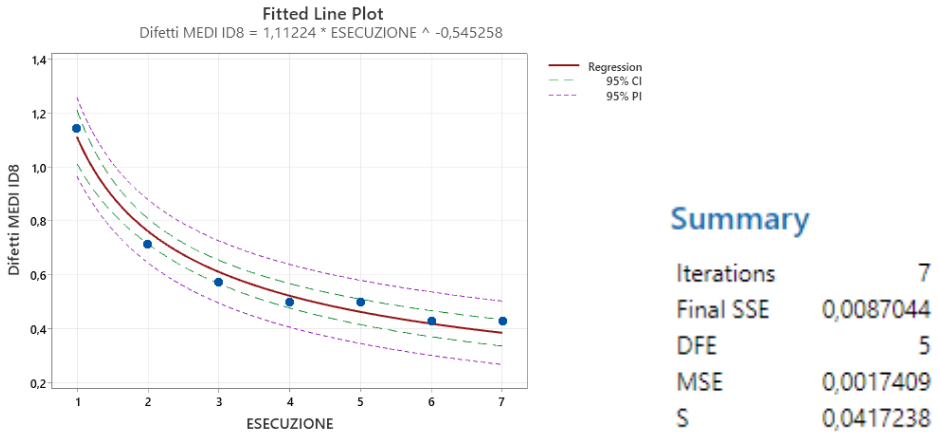


Figura 181 - Curva di Regressione e principali statistiche riassuntive di una Regressione con andamento a Potenza Difetti Medi di Disassemblaggio vs Esecuzione della molecola ID8 nelle prove Ripetute

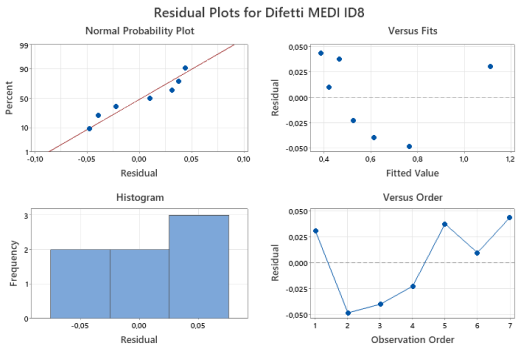


Figura 182 - Output dei Grafici dei Residui restituiti dal modello con andamento a Potenza

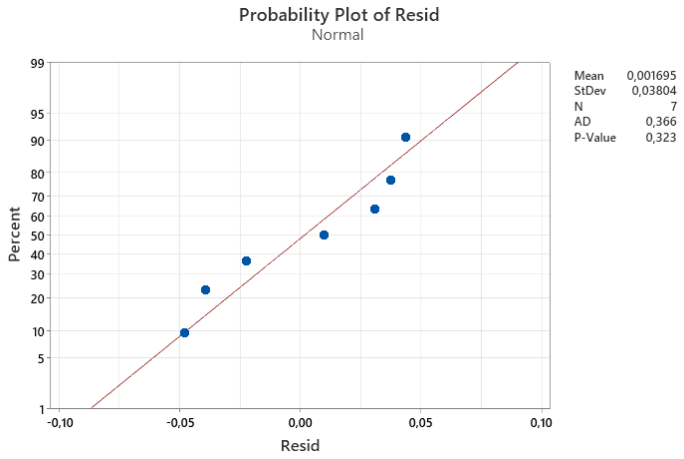


Figura 183 - Risultato test di Anderson-Darling condotto sui Residui ottenuti da un modello con andamento a Potenza Difetti Medi di Disassemblaggio vs Esecuzione della Molecola ID8 nelle prove Ripetute. Essendo il P-value > 0,05 allora si può confermare che i Residui seguono una distribuzione Normale e quindi la bontà del modello

C.2.6 Tempi Medi Totali di Disassemblaggio

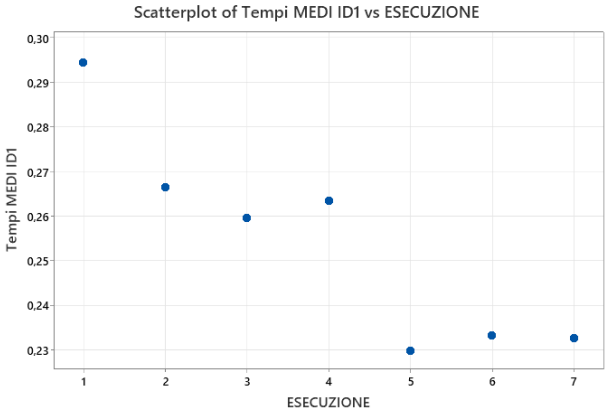


Figura 184 - Grafico di Dispersione o Scatterplot Tempi Medi di Disassemblaggio vs Esecuzione della molecola ID1 nelle prove Ripetute

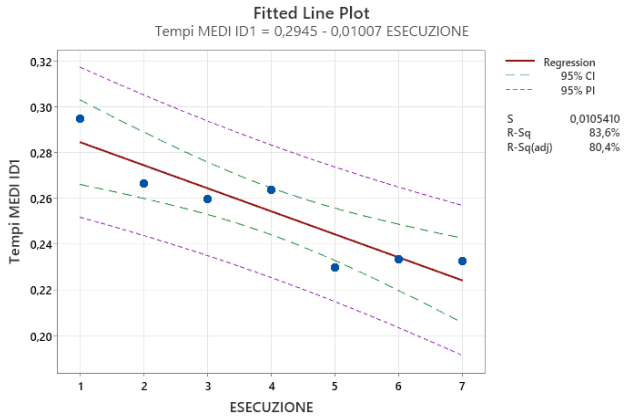


Figura 185 - Curva di Regressione e principali statistiche riassuntive di una Regressione Lineare Tempi Medi di Disassemblaggio vs Esecuzione della molecola ID1 nelle prove Ripetute

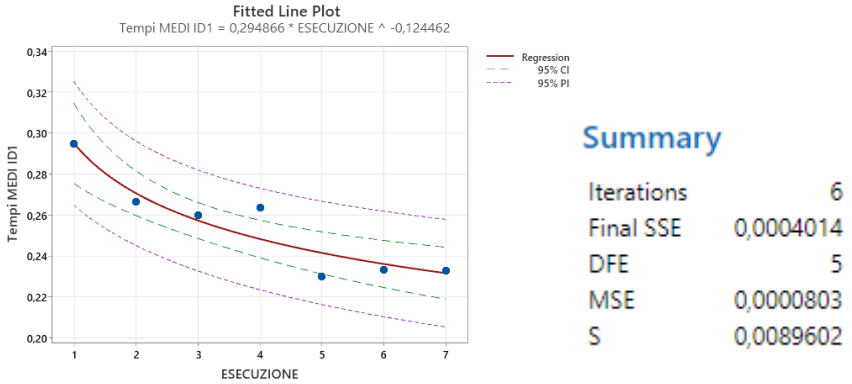


Figura 186 - Curva di Regressione e principali statistiche riassuntive di una Regressione con andamento a Potenza Tempi Medi di Disassemblaggio vs Esecuzione della molecola ID1 nelle prove Ripetute

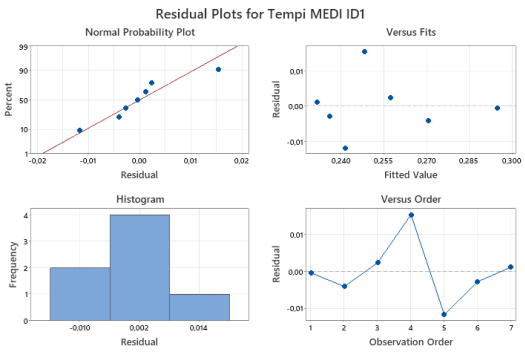


Figura 187 - Output dei Grafici dei Residui restituiti dal modello con andamento a Potenza

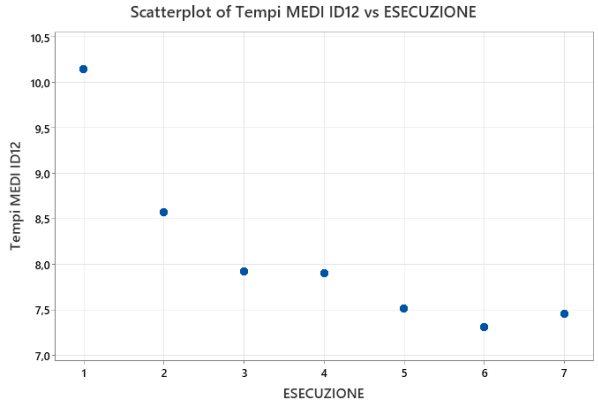


Figura 188 - Grafico di Dispersione o Scatterplot Tempi Medi di Disassemblaggio vs Esecuzione della molecola ID12 nelle prove Ripetute

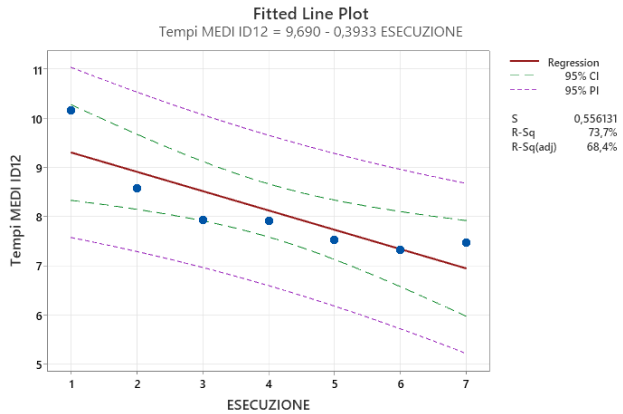


Figura 189 - Curva di Regressione e principali statistiche riassuntive di una Regressione Lineare Tempi Medi di Disassemblaggio vs Esecuzione della molecola ID12 nelle prove Ripetute

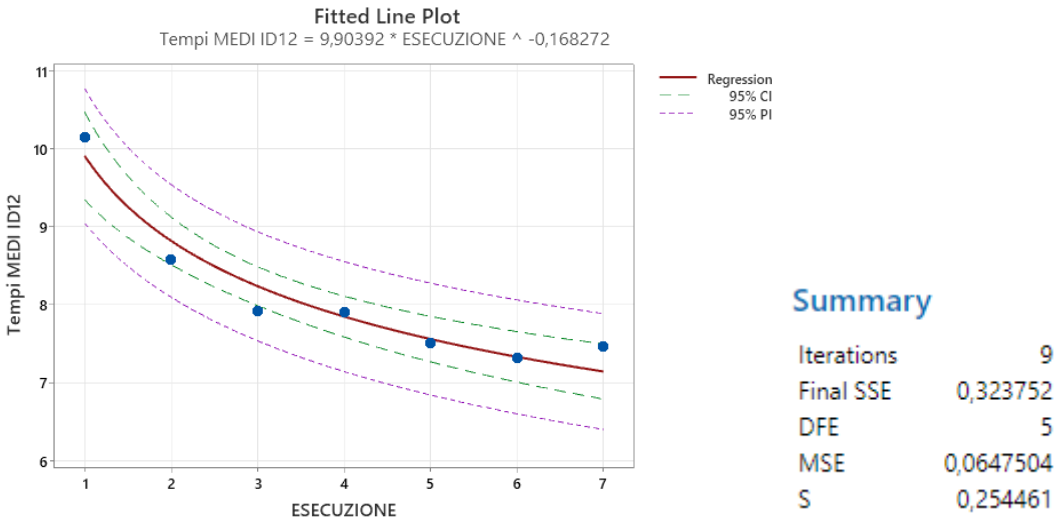


Figura 190 - Curva di Regressione e principali statistiche riassuntive di una Regressione con andamento a Potenza Tempi Medi di Disassemblaggio vs Esecuzione della molecola ID12 nelle prove Ripetute

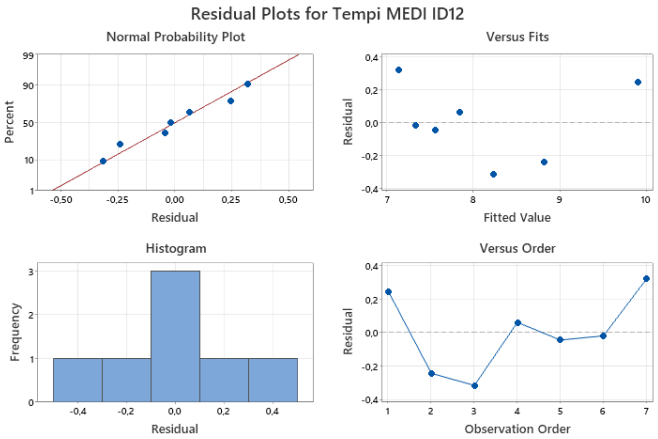


Figura 191 - Output dei Grafici dei Residui restituiti dal modello con andamento a Potenza

BIBLIOGRAFIA

Alkan B., 2017-2018, *A Method to Assess Assembly Complexity of Industrial Products in Early Design Phase*, IEEE Access 6: 989–999

Alkan B., 2018, *Complexity in manufacturing systems and its measures: a literature review*, European J. Industrial Engineering, Vol. 12, No. 1

Alkan B. & Harrison, R., 2019*, A virtual engineering based approach to verify structural complexity of component-based automation systems in early design phase*, Journal of Manufacturing Systems 53: 18-31

Alkan B., 2019, *An experimental investigation on the relationship between perceived assembly complexity and product design complexity*, International Journal of Interactive Design and Manufacturing 13: 1145-1157

Falck A. C., Örtengrenb R., Rosenqvist M. & Söderberg R., 2016, *Criteria for assessment of basic manual assembly complexity*, Elsevier 44: 424–428

Falck A. C., Örtengrenb R., Rosenqvist M. & Söderberg R., 2017, *Basic complexity criteria and their impact on manual assembly quality in actual production*, Elsevier 58: 117-128

Falck A. C., Tarrar M., Mattsson S., Andersson L., Rosenqvist M. & Söderberg R., 2017, *Assessment of manual assembly complexity: a theoretical and empirical comparison of two methods*, International Journal of Production Research 55: 7237-7250

Galetto M, Verna E. & Genta G, 2020*, Accurate estimation of prediction models for operator-induced defects in assembly manufacturing processes*, Quality Engineering 32: 595-613

Krugh M., Antani K., Mears L. & Schulte J., 2016, *Prediction of Defect Propensity for the Manual Assembly of Automotive Electrical Connectors*, Elsevier 5: 144–157

Mattsson M., Tarrar M. & Fast-Berglund A., 2016, *Perceived production complexity – understandingmore than parts of a system*, International Journal of Production Research 54: 6008-6016

McLeod S., 2008*, Likert Scale Definition, Examples and Analysis*, https://www.simplypsychology.org/likert-scale.html

Menard, S., 2001, *Applied Logistic Regression Analysis*, John Wiley & Sons, Inc., Hoboken, New Jersey 07-106

Minitab Documentation https://support.minitab.com/en-us/minitab

Sinha K., 2014, *Structural Complexity and its Implications for Design of Cyber – Physical Systems*, Massachusetts Institute Of Technology

Su Q., Liu L. & Whitney D.E., 2011, *A Systematic Study of the Prediction Model for Operator-Induced Assembly Defects Based on Assembly Complexity Factors* 40: 107–120

Tae Kyun Kim, 2017, *Understanding one-way ANOVA using conceptual figures*, Korean Journal of Anesthesiology 70: 22-26