

POLITECNICO DI TORINO

Corso di Laurea magistrale in Ingegneria Meccanica



TESI DI LAUREA

Valutazione dello step energetico di discretizzazione di batteria per l'ottimizzazione della traiettoria di controllo di un veicolo ibrido

Relatore:
Daniela Anna Misul

Correlatore:
Alessia Musa

Candidato:
Elisabetta Fogliasso

Anno accademico 2018/2019

Ai nonni

Sommario

L'obiettivo del seguente elaborato è l'analisi dei risultati forniti dal tool di ottimizzazione HEVbox. Il codice HEVbox è un tool di ottimizzazione creato precedentemente da altri studenti del politecnico ed è in grado di simulare il comportamento di un veicolo ibrido su una missione prefissata andando inoltre ad ottimizzare i controlli tramite l'algoritmo dynamic programming. La dipendenza dei risultati numerici dal tipo di discretizzazione che si effettua su variabili che sarebbero in realtà uno spazio continuo fa sì che vengano prodotti errori numerici. Scopo di questo elaborato è l'utilizzo del tool per la produzione dei risultati e l'analisi dei risultati stessi in modo da relazionare l'errore numerico sul risultato finale alla discretizzazione utilizzata in particolare per definire la batteria del veicolo. Questo lavoro, in particolare, è suddiviso in più fasi: la raccolta dei dati, l'analisi dei dati ottenuti, l'elaborazione dei dati tramite strumenti statistici quali regressione semplice e multilineare. L'elaborazione dei dati svolta in questa tesi ha permesso relazionare l'errore relativo di CO₂ allo step energetico utilizzato, ovvero alla griglia di discretizzazione utilizzata. Il lavoro svolto ha permesso inoltre di determinare l'influenza dei diversi parametri di design del veicolo sull'errore relativo.

Sommario

INTRODUZIONE.....	8
1. Il cambiamento climatico	10
1.1 Crescita della temperatura globale.....	10
1.2 Il biossido di carbonio CO ₂	12
1.3 Le emissioni nel settore dei trasporti.....	14
2. Veicoli tradizionali e veicoli ibridi	17
2.1 I veicoli tradizionali.....	17
2.2 Veicoli a trazione elettrica	18
3. Codice HEVbox.....	23
3.1 Variabili utilizzate dal tool.....	23
3.3 Step implementati nel codice.....	25
4. Dynamic Programming	27
5. Cenni di Statistica	29
5.1 Distribuzione di Fisher, analisi della varianza e significato del p-value.....	29
5.2 La regressione semplice.....	30
5.3 La regressione multipla	32
5.4 Tecniche statistiche per il machine learning	34
6. Andamento dell'errore relativo di CO ₂ in funzione dello step energetico	36
6.1 Raccolta dei dati tramite simulazioni	37
6.2 Fitting degli andamenti ottenuti con polinomio di secondo grado	41
6.3 Confronto dei dati con il un polinomio dai coefficienti medi	44
6.4 Eliminazione degli outliers	48
6.5 Osservazioni preliminari.....	54
6.6 Ricerca di differenti equazioni	56
6.7 Fitting degli andamenti ottenuti con polinomio di terzo grado	57
7. Regressioni multilineari con superficie di secondo grado.....	61
7.1 Regressioni multiple considerando i dati completi	61
7.2 Regressioni multiple a tre variabili	69
8. Regressioni multilineari con superficie di terzo grado.....	71
8.1 Regressioni multiple considerando i dati completi	71
8.2 Regressioni multiple senza outliers	78
8.3 Regressioni multiple a tre variabili	85

8.4 Osservazioni	88
9. Test delle superfici di regressione.....	89
9.1 Superfici di secondo grado.....	89
9.2 Superfici di terzo grado	92
10. Regressioni multiple per l'architettura p3	95
10.1 Regressioni multiple dei dati completi.....	96
10.2 Regressioni multiple a tre variabili	101
11. Influenza degli altri parametri di design sull'errore relativo	105
11.1 Regressioni multilineari	105
11.2 Stepwise selection	114
11.3. Osservazioni	119
12. L'influenza della missione sull'errore relativo	121
12.1 Calcolo della velocità media e del numero di accelerazioni	122
12.2 Confronto tra WHVC e ciclo ARTEMIS	123
13. Possibili sviluppi futuri.....	128
Ringraziamenti.....	129
Bibliografia.....	130

INTRODUZIONE

La crescita globale della domanda di petrolio è legata per la maggior parte al settore dei trasporti. Supponendo che non avvengano avanzamenti tecnologici di svolta, i combustibili liquidi continueranno a fornire la maggior parte dell'energia totale consumata: il settore dei trasporti rappresenta infatti l'86,4% dell'aumento totale della richiesta di petrolio dal 2012 al 2035. Le automobili, in particolare, sono la principale componente della domanda di petrolio nel settore dei trasporti e si prevede che rimanga tale anche nel 2035 nonostante i trasporti merci su strada e il settore aeronautico stiano crescendo con maggior velocità. Oggi non esistono soluzioni ottimali per sostituire il petrolio come fonte principale di energia per il trasporto su strada. Sarà quindi necessario adottare in futuro più soluzioni coesistenti in modo da creare un insieme di fonti energetiche per il trasporto. Il risparmio energetico e le emissioni sono e continueranno ad essere di primario interesse nel futuro dei sistemi di trasporto: il trasporto sostenibile e le risorse di energia rinnovabile sono l'unica soluzione a lungo termine. Le emissioni globali dovute al settore dei trasporti sono infatti aumentate solo dello 0,6% nel 2018 rispetto all'1,6% annuali dello scorso decennio, questo grazie al miglioramento delle efficienze e all'elettrificazione. In particolare, il settore dei trasporti è responsabile del 24% delle emissioni dirette di CO₂. I veicoli da strada come automobili, camion e autobus contribuiscono per i ¾ delle emissioni di CO₂ totali dovute al settore dei trasporti.

I trasporti rappresentano almeno un quarto delle emissioni di gas serra europei i quali sono la causa principale dell'inquinamento dell'aria. Le emissioni inquinanti derivanti dai trasporti che danneggiano la nostra salute devono essere drasticamente ridotte. Avendo come obiettivo il risparmio energetico e la riduzione delle emissioni inquinanti, l'unione Europea ha posto stretti vincoli nel nuovo programma per la riduzione delle emissioni di CO₂. La media delle emissioni omologate della gamma di un qualsiasi costruttore dovrà passare dai 95 g/km massimi del 2021 agli 80 g/kg previsti nel 2025: un taglio del 15% che anticipa quello del 37,5% previsto per il 2030. Quello emesso dall'unione Europea sembra essere un cronoprogramma molto rigido, specie se si considera la media delle emissioni omologate delle auto di nuova immatricolazione dello scorso anno, il quale toccava quota 118 g/km.

I nuovi limiti rendono praticamente indispensabile l'adozione di tecnologie di elettrificazione, cioè l'introduzione massiccia di veicoli ibridi ed elettrici i quali garantiscono da un lato la riduzione delle emissioni inquinanti e di CO₂ e dall'altro le stesse performance richieste dagli utenti. In quest'ottica le tecnologie di elettrificazione applicate al settore dei trasporti possono rappresentare una soluzione a medio-lungo termine al fine di ridurre l'inquinamento atmosferico. Quest'ultimo è definito come uno stato di alterazione delle condizioni naturali dell'aria e può esser causato da fattori differenti, noti come principali agenti inquinanti (fisici, chimici e biologici). Gli effetti dell'inquinamento atmosferico sono sempre più noti e tangibili, soprattutto nei grandi centri metropolitani in cui è particolarmente sentito il problema dello smog. Le sostanze inquinanti vanno ad impattare sia sull'ambiente sia sulla salute dell'uomo: se per l'uomo le sostanze inquinanti vanno ad interferire con l'attività polmonare, per l'ambiente queste ultime favoriscono effetti come il buco dell'ozono, l'effetto serra e le piogge acide. Limitare quindi la produzione sia di agenti inquinanti sia di CO₂ è una sfida importante per salvaguardare l'ambiente in cui viviamo.

In quest'ottica, il lavoro svolto in questa ha come obiettivo finale la previsione e minimizzazione delle emissioni di CO₂ per un veicolo ibrido. In questa tesi infatti, viene fatto uso di un tool di ottimizzazione realizzato precedentemente a questo elaborato e vengono analizzati i risultati forniti dall'algoritmo dynamic programming il quale ha l'obiettivo di ottimizzare i controlli per il veicolo. Il tool di ottimizzazione ha l'obiettivo di simulare un veicolo ibrido su un ciclo prefissato e scelto dall'utente. Alla base dell'algoritmo che costituisce il tool vi è la necessità di andare a rappresentare in ambiente di calcolo variabili reali alcune delle quali sono effettivamente variabili continue, come per esempio l'asse temporale o la SOC window della batteria del veicolo. Tali variabili continue devono necessariamente essere rese variabili discrete per poter essere elaborate all'interno dell'ambiente di calcolo. L'introduzione della discretizzazione produce un'approssimazione e quindi un errore numerico sul risultato finale. L'analisi dei risultati prodotti dal toolbox ha permesso di correlare la discretizzazione utilizzata all'errore relativo percentuale che si ottiene tramite la simulazione. La raccolta dei dati e il successivo uso di strumenti statistici hanno permesso di ricavare relazioni tra l'errore relativo, la discretizzazione della batteria e diversi parametri di design del veicolo.

1. Il cambiamento climatico

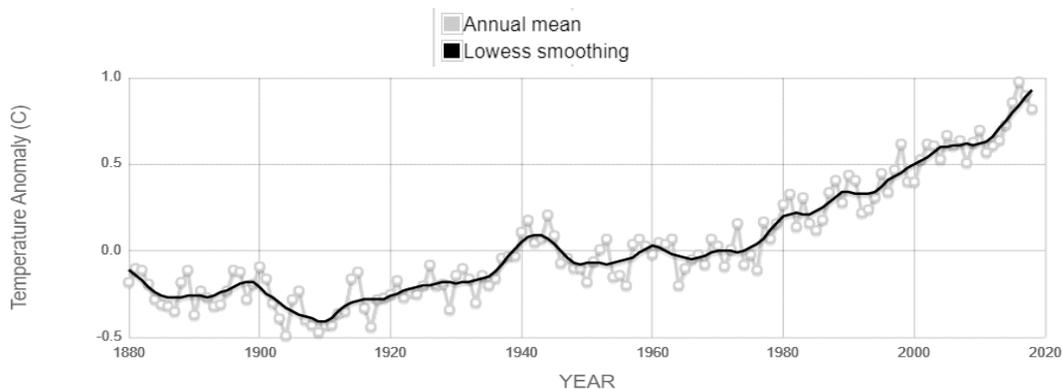
Il clima terrestre è cambiato nel corso dei secoli: negli ultimi 650.000 anni si sono verificati sette cicli di avanzamento e ritirata dei ghiacciai con un brusco termine dell'ultima era glaciale circa 7.000 anni fa il quale ha dato inizio all'attuale era climatica insieme alla civilizzazione umana. La maggior parte dei cambiamenti climatici sono attribuiti a piccole variazioni dell'orbita terrestre che variano la quantità di energia solare che il nostro pianeta riceve.

Il trend attuale riguardante il riscaldamento globale è particolarmente significativo poiché è per la maggior parte dovuto all'attività umana dalla metà del ventesimo secolo. I satelliti artificiali orbitanti insieme ad altre tecnologie hanno permesso di ottenere differenti tipi di informazioni sul nostro pianeta e sul suo clima su scala globale. I dati raccolti attraverso gli anni hanno rivelato segnali di cambiamento climatico ed inoltre a metà del diciannovesimo secolo è stata dimostrata la capacità di trattenimento del calore da parte del biossido di carbonio e di altri gas, essi possono inoltre influenzare il trasferimento di energia infrarossa attraverso l'atmosfera. Non c'è dubbio sul fatto che l'aumento dei gas serra sia una delle cause del riscaldamento globale.

1.1 Crescita della temperatura globale

Dal tardo diciannovesimo secolo la media della temperatura superficiale terrestre è cresciuta di circa 0,9 gradi Celsius, incremento dovuto soprattutto all'aumento del diossido di carbonio ed altre emissioni di gas in atmosfera dovute all'attività umana. La maggior parte del riscaldamento globale si è verificata negli scorsi 35 anni con i cinque anni più caldi a partire dal 2010. Non solo il 2016 è risultato essere l'anno più caldo ma otto dei dodici mesi sono stati registrati come i più caldi per quei rispettivi mesi.

Il grafico riportato in figura 1.1 illustra il cambiamento della temperatura sulla superficie terrestre rispetto alla media della temperatura degli anni tra il 1951 e il 1980. Si può notare come 19 degli anni più caldi si sono verificati dopo il 2001 con l'eccezione dell'anno 1998. L'anno 2016 è registrato come l'anno più caldo dalla NASA.



Source: climate.nasa.gov

Figura 1.1 – Cambiamento relativo della temperatura. Data source: NASA/GISS

In figura 1.2 è riportata la variazione della temperatura globale secondo i dati NASA/GISS confrontando l'anno 1884 con l'anno 2018. Si può notare una marcata variazione di temperatura a livello globale.

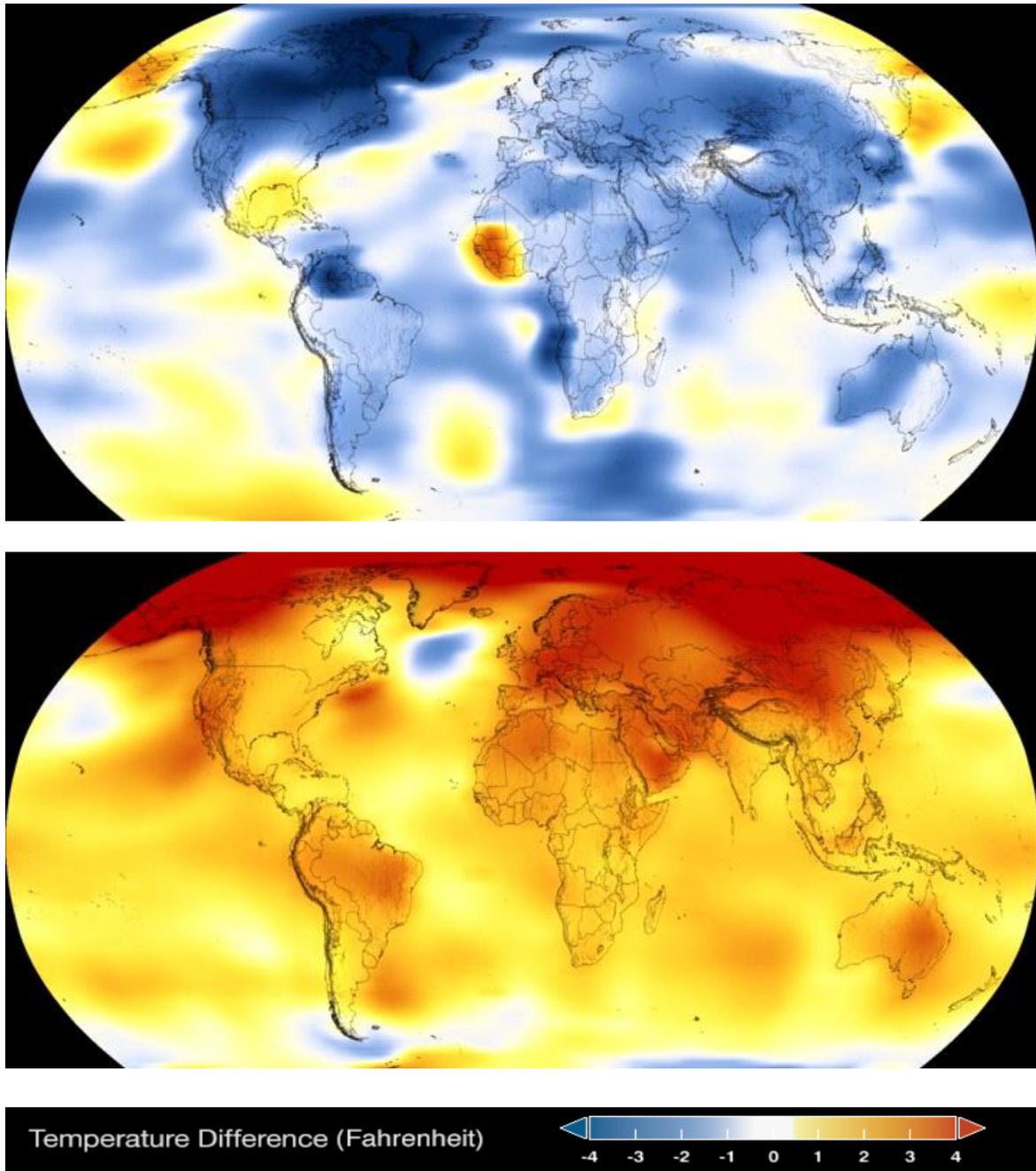


Figura 1.2 – Confronto delle temperature tra 1884 e 2018 – Data Source: NASA/GISS

1.2 Il biossido di carbonio CO₂

Il biossido di carbonio è un importante gas serra emesso in atmosfera tramite attività umane quali la deforestazione e la combustione di combustibili fossili ma anche tramite processi naturali come la respirazione umana e le eruzioni vulcaniche. Nel 2018 la media globale di CO₂ ammonta a 407,4 ppm (parts per million), tale valore supera i livelli raggiunti negli ultimi 800.000 anni, infatti l'ultima volta che i livelli di biossido di carbonio nell'atmosfera sono stati così alti fu più di 3 milioni di anni fa quando la temperatura era di 2-3 gradi celsius maggiore rispetto a quella dell'era preindustriale. Le concentrazioni di biossido di carbonio stanno crescendo a causa dei combustibili fossili utilizzati come fonte di energia per le attività umane. I combustibili fossili come carbone e petrolio contengono il carbonio estratto dall'atmosfera attraverso la fotosintesi nell'arco di milioni di anni. L'attività umana sta restituendo all'atmosfera questa quantità di carbonio in qualche centinaio di anni.

Secondo il report sul clima prodotto nel 2018 dalla NOAA (National Oceanic Atmospheric Administration) e dalla American Meteorological Society, tra l'anno 2017 e l'anno 2018 si è verificato un incremento della concentrazione globale di biossido di carbonio di 2,5 ppm, aumento simile a quello verificatosi tra il 2016 e il 2017 di circa 2,2 ppm. Tale crescita di concentrazione è importante se paragonata a quella degli anni sessanta, in cui l'incremento si aggirava intorno alle 0,6 ppm annue. Il tasso annuale di aumento di concentrazione di biossido di carbonio atmosferico degli ultimi 60 anni è circa 100 volte più alto rispetto al precedente aumento naturale come verificatosi alla fine dell'ultima era glaciale tra 11,000 e 17,000 anni fa.

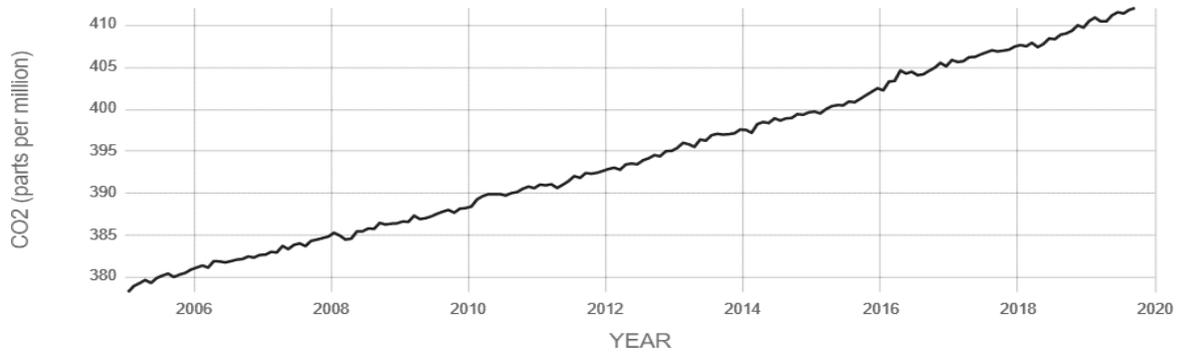
Il biossido di carbonio è un gas serra con la capacità di assorbire calore e ha una grande importanza per il trattenimento del calore terrestre. La superficie terrestre e la superficie oceanica rilasciano infatti energia termica poiché riscaldate dai raggi solari. Diversamente dall'ossigeno o dall'azoto, che compongono la maggior parte dell'atmosfera, i gas serra assorbono il calore emanato dalle superfici e lo rilasciano gradualmente nel tempo. Senza gli effetti naturali dei gas serra, la temperatura annuale media terrestre sarebbe drasticamente troppo bassa ma l'aumento dei gas serra ha estremizzato questo effetto aumentando le temperature medie globali. Il biossido di carbonio è uno dei più importanti gas serra, assorbe meno calore per molecola rispetto al metano o agli ossidi di azoto ma è più abbondante e rimane in atmosfera per molto più a lungo. L'aumento di CO₂ in atmosfera è responsabile di circa i due terzi del totale dello sbilanciamento energetico causa dell'aumento delle temperature.

Un'altra ragione alla base dell'importanza del biossido di carbonio è che si dissolve negli oceani, poiché reagisce con le molecole d'acqua producendo acido carbonico e abbassando il pH oceanico. Dall'inizio della rivoluzione industriale il livello di pH degli oceani è sceso da 8,21 a 8,10, fenomeno chiamato acidificazione oceanica. Un calo dello 0,1 potrebbe non sembrare significativo ma è necessario ricordare che il pH ha una scala logaritmica per cui una variazione dello 0,1 significa un incremento del 30% di acidità la quale interferisce negativamente con le attività della fauna e della flora marina.

I naturali aumenti della concentrazione di biossido di carbonio hanno periodicamente surriscaldato la temperatura terrestre durante i cicli glaciali negli ultimi milioni di anni. I periodi caldi interglaciali iniziavano con un piccolo aumento della luce solare dovuto ad una piccola oscillazione dell'asse di rotazione terrestre o dal percorso della sua orbita intorno al sole. La luce solare aggiuntiva ha causato un aumento delle temperature riscaldando gli oceani i quali hanno rilasciato biossido di carbonio. Questo composto in atmosfera ha amplificato l'aumento di temperatura iniziale.

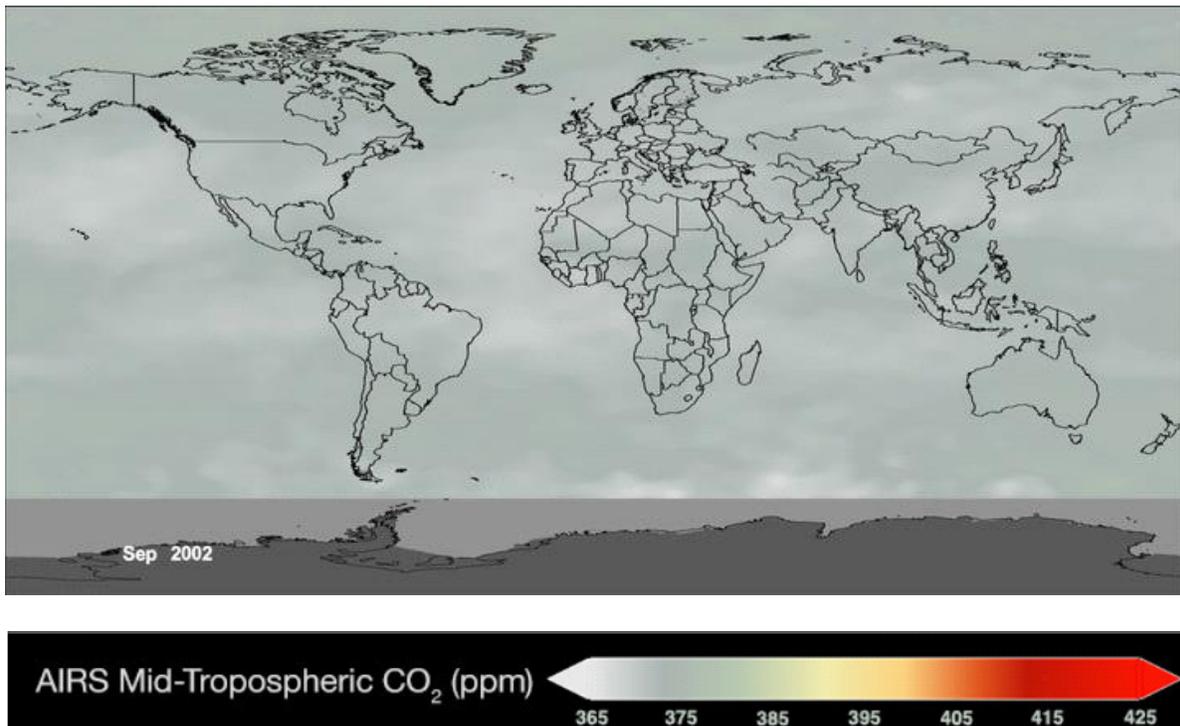
Durante i cicli glaciali degli scorsi milioni di anni i livelli di biossido di carbonio non hanno mai superato i 300 ppm. Prima della rivoluzione industriale infatti la media globale era di 280 ppm. Da quando iniziarono i costanti monitoraggi del Mauna Loa Volcanic Observatory nel 1958 il livello di biossido di carbonio globale ammontava già a 315 ppm. Nel maggio 2013 le misurazioni superarono i 400 ppm per la prima volta. Se la domanda di energia continua a salire e questa viene soddisfatta tramite combustibili fossili il biossido di carbonio atmosferico potrebbe superare i 900 ppm entro la fine di questo secolo.

La figura 1.3 mostra la variazione della concentrazione di CO₂ secondo i dati del Mauna Loa Observatory delle Hawaii mentre in figura 1.4 è mostrata la variazione della concentrazione del diossido di carbonio tra il 2002 e il 2016.



Source: climate.nasa.gov

Figura 1.3 – CO₂ ppm – Data Source: NOAA, Credits: NASA



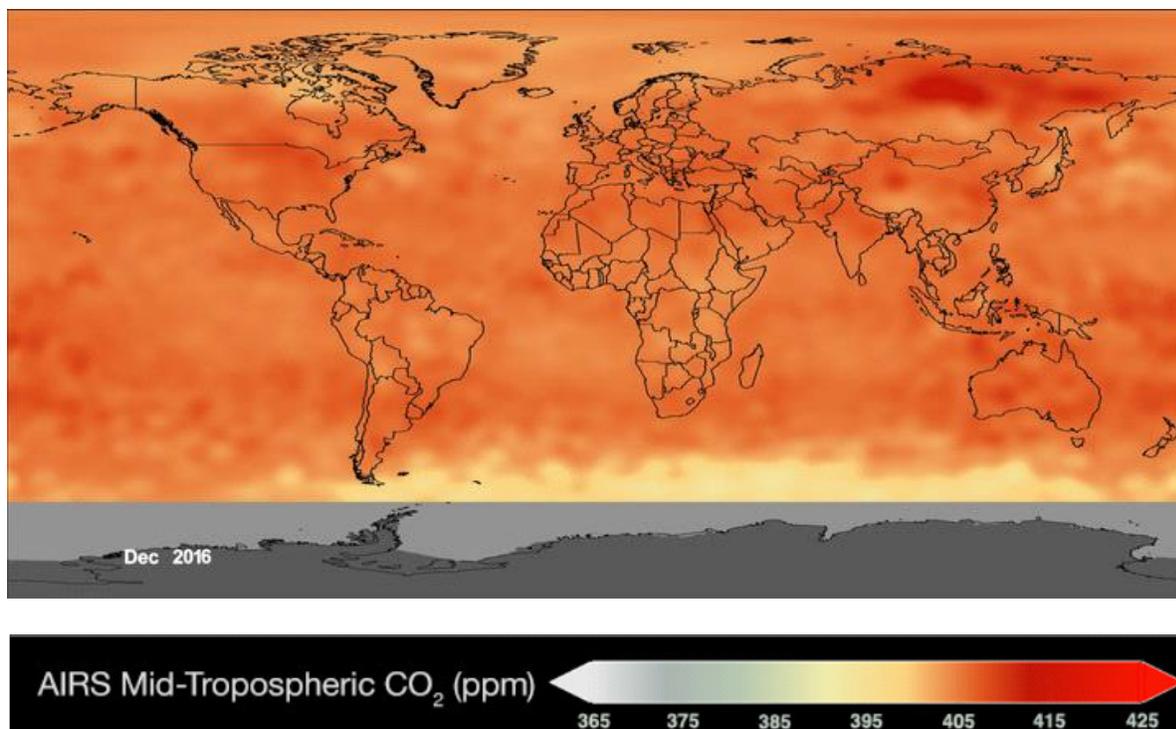


Figura 1.3 – Concentrazione CO₂ 2002 – 2016. Data Source Atmospheric Infrared Sounder (AIRS), Credits: NASA

1.3 Le emissioni nel settore dei trasporti

Il settore dei trasporti rappresenta almeno un quarto delle emissioni dei gas serra in Europa e rappresenta una delle cause principali dell'inquinamento dell'aria. Il settore dei trasporti infatti, non ha visto lo stesso graduale declino delle emissioni come gli altri settori: le emissioni hanno iniziato ad essere ridotte solo nel 2007 e rimangono ancora maggiori rispetto al 1990. In questo settore, il trasporto su strada è di gran lunga la maggiore fonte di emissioni totalizzando più del 70% delle emissioni di gas serra nel settore dei trasporti nel 2014. In figura 1.4 sono visibili le emissioni dei diversi settori.

Con un cambiamento globale verso un'economia a bassa emissione di carbonio già in atto, la strategia di mobilità a basse emissioni della commissione adottata nel luglio 2016 mira ad assicurarsi che l'Europa rimanga competitiva e capace di rispondere alla crescente necessità di mobilità di persone e beni.

La risposta Europea alla sfida della riduzione delle emissioni nel settore dei trasporti è un irreversibile spostamento verso la mobilità a basse emissioni. Entro i prossimi trenta anni le emissioni di gas serra dei trasporti dovranno essere almeno il 60% inferiori rispetto a quelle del 1990 in modo da arrivare nei decenni successivi a zero emissioni. Le emissioni di inquinanti devono infatti essere drasticamente ridotte senza ritardo. La strategia Europea integra un insieme di misure di supporto per la transizione verso un'economia a basse emissioni di carbonio e supporta lavoro, crescita, investimenti ed innovazione. La strategia ha come obiettivo il beneficio per i cittadini europei tramite il miglioramento della qualità dell'aria, la riduzione del rumore, la riduzione dei livelli di traffico e il miglioramento della sicurezza. I consumatori potranno beneficiare di veicoli a basso consumo di energia, di infrastrutture migliori e di

migliori connessioni tra metodi di trasporto e sicurezza grazie all'inserimento di nuove tecnologie digitali.

L'unione Europea ha identificato tre aree prioritarie su cui agire:

- L'aumento dell'efficienza dei sistemi di trasporto, facendo uso di tecnologie digitali, prezzi intelligenti ed un ulteriore incoraggiamento al passaggio verso le modalità di trasporto a basse emissioni;
- Velocizzare lo sviluppo delle energie alternative a basse emissioni per il trasporto come l'utilizzo di biocombustibili, elettricità, idrogeno e combustibili sintetici rinnovabili.
- Puntare a veicoli a zero emissioni. Mentre ulteriori miglioramenti al motore a combustione interna saranno necessari, l'Europa deve accelerare la transizione verso veicoli a basse o zero emissioni.

Città ed autorità locali avranno un ruolo cruciale nell'attuare questa strategia, stanno già implementando incentivi per veicoli ed energie alternative a basse emissioni incoraggiando l' "active travel", i trasporti pubblici e i car sharing.

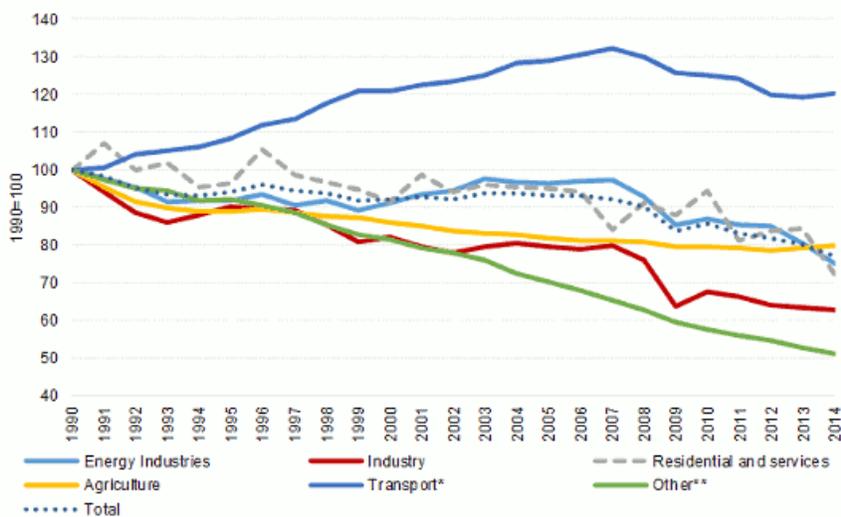


Figura 1.4 – Emissioni per differenti settori – Source: EEA (European Environment Agency)

Norme su emissioni CO2 per autovetture e furgoni

L'Unione Europea ha adottato misure per la riduzione delle emissioni di CO2 per furgoni ed autovetture con limiti ancora più rigorosi nel 2019. Le norme rivedute contengono disposizioni più rigorose sulle emissioni di CO2 per le autovetture e i veicoli commerciali. Gli stati membri hanno infatti approvato l'accordo provvisorio raggiunto il 17 dicembre 2018 dai rappresentanti della presidenza e del Parlamento. In base alle nuove norme, dal 2030 le nuove auto genereranno emissioni di CO2 inferiori in media del 37,5% rispetto ai livelli del 2021. Tra il 2025 e il 2029 autovetture e furgoni dovranno generare emissioni di CO2 inferiori del 15%, le emissioni di CO2 dei nuovi furgoni dovranno essere ridotte in media, a partire dal 2030, del 31% rispetto al 2021. In figura 1.5 è riportata l'infografica con il taglio delle emissioni sul settore dei trasporti.



Figura 1.5 – Infografica sui trasporti stradali: taglio sulle emissioni di CO2

Il Parlamento e il Consiglio hanno approvato un meccanismo per andare a stimolare la vendita di veicoli a basse e a zero emissioni come le auto completamente elettriche ed i veicoli ibridi ricaricabili, sulla base dell'approccio suggerito dalla Commissione nella proposta iniziale. I costruttori che rispettano determinati livelli di riferimento saranno ricompensati con obiettivi per le emissioni di CO2 meno severi. I livelli di riferimento per il 2025 saranno fissati al 15% per auto e furgoni, per il 2030 al 35% per le auto e al 30% per i furgoni. Sono stati mantenuti incentivi specifici per le autovetture a basse e a zero emissioni concordati nell'orientamento generale del Consiglio. Va inoltre ricordato che sono state concordate norme più severe per la transizione dalla procedura di prova NEDC alla più accurata procedura WLTP come base per il calcolo degli obiettivi specifici in materia di emissioni per i costruttori. Sarà inoltre dedicata maggiore attenzione al controllo delle emissioni reali di guida. La commissione monitorerà infatti la rappresentatività reale dei valori delle emissioni di CO2 sulla base dei dati registrati dai misuratori del consumo di carburante installati sulle auto e sui furgoni nuovi. Sono infine previste disposizioni specifiche sulle prove di conformità in servizio e sulla individuazione di strategie per il miglioramento artificiale delle prestazioni di emissioni di auto e furgoni.

Le misure e gli obiettivi proposti si basano sul quadro di politica climatica ed energetica ad orizzonte 2030 e inoltre sulla strategia dell'Unione dell'energia la quale punta ad una riduzione delle emissioni dei trasporti e del consumo energetico.

2. Veicoli tradizionali e veicoli ibridi

2.1 I veicoli tradizionali

I veicoli tradizionali con motore a combustione interna (ICE) forniscono buone prestazioni grazie ai vantaggi dell'alta densità energetica dei combustibili a petrolio, il motore a combustione interna infatti perette di convertire l'energia chimica posseduta dal combustibile in lavoro meccanico reso disponibile all'albero motore. Il motore a combustione interna è infatti, per questa ed altre ragioni, la macchina motrice più diffusa ed utilizzata al mondo per la trazione di tipo stradale. Sicuramente entrano in gioco altre motivazioni tra cui l'alto valore del rapporto potenza-volume ma anche la possibilità di alimentazione con differenti combustibili. Tali caratteristiche hanno reso possibile la sua diffusione per la trazione dei veicoli che sono quindi diventati i veicoli di tipo tradizionale.

D'altra parte, i veicoli che possiedono ICE tradizionali portano come svantaggi un alto consumo energetico e l'inquinamento ambientale. Essi producono infatti emissioni inquinanti quali il particolato, NO_x, HC e CO nonostante l'uso di differenti tecnologie quali EGR (Exhaust Gas Recirculation) ed SCR (Selective Catalytic Reduction). Non di minore importanza sono le emissioni di CO₂ provocate da questo tipo di macchina termica. C'è inoltre da tenere conto del fatto che generalmente i motori sono sovradimensionati rispetto alla potenza media richiesta, inoltre i veicoli dotati di motori a combustione interna hanno quindi un alto consumo di carburante, questo è dovuto ad alcune principali motivazioni:

- Le caratteristiche del rendimento non combaciano con le reali necessità operative;
- La dissipazione dell'energia cinetica del veicolo durante le frenate, specialmente nelle aree urbane;
- La bassa efficienza della trasmissione idraulica nelle automobili attuali anche in questo caso nelle aree urbane;
- Non può essere raggiunta una strategia precisa nel caso di cambio manuale poiché dipende fortemente dal comportamento del guidatore. Questi limiti possono essere parzialmente evitati tramite il cambio automatico (AMT) o tramite un sistema CVT.

Per rimediare a questa serie di problematiche lo sviluppo tecnologico si sta indirizzando verso i veicoli con propulsione ibrida. I veicoli ibridi combinano due o più fonti di potenza che può direttamente o indirettamente generare la propulsione del veicolo. La fonte di energia primaria è generalmente l'energia chimica accumulata nel combustibile del motore a combustione interna ed in base alla natura della fonte secondaria di potenza, "ibrido" può significare ibrido elettrico (HEV), ibrido meccanico, ibrido idraulico, cella di combustibile ibrida eccetera. La trazione di tipo ibrido elettrico consente di sfruttare il motore a combustione interna nelle condizioni di funzionamento che garantiscono rendimenti più elevati insieme ad un sistema di accumulo che eroga picchi di potenza necessari quando richiesti oppure che immagazzina la potenza in eccesso durante le fasi di frenatura. Grazie a queste caratteristiche i veicoli ibridi sono oggi competitivi e rappresentano un'ottima alternativa rispetto ai veicoli di tipo tradizionale.

2.2 Veicoli a trazione elettrica

I veicoli a trazione elettrica sono classificati in primo luogo in BEVs (battery electric vehicles) ed HEVs (hybrid electric vehicles). Nei seguenti paragrafi vengono analizzate le differenze tra questi due tipi di veicoli.

BEVs

I veicoli elettrici con batteria (BEVs, battery electric vehicles) utilizzano come unica fonte di potenza la batteria in dotazione. Questi veicoli sono dotati di un inverter DC/AC il quale ha il compito di convertire la corrente continua proveniente dalle batterie in corrente alternata, la potenza elettrica viene successivamente trasmessa dall'inverter alla macchina elettrica la quale converte tale potenza in potenza meccanica resa disponibile per la propulsione del veicolo. La trazione di questo tipo di veicoli è quindi completamente elettrica. La macchina elettrica (o le macchine elettriche) in dotazione a questi veicoli sono di tipo reversibile, questo significa che possono funzionare sia da motore che da generatore. Nel primo caso la macchina elettrica converte l'energia elettrica in energia meccanica, nel secondo caso l'energia meccanica è convertita in energia elettrica utilizzata per la ricarica delle batterie, questo avviene durante la fase di frenata (regenerative braking) in cui l'energia cinetica del veicolo viene recuperata e utilizzata per la ricarica.

È facile comprendere come questi veicoli abbiano dei vantaggi importanti rispetto ai tradizionali veicoli a combustione interna, tra cui l'alta efficienza energetica e zero emissioni inquinanti. In realtà si parla di zero emissioni inquinanti riferendosi alle emissioni tank-to-wheel ma è ragionevole riferirsi anche alle emissioni wheel-to-wheel le quali tengono in considerazione gli inquinanti e la CO₂ emessi durante il processo di produzione dell'energia per la ricarica della batteria stessa. Nonostante ciò le prestazioni sono molto meno competitive rispetto ad un motore a combustione interna soprattutto a causa della durata operativa della batteria ed in particolare a causa del basso contenuto energetico delle batterie rispetto all'alto contenuto energetico del combustibile.

HEVs

I veicoli ibridi elettrici (HEV) i quali usano due fonti di potenza, una primaria ed una secondaria, hanno i vantaggi sia dei veicoli tradizionali sia dei veicoli BEVs e superano i loro svantaggi. Questi presentano infatti sia un motore a combustione interna sia un motore elettrico, alimentato da energia elettrica proveniente da delle batterie, anche in questo caso vi è la presenza di un inverter DC/AC. In particolare, i veicoli di tipo HEV includono i seguenti componenti principali:

- ICE: i veicoli ibridi più comuni sul mercato odierno sono alimentati da un motore a combustione interna. Comunque, il motore su un veicolo ibrido è più piccolo rispetto a quello di un veicolo tradizionale e sfrutta tecnologie avanzate per ridurre emissioni ed aumentare l'efficienza;
- Serbatoio: in un veicolo ibrido, è il congegno primario di stoccaggio di energia a bordo;
- Macchina elettrica: in un veicolo ibrido possono essere presenti una o più macchine elettriche. L'elettronica avanzata gli permette di fungere sia da motori che da generatori in modo da poter sia attingere all'energia di batteria per accelerare il veicolo sia recuperare energia dalla vettura per rifornire le batterie;

- Batterie: sono il congegno secondario di stoccaggio di energia a bordo;
- Trasmissione: la trasmissione è un componente chiave dei veicoli ibridi poiché consente di fornire la potenza dalle sorgenti alle ruote.

I veicoli di tipo HEV possono funzionare in differenti modalità:

- Puro elettrico: la propulsione è affidata completamente alla macchina elettrica la quale converte l'energia proveniente dalle batterie in energia meccanica; si utilizza nel caso in cui non vi è un'elevata richiesta di potenza e si ha un livello di carica della batteria sufficientemente elevato. Il motore termico non è funzionante in questa fase che è ottimale per tratti urbani.
- Puro termico: la propulsione è affidata completamente dal motore a combustione interna; solitamente si utilizza questa modalità nei tragitti extraurbani ed autostradali con richiesta continua di potenza.
- Power split: la propulsione è affidata in parte alla macchina termica ed in parte alla macchina elettrica, entrambe le macchine provvedono a fornire una parte di energia meccanica totale necessaria alla trazione del veicolo. La somma dell'energia meccanica proveniente dalla macchina termica e dalla macchina elettrica rappresenta l'energia totale necessaria alla trazione;
- Battery charging: la propulsione è affidata al motore a combustione interna il quale provvede anche a fornire energia meccanica in eccesso la quale viene utilizzata per la ricarica della batteria.

È chiaro che a differenza dei veicoli BEVs le emissioni saranno diverse da zero ma vi sono diversi aspetti positivi quali:

- Frenata rigenerativa: quando il veicolo sta frenando, l'energia cinetica può essere recuperata da un generatore ed immagazzinata a bordo;
- Idling reduction: in base alla taglia della fonte secondaria di potenza il motore può essere spento alle fermate o in condizioni di basse velocità;
- Miglioramento dell'efficienza dell'ICE: la fonte di potenza secondaria aiuta l'ICE impedendogli di lavorare in condizioni di inefficienza;
- ICE Downsizing/Downspeeding: grazie alla presenza della fonte di potenza secondaria si può scegliere un ICE più piccolo senza compromettere le performance.
- Eliminare o mitigare le perdite di frizione non impegnando il motore fino a che le velocità sono accoppiate;
- Gli accessori e gli ausiliari sono alimentati dal motore elettrico;
- In aggiunta a considerazioni energetiche, che sono la maggior motivazione per gli ibridi, possono essere ottenuti ulteriori benefici avendo più controllo sul punto operativo del motore e sui transitori. Questo influenza direttamente sia le emissioni che la guidabilità del veicolo.

Attualmente, i veicoli ibridi sono più costosi di quelli tradizionali, più pesanti a causa del peso addizionale del motore elettrico e del sistema di stoccaggio dell'energia. L'affidabilità di tali veicoli è ancora argomento di studio oggi e sono inoltre richiesti sistemi di controllo più complessi per ottimizzare il rendimento del carburante.

Plug-in HEVs (PHEVs)

Un veicolo ibrido Plug-In è molto simile ad un ibrido convenzionale ma ha la capacità di essere collegato via cavo per la ricarica della batteria. I veicoli ibridi plug-in sono progettati per distanze relativamente lunghe con poco o nessun utilizzo dell'ICE (all-electric range), questo significa inoltre che la taglia della

batteria è generalmente maggiore rispetto agli HEVs tradizionali. Inoltre, anche prima che la carica sia completamente terminata il motore può procurare potenza aggiuntiva per la ricarica della batteria o per l'accelerazione. La differenza principale tra i veicoli plug-in e i veicoli ibridi tradizionali è che quelli di tipo plug-in possono utilizzare il motore elettrico come fonte primaria di potenza ed utilizzare l'ICE per fornire potenza aggiuntiva.

Le caratteristiche appena descritte fanno comprendere come la tecnologia ibrida sia adatta a veicoli destinati alle aree urbane poiché è possibile ridurre fortemente i consumi di combustibile e le emissioni di CO₂. In figura 2.1 sono riportate le caratteristiche dei diversi tipi di veicoli a trazione elettrica finora descritti.

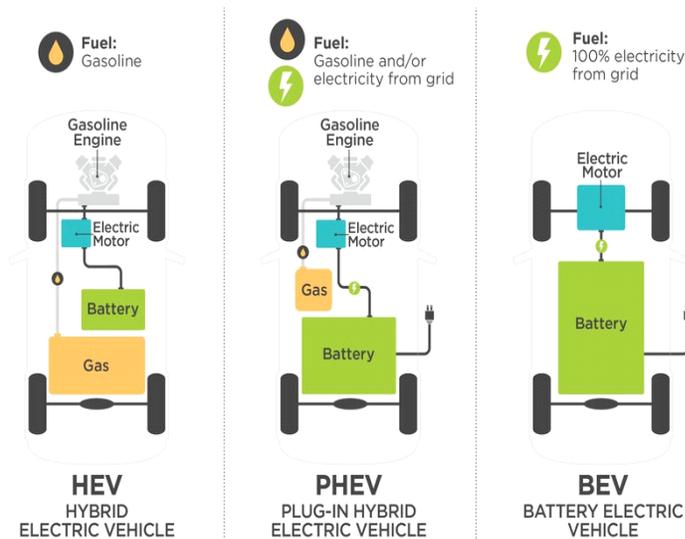


Figura 2.1 – Caratteristiche HEV, PHEV e BEV

HEVs: serie e parallelo

I veicoli HEVs appena descritti possono essere classificati in base al tipo di architettura utilizzata, si parla infatti di architettura serie e parallelo.

La configurazione tipica di un veicolo ibrido elettrico con architettura serie prevede che il motore elettrico per la trazione del veicolo sia alimentato da delle batterie e/o dal motore a combustione interna collegato ad un generatore. Il motore elettrico converte la potenza elettrica in input in potenza meccanica utile alla trazione del veicolo, viene infatti resa disponibile alle ruote tramite una catena di trasmissione. Il motore termico non è connesso meccanicamente alle ruote, il generatore provvede a convertire la potenza meccanica in potenza elettrica. Motore termico e generatore accoppiati sono anche utili a fornire potenza aggiuntiva al motore elettrico quando la domanda di potenza è elevata oppure quando la domanda di potenza è bassa sono utili alla ricarica della batteria. La trasmissione meccanica ha solitamente un rapporto di trasmissione fissato. È necessario ricordare che nell'architettura serie il motore termico rappresenta l'unica soluzione per la ricarica delle batterie. Bisogna inoltre tenere conto del fatto che le conversioni di potenza sono numerose, quindi l'efficienza totale di conversione è bassa nonostante

la macchina elettrica sia caratterizzata da alte efficienze. In questo tipo di architettura il motore ed il generatore sono solitamente macchine a corrente alternata quindi è necessario utilizzare un inverter DC/AC. Poiché gli inverter sono macchine elettriche reversibili, il convertitore di potenza può essere usato per avviare l'ICE siccome il convertitore di potenza associato al motore elettrico può essere usato per la frenata rigenerativa. Le performances del veicolo sono completamente determinate dalla taglia e dalle caratteristiche del motore elettrico. La taglia del motore elettrico viene determinata allo stesso modo rispetto ad un veicolo elettrico classico. Comunque, il controllo del sistema di trazione è differente da quello di un veicolo di tipo puro elettrico poiché viene aggiunta un'unità addizionale motore/generatore. In figura 2.2 è riportato uno schema essenziale che permette di comprendere la configurazione di un'architettura serie.

L'architettura serie permette di ottimizzare l'efficienza del motore. Poiché l'ICE non è direttamente connesso alle ruote, esso può funzionare al suo punto di maggiore efficienza e può essere spento quando non necessario. La necessità di convertire la potenza meccanica uscente dal motore termico in potenza elettrica produce significative inefficienze se si utilizza l'ICE come fonte primaria. Questo tipo di architettura sfrutta comunque un motore termico di piccola taglia ed una grande taglia della batteria, cioè permette di avere emissioni davvero basse ed eccellenti efficienze soprattutto in ambiente cittadino. C'è comunque da ricordare che questo tipo di architettura utilizza un sistema di stoccaggio di energia molto grande e pesante, l'efficienza cala quando si affrontano lunghe tratte ad alta velocità e la taglia dell'ICE deve comunque essere abbastanza grande da garantire le accelerazioni richieste.

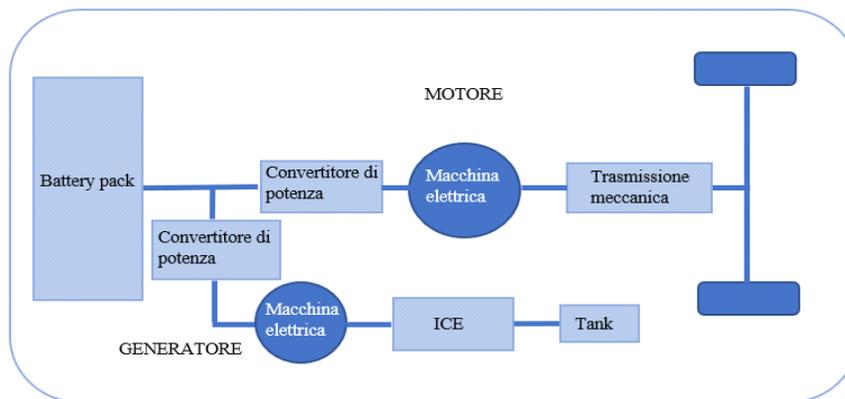


Figura 2.2 – Series HEV

La configurazione parallelo ha caratteristiche che permettono sia al motore termico sia a quello elettrico di fornire la loro energia meccanica direttamente alle ruote, entrambi i motori infatti sono connessi alle ruote tramite un sistema di trasmissione. I maggiori vantaggi dell'architettura parallelo rispetto a quella di tipo serie sono che innanzitutto non è richiesto un generatore, inoltre il motore elettrico è più piccolo ed infine non è necessario convertire più volte la potenza proveniente dal motore termico e diretta alle ruote. Per questi motivi l'efficienza totale può essere maggiore, d'altra parte però il controllo di questo tipo di architettura è più complesso rispetto a quello dell'architettura serie.

Esistono inoltre gli ibridi con architettura complessa che combina le caratteristiche dell'architettura parallelo con quella dell'architettura serie in modo da estrapolare da entrambe i benefici corrispondenti. In figura 2.3 è riportato uno schema per l'architettura in parallelo.

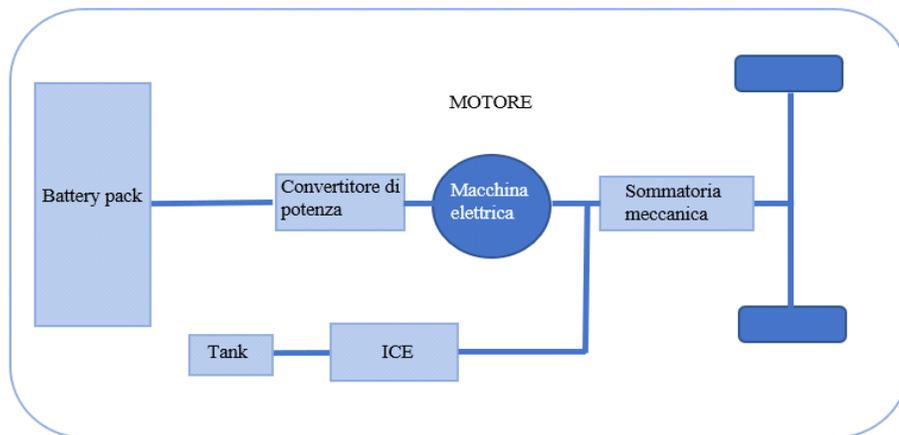


Figura 2.3 – Parallel HEV

Classificazione HEVs in base alla posizione della macchina elettrica

In base alla posizione assunta dalla macchina elettrica si possono classificare i veicoli HEV come segue.

P1: la macchina elettrica è sempre connessa direttamente al motore termico; In particolare esiste la configurazione P1f in cui si ha il generatore elettrico situato sul giro-cinghia del motore elettrico, la configurazione P1r invece è caratterizzata dalla macchina elettrica posizionata all'uscita dell'albero motore, in questo secondo caso si ottiene una frenata rigenerativa più efficiente;

P2: la macchina elettrica è posizionata in mezzo all'unità motore termico-trasmissione con la possibilità di disaccoppiarla dal motore termico tramite una frizione dedicata per cui è consentita una modalità di trazione pura elettrica.

P3: la macchina elettrica è posizionata tra la trasmissione e l'unità differenziale (double shaft);

P4: la macchina elettrica è posizionata sull'asse secondario mentre quella termica è posizionata sull'asse primario. È tipicamente collegata al differenziale tramite una trasmissione dedicata. (double drive). Si ha una elevata efficienza della frenata rigenerativa ed inoltre durante la fase di trazione è possibile recuperare l'energia in eccesso.

3. Codice HEVbox

Per poter svolgere il lavoro di questa tesi è stato utilizzato di un tool di ottimizzazione creato precedentemente. Il tool è stato scritto tramite l'ambiente di calcolo Matlab ed ha come obiettivo la simulazione e l'ottimizzazione di un veicolo ibrido lungo una missione prefissata. Il tool utilizzato è infatti un tool di ottimizzazione il quale permette di individuare la strategia di controllo ottimale che può essere adottata su un determinato ciclo per un determinato veicolo grazie alla minimizzazione di una funzione di costo. La funzione di costo da minimizzare riguarda le emissioni di CO₂ ttw.

3.1 Variabili utilizzate dal tool

Variabili di input

Le variabili in input fornite al tool di ottimizzazione riguardano le informazioni sulla missione che deve essere affrontata dal veicolo. Vengono quindi inserite come variabili iniziali:

- La velocità per ogni istante di tempo richiesta dalla missione;
- La pendenza della strada per ogni istante di tempo della missione;
- Il numero di marcia usato per ogni istante di tempo durante la missione (nel caso di un veicolo ibrido tale variabile non è un vincolo per cui l'utente è libero di scegliere il numero di marcia da utilizzare).

Variabili di controllo

Le variabili di controllo sono due:

- Numero di marcia
- Power flow

Il numero di marcia si riferisce al rapporto di trasmissione usato tramite il cambio mentre il power flow indica il modo in cui viene suddivisa la potenza richiesta per la propulsione del veicolo. Il tool considera quattro tipi di power flow:

- Modalità puro elettrico: la potenza richiesta per la propulsione del veicolo è completamente fornita dalla macchina elettrica.
- Modalità puro termico: la potenza richiesta per la propulsione del veicolo è completamente fornita dal motore termico.
- Power split: la potenza richiesta viene fornita in parte dalla macchina elettrica ed in parte dal motore termico.
- Battery charging: il motore termico produce la potenza richiesta ed inoltre produce potenza aggiuntiva impiegata per la ricarica delle batterie del veicolo.

Sub-variabili di controllo

Ad ogni variabile di controllo viene associata una sub-variabile di controllo che la quantifica matematicamente, per esempio per quanto riguarda il numero di marcia, la sub-variabile di controllo è il rapporto di trasmissione.

Variabili di stato

Tali variabili vanno a descrivere completamente lo stato del sistema considerato per ogni istante di tempo durante la missione considerata. Si considerano due variabili di stato:

- Engine state: ovvero lo stato del motore termico, che può essere acceso o spento.
- Battery state of charge: ovvero lo stato di carica della batteria che si riferisce al livello di carica delle batterie del veicolo ibrido per ogni specifico istante di tempo. Tale variabile si può modificare in modo continuo dal 40% all'80% della taglia della batteria per un veicolo ibrido non plug-in e dal 15% al 90% della taglia della batteria per un veicolo ibrido plug-in. Il codice richiede un'approssimazione di questa variazione continua tramite una griglia di discretizzazione che, se sufficientemente fine, fornisce risultati approssimati ma molto simili al risultato che si otterrebbe se fosse possibile utilizzare uno spazio continuo per la SOC window.

Esistono inoltre le variabili “score” e le variabili intermedie. Le prime sono quelle utilizzate per andare a costruire la funzione di costo da minimizzare mentre le seconde sono quelle che devono essere calcolate a partire dalle variabili di input e che sono utili per il calcolo delle score variables.

3.2 Approccio computazionale

Il codice HEVbox si basa su due metodi di calcolo fondamentali per poter simulare il veicolo ibrido lungo una missione:

- L'asse temporale viene discretizzato in intervalli, per cui, le variabili di controllo assumono valori costanti sul medesimo intervallo di tempo discreto. Le altre variabili invece sono descritte per un intervallo di tempo tramite due valori ad inizio e fine intervallo di tempo.
- Le variabili vengono calcolate tramite l'uso di matrici di configurazione.

Matrici di configurazione

Lo spazio delle configurazioni è definito come:

$$S_{Conf} = S_{GN} \times S_{PF} \times S_{ES} \quad (3.1)$$

Il numero di possibili configurazioni viene definito come:

$$N_{Conf} = N_{GN} \times N_{PF} \times N_{ES} \quad (3.2)$$

In cui N_{GN} è il numero di marcia per la trasmissione, N_{PF} è il numero di powerflow mentre N_{ES} rappresenta il numero degli stati del motore.

La missione è discretizzata nel tempo in intervalli per cui una matrice di configurazione associata ad una specifica variabile è una matrice che contiene il valore di quella specifica variabile per ogni intervallo di tempo per ogni configurazione possibile.

Per ogni variabile di input, intermedia e di output vi sono associate due matrici di configurazione:

- La prima contiene i valori delle variabili all'inizio di ogni intervallo di tempo per ogni configurazione;
- La seconda contiene i valori delle variabili alla fine di ogni intervallo di tempo per ogni configurazione.

Questo tipo di approccio è molto efficiente e permette di calcolare in una fase di pre-processing i valori di tutte le variabili per ogni istante di tempo per tutte le possibili combinazioni di controlli che possono essere scelte.

3.3 Step implementati nel codice

Con l'obiettivo di andare a simulare un veicolo ibrido lungo un ciclo predefinito, sono implementati nel tool di ottimizzazione una serie di step che verranno elencati in seguito.

Il primo step consiste nell'inserimento da parte dell'utente delle informazioni sul tipo di veicolo, sull'architettura, sul tipo di motore termico e sul range elettrico insieme ai parametri di design. Successivamente le informazioni fornite verranno elaborate al fine di generare mappe che descrivono il funzionamento dei differenti componenti per l'architettura ibrida, vengono specificate invece dall'utente le grandezze legate al cambio e all'inverter. Vengono inoltre acquisite le informazioni sul ciclo da eseguire come il profilo di velocità. A questo punto il tool acquisisce le informazioni sulle variabili di controllo che possono essere utilizzate durante la simulazione (numero di marcia, power flows). Per ogni variabile vengono create le matrici di configurazione come indicato al paragrafo 3.2 e vengono scritte le equazioni del modello del veicolo all'inizio e alla fine di ogni istante di tempo.

Lo step successivo consiste in un "feasibility check" su ogni macchina per richiesta di velocità e per richiesta di potenza mentre per le batterie il check viene eseguito solamente sulla potenza elettrica in input o in output. In particolare, per quanto riguarda il feasibility check sulla velocità è necessario ricordare che la velocità richiesta per ogni componente specifico dipende dalla velocità istantanea del veicolo ma anche dal rapporto di trasmissione del final drive e dal rapporto di trasmissione del cambio ed eventualmente dal rapporto di trasmissione dei dispositivi utilizzati per accoppiare il motore termico con le macchine elettriche. La velocità richiesta ad un componente viene calcolata e viene verificato che quest'ultima sia inferiore alla velocità massima concessa dal componente. La potenza richiesta da un componente dipende invece sia dalla velocità del veicolo sia dall'accelerazione del veicolo stesso insieme ad altri parametri quali area frontale del veicolo, coefficiente di attrito eccetera. Anche in questo caso si va a verificare che la potenza richiesta per ogni istante di tempo sia inferiore a quella erogabile dai diversi componenti. L'ultimo check effettuato è quello sulla batteria e va a verificare che la potenza elettrica gestita dalla batteria sia minore del valore massimo.

A questo punto viene eseguito l'ultimo step, ovvero la fase di ottimizzazione, eseguita grazie all'utilizzo dell'algoritmo Dynamic Programming, descritto nel dettaglio al capitolo successivo.

4. Dynamic Programming

L'algoritmo dynamic programming (DP) ha lo scopo di ottimizzare il design del veicolo e la strategia operativa. Questo metodo permette di ricavare una soluzione ottimale assoluta del sistema quando si è in presenza di un problema discretizzato su N stadi con un numero elevato di fasi decisionali. La DP ha l'obiettivo di minimizzare la *funzione di costo*, ovvero la risposta del sistema che è caratterizzato da uno *stato*, *decisioni* e da un *obiettivo*. La DP va a fornire una regola ottimale per ogni stadio cioè fornisce la decisione o il controllo da attuare per raggiungere l'obiettivo una volta che è stata stabilita la funzione obiettivo. Tale algoritmo si basa sul principio di ottimalità formulato da R.Bellman:

“An optimal policy has the property that whatever the initial state and initial decision are, the remaining decision must constitute an optimal policy with regard to the state resulting from the first decision”

Le caratteristiche principali del modello utilizzato sono due:

- Il sistema dinamico a tempo discreto;
- Funzione cost-to-go additiva nel tempo.

Sistema dinamico a tempo discreto

In un sistema dinamico lo stato del sistema stesso può evolversi nel tempo seguendo leggi fisiche. L'evoluzione nel tempo del sistema dinamico è caratterizzata da tre tipi differenti di variabili quali:

- Variabili di stato: tali variabili assumono un valore per ogni istante di tempo. I valori assunti potrebbero essere limitati ed inoltre il sistema viene descritto tramite un vettore m-dimensionale del tipo:

$$\bar{x} = \{x_1, x_2 \dots x_m\} \quad (4.1)$$

- Decisioni: i possibili controlli che possono essere applicati al sistema vengono rappresentati dalle decisioni per ogni istante di tempo. A seconda della decisione presa si avrà una differente evoluzione del sistema. In generale sono descritte da n variabili di controllo e vengono descritte tramite un vettore n-dimensionale del tipo:

$$\bar{u} = \{u_1, u_2 \dots u_n\} \quad (4.2)$$

- Variabili di disturbo o esterne: sono variabili esterne al sistema con cui interagiscono ma che non possono essere controllate. Sono descritte tramite un vettore a-dimensionale del tipo:

$$\bar{w} = \{w_1, w_2 \dots w_a\} \quad (4.3)$$

Il sistema dinamico a tempo discreto viene quindi definito come una funzione delle variabili di stato, delle decisioni e delle variabili esterne, tale funzione rappresenta la legge fisica che determina l'evoluzione nel tempo del sistema.

Funzione obiettivo

La funzione obiettivo viene chiamata funzione cost-to-go ed è la funzione che viene minimizzata sulla missione. Tale funzione è additiva, il che significa che il costo viene accumulato nel tempo. Viene definita una sub-funzione che riguarda la fattibilità dei controlli ed i limiti delle variabili di stato. Tale funzione può essere definita come:

$$J_{tot} = C_N(x_N) + \sum_{k=0}^{N-1} C_k(x_k, u_k, w_k) \quad \text{con } k = 0, 1, \dots, N-1 \quad (4.4)$$

In cui C_N indica il costo finale del processo e con C_k il costo istantaneo di transizione riferito al singolo istante di tempo.

L'ottimizzazione della funzione obiettivo viene risolta tramite due fasi: backward e forward. Durante la fase di backward la DP opera dall'ultimo stadio del sistema procedendo all'indietro fino a raggiungere lo stadio iniziale. La politica ottimale va a definire per ogni istante di tempo qual è la decisione migliore per ogni possibile stato del sistema che si potrebbe verificare in modo da minimizzare la funzione di costo. Durante la fase di forward invece la DP opera partendo dallo stadio iniziale e procedendo fino a raggiungere l'ultimo stadio. Per ogni stadio andrà ad attuare le politiche determinate in fase di backward. Il set di controlli che vanno a minimizzare la funzione di costo vengono restituiti come risultato della fase di forward.

La DP permette quindi il calcolo della strategia di controllo ottimale, in particolare permette di calcolare marcia e power flow che vanno a minimizzare la funzione di costo. Nella fase di backward viene determinata la sequenza dei numeri di marcia e del power flow per tutte le possibili combinazioni di engine state e battery state of charge per ogni intervallo di tempo considerato andando a minimizzare la funzione espressa dall'equazione 4.5

$$\psi^* = \begin{cases} \psi(k) + \varepsilon & \text{if } SOC_{min} < SOC(k) < SOC_{max} \wedge \varphi(k+1) \\ \text{infinite} & \end{cases} \quad (4.5)$$

In cui ε rappresenta le emissioni di CO₂ sull'intervallo di tempo. $SOC(k)$ rappresenta lo state of charge della batteria mentre φ considera se è possibile una combinazione di numero di marcia e powerflow per un determinato intervallo di tempo discreto.

A questo punto inizia la fase di forward che va ad applicare la strategia ottimale così ottenuta dall'istante di tempo iniziale fino all'istante di tempo finale. L'algoritmo infatti legge la combinazione ottimale del numero di marcia e power flow ed attua i controlli in modo da calcolare lo stato del motore e lo state of charge della batteria all'inizio dell'intervallo di tempo successivo.

Questo algoritmo permette quindi di definire la strategia di controllo ottimale che va a minimizzare le emissioni del veicolo su un determinato ciclo.

5. Cenni di Statistica

In questa tesi sono stati utilizzati diversi strumenti statistici per andare ad analizzare i risultati ottenuti tramite le simulazioni del codice HEVBox. Scopo di questo capitolo è di riassumere brevemente i fondamenti teorici a riguardo degli strumenti matematici utilizzati.

5.1 Distribuzione di Fisher, analisi della varianza e significato del *p-value*

Distribuzione di Fisher

R. A. Fisher studiò i vantaggi legati alla valutazione del rapporto delle varianze. In particolare, studiò una forma di test, legata al calcolo della varianza della popolazione, fatto in modi diversi che includono o escludono i fattori che si ritiene influenzino il risultato. Si prende in esame come esempio un caso semplice di una serie di misure fatte da operatori differenti: il comportamento degli operatori è rappresentato dal valore medio che essi ottengono su un insieme di misurazioni. Le differenze dei diversi valori medi sono dovute al fatto che l'azione di misura di ogni operatore ha un'influenza sistematica sul risultato di misura oppure corrispondono a differenze normali dovute al caso e quindi legate a fattori aleatori?

Rimane palese il fatto che la varianza calcolata sui dati prodotti da ogni singolo operatore non conterrà l'effetto della differenza tra gli operatori (varianza entro i gruppi) mentre quella calcolata usando i valori medi dei diversi operatori conterrà fortemente l'effetto della loro differenza (varianza tra i gruppi).

Il rapporto delle varianze calcolato inserendo a numeratore la varianza calcolata con la presenza del fattore in esame ed inserendo a denominatore quella in cui si è cercato di eliminare tutti i possibili fattori di disturbo può indicare se il fattore preso in esame è effettivamente significativo. Per trarre tale indicazione è necessario confrontare il risultato ottenuto con la naturale variabilità dei rapporti tra varianze descritta dalla distribuzione di Fisher riportata in formula 5.1:

$$F = \frac{\frac{s_1^2}{\sigma_1^2}}{\frac{s_2^2}{\sigma_2^2}} \quad (5.1)$$

La distribuzione di Fisher permette quindi di andare a determinare l'intervallo fiduciario per la stima dei rapporti delle varianze con un determinato rischio di errore. Tale proprietà viene utilizzata nell'analisi della varianza che consente di evidenziare la significatività dei fattori coinvolti.

*Analisi della varianza (ANOVA) e significato del *p-value**

Il problema da affrontare è se, prendendo i risultati sperimentali non globalmente ma organizzati in modo strutturato che li suddivida secondo il livello corrispondente al fattore che si vuole analizzare, si possa distinguere quando le differenze che si sono riscontrate siano dovute al caso oppure ad effetti sistematici.

Il metodo dei test d'ipotesi va a formulare un'ipotesi collegata al comportamento del fattore che si vuole analizzare per cui è necessario andare a stilare l'ipotesi di partenza o ipotesi nulla. L'analisi della varianza sfrutta il fatto che la varianza della popolazione da cui provengono i campioni può essere stimata come

varianza entro i gruppi e varianza tra i gruppi. Se tutti i campioni fossero tratti dalla stessa popolazione i due metodi usati per la stima dovrebbero dare approssimativamente lo stesso risultato. Si può concludere in situazioni del genere che probabilmente i campioni sono tratti da un'una popolazione unica mentre in caso contrario l'ipotesi nulla viene rifiutata.

L'analisi della varianza consiste quindi nel calcolo della statistica F. Se tale valore è elevato la variabilità tra le medie campionarie è maggiore di quella attesa sulla base della variabilità all'interno dei singoli campioni, pertanto l'ipotesi che tutti i campioni appartengano alla stessa popolazione deve essere rifiutata. È necessario però chiedersi quanto deve essere elevato il valore della statistica F per poter rifiutare l'ipotesi iniziale, per tale motivo si introduce il concetto di valore p.

Come detto precedentemente, quando si effettua un test di significatività statistica, si assume inizialmente l'ipotesi zero, secondo la quale non esiste differenza tra i gruppi riguardo al parametro considerato. Secondo tale ipotesi, i gruppi sono uguali e la differenza osservata va attribuita al caso per cui l'ipotesi zero può essere vera oppure falsa. Per decretare la verità o la falsità di tale ipotesi è necessario condurre un test statistico. Se il test consiglia di rifiutare l'ipotesi zero a differenza osservata viene dichiarata statisticamente significativa. Se invece il test consiglia di accettare l'ipotesi zero allora la differenza è statisticamente non significativa.

I risultati di un test statistico non hanno un valore di certezza assoluta ma esprimono una probabilità, pertanto la decisione di respingere l'ipotesi zero è probabilmente giusta ma potrebbe essere errata. Si introduce quindi il livello di significatività del test.

Il livello di significatività adottato è del 5% ed è espressa dal valore p. Il valore p è la probabilità di ottenere un risultato altrettanto estremo o più estremo di quello osservato se la diversità è interamente dovuta alla sola variabilità campionaria, assumendo quindi che l'ipotesi iniziale nulla sia vera. Un valore p che si avvicina a 0 testimonia una bassa probabilità che la differenza osservata possa essere ascritta al caso. Se $p < 0.05$ l'evidenza empirica è fortemente contraria all'ipotesi nulla che quindi non va rifiutata. In tal caso si dice che i dati osservati sono statisticamente significativi.

5.2 La regressione semplice

Andando ad osservare l'andamento globale dei dati è possibile identificare fattori sistematici, tale esame è fondamentale e spesso risolutivo poiché svela relazioni interne non evidenziate in altra maniera. La regressione ha un'ampia utilizzazione ogni qual volta si desidera di identificare la relazione funzionale tra una o più variabili indipendenti, spesso indicate con la lettera X ed una variabile dipendente spesso indicata con la lettera Y. Il modello di regressione lineare semplice viene espresso come indicato in equazione 5.2:

$$Y_i = \beta_0 + \beta_1 X_1 + \epsilon_i \quad (5.2)$$

In cui si ha:

β_0 = intercetta

β_1 = inclinazione

ϵ_i = errore casuale in Y corrispondente all'i-esima osservazione

Le operazioni da eseguire possono essere schematizzate in fasi successive:

- Presentazione grafica dei dati sperimentali o considerazioni sulla base dell'esperienza per fare una scelta dell'ipotesi di un modello matematico che rappresenti l'andamento osservato
- Calcolo dei parametri del modello matematico dei minimi quadrati
- Valutazione dei residui, cioè degli scarti tra valori sperimentali e valori calcolati con il modello matematico per decidere se il modello adottato è adeguato.

Mentre la seconda fase è descritta in modo preciso da un procedimento matematico correttamente utilizzabile in moltissimi casi, la prima e l'ultima fase sono affidate all'esperienza di chi analizza i dati sperimentali. È consigliabile andare ad utilizzare la rappresentazione grafica, che può indicare la forma dell'andamento generale e consente di valutare se essa è rappresentabile da una retta, da curve come parabole, esponenziali eccetera. È sicuramente necessario utilizzare tutte le informazioni sul fenomeno fisico descritto dai dati sperimentali in modo da scegliere ove possibile un modello matematico corrispondente eventualmente a leggi fisiche coinvolte. Il modello matematico adottato propone solo l'andamento della linea che descrive i dati e cioè costituito da un'equazione generica che contiene dei parametri da determinare per identificare la curva di quella famiglia che meglio si adatta ai dati sperimentali. Nel modello si individuano quindi la variabile dipendente, la variabile indipendente (o più variabili indipendenti). Nella seconda si va ad identificare i parametri, ovvero si utilizza il metodo dei minimi quadrati per andare a determinare il valore dei parametri. Tale metodo può essere applicato con facilità se la funzione matematica utilizzata è lineare nei parametri, da cui il nome di regressione lineare. Nella terza fase è necessario valutare se il modello scelto è adatto ai dati sperimentali. Se il modello matematico scelto corrisponde al fattore sistematico presente allora i residui, cioè la differenza tra i valori dei dati sperimentali e quelli corrispondenti calcolati con il modello matematico, sono dovuti solo ai fattori accidentali, quindi assumono una distribuzione normale. Se tale ipotesi non viene esclusa da un test di normalità allora si può accettare il modello scelto. L'andamento dei residui è un utile indicatore sulla bontà del modello adottato infatti i residui mantengono gli effetti accidentali presenti e si dispongono con andamento aleatorio. Per una valutazione più rigorosa della bontà del modello scelto si può fare un test di normalità.

Un modo di procedere adottato frequentemente per la sua semplicità consiste nella scelta di una funzione polinomiale come modello e aumentarne il grado per ottenere un migliore accordo con l'andamento sperimentale. Su tale modo di procedere bisogna:

- Tener conto del comportamento fisico di quanto si vuole descrivere. Il modello matematico deve essere adatto al fenomeno da descrivere. Il modello polinomiale è valido in modo ampio per fenomeni descritti con un opportuno sviluppo in serie di potenze.
- Un eccessivo aumento del grado del polinomio porta ad un inseguimento dei fattori accidentali.

5.3 La regressione multipla

Finora è stato preso in considerazione il modello di regressione lineare semplice, in cui una sola variabile indipendente o esplicativa X viene utilizzata per prevedere il valore della variabile dipendente o risposta Y . Spesso si può ottenere un modello migliore prendendo in considerazione più di una variabile esplicativa: si introduce quindi il modello di regressione multipla in cui si fa ricorso a più variabili esplicative per effettuare previsioni su una variabile dipendente.

Nel caso di p variabili esplicative il modello di regressione multipla assume la seguente espressione:

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \dots + \beta_n X_{ni} + \epsilon_i \quad (5.3)$$

In cui si hanno:

β_0 = intercetta

β_n = inclinazione di Y rispetto alla variabile X_n tenendo costanti le altre variabili esplicative

ϵ_i = errore in corrispondenza dell'osservazione i

È necessario verificare se esiste una relazione significativa tra la variabile dipendente e l'insieme delle variabili esplicative. Dal momento che siamo in presenza di più di una variabile esplicativa, l'ipotesi nulla va specificata nella seguente maniera:

H_0 : Non vi è una relazione significativa tra la variabile dipendente e le variabili esplicative (cioè i coefficienti dell'equazione sono uguali a zero). Tale problema di verifica di ipotesi viene risolto ricorrendo al test F .

Test F sull'intero modello nel modello di regressione multipla

La statistica F è data dal rapporto tra la media dei quadrati della regressione (MQR) e la media dei quadrati dell'errore (MQE). La formula è riportata nell'equazione 5.4

$$F = \frac{MQR}{MQE} \quad (5.4)$$

F è la statistica test F avente una distribuzione F con p e $n-p-1$ gradi di libertà indicando con p il numero di variabili esplicative nel modello di regressione.

La regola decisionale in questo caso è

- Rifiutare H_0 se $F > F_U$ in cui F_U è il valore critico sulla coda di destra di una distribuzione F con p e $n-p-1$ gradi di libertà, altrimenti accettare H_0 .

Se il livello di significatività scelto, o p -value, è 0.05 si ricava il valore critico di F . Per cui, se il p -value è inferiore a 0.05 si può rifiutare H_0 e dedurre che esiste una relazione tra almeno una variabile esplicativa la variabile dipendente.

Test t per la verifica di ipotesi sull'inclinazione nel modello di regressione multipla

Viene calcolata la statistica t come indicato nell'equazione 5.5:

$$t = \frac{b_k}{S_{b_k}} \quad (5.5)$$

In cui p è il numero di variabili esplicative, b_k è l'inclinazione di Y rispetto alla variabile k tenendo costante le altre variabili, S_{b_k} è l'errore standard del coefficiente di regressione b_k e t è la statistica test con distribuzione t con n-p-1 gradi di libertà. Pertanto se vogliamo stabilire se una variabile ha un effetto significativo sulla variabile Y si formula l'ipotesi nulla in tal modo:

H₀: il singolo coefficiente è pari a zero.

Si calcola quindi la statistica t, se il p-value corrispondente è inferiore di 0.05 si rifiuta l'ipotesi nulla e si può concludere che vi è una relazione significativa tra la variabile considerata e la variabile dipendente.

Riassumendo

Esistono due tipi di verifica d'ipotesi d'interesse immediato:

- Un test di significatività complessiva della regressione:
H₀: tutti i coefficienti sono uguali a zero.
Il test più comune per ipotesi di questo tipo è un test F.
- Un test sulla significatività di ciascun singolo coefficiente:
H₀: il singolo coefficiente è pari a zero
Il test più comune per ipotesi di questo tipo è il test t.

Il coefficiente di determinazione

Il coefficiente di determinazione consente di valutare la bontà del modello di regressione stimato. Nel modello di regressione multipla, dal momento che si è in presenza di almeno due variabili esplicative, il coefficiente di determinazione rappresenta la proporzione di variabilità della Y spiegata dalle variabili esplicative. Come per la regressione semplice, viene calcolato R^2 per stimare la proporzione di varianza della variabile risposta spiegata dalla regressione tramite l'espressione riportata nell'equazione 5.6:

$$R^2 = \frac{\text{somma dei quadrati dei residui}}{\text{somma dei quadrati totale}} \quad (5.6)$$

Tuttavia, alcuni ricercatori ritengono che quando si ricorre ad un modello di regressione multipla sia opportuno fare uso di un indice che tenga conto anche del numero di variabili esplicative incluse nel modello e dell'ampiezza del campione. Si calcola per cui R^2 corretto come indicato all'equazione 5.7:

$$R^2 = 1 - \frac{n-1}{n-p-1} (1 - R^2) \quad (5.7)$$

In cui p è il numero di variabili esplicative incluse nel modello.

R^2 non diminuisce mai quando una nuova variabile viene aggiunta al modello, ciò può essere uno svantaggio se si desidera confrontare modelli con un numero di variabili esplicative diverso. Per tale motivo si va ad analizzare corretto che mostra la proporzione di variabilità di Y spiegata da tutte le variabili indipendenti X , corretta per il numero di variabili utilizzate, per cui, aggiungendo una variabile nel modello corretto può anche diminuire.

Riassumendo:

R^2 ed corretto spiegano se le variabili esplicative sono idonee a prevedere i valori della variabile dipendente. Non esprimono se

- Una variabile inclusa è statisticamente significativa
- Il modello è ben specificato
- Il gruppo di variabili esplicative è più appropriato

5.4 Tecniche statistiche per il machine learning

La tecnica Cross validation

La tecnica cross validation viene utilizzata in applicazioni di machine learning in modo da stimare la bontà di un modello costruito. Questa tecnica permette di decidere se il modello è in generale performante per prendere decisioni su dati non utilizzati per la costruzione del modello stesso. La cross validation (o validazione incrociata) viene inoltre utilizzata per eliminare il problema dell'overfitting nei training set ed è detta anche k-fold validation. La convalida incrociata prevede di suddividere il training set in k parti di uguale dimensione, una delle k parti viene utilizzata come validation set mentre le restanti parti vengono utilizzate come training dataset. La procedura viene ripetuta per ognuna delle k parti del training set.

Algoritmo Stepwise

La regressione stepwise è infatti un metodo di regressione automatico, consigliabile in studi esplorativi in presenza di molteplici variabili per la stima di modelli lineari multivariati. Tale tipo di regressione nasce dalla necessità di selezionare un sottoinsieme ottimo tra un numero di variabili esplicative per la costruzione di un modello efficiente. Qualora si debba effettivamente andare a scegliere un modello di regressione per un insieme limitato di variabili predittive si gestisce manualmente l'operazione andando a considerare tutte le possibili combinazioni di variabili esplicative. Se il numero delle variabili esplicative aumenta, il metodo stepwise è decisamente più efficiente per la risoluzione di questo tipo di problema.

Quando si ha un grande numero di predittori non è facile determinare il subset migliore. La selezione del subset migliore può infatti riscontrare problemi statistici quando si ha un alto numero di predittori. Più è grande il numero di variabili, maggiori sono le possibilità di ottenere modelli che sembrano adattarsi al training dataset ma nonostante ciò questi possono non avere significatività predittiva su dati futuri a causa

di problemi di overfitting ed alta varianza dei coefficienti stimati. Per questi motivi i metodi stepwise sono alternative ottimali per la selezione del subset di variabili ottimale.

Forward Stepwise selection

La selezione forward stepwise è un algoritmo efficiente che permette di selezionare il migliore subset di predittori. Nonostante la migliore procedura di selezione dei predittori considera tutti i possibili modelli contenenti i subset dei p predittori, la selezione stepwise forward considera un set inferiore di modelli. In particolare, ad ogni step la variabile che fornisce il più grande miglioramento addizionale al modello viene aggiunta. L'algoritmo funziona come segue:

- Sia M_0 il modello nullo, non contenente predittori
- Per $k=0, \dots, p-1$:
 - a) Si considerano tutti i $p-k$ modelli che aumentano i predittori in M_k con un predittore addizionale
 - b) Si sceglie il migliore tra questi $p-k$ modelli. Il migliore viene determinato tramite il valore maggiore di R^2 .
- 3. Si seleziona un singolo modello tra i modelli M_0, \dots, M_p utilizzando il corretto

6. Andamento dell'errore relativo di CO2 in funzione dello step energetico

Come descritto nei capitoli precedenti per lo svolgimento di questo lavoro è stato utilizzato un tool di ottimizzazione in grado di simulare un veicolo ibrido. Al capitolo 3 è stato illustrato l'utilizzo da parte di questo tool di variabili ottenute andando a discretizzare spazi continui come l'asse temporale e SOC window della batteria. Andando a discretizzare uno spazio continuo si va ad effettuare un'approssimazione e processando successivamente le variabili approssimate si può commettere un errore più o meno grande. La discretizzazione è purtroppo un passaggio inevitabile per poter eseguire l'algoritmo ma come detto va ad introdurre un'approssimazione e quindi degli errori di calcolo. Da un lato, più è grossolana la griglia di discretizzazione, maggiore è l'errore numerico che si va a commettere tramite la simulazione, d'altra parte invece una griglia di discretizzazione molto fine permette di ottenere risultati affetti da errori numerici molto piccoli ma va ad aumentare drasticamente i tempi di calcolo. Fortunatamente si è notata una tendenza asintotica del risultato numerico: infittendo infatti la griglia di discretizzazione l'errore numerico va a diminuire progressivamente. La discretizzazione della batteria è quindi un parametro che va ad influire pesantemente sul risultato numerico finale: il numero di livelli di SOC è infatti il numero di punti della griglia utilizzata per andare a discretizzare il range di variazione dello state of charge della batteria. L'obiettivo principale di questo lavoro è cercare correlazioni tra l'errore numerico e la griglia di discretizzazione utilizzata.

Discretizzazione della batteria

La discretizzazione della batteria va ad approssimare uno spazio continuo tramite una griglia formata da più punti. Viene definito h lo step energetico di discretizzazione scelto, il quale va a dividere la SOC window in N intervalli discreti di uguale dimensione. L'errore prodotto da tale discretizzazione si verifica nella fase di ottimizzazione tramite l'algoritmo dynamic programming e dipende dallo step energetico e quindi, in ultima analisi, dal numero di livelli di SOC scelto.

In questo capitolo vengono svolte più simulazioni utilizzando il tool di ottimizzazione andando a variare la griglia di discretizzazione della batteria in modo da verificare l'errore relativo corrispondente ad ogni livello di discretizzazione. L'obiettivo finale è quello di andare a relazionare l'errore relativo alla griglia di discretizzazione scelta. La prima analisi è stata svolta considerando un veicolo ibrido con architettura p2 ed è stato scelto il ciclo guida WHVC.

Ciclo guida utilizzato

In questa prima fase il ciclo guida utilizzato è il ciclo WHVC, riportato in figura 6.1. Il ciclo guida WHVC è stato studiato per la raccolta di dati a riguardo del modello di guida reale per i veicoli heavy duty. Il test viene effettuato tramite un banco dinamometrico con l'obiettivo di confrontare le emissioni dei veicoli. Il test per il ciclo WHVC dura 1800 secondi ed è suddiviso in tre differenti tratti che rappresentano diverse condizioni per la guida: tratto urbano, tratto rurale e tratto autostradale, in particolare si definiscono i tratti in base a durata, velocità massima e velocità media:

- Tratto urbano: tratto di durata 900 secondi con velocità media 21,3 km/h, velocità massima di 66,2 km/h;
- Tratto rurale: tratto di durata 481 secondi, velocità media 43,6 km/h, velocità massima 75,9 km/h;
- Tratto autostradale: tratto di durata 419 secondi, velocità media 76,7 km/h, velocità massima di 87,8 km/h.

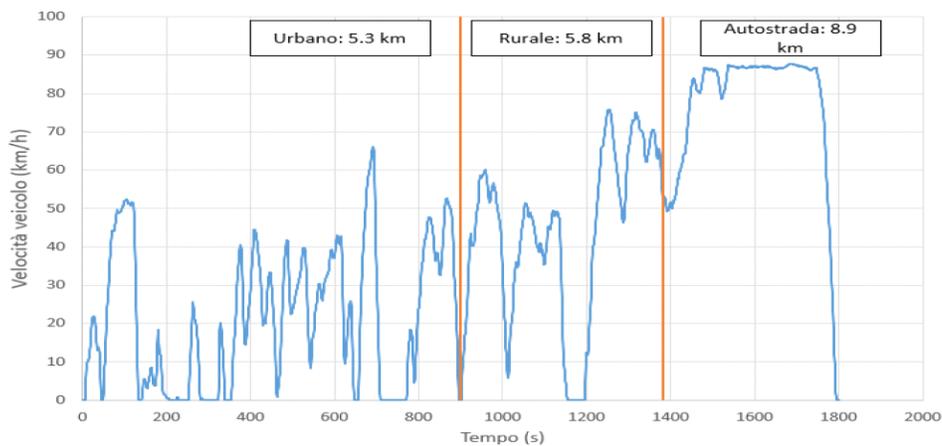


Figura 6.1 – Ciclo guida WHVC

6.1 Raccolta dei dati tramite simulazioni

La fase iniziale di questo lavoro consiste nella raccolta dei dati tramite simulazioni. È stato utilizzato il codice HEVBox scritto per l'ambiente matlab per raccogliere i dati necessari. In questa fase preliminare infatti l'obiettivo è quello di andare a definire ed analizzare per l'architettura p2 una serie di layout differenti in cui si vanno a modificare i parametri di design quali:

- Engine Displacement: cilindrata del motore termico
- PE Ratio: rapporto tra la potenza massima della macchina elettrica e l'energia della batteria
- Em1Power: taglia della macchina elettrica
- Em1SpRatio ed FDSpRatio: rapporti di trasmissione del riduttore e del final drive

- Crate Dis max: rapporto tra corrente massima capacità nominale durante la fase di scarica
- Crate Char max: rapporto tra la corrente massima e capacità nominale durante la fase di carica

In tabella 6.1 sono riportati i layout analizzati.

Tabella 6.1 – Layout analizzati

	EngDisp [l]	PEratio [-]	EM1Power [kW]	EM1SpRatio [-]	FDSpRatio [-]	CrateDis max [-]	CrateChar max [-]
1	2,997	4,4434	111,53	3,4116	3,4971	6,1699	7,5957
2	3,1563	5,3003	107,79	3,2535	2,9016	9,2681	3,9995
3	2,8626	5,3069	115,76	3,6479	4,684	11,835	10,656
4	2,4879	5,52	125,73	2,3768	2,6189	8,189	3,9556
5	3,2583	2,1582	105,38	3,2681	3,585	11,15	9,0996
6	3,5714	8,7368	102,15	5,4974	2,6335	11,91	8,2622
7	3,9025	3,3271	93,384	2,3025	3,0093	10,218	4,4658
8	2,4023	7,0823	133,19	2,9472	3,1869	8,6614	3,8494
9	3,8563	6,7549	101,38	3,1741	3,3169	6,2627	9,3389
10	3,4814	9,207	111,74	4,0439	4,7949	11,941	9,3633
11	3,7975	5,0718	103,48	3,7013	3,6721	10,176	3,4038
12	3,5362	2,7866	110,46	2,9835	3,385	5,1958	2,9937
13	3,3427	9,0137	115,63	4,437	2,5244	7,8496	8,7285
14	4,1729	10,191	93,691	3,3057	2,334	5,3008	8,8164
15	4,2362	7,3701	95,435	4,3464	2,5698	4,0654	9,9541
16	4,2133	3,6194	96,075	3,101	3,2777	10,898	7,8435
17	4,0827	2,512	99,562	4,7109	2,2845	8,3879	3,9666
18	3,603	8,9697	94,409	3,509	3,604	11,282	7,0537
19	3,5106	2,1846	97,28	3,762	2,4146	7,708	7,9717
20	3,3799	2,5537	100,77	4,2849	4,6382	11,741	7,3174

Per ogni layout sono state svolte le simulazioni per differenti livelli di SOC, inizialmente uguali per tutti i layout, al fine di ricavare gli andamenti di CO₂ e di consumo di combustibile. Il livello di SOC dell'ultima simulazione è stato calcolato tramite i seguenti passaggi: in primis si è calcolata la taglia della batteria per ogni layout tramite la 6.1:

$$E_{batt} = \frac{EM1\ Power}{PE\ ratio} \quad [kWh] \quad (6.1)$$

Sono stati definiti gli step energetici h i quali assumono valori da 0,1 a 0,9 e di conseguenza sono stati calcolati i livelli di SOC corrispondenti tramite la 6.2:

$$N\ SOC_l = \frac{0,4\ E_{batt}}{h} - 1 \quad (6.2)$$

In tal modo è possibile ottenere il numero di livelli di SOC a cui corrisponde uno step energetico inferiore a 0,2. È stata quindi svolta l'ultima simulazione per ogni layout con il numero di SOC ottenuto. Si ottengono così i valori di CO₂ ttw (tank to wheel) e di fc (fuel consumption) al variare dei livelli di SOC. A questo punto è possibile calcolare l'errore relativo di CO₂ e il nuovo step energetico come indicato dalla 6.3:

$$Step\ energetico = \frac{0,4\ E_{batt}}{N\ SOC - 1} \quad (6.3)$$

L'errore relativo di CO₂ viene calcolato come indicato invece dalla 6.4:

$$ErrRelCO_2 = \frac{CO_{2MAX} - CO_2}{CO_{2MAX}} * 100 \quad (6.4)$$

La tabella 6.2 riassume i valori ottenuti per il layout 1 a titolo di esempio.

<i>Tabella 6.2 – Layout 1</i>				
SOC	CO2 ttw [g/km]	fc	Relative error [%]	Step Energetico
251	10000	0	-2518,657938	40,16023766
501	207,6637	1,3867	45,61998036	20,08011883
751	253,6461	1,6938	33,57876268	13,38674589
1001	295,1794	1,9711	22,70261211	10,04005941
1501	329,0022	2,197	13,84557774	6,693372943
2001	333,3522	2,226	12,70646154	5,020029707

3001	359,2523	2,399	5,924111293	3,346686471
4001	364,1852	2,4319	4,632353519	2,510014853
5001	368,7027	2,4621	3,449374795	2,008011883
6001	371,5802	2,4813	2,695855974	1,673343236
7001	375,5402	2,5077	1,65886743	1,434294202
8001	376,0378	2,511	1,528563011	1,255007427
10001	377,5878	2,5214	1,122671031	1,004005941
12001	378,9831	2,5307	0,757289689	0,836671618
15001	379,625	2,535	0,589198036	0,669337294
60001	381,875	2,55	0	0,167334324

Per ogni layout vengono così ottenuti gli andamenti di CO₂ ttw in funzione del numero di livelli di SOC utilizzati, l'andamento di fc in funzione anche in questo caso del numero di livelli di SOC ed infine sono stati ricavati gli andamenti dell'errore relativo di CO₂ in funzione dello step energetico.

Vengono riportati in figura 6.2 a titolo di esempio gli andamenti ottenuti per il layout 1.

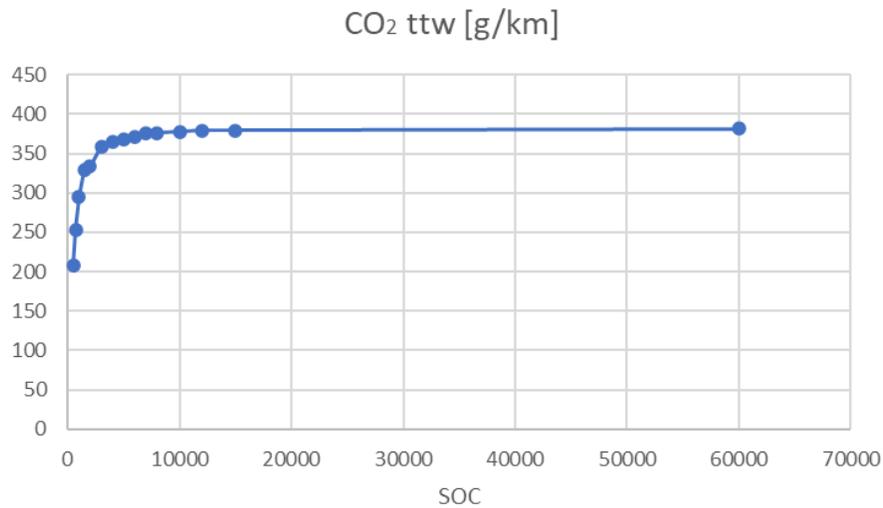


Figura 6.2 – Andamenti ottenuti per il layout 1

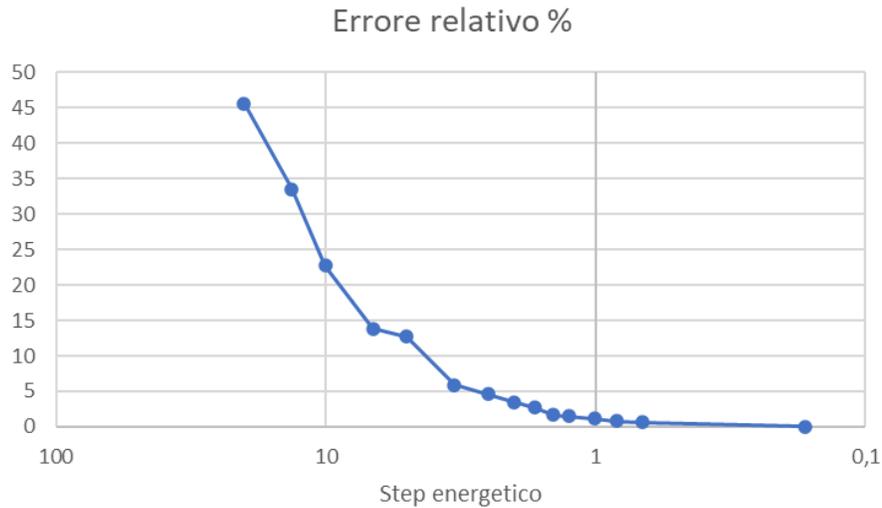
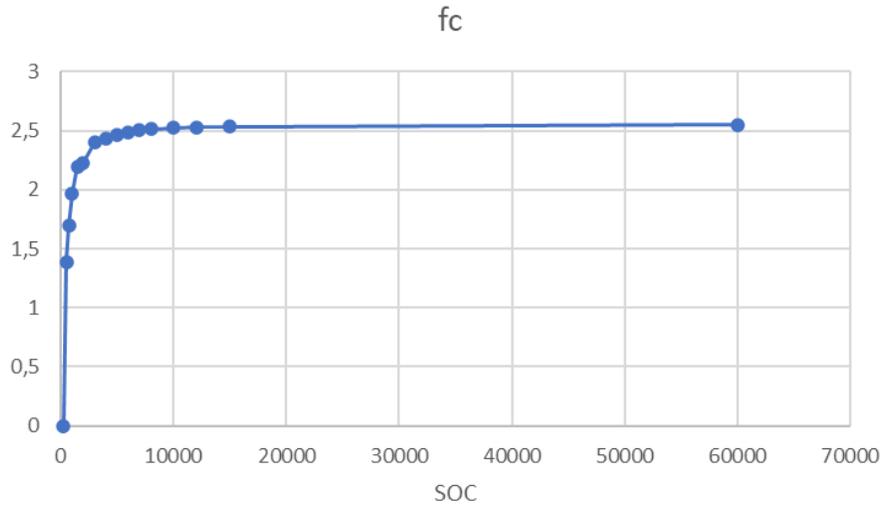


Figura 6.2 – Andamenti ottenuti per il layout 1

6.2 Fitting degli andamenti ottenuti con polinomio di secondo grado

L'obiettivo di questa seconda fase è quello di individuare i coefficienti di un polinomio in grado di riprodurre l'andamento dei dati ottenuti tramite le simulazioni. In particolare si vuole riprodurre l'andamento dell'errore relativo in funzione dello step energetico. Si ipotizza di riprodurlo tramite un polinomio di grado n riportato nell'equazione 6.5:

$$y(x) = p_1x^n + p_2x^{n-1} + \dots + p_nx + p_{n+1} \quad (6.5)$$

in cui y rappresenta l'errore relativo mentre x rappresenta lo step energetico. Per fare questa operazione è stato creato uno script matlab che esegue le seguenti operazioni:

- Importazione dei dati da excel
- Valutazione dei coefficienti p_1, p_{n+1}
- Plot dei dati ottenuti da simulazione HEVbox confrontati con la funzione ottenuta tramite fitting
- Salvataggio dei coefficienti

L'analisi è stata svolta inizialmente considerando un secondo grado per il polinomio per cui è stata utilizzata l'equazione 6.6:

$$y(x) = p_1x^2 + p_2x + p_3 \quad (6.6)$$

Viene riportato in figura 6.3 l'andamento ottenuto per il layout 1.

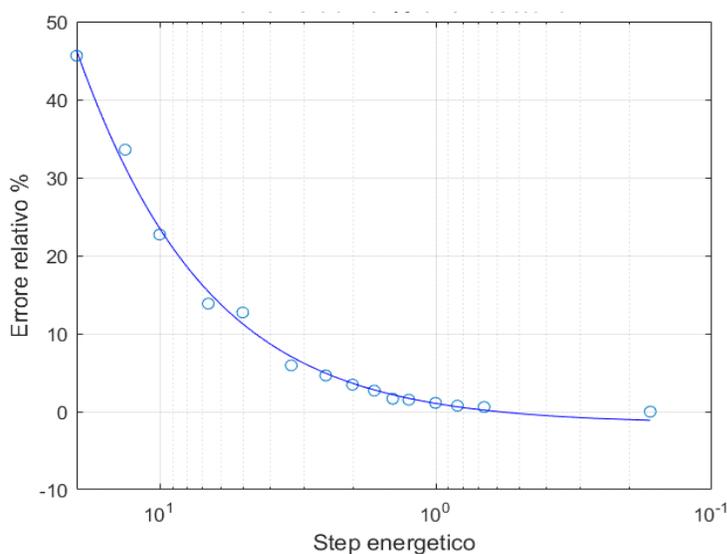


Figura 6.3 – Errore Relativo % CO2 ttw – layout 1

In tabella 6.3 sono riassunti i valori dei coefficienti ottenuti per i 20 layout analizzati in queste fasi.

<i>Tabella 6.3 – Coefficienti ottenuti per il polinomio di secondo grado</i>			
Layout	p_1	p_2	p_3
1	-0,0112	2,6064	-1,5485
2	0,0019	2,4281	-1,1102
3	0,0006	2,3056	-1,2046
4	-0,0096	2,5810	-1,3659

5	-0,0190	2,7632	-1,9513
6	0,0020	2,1687	-1,0871
7	-0,0172	2,6818	-1,6117
8	0,0228	2,1871	-0,7082
9	-0,0163	2,5558	-1,4257
10	0,0046	1,9617	-1,0707
11	-0,0153	2,5605	-1,6562
12	-0,0142	2,4707	-1,4668
13	-0,0057	2,5329	-1,2198
14	-0,0123	2,3389	-0,9940
15	-0,0134	2,3858	-1,2517
16	-0,0170	2,6759	-1,7822
17	-0,0187	2,7578	-2,2184
18	0,0025	2,2138	-0,9461
19	-0,0195	2,7826	-1,9732
20	-0,0150	2,5197	-2,3361

In tabella 6.4 sono riassunti i valori massimi e minimi dei coefficienti calcolati in modo da poterne determinare il range di variabilità.

Tabella 6.4 – Range di variabilità dei coefficienti di secondo grado

Coefficiente	p_1	p_2	p_3
<i>minimo</i>	-0,0195	1,9617	-2,3361
<i>massimo</i>	0,0228	2,7826	-0,7082

Finora sono state considerate tutte le architetture analizzate per la determinazione dei tre coefficienti. In particolare sono stati presi in considerazione anche i layout 5 e 19 per i quali si utilizza (per motivi prettamente pratici) uno step energetico finale differente rispetto agli altri layout. Con l'obiettivo di determinare se esiste un'unica equazione polinomiale che rappresenti l'andamento generale per l'architettura p2 sono stati esclusi dall'analisi i layout 5 e 19. In tabella 6.5 sono riportati i range di variabilità dopo l'esclusione dei layout 5 e 19, vengono inoltre riportati i valori medi.

<i>Tabella 6.5 – Range di variabilità dei coefficienti di secondo grado</i>			
<i>coefficienti</i>	<i>p₁</i>	<i>p₂</i>	<i>p₃</i>
<i>minimo</i>	-0,0195	1,9617	-2,2184
<i>massimo</i>	0,0228	2,7826	-0,7082
<i>media</i>	-0,0076	2,4443	-1,3689
<i>range</i>	0,0042	0,8209	1,5102

In figura 6.4 è riportato il grafico che mette in relazione il valore dei coefficienti ricavati con il grado della variabile a cui sono associati.

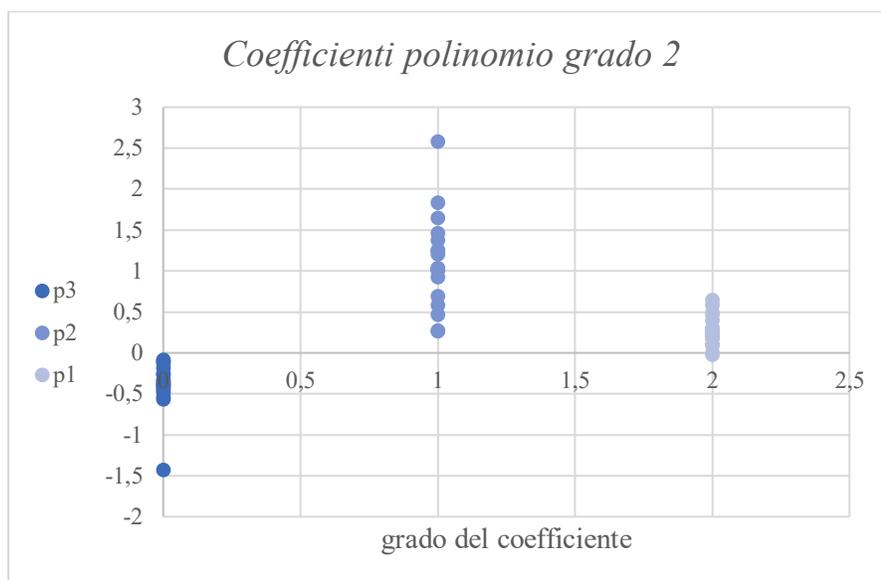


Figura 6.4 – Coefficienti dei polinomi di secondo grado

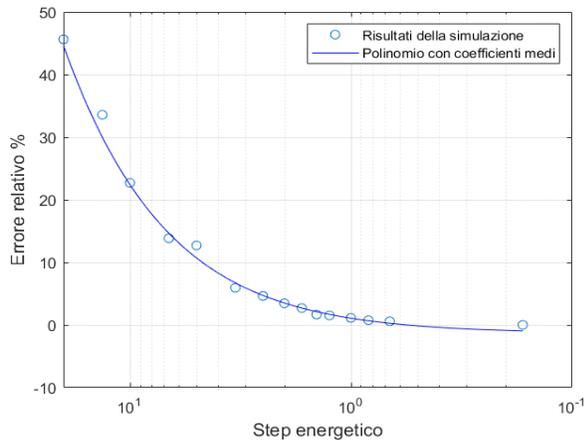
In prima analisi si può osservare come il coefficiente di secondo grado abbia un range di variazione ridotto rispetto ai range degli altri coefficienti.

6.3 Confronto dei dati con il un polinomio dai coefficienti medi

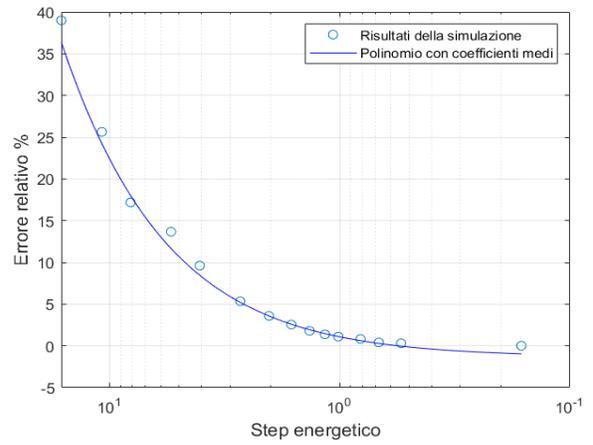
Con l’obiettivo di ottenere un’unica equazione che possa essere adatta per ogni layout per l’architettura p2, si è tentato di confrontare l’andamento del polinomio di secondo grado considerando i coefficienti medi con i dati ottenuti dalle simulazioni per poter capire se questi ultimi vengono ben rappresentati. Si è considerato quindi come polinomio medio un polinomio del tipo indicato nell’equazione 6.7:

$$y(x) = p_{1\ med}x^2 + p_{2\ med}x + p_{3\ med} \quad (6.7)$$

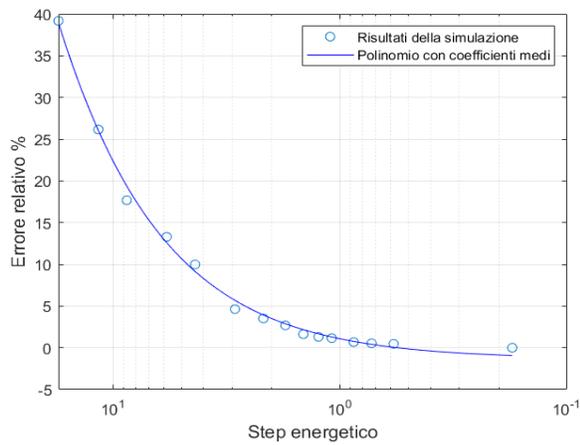
Sono stati calcolati due tipi differenti di coefficienti medi utilizzando prima la media aritmetica e successivamente utilizzando la root mean square. In figura 5.5 sono riportati i confronti tra i risultati delle simulazioni e il polinomio medio calcolato con media aritmetica insieme ai rispettivi valori di R².



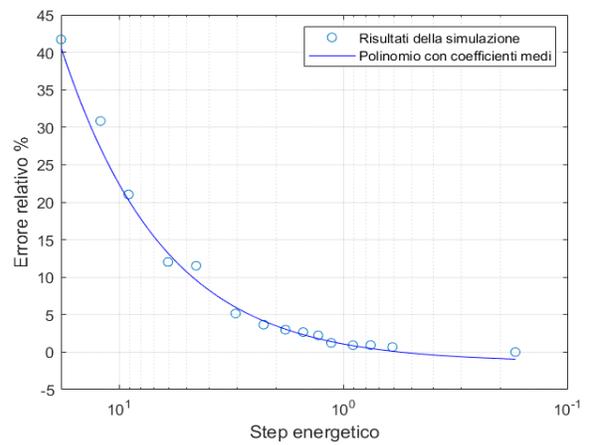
$R^2 = 0,9948$ – Layout 1



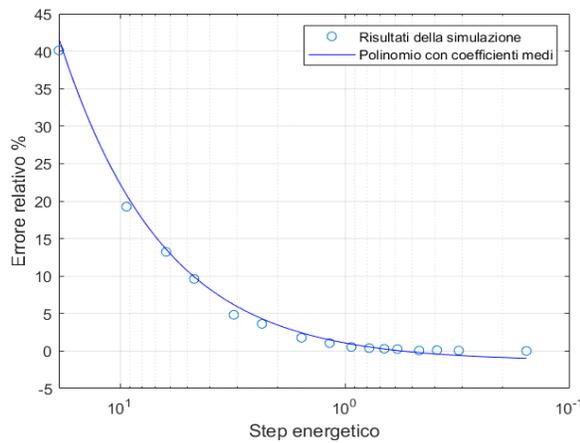
$R^2 = 0,9960$ – Layout 2



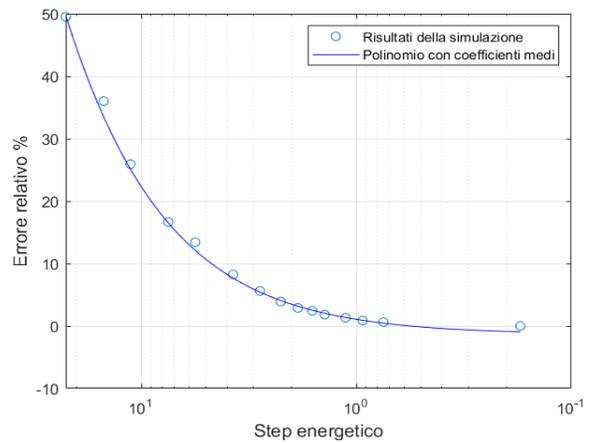
$R^2 = 0,9963$ – Layout 3



$R^2 = 0,9932$ -Layout 4

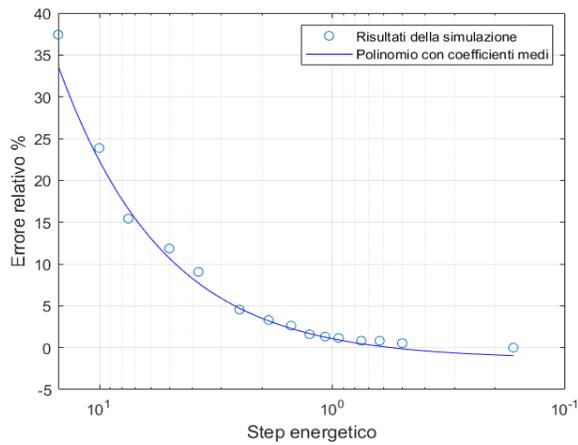


$R^2 = 0,9976$ – Layout 6

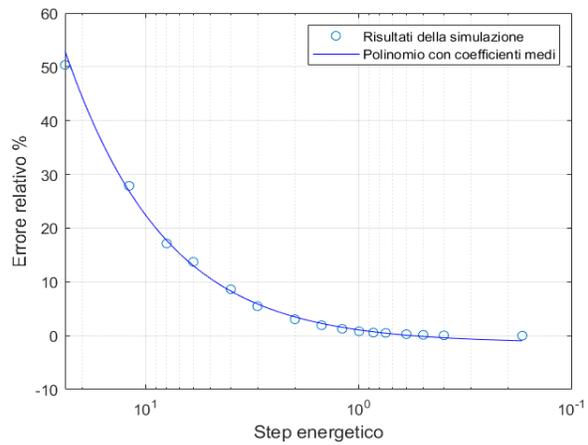


$R^2 = 0,9977$ – Layout 7

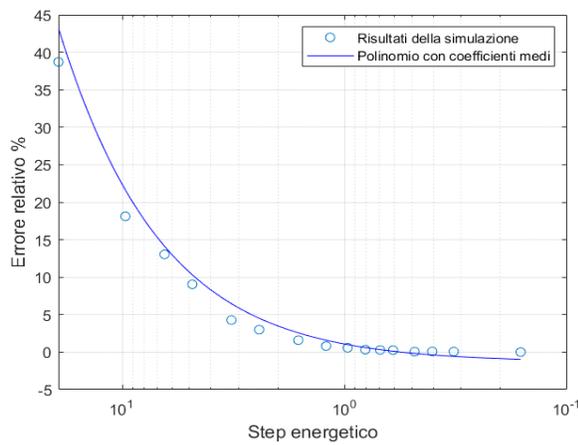
Figura 6.5 – Confronto risultati delle simulazioni con polinomio dai coefficienti medi



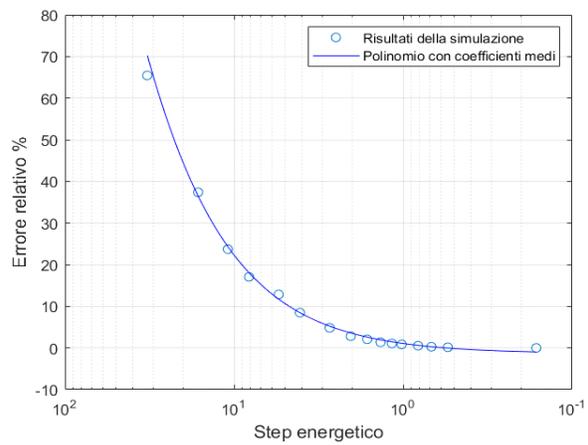
$R^2 = 0,9937$ – Layout 8



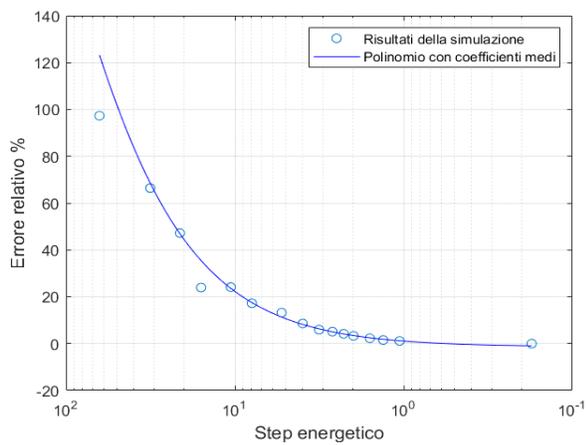
$R^2 = 0,9976$ – Layout 9



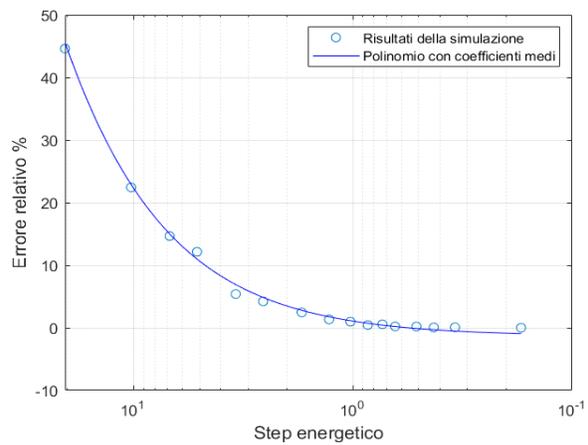
$R^2 = 0,9958$ – Layout 10



$R^2 = 0,9971$ – Layout 11

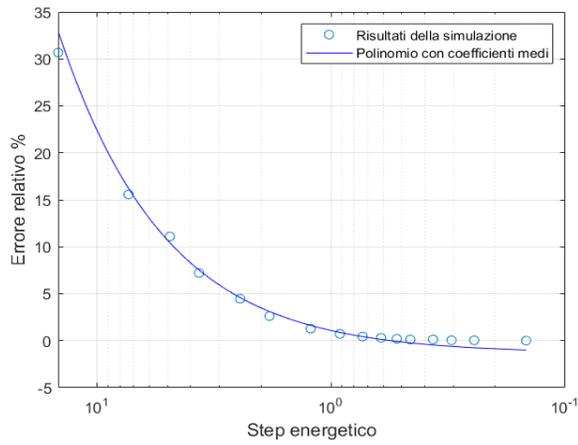


$R^2 = 0,9803$ – Layout 12

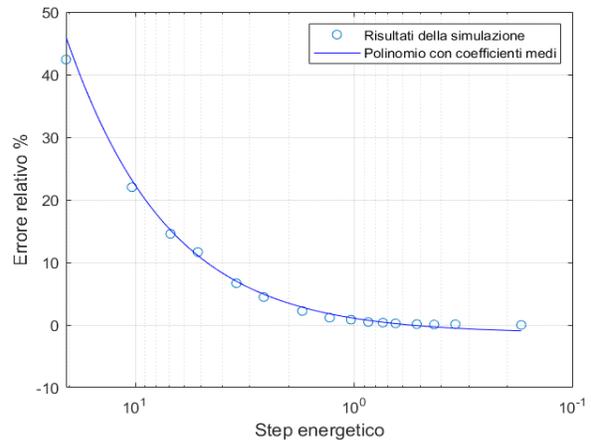


$R^2 = 0,9973$ – Layout 13

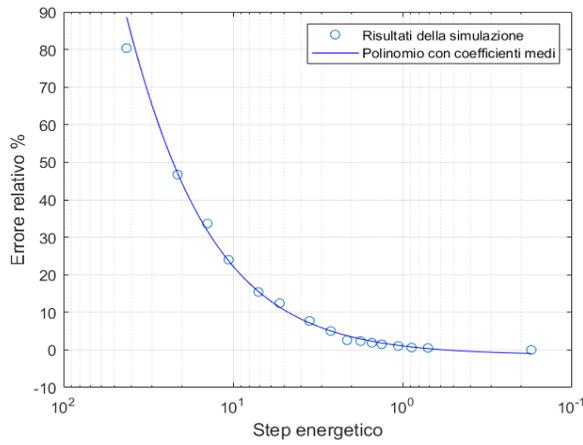
Figura 6.5 – Confronto risultati delle simulazioni con polinomio dai coefficienti medi



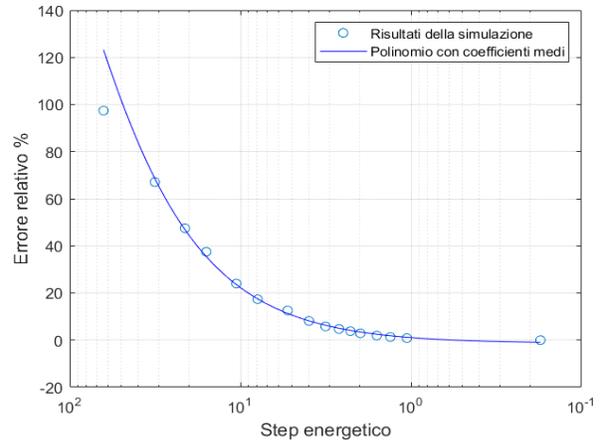
$R^2 = 0,9972$ – Layout 14



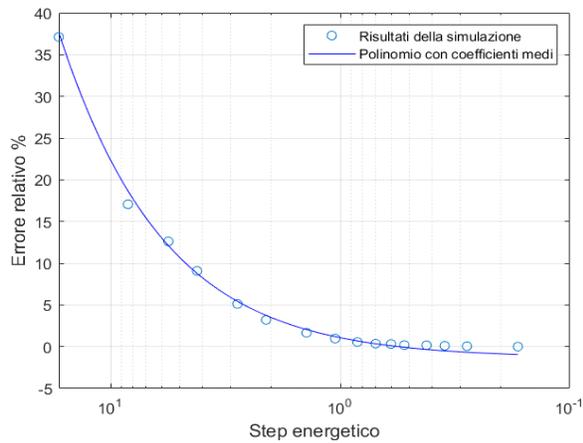
$R^2 = 0,9981$ – Layout 15



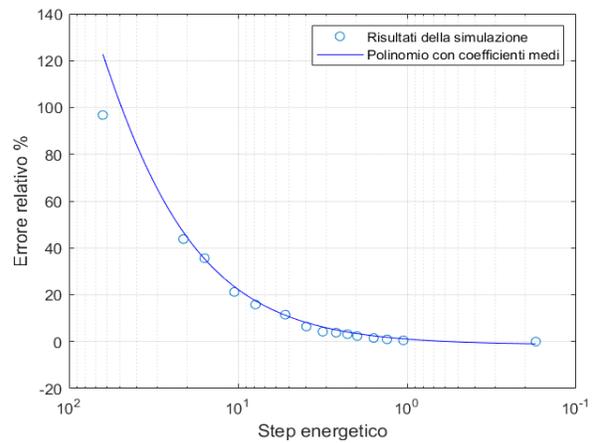
$R^2 = 0,9952$ – Layout 16



$R^2 = 0,9810$ – Layout 17



$R^2 = 0,9970$ – Layout 18



$R^2 = 0,9900$ – Layout 20

Figura 6.5 – Confronto risultati delle simulazioni con polinomio dai coefficienti medi

Viene riportato in figura 6.6 il confronto tra i risultati delle simulazioni e il polinomio medio i cui coefficienti sono stati calcolati con la media RMS. Viene riportato solamente il layout 1: questo tipo di media è stata esclusa poiché non tiene conto dei segni dei coefficienti.

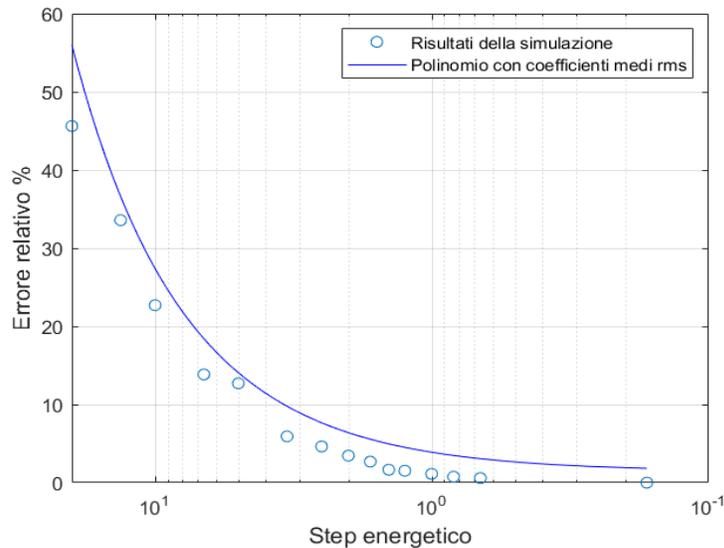


Figura 6.6 - Confronto risultati dalla simulazione con polinomio dai coefficienti medi RMS – Layout 1

6.4 Eliminazione degli outliers

L'analisi appena svolta sui dati completi viene ripetuta una volta eliminati gli outliers. Osservando infatti gli andamenti ottenuti, è facile notare come i punti ad errore relativo maggiore si discostino maggiormente dagli andamenti medi. Per tale motivo si è cercato di migliorare il fitting dei punti risultato delle simulazioni eliminando i punti outliers. Osservando inoltre le curve CO₂ – SOC è evidente che la zona di interesse va dal dopo ginocchio della curva fino al plateau, sono stati considerati quindi questi dati ed esclusi quelli al di fuori della zona d'interesse.

Come per i dati completi, è stato svolto il fitting dei dati di tutti i layout utilizzando un polinomio di secondo grado come mostrato nell'equazione 6.6. In figura 6.7 è riportato il risultato ottenuto per il layout 1. Per tutti i layout sono stati ottenuti i coefficienti per l'equazione polinomiale di secondo grado. I coefficienti vengono riassunti dalla tabella 6.6.

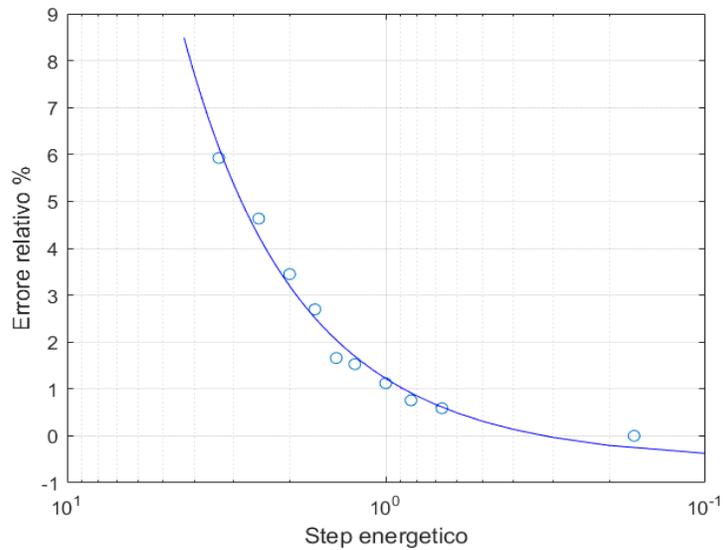


Figura 6.7 – Errore Relativo % CO2 ttw – Layout 1

Tabella 6.6 – Coefficienti polinomio di secondo grado senza outliers

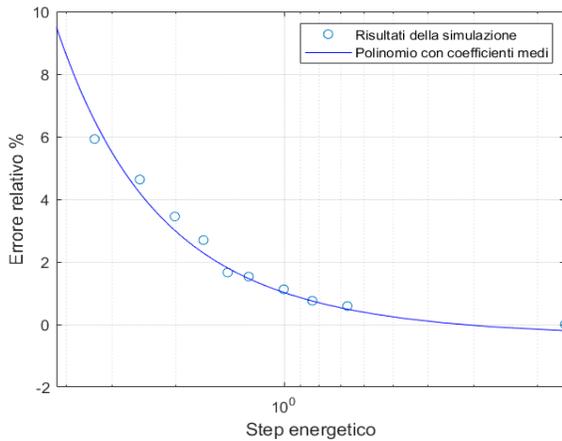
Layout	p_1	p_2	p_3
1	0,1047	1,6501	-0,5391
2	0,3999	1,0354	-0,3201
3	0,1825	1,2556	-0,3882
4	0,0042	1,8403	-0,4568
6	0,2548	0,9302	-0,3472
7	0,3016	1,2068	-0,3735
8	0,2313	1,3800	-0,2631
9	0,3080	1,0325	-0,3958
10	0,2296	0,6933	-0,2539
11	0,2920	1,0159	-0,4285
12	-0,0202	2,5837	-1,4287
13	0,0975	1,4673	-0,5604
14	0,6473	0,2775	-0,0957

15	0,5846	0,2705	-0,1212
16	0,4944	0,4696	-0,0827
17	0,2334	1,2517	-0,4866
18	0,4719	0,5819	-0,1848
20	0,1830	1,0132	-0,4061

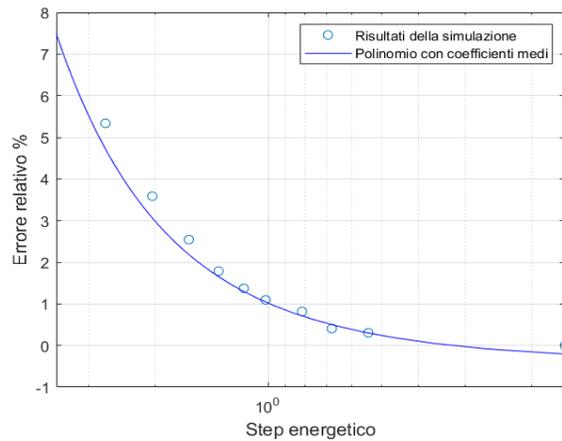
Per poter analizzare i coefficienti è stato determinato il massimo ed il minimo valore per ogni coefficiente, è stata calcolata la media aritmetica, la media RMS ed il range di variabilità dei coefficienti. Tali valori sono riportati in tabella 6.7.

Coefficienti	p_1	p_2	p_3
minimo	-0,0202	0,2705	-1,4287
massimo	0,6473	2,5837	-0,0827
media	0,2773	1,1366	-0,3984
media rms	0,3307	1,2686	0,4911
Range	0,6675	2,3132	1,346

Come svolto per i dati completi anche in questo caso è stato fatto il confronto tra i dati risultato delle simulazioni ed un polinomio dai coefficienti medi indicato nella 5.7, calcolati con media aritmetica e media RMS. In figura 6.8 sono riportati i risultati ottenuti per tutti i layout presi in considerazione.

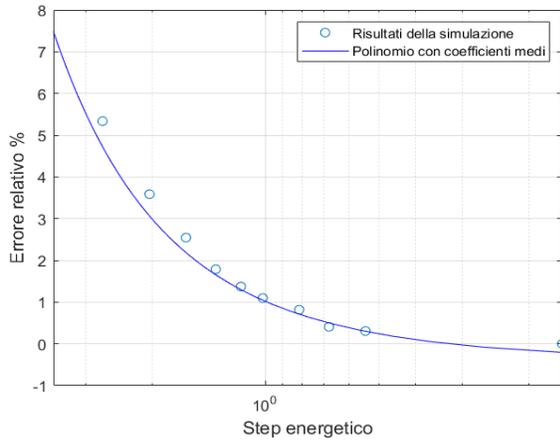


$$R^2 = 0,9766 - \text{Layout 1}$$

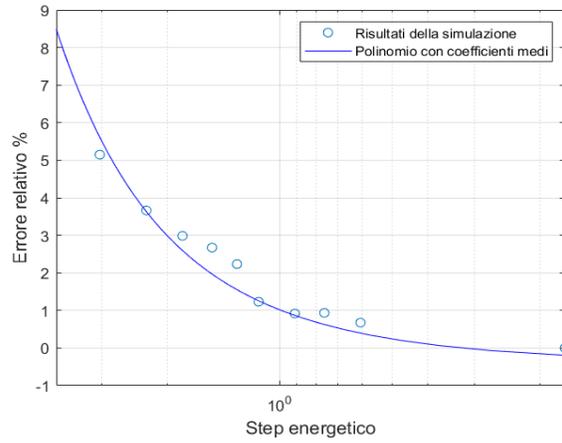


$$R^2 = 0,9951 - \text{Layout 2}$$

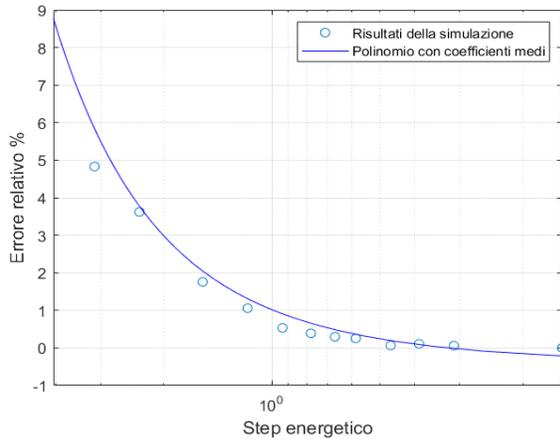
Figura 6.8 – Confronto risultati delle simulazioni con polinomio dai coefficienti medi



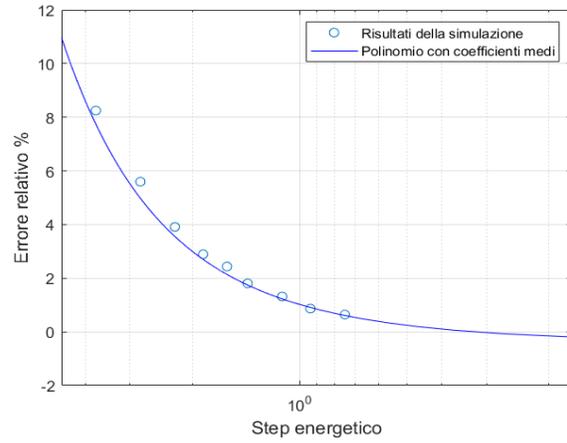
$R^2 = 0,9812$ – Layout 3



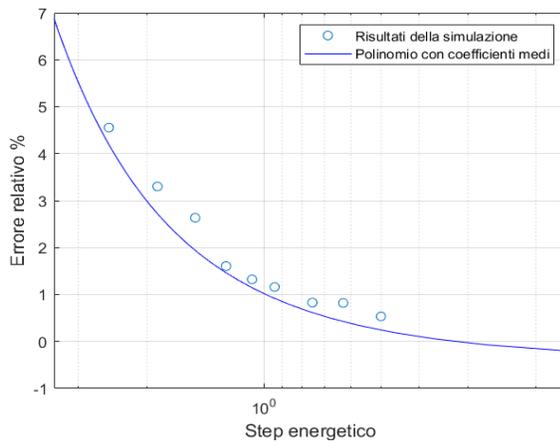
$R^2 = 0,9641$ -Layout 4



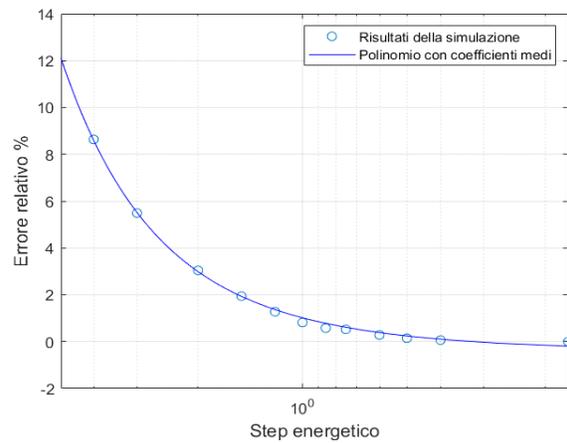
$R^2 = 0,9877$ – Layout 6



$R^2 = 0,9977$ – Layout 7

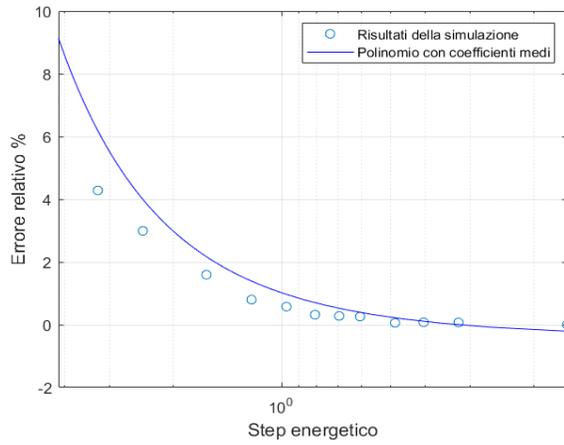


$R^2 = 0,9865$ – Layout 8

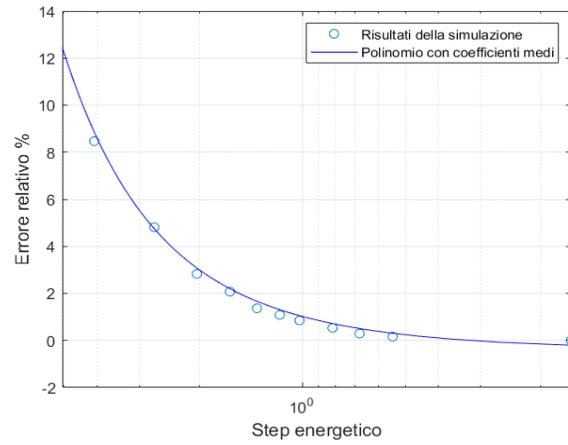


$R^2 = 0,9983$ – Layout 9

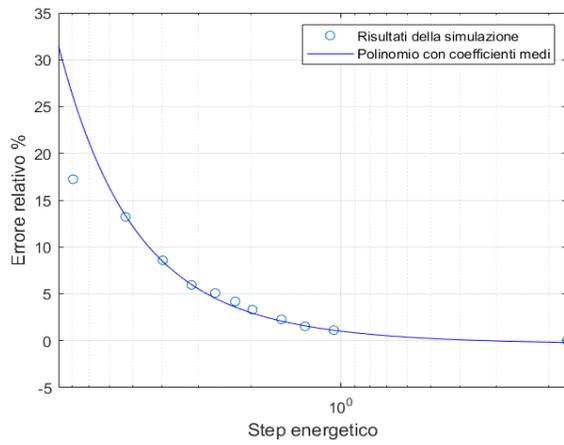
Figura 6.8 – Confronto risultati delle simulazioni con polinomio dai coefficienti medi



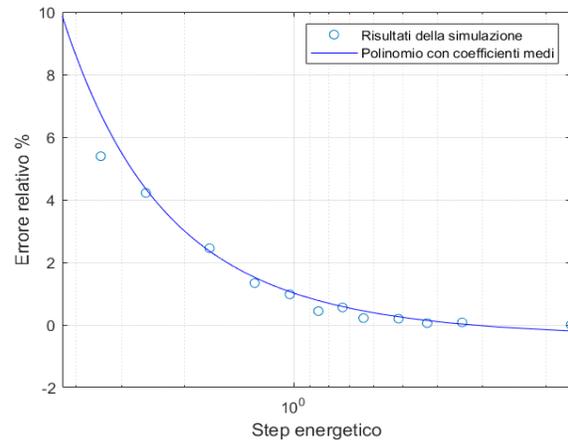
$R^2 = 0,9925$ – Layout 10



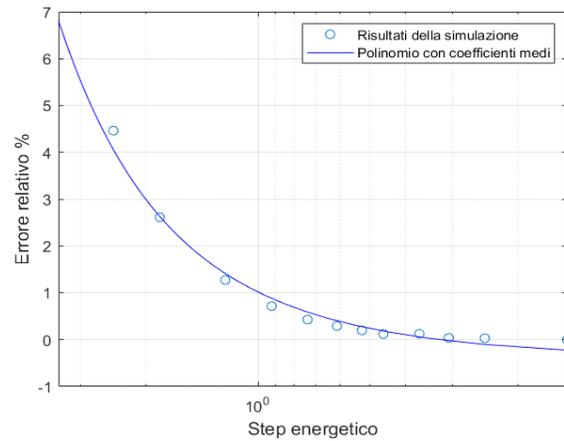
$R^2 = 0,9964$ – Layout 11



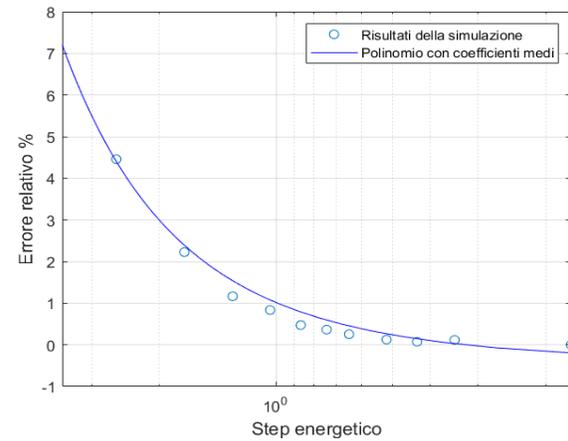
$R^2 = 0,9390$ – Layout 12



$R^2 = 0,9806$ – Layout 13

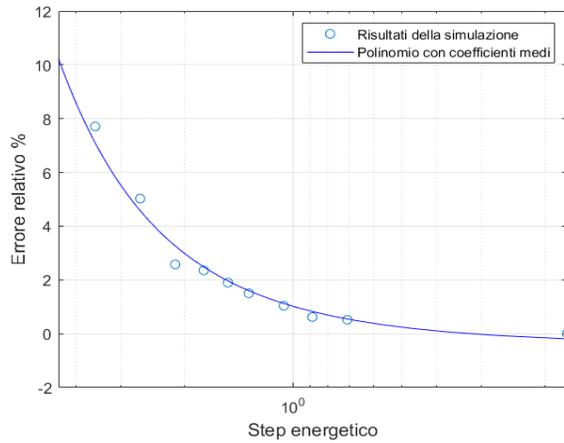


$R^2 = 0,9857$ – Layout 14

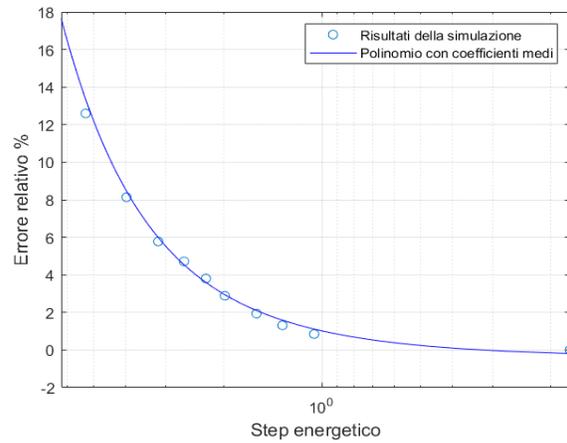


$R^2 = 0,9828$ – Layout 15

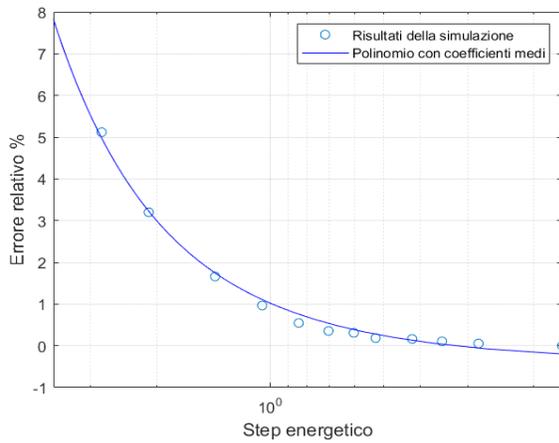
Figura 6.8 – Confronto risultati delle simulazioni con polinomio dai coefficienti medi



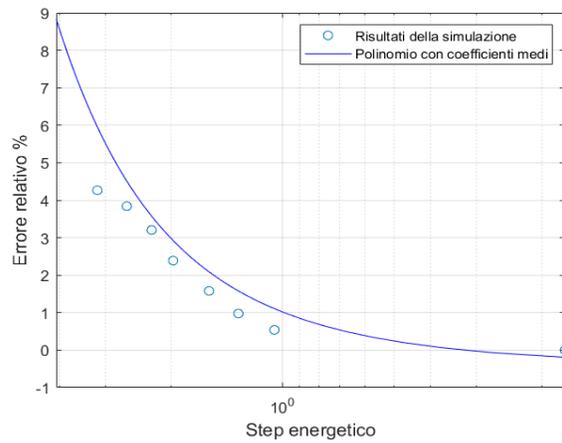
$R^2 = 0,9820$ – Layout 16



$R^2 = 0,9969$ – Layout 17



$R^2 = 0,9931$ – Layout 18



$R^2 = 0,9641$ – Layout 20

Figura 6.8 – Confronto risultati delle simulazioni con polinomio dai coefficienti medi

Viene riportato in figura 6.9 il confronto tra i risultati delle simulazioni e il polinomio medio i cui coefficienti sono stati calcolati con la media RMS. Si riporta solamente il layout 1 a titolo di esempio poiché questo tipo di media è stata ritenuta meno utile rispetto alla media aritmetica poiché non tiene conto dei segni dei coefficienti.

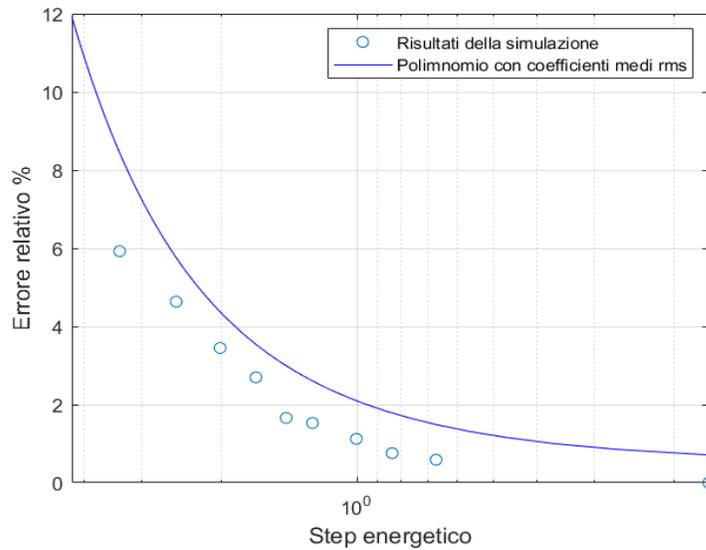


Figura 6.9 - Confronto risultati dalla simulazione con polinomio dai coefficienti medi RMS – Layout 1

6.5 Osservazioni preliminari

Osservando gli errori quadratici medi insieme agli andamenti grafici si può dedurre che la media migliore per riprodurre l'andamento dei dati risultato delle simulazioni sia la media aritmetica. La differenza sostanziale è data dal valore del coefficiente p3, negativo per tutte le architetture, il quale viene reso positivo dalla media RMS, per cui l'equazione del polinomio si discosta maggiormente dai risultati delle simulazioni. Viene quindi escluso questo secondo tipo di media sia per i dati completi sia per i dati da cui sono stati esclusi gli outliers. Osservando gli errori quadratici medi riportati nella tabella 6.8 si può osservare che per quasi tutte le architetture questo è più elevato quando si tiene conto anche degli outliers.

<i>Tabella 6.8</i>		
Layout	R^2	R^2 senza outliers
1	0,9948	0,9766
2	0,9960	0,9951
3	0,9963	0,9812
4	0,9932	0,9641
6	0,9976	0,9877
7	0,9977	0,9977
8	0,9937	0,9865

9	0,9976	0,9983
10	0,9958	0,9925
11	0,9971	0,9964
12	0,9803	0,9390
13	0,9973	0,9806
14	0,9972	0,9857
15	0,9981	0,9828
16	0,9952	0,9820
17	0,9810	0,9969
18	0,9970	0,9931
20	0,9900	0,9641

Confrontando l'andamento del polinomio dai coefficienti medi con i punti dei singoli layout si può osservare quali di questi vengono meglio rappresentati dal polinomio e in quali casi la curva si allontana maggiormente dai punti risultato delle simulazioni. In particolare si può notare che i punti risultato delle simulazioni si distanziano di più dal polinomio con coefficienti medi per i layout 4,8,10 e 20. In particolare, il layout 20 è il layout che più si allontana dall'andamento medio, come visibile in figura 6.10.

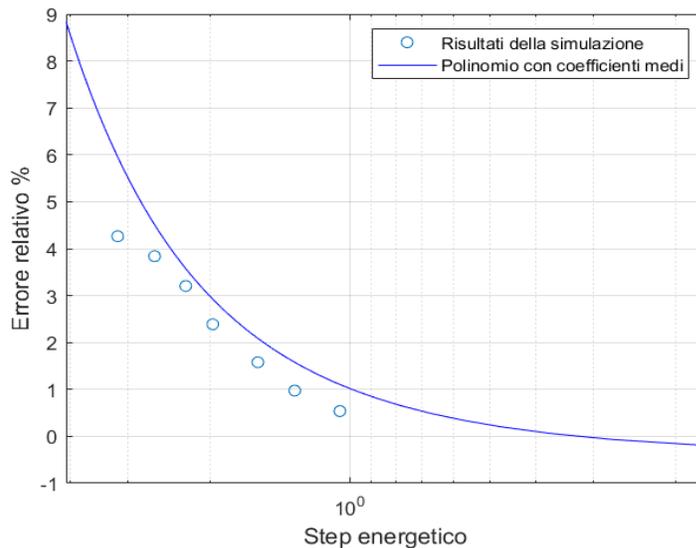


Figura 6.10 – Layout 20

Si nota che tale layout ha una taglia della batteria di 38,46 kWh. Questo layout viene quindi confrontato con il layout 12 il quale presenta taglia della batteria confrontabile. In tabella 6.9 vengono riportati i parametri di design per i layout 12 e 20 in modo da poterli confrontare.

Tabella 6.9 – Layout 12 e 20

Layout	PERat [-]	EM1P [kWh]	Em1SpR [-]	FDSpR [-]	CrateDis [-]	CrateChar [-]	EngDisp [l]	Ebatt [kWh]
12	2,79	110,46	2,98	3,385	5,20	2,99	3,5362	36,64
20	2,55	100,77	4,28	4,6.4	11,74	7,32	3,3799	39,46

Con questa prima analisi si esclude quindi che il parametro che va ad influenzare la distanza dei risultati delle simulazioni dal polinomio media sia unicamente la taglia della batteria.

6.6 Ricerca di differenti equazioni

Potrebbero esserci altri tipi di curve che rappresentano meglio l'andamento dei dati. Si è pensato quindi di fittare i punti con un'equazione esponenziale come quella espressa nell'equazione 6.8

$$y(x) = ae^{-bx} \quad (6.8)$$

Si riporta in figura 6.11 l'andamento ottenuto per il layout 1 confrontato all'andamento ottenuto tramite l'equazione polinomiale: è facile notare come l'equazione polinomiale sia più adatta a fittare i dati per cui l'equazione di tipo esponenziale è stata esclusa. Allo stesso modo si è pensato di utilizzare l'equazione di un'iperbole ma anche in questo caso l'equazione polinomiale si è rivelata più adatta.

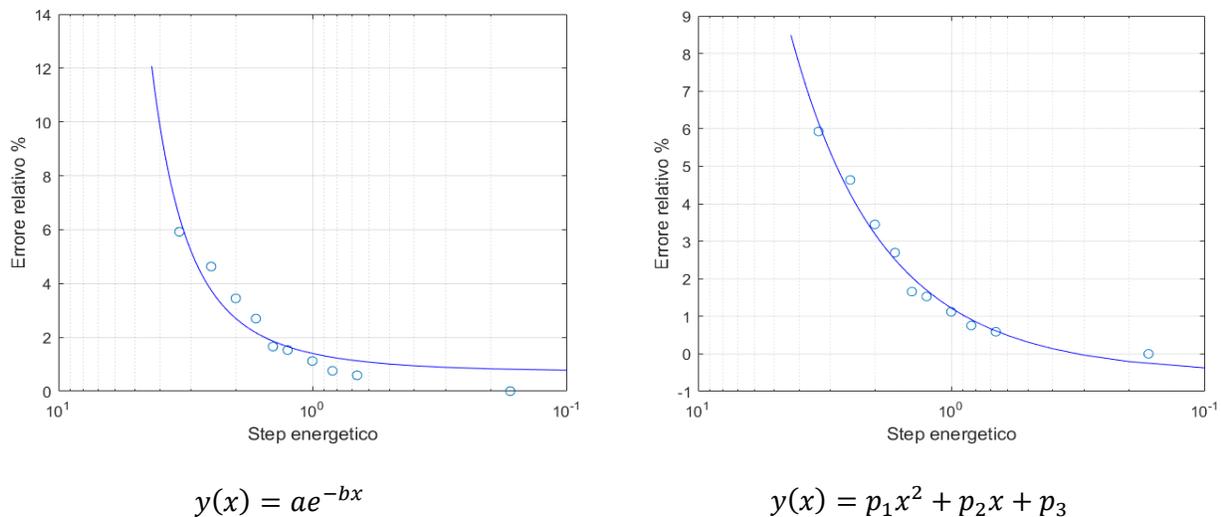


Figura 6.11 – Confronto tra equazione polinomiale ed equazione esponenziale per il layout 1

6.7 Fitting degli andamenti ottenuti con polinomio di terzo grado

L'obiettivo di questa fase è quello di andare a svolgere un'analisi analoga a quella svolta al paragrafo 5.2 andando a modificare l'equazione polinomiale scelta. L'analisi di questo paragrafo prevede infatti l'utilizzo di un polinomio di terzo grado come riportato nell'equazione 6.9:

$$y(x) = p_1x^3 + p_2x^2 + p_3x + p_4 \quad (6.9)$$

Sono quindi stati analizzati i 20 layout considerati ai paragrafi precedenti e sono stati ricavati i coefficienti per un polinomio di grado 3. Tali coefficienti vengono riportati in tabella 6.10.

Tabella 6.10 – Coefficienti per il polinomio di terzo grado

<i>Layout</i>	p_1	p_2	p_3	p_4
1	-0,0039	0,1038	1,7989	-0,6532
2	0,0016	-0,0374	2,6517	-1,3119
3	0,0003	0,0086	2,2567	-1,1573
4	-0,0061	0,1537	1,5384	-0,3138
6	-0,0020	0,0543	1,8808	-0,8829
7	-0,0018	0,0423	2,2144	-1,0331
8	0,0020	-0,0208	2,4171	-0,9007
9	-0,0019	0,0437	2,1068	-1,0186
10	-0,0020	0,0578	1,6579	-0,8471
11	-0,0009	0,0244	2,1809	-1,1912
12	-0,0004	0,0203	1,8301	0,0467
13	-0,0018	0,0465	2,0377	-0,9746
14	-0,0051	0,0916	1,888	-0,7418
15	-0,001	0,0150	2,2128	-1,1158
16	-0,0001	-0,0111	2,5842	-1,6651
17	-0,0001	-0,0100	2,5970	-1,8386
18	-0,0008	-0,0170	2,3107	-1,0082
20	-0,0001	-0,0066	2,5654	-1,4824

In figura 6.12 è riportato il grafico che mette in relazione il valore dei coefficienti ricavati con il grado della variabile a cui sono associati. È utile osservare che il coefficiente di grado 3 ha un range di variazione molto piccolo rispetto agli altri coefficienti. È possibile quindi che possa esistere una terza dimensione per andare a rappresentare l'andamento dell'errore relativo ovvero che questo possa dipendere oltre che dallo step energetico da altre variabili. La distribuzione dei coefficienti visibile nel grafico fa quindi pensare che possa esistere una dipendenza tra i coefficienti ed ulteriori variabili oltre allo step energetico.

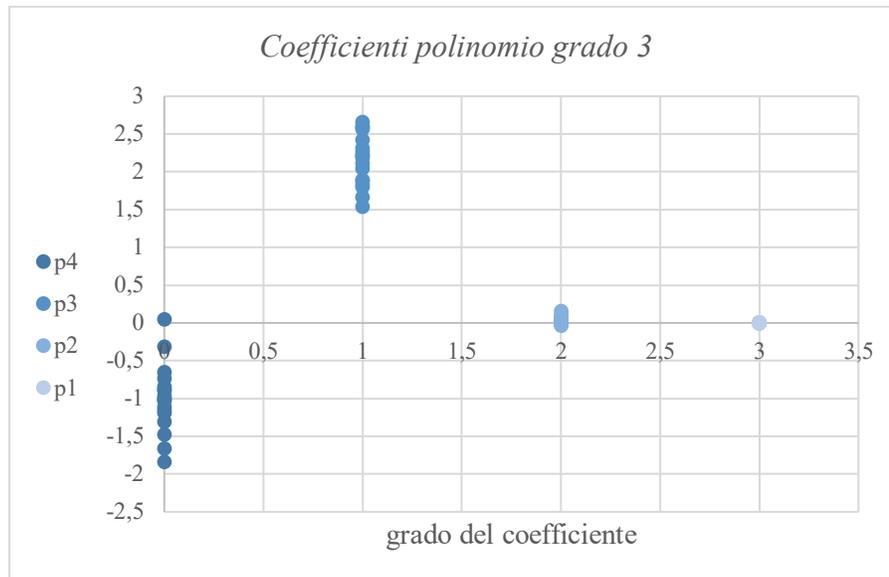


Figura 6.12 – Coefficienti per i polinomi di grado 3

Come svolto ai paragrafi precedenti sono stati ottenuti i coefficienti utilizzando i dati da cui sono stati esclusi gli outliers. In tabella 6.11 vengono quindi riportati i coefficienti ottenuti.

Layout	p_1	p_2	p_3	p_4
1	-0,2653	1,4874	-0,2494	0,0574
2	-0,2208	1,3420	-0,0350	-0,0335
3	-0,3108	1,6048	-0,4776	0,1096
4	-0,1208	0,5793	1,1136	-0,2424
6	-0,2923	1,6410	-0,7645	0,1314
7	-0,0922	0,8366	0,3915	-0,0932
8	-0,2350	1,1648	0,3863	-0,0093

9	-0,0494	0,6070	0,5686	-0,2310
10	-0,1648	1,0410	-0,3364	0,0481
11	-0,1041	0,9411	-0,0195	-0,0464
12	-0,0590	0,6825	0,4847	-0,0431
13	-0,2625	1,4635	-0,3648	0,0074
14	-0,0750	0,9292	0,0026	-0,0330
15	-0,1764	1,2988	-0,4870	0,0737
16	-0,0035	0,5134	0,4419	-0,0735
17	-0,0298	0,4742	0,7486	-0,2612
18	-0,1333	1,1045	-0,0583	-0,0177
20	-0,1395	0,9542	0,3750	-0,1360

Anche i coefficienti ottenuti grazie al fitting con esclusione degli outliers sono stati raggruppati in base al grado del coefficiente in modo da verificare il range di variabilità.

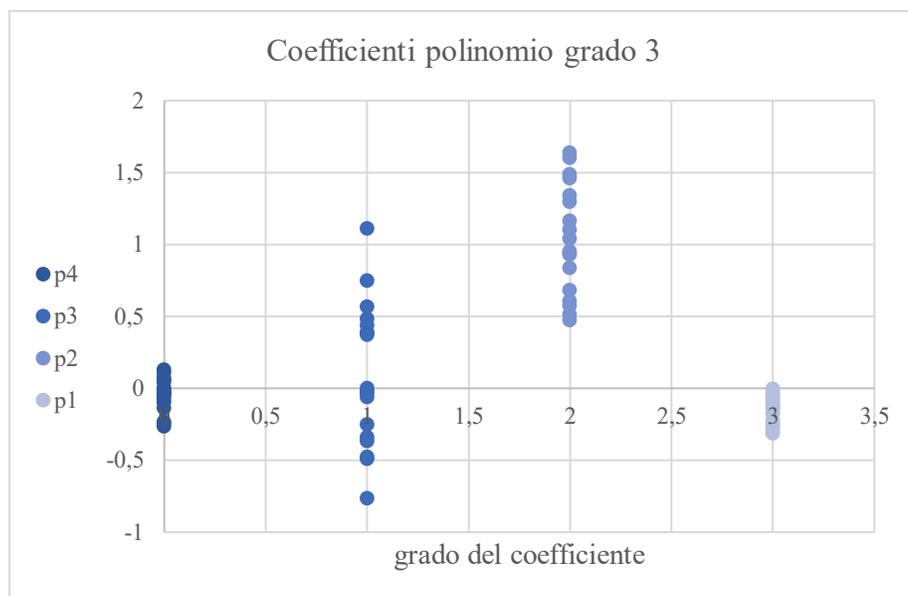


Figura 6.12 – Coefficienti per i polinomi di grado 3

In questa fase preliminare si è quindi tentato di riprodurre l'andamento dell'errore relativo di CO₂ in funzione dello step energetico per i diversi layout analizzati. Sono stati considerati due tipi differenti di polinomi: polinomio di secondo grado e polinomio di terzo grado. Entrambe le equazioni riescono a riprodurre l'andamento dell'errore relativo ma in particolare dall'analisi dei coefficienti di terzo grado è possibile dedurre che esistono ulteriori variabili legate all'errore relativo. Per il polinomio di secondo grado sono stati confrontati i valori dell'errore quadratico medio nei casi di dati con e senza outliers e si è potuto constatare che è maggiore nel caso di dati completi.

Si è potuto inoltre constatare che non è possibile ricavare un'unica equazione che vada a rappresentare tutti i possibili layout per l'architettura p2 poiché al variare dei parametri i punti risultato delle simulazioni si discostano o si avvicinano al polinomio ottenuti con i coefficienti medi. Vista la possibile dipendenza dell'errore relativo da ulteriori variabili è necessario andare ad utilizzare strumenti matematici che permettano di verificare effettivamente tali relazioni, da qui l'idea di utilizzare la regressione multipla.

7. Regressioni multilineari con superficie di secondo grado

L'analisi preliminare svolta ai capitoli precedenti ha portato alla luce il legame tra errore relativo di CO2 e step energetico. Le analisi svolte hanno permesso di determinare per ogni layout analizzato i coefficienti di un'equazione polinomiale di secondo grado ed i coefficienti di un'equazione polinomiale di terzo grado. La distribuzione di tali coefficienti in base al loro grado fa emergere la possibilità che l'errore relativo dipenda dallo step energetico e da ulteriori variabili, ovvero dalle variabili di design del veicolo. Sorge così la necessità di utilizzare le regressioni multilineari che permettono di descrivere l'andamento di una variabile dipendente grazie all'utilizzo di più variabili dipendenti o esplicative. In primo luogo sono state considerate come variabili esplicative, oltre allo step energetico, la taglia della macchina elettrica, la taglia della macchina elettrica rapportata ai C Rate di carica e di scarica ed infine la taglia della batteria. Le regressioni multilineari sono state svolte tramite l'ambiente di calcolo Matlab che ha permesso non solo di ottenere le equazioni delle diverse superfici ma anche di visualizzare le superfici ottenute in modo da poterle confrontare con i dati. In questo capitolo le regressioni multilineari sono state svolte inserendo i termini fino al secondo grado per lo step energetico.

7.1 Regressioni multiple considerando i dati completi

Regressione multipla: taglia della macchina elettrica

Il primo tentativo di regressione multipla svolto considera come variabili esplicative lo step energetico e la potenza della macchina elettrica. Tramite Matlab è stata eseguita una regressione multipla in modo da correlare errore relativo a step energetico e potenza della macchina elettrica. Si ottiene una superficie su cui sono distribuiti i punti ottenuti tramite simulazione. In particolare la superficie è espressa come nell'equazione 7.1:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_1^2 + \beta_4 x_1 x_2 \quad (7.1)$$

In cui x_1 rappresenta lo step energetico ed x_2 rappresenta la potenza del motore elettrico. In figura 7.1 si possono visualizzare le superfici ottenute tramite una vista tridimensionale ed una vista secondo il piano $y-x_1$. Entrambe le viste sono state riportate utilizzando l'asse logaritmico per lo step energetico. La tabella 7.1 riporta i coefficienti ricavati insieme ai rispettivi valori p. Si riporta inoltre per questo tipo di regressione il valore p per il test F, il coefficiente di determinazione, il coefficiente di determinazione corretto ed il valore di RMSE.

Root Mean Squared Error	1,22
R-squared	0,995
Adjusted R-Squared	0,995
p-value	0,000

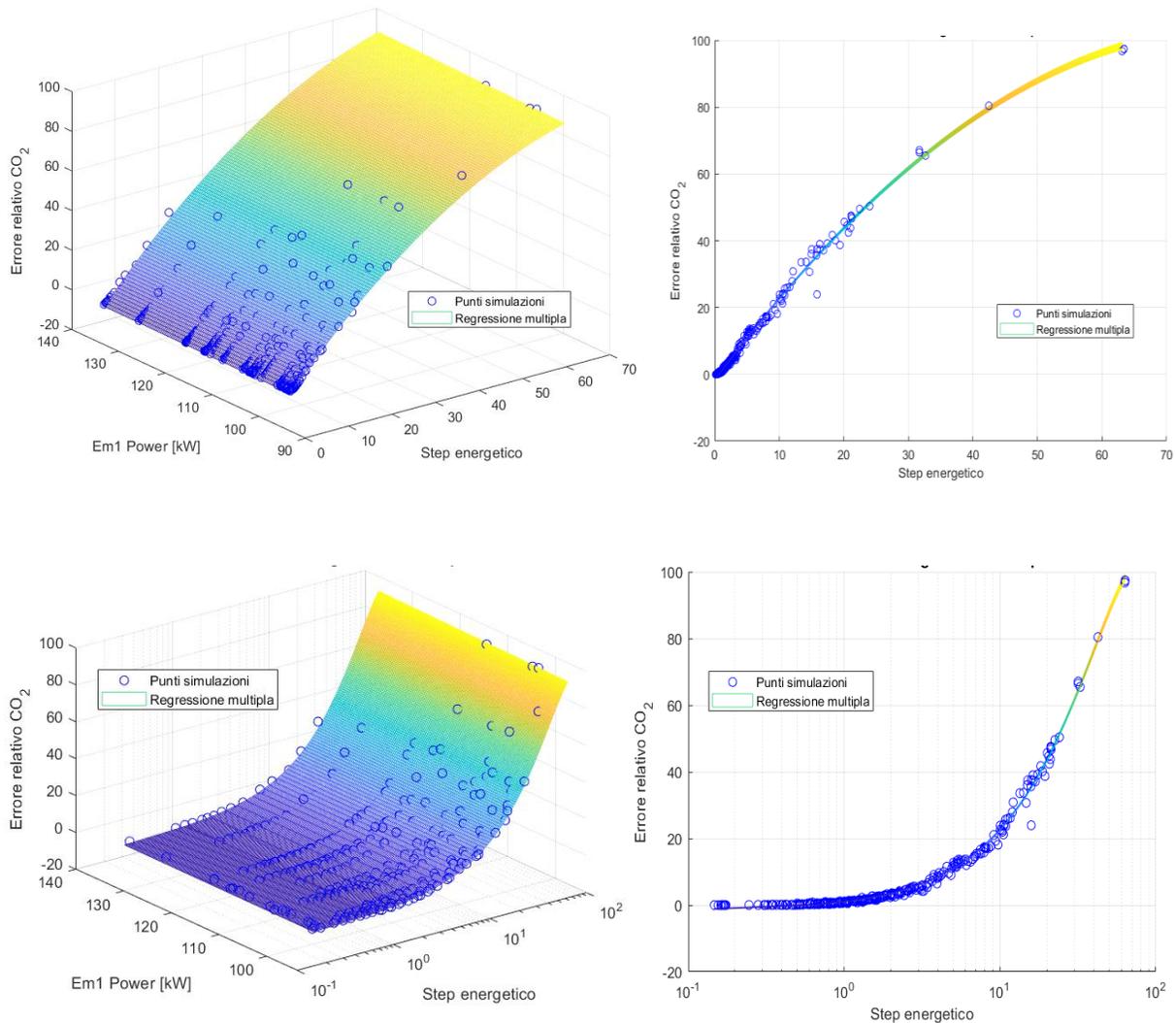


Figura 7.1 – Regressione multipla, viste in tre dimensioni e vista del piano y-x₁.

Tabella 7.1 – Coefficienti stimati

Termine	Valore	pValue
Intercetta	-2,1742	0,0130
X ₁	2,4848	0,0000
X ₂	0,0059	0,4728
X ₁ ²	-0,0158	0,0000
X ₁ X ₂	0,0008	0,4747

La regressione permette quindi di ottenere i coefficienti di una superficie polinomiale di secondo grado rispetto allo step energetico e di primo grado rispetto alla potenza della macchina elettrica. Sulla base del test F il modello va a riprodurre bene i dati. I valori di R² esprimono la bontà del modello utilizzato. La variabile più significativa è lo step energetico.

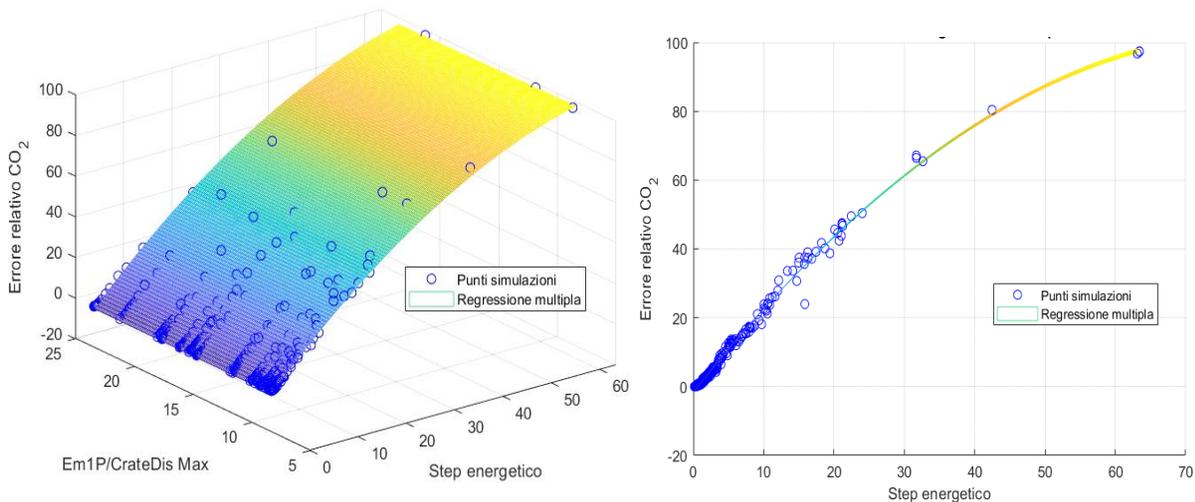
Regressione multipla: taglia della macchina elettrica rapportata a C rate di scarica

È stata eseguita la regressione multipla per descrivere l'errore relativo in base a step energetico ed EMIP/Crate di scarica: si va a rapportare la potenza della macchina elettrica al C rate di scarica in modo da ottenere la seconda variabile esplicativa per la descrizione dell'errore relativo oltre allo step energetico. L'equazione utilizzata per la regressione è la medesima espressa nell'equazione 7.1, in questo caso x_1 rappresenta lo step energetico ed x_2 rappresenta la variabile potenza della macchina elettrica rapportata al C Rate di scarica. In figura 7.2 si possono visualizzare le superfici ottenute tramite una vista tridimensionale ed una vista secondo il piano $y-x_1$. Entrambe le viste sono state riportate utilizzando l'asse logaritmico per lo step energetico.

La tabella 7.2 riporta i coefficienti ricavati insieme ai rispettivi valori p.

Si riporta inoltre per questo tipo di regressione il valore p per il test F, il coefficiente di determinazione, il coefficiente di determinazione corretto ed il valore di RMSE.

Root Mean Squared Error	1,23
R-squared	0,995
Adjusted R-Squared	0,995
p-value	0,000



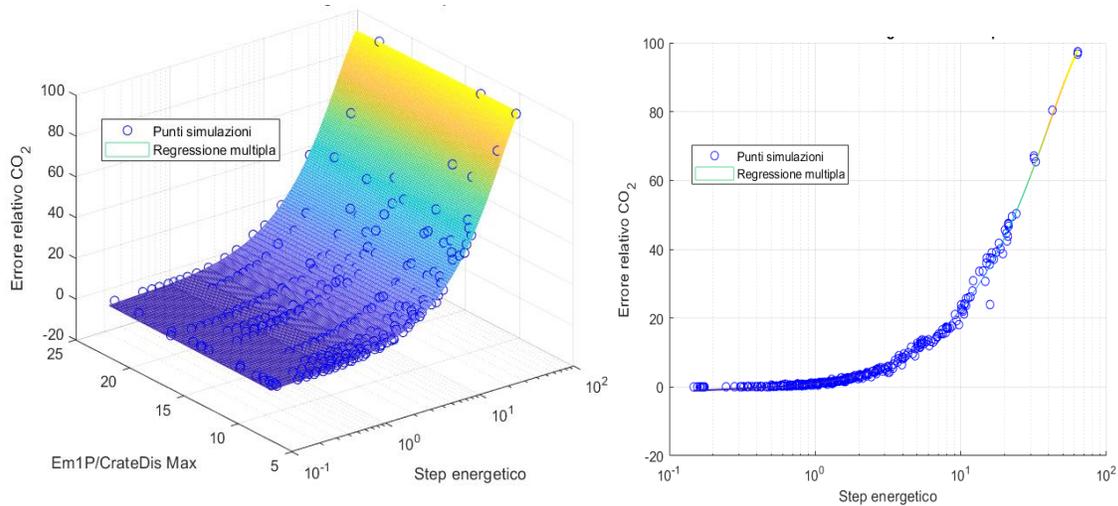


Figura 7.2 – Regressione multipla, viste in tre dimensioni e viste del piano $y-x_1$.

Tabella 7.2 – Coefficienti stimati

Termine	Valore	pValue
Intercetta	-1,8028	0,0000
X ₁	2,5916	0,0000
X ₂	0,0192	0,2947
X ₁ ²	-0,0158	0,0000
X ₁ X ₂	-0,0016	0,3258

La regressione permette quindi di ottenere i coefficienti di una superficie polinomiale di secondo grado rispetto allo step energetico e di primo grado rispetto alla potenza della macchina elettrica rapportata a C rate di scarica. Sulla base del test F il modello va a riprodurre bene i dati, si possono inoltre osservare i valori di R² che esprimono la bontà del modello utilizzato. Anche in questo caso la variabile più significativa è lo step energetico.

Regressione multipla: taglia della macchina elettrica rapportata a C rate di carica

È stata eseguita la regressione multipla per descrivere l'errore relativo in base a step energetico ed EM1P/Crate di carica: si va a rapportare la potenza della macchina elettrica al C rate di carica in modo da ottenere la seconda variabile esplicativa per la descrizione dell'errore relativo oltre allo step energetico. L'equazione utilizzata per la regressione è la medesima espressa nell'equazione 7.1, in questo caso x₁ rappresenta lo step energetico ed x₂ rappresenta la variabile potenza della macchina elettrica rapportata al C Rate di scarica. In figura 7.3 si possono visualizzare le superfici ottenute tramite una vista tridimensionale ed una vista secondo il piano $y-x_1$. Entrambe le viste sono state riportate utilizzando l'asse logaritmico per lo step energetico.

La tabella 7.3 riporta i coefficienti ricavati insieme ai rispettivi valori p.

Si riporta inoltre per questo tipo di regressione il valore p per il test F, il coefficiente di determinazione, il coefficiente di determinazione corretto ed il valore di RMSE.

Root Mean Squared Error	1,22
R-squared	0,995
Adjusted R-Squared	0,995
p-value	0,000

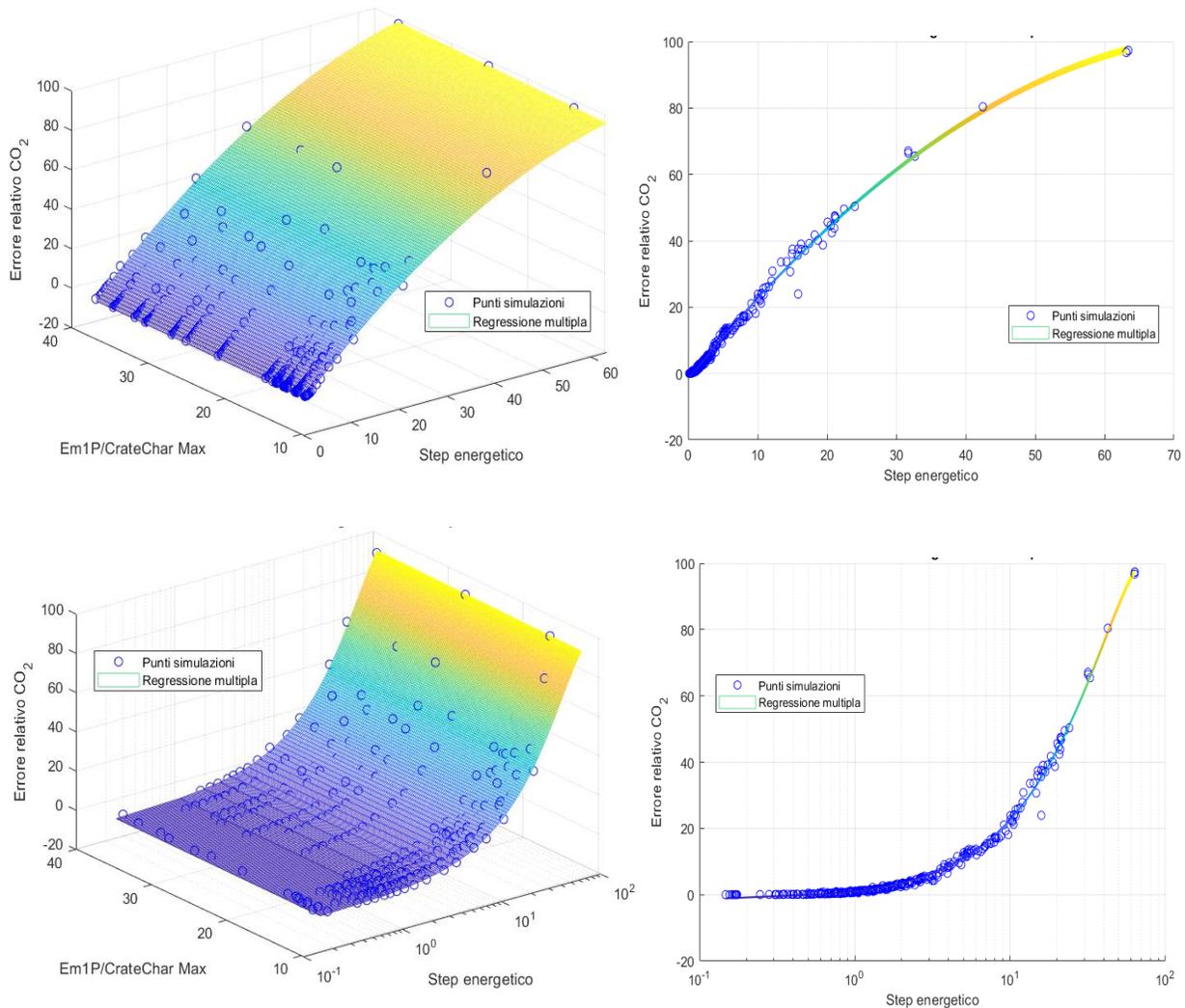


Figura 7.3 – Regressione multipla, viste in tre dimensioni e viste del piano y-x₁.

Tabella 7.3 – Coefficienti stimati

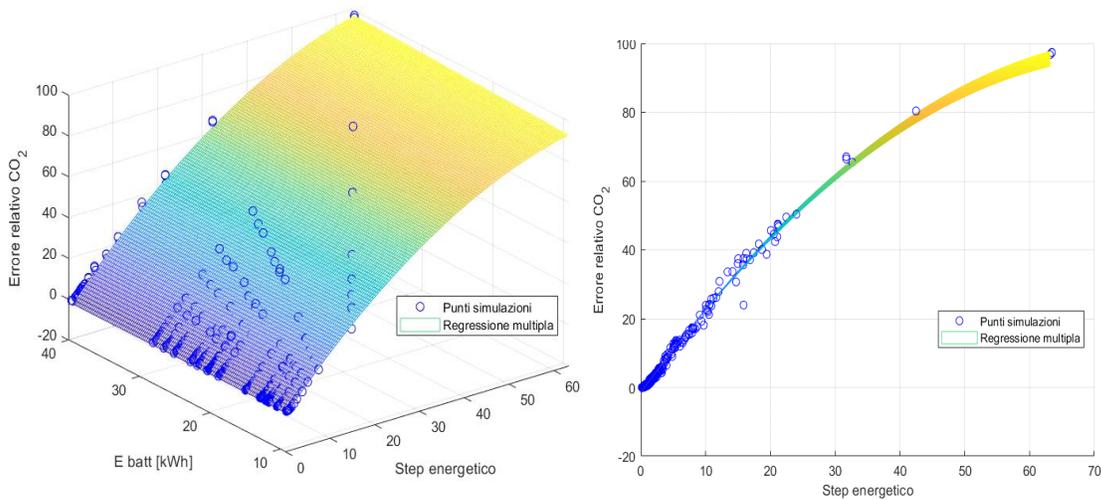
Termine	Valore	pValue
Intercetta	-1,7587	0,0000
X ₁	2,5495	0,0000
X ₂	0,0116	0,2147
X ₁ ²	-0,0159	0,0000
X ₁ X ₂	0,0010	0,2450

La regressione permette quindi di ottenere i coefficienti di una superficie polinomiale di secondo grado rispetto allo step energetico e di primo grado rispetto alla potenza della macchina elettrica rapportata a Crate di carica. Sulla base del test F il modello va a riprodurre bene i dati. I valori di R^2 esprimono la bontà del modello utilizzato. La variabile statisticamente più significativa per il modello è anche in questo caso lo step energetico.

Regressione multipla: taglia della batteria

In questo paragrafo è stata eseguita la regressione multipla per descrivere l'errore relativo in base a step energetico e taglia della batteria. L'equazione utilizzata per la regressione è espressa nell'equazione 7.1, in questo caso x_1 rappresenta lo step energetico ed x_2 rappresenta la variabile taglia della batteria. In figura 7.4 si possono visualizzare le superfici ottenute tramite una vista tridimensionale ed una vista secondo il piano $y-x_1$. Entrambe le viste sono state riportate utilizzando l'asse logaritmico per lo step energetico. La tabella 7.4 riporta i coefficienti ricavati insieme ai rispettivi valori p. Si riporta inoltre per questo tipo di regressione il valore p per il test F, il coefficiente di determinazione, il coefficiente di determinazione corretto ed il valore di RMSE.

Root Mean Squared Error	1,22
R-squared	0,995
Adjusted R-Squared	0,995
p-value	0,000



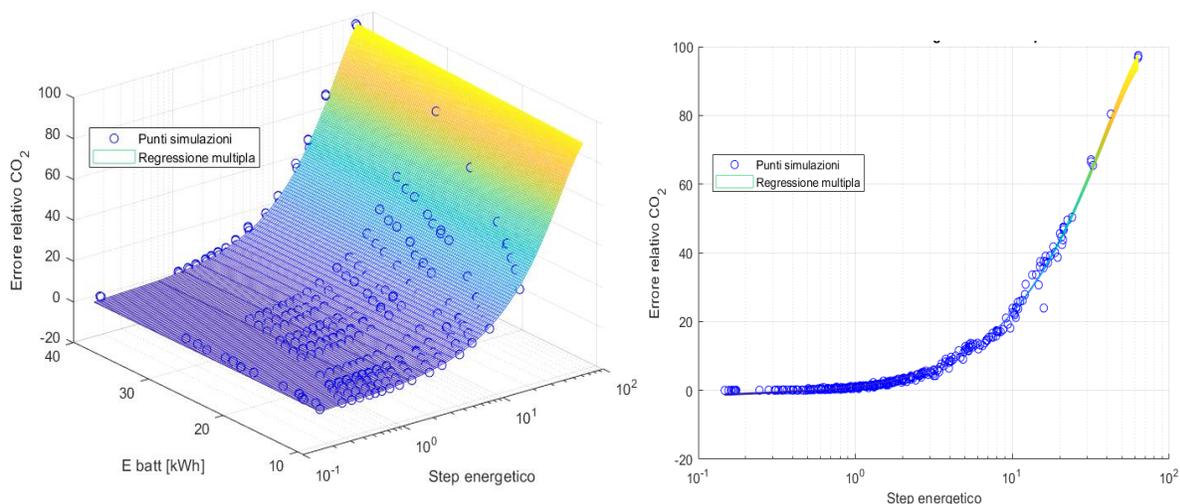


Figura 7.4 – Regressione multipla, viste in tre dimensioni e viste del piano $y-x_1$.

Tabella 7.4 – Coefficienti stimati

Termine	Valore	pValue
Intercetta	-1,2602	0,0000
x_1	2,5277	0,0000
x_2	-0,0152	0,1217
x_1^2	-0,0167	0,0000
$x_1 x_2$	0,0026	0,0348

La regressione permette quindi di ottenere i coefficienti di una superficie polinomiale di secondo grado rispetto allo step energetico e di primo grado rispetto alla taglia della batteria. Sulla base del test F il modello va a riprodurre bene i dati. I valori di R^2 esprimono la bontà del modello utilizzato. La variabile più significativa è lo step energetico, sembra inoltre essere significativo il termine di interazione tra le due variabili.

Regressione multipla: taglia della batteria rapportata a C rate di scarica

In questo caso la regressione eseguita tiene conto dello step energetico e della taglia della batteria rapportata al valore di C rate di scarica. La superficie, i cui coefficienti sono il risultato della regressione, presenta l'equazione indicata dalla 7.1. In questo caso x_1 rappresenta lo step energetico ed x_2 rappresenta la variabile taglia della batteria rapportata a C rate di scarica. In figura 7.10 si possono visualizzare le superfici ottenute tramite una vista tridimensionale ed una vista secondo il piano $y-x_1$. Entrambe le viste sono state riportate utilizzando l'asse logaritmico per lo step energetico. La tabella 7.10 riporta i coefficienti ricavati insieme ai rispettivi valori p. Si riporta inoltre per questo tipo di regressione il valore p per il test F, il coefficiente di determinazione, il coefficiente di determinazione corretto ed il valore di RMSE.

Root Mean Squared Error	1,23
R-squared	0,995
Adjusted R-Squared	0,995

p-value	0,000
---------	-------

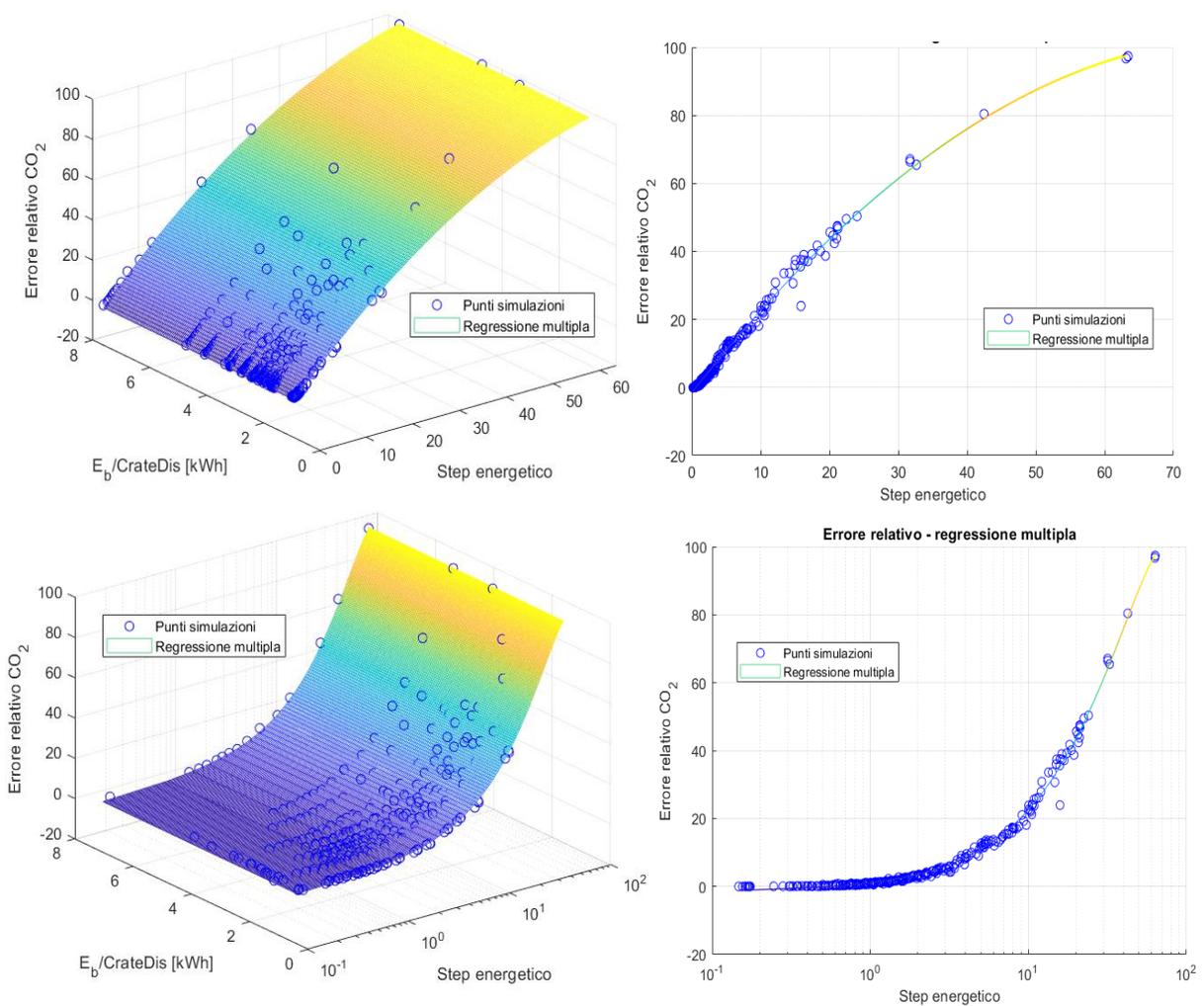


Figura 7.5 – Regressione multipla, viste in tre dimensioni e viste del piano y-x1.

Tabella 7.5 – Coefficienti stimati

Termine	Valore	pValue
Intercetta	-1,5059	0,0000
X1	2,5675	0,0000
X2	-0,0179	0,7542
X1 ²	-0,0158	0,0000
X1X2	0,0016	0,7324

La regressione permette quindi di ottenere i coefficienti di una superficie polinomiale di secondo grado rispetto allo step energetico e di primo grado rispetto alla taglia della batteria rapportata a C rate di scarica. Sulla base del test F il modello va a riprodurre bene i dati. I valori di R² esprimono la bontà del modello utilizzato. Anche in questo caso la variabile più significativa è legata allo step energetico.

7.2 Regressioni multiple a tre variabili

Regressione multipla: taglia della macchina elettrica e taglia della batteria

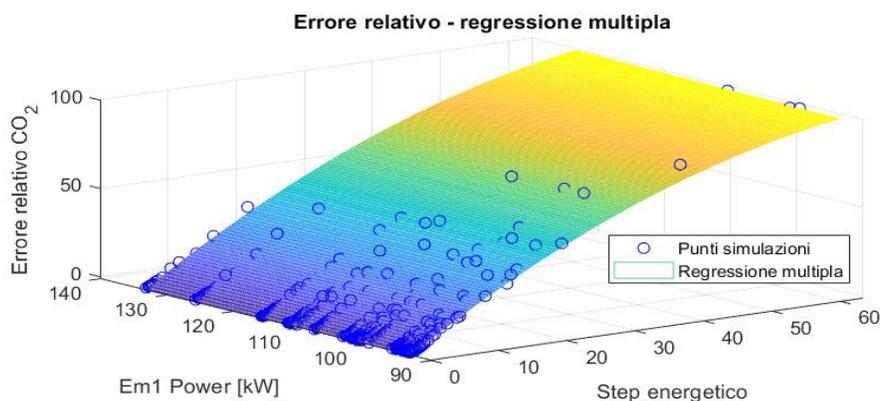
Poiché ai paragrafi precedenti sono state evidenziate le dipendenze dell'errore relativo da parametri quali taglia della macchina elettrica e taglia della batteria potrebbe essere interessante ricercare un'unica equazione in grado di descrivere l'andamento dell'errore relativo a partire da tre variabili esplicative: step energetico, potenza della macchina elettrica, taglia della batteria. Per tale tipo di regressione viene utilizzata l'equazione 8.2:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_1^2 + \beta_5 x_1 x_2 + \beta_6 x_1 x_3 + \beta_7 x_2 x_3 \quad (8.2)$$

In cui x_1 rappresenta lo step energetico, x_2 rappresenta la potenza della macchina elettrica ed x_3 rappresenta la taglia della batteria. Si riportano per questo tipo di regressione il valore p per il test F, il coefficiente di determinazione, il coefficiente di determinazione corretto ed il valore di RMSE.

Root Mean Squared Error	1,22
R-squared	0,995
Adjusted R-Squared	0,995
p-value	0,000

In figura 8.11 vengono riportate le viste in tre dimensioni della superficie di regressione ottenuta. La tabella 8.11 riporta il valore dei coefficienti ottenuti insieme ai pValue.



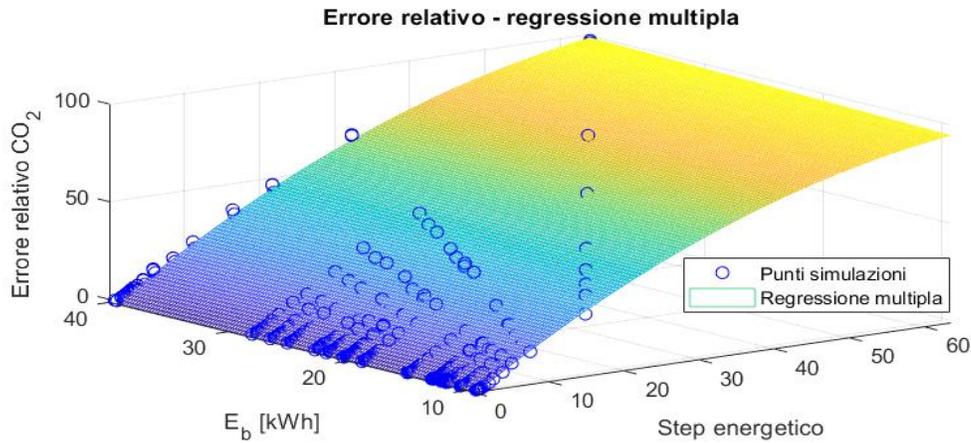


Figura 7.6 – Regressione multipla a tre variabili

Tabella 7.6 – Coefficienti stimati

<i>Termine</i>	<i>Valore</i>	<i>pValue</i>
Intercetta	-2,5645	0,2790
X ₁	2,4530	0,0000
X ₂	0,0125	0,5879
X ₃	0,0139	0,9065
X ₁ ²	-0,0166	0,0000
X ₁ X ₂	0,0007	0,5458
X ₁ X ₃	0,0026	0,0331
X ₂ X ₃	-0,0003	0,8037

Sulla base del test F il modello va a riprodurre bene i dati. I valori di R² esprimono la bontà del modello utilizzato. I termini più significativi sono lo step energetico insieme al termine di interazione tra step energetico e taglia della batteria.

8. Regressioni multilineari con superficie di terzo grado

Ai capitoli precedenti è stata svolta un'analisi preliminare che ha portato alla luce il legame tra errore relativo di CO₂ e step energetico. Le analisi svolte hanno permesso di determinare per ogni layout analizzato i coefficienti di un'equazione polinomiale di secondo grado ed i coefficienti di un'equazione polinomiale di terzo grado. La distribuzione di tali coefficienti in base al loro grado fa emergere la possibilità di trovare legami tra l'errore relativo ed altre variabili di design. Sorge così la necessità di utilizzare le regressioni multilineari che permettono di descrivere l'andamento di una variabile dipendente grazie all'utilizzo di più variabili dipendenti o esplicative. In primo luogo sono state considerate come variabili esplicative, oltre allo step energetico, la taglia della macchina elettrica, la taglia della macchina elettrica rapportata ai C Rate di carica e di scarica ed infine la taglia della batteria. Le regressioni multilineari sono state svolte tramite l'ambiente di calcolo Matlab che ha permesso non solo di ottenere le equazioni delle diverse superfici ma anche di visualizzare le superfici ottenute in modo da poterle confrontare con i dati. In questo capitolo le regressioni multilineari sono state svolte inserendo i termini fino al terzo grado per lo step energetico.

8.1 Regressioni multiple considerando i dati completi

Regressione multipla: taglia della macchina elettrica

Il primo tentativo di regressione multipla svolto considera come variabili esplicative lo step energetico e la potenza della macchina elettrica. Tramite Matlab è stata eseguita una regressione multipla in modo da correlare errore relativo a step energetico e potenza della macchina elettrica. Si ottiene una superficie su cui sono distribuiti i punti ottenuti tramite simulazione. In particolare la superficie è espressa come nell'equazione 8.1:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_1^2 + \beta_4 x_1 x_2 + \beta_5 x_1^3 + \beta_6 x_1^2 x_2 \quad (8.1)$$

In cui x_1 rappresenta lo step energetico ed x_2 rappresenta la potenza del motore elettrico. In figura 8.1 si possono visualizzare le superfici ottenute tramite una vista tridimensionale ed una vista secondo il piano $y-x_1$. Entrambe le viste sono state riportate utilizzando l'asse logaritmico per lo step energetico.

La tabella 8.1 riporta i coefficienti ricavati insieme ai rispettivi valori p.

Si riporta inoltre per questo tipo di regressione il valore p per il test F, il coefficiente di determinazione, il coefficiente di determinazione corretto ed il valore di RMSE.

Root Mean Squared Error	1,18
R-squared	0,995
Adjusted R-Squared	0,995
p-value	0,000

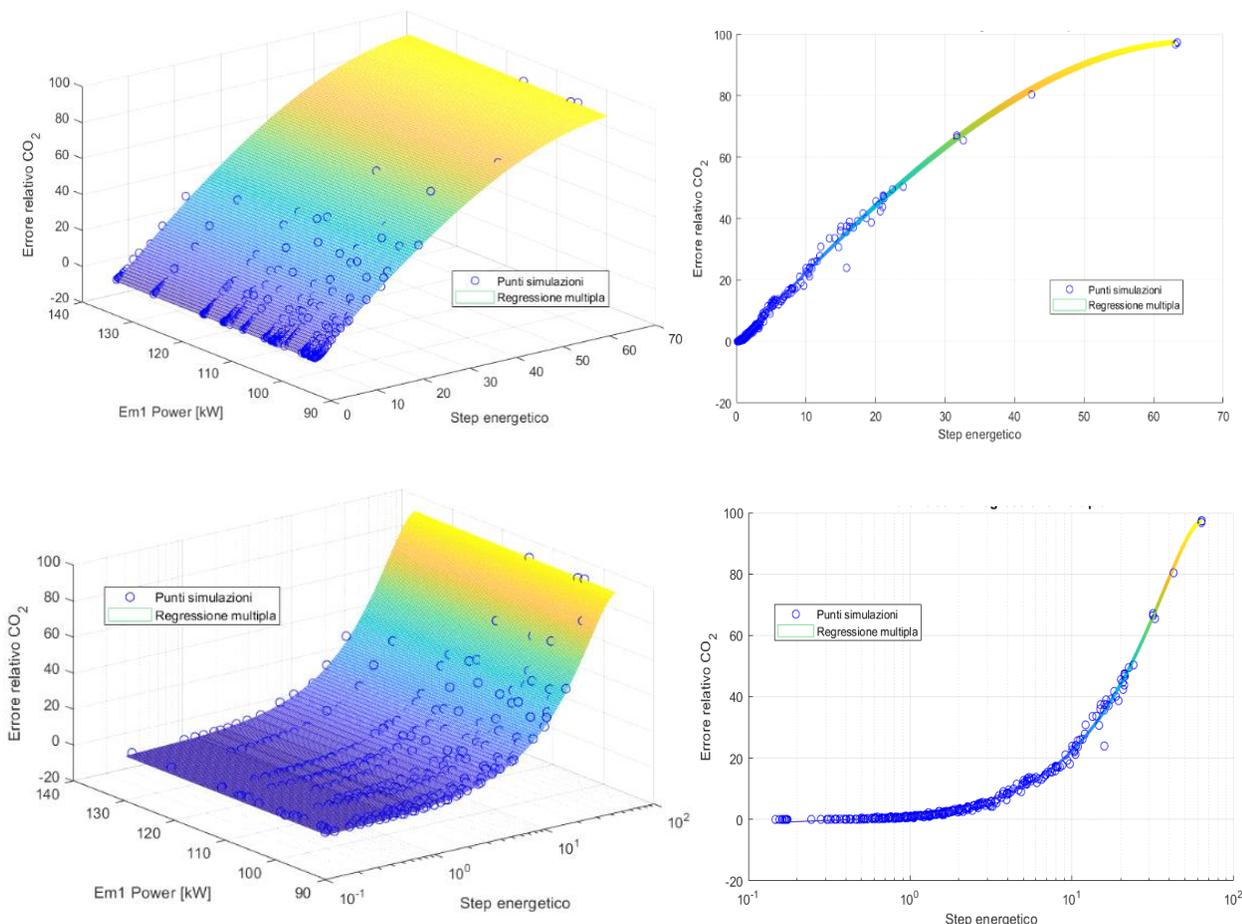


Figura 8.1 – Regressione multipla, viste in tre dimensioni e vista del piano y-x₁.

Tabella 8.1 – Coefficienti stimati

Termine	Valore	pValue
Intercetta	-1,3863	0,1390
X ₁	2,097	0,0000
X ₂	0,0009	0,9183
X ₁ ²	-0,0011	0,8575
X ₁ X ₂	0,0029	0,1284
X ₁ ³	-0,0001	0,0000
X ₁ ² X ₂	-0,00003	0,4713

La regressione multipla ha permesso di visualizzare la superficie su cui giacciono i punti di errore relativo rispetto a step energetico e potenza del motore elettrico. Si nota che i punti vengono rappresentati fedelmente dalla superficie soprattutto per valori di errore relativo inferiore a 10. Si nota osservando i valori p che i termini significativamente statistici sono i termini legati allo step energetico.

Regressione multipla: taglia della macchina elettrica rapportata a C rate di scarica

È stata eseguita la regressione multipla per descrivere l'errore relativo in base a step energetico ed EM1P/Crate di scarica: si va a rapportare la potenza della macchina elettrica al C rate di scarica in modo da ottenere la seconda variabile esplicativa per la descrizione dell'errore relativo oltre allo step energetico. L'equazione utilizzata per la regressione è la medesima espressa nell'equazione 8.1, in questo caso x_1 rappresenta lo step energetico ed x_2 rappresenta la variabile potenza della macchina elettrica rapportata al C Rate di scarica. In figura 8.2 si possono visualizzare le superfici ottenute tramite una vista tridimensionale ed una vista secondo il piano $y-x_1$. Entrambe le viste sono state riportate utilizzando l'asse logaritmico per lo step energetico. La tabella 8.2 riporta i coefficienti ricavati insieme ai rispettivi valori p. Si riporta inoltre per questo tipo di regressione il valore p per il test F, il coefficiente di determinazione, il coefficiente di determinazione corretto ed il valore di RMSE.

Root Mean Squared Error	1,19
R-squared	0,995
Adjusted R-Squared	0,995
p-value	0,000

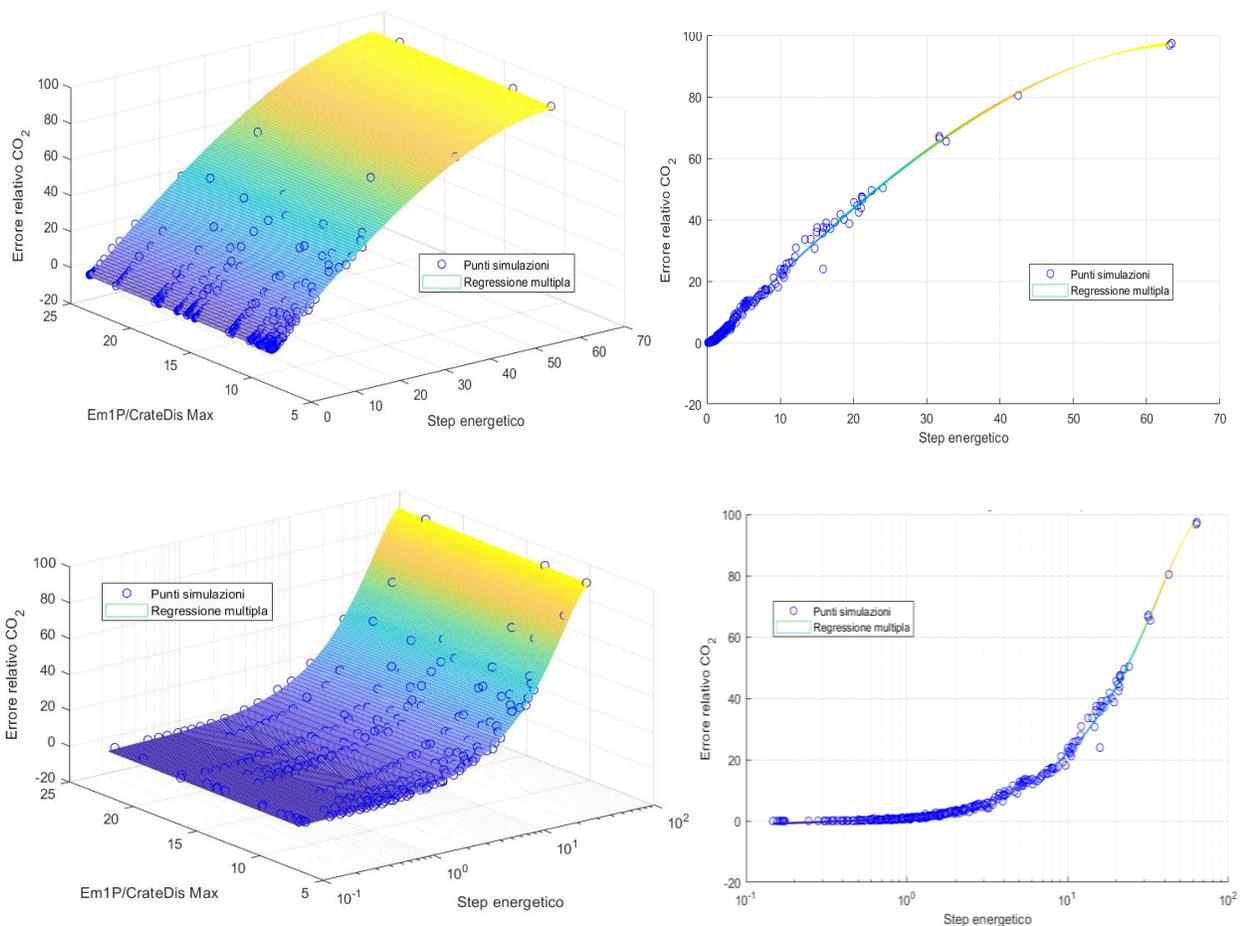


Figura 8.2 – Regressione multipla, viste in tre dimensioni e viste del piano $y-x_1$.

Tabella 8.2 – Coefficienti stimati

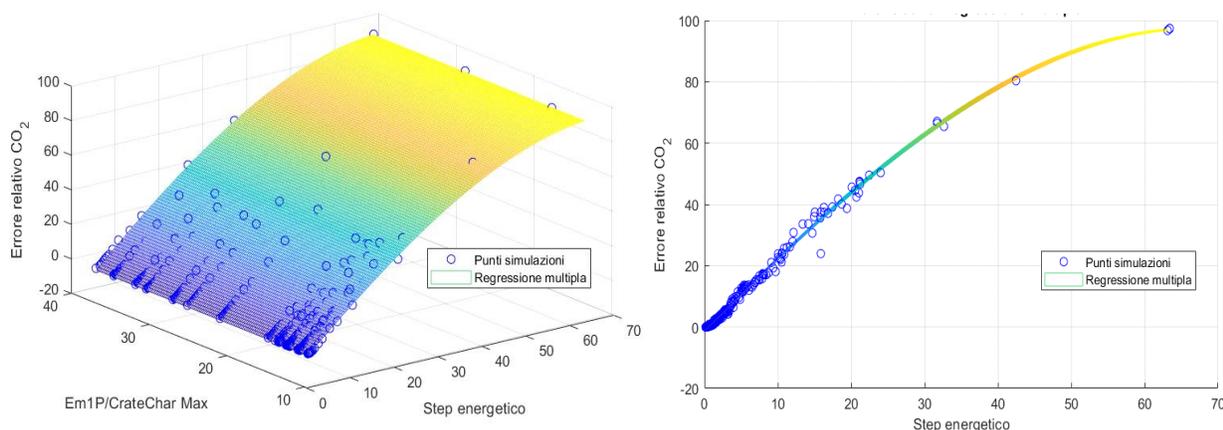
Termine	Valore	pValue
Intercetta	-1,7136	0,0000
X ₁	2,4994	0,0000
X ₂	0,0307	0,1348
X ₁ ²	-0,0073	0,0038
X ₁ X ₂	-0,0059	0,1214
X ₁ ³	-0,0001	0,0000
X ₁ ² X ₂	0,0001	0,1579

La regressione multipla ha permesso di visualizzare la superficie su cui giacciono i punti di errore relativo rispetto a step energetico e potenza del motore elettrico. Si nota che i punti vengono rappresentati fedelmente dalla superficie soprattutto per valori di errore relativo inferiore a 10. Si nota osservando i valori p che i termini significativamente statistici sono i termini legati allo step energetico anche per questo secondo tipo di regressione.

Regressione multipla: taglia della macchina elettrica rapportata a C rate di scarica

In questo paragrafo è stata eseguita la regressione multipla per descrivere l'errore relativo in base a step energetico e potenza della macchina elettrica questa volta rapportata al valore C Rate di carica. L'equazione utilizzata per la regressione è la medesima espressa nell'equazione 8.1, in questo caso x₁ rappresenta lo step energetico ed x₂ rappresenta la variabile potenza della macchina elettrica rapportata al C Rate di carica. In figura 8.3 si possono visualizzare le superfici ottenute tramite una vista tridimensionale ed una vista secondo il piano y-x₁. Entrambe le viste sono state riportate utilizzando l'asse logaritmico per lo step energetico. La tabella 8.3 riporta i coefficienti ricavati insieme ai rispettivi valori p. Si riporta inoltre per questo tipo di regressione il valore p per il test F, il coefficiente di determinazione, il coefficiente di determinazione corretto ed il valore di RMSE.

Root Mean Squared Error	1,18
R-squared	0,995
Adjusted R-Squared	0,995
p-value	0,000



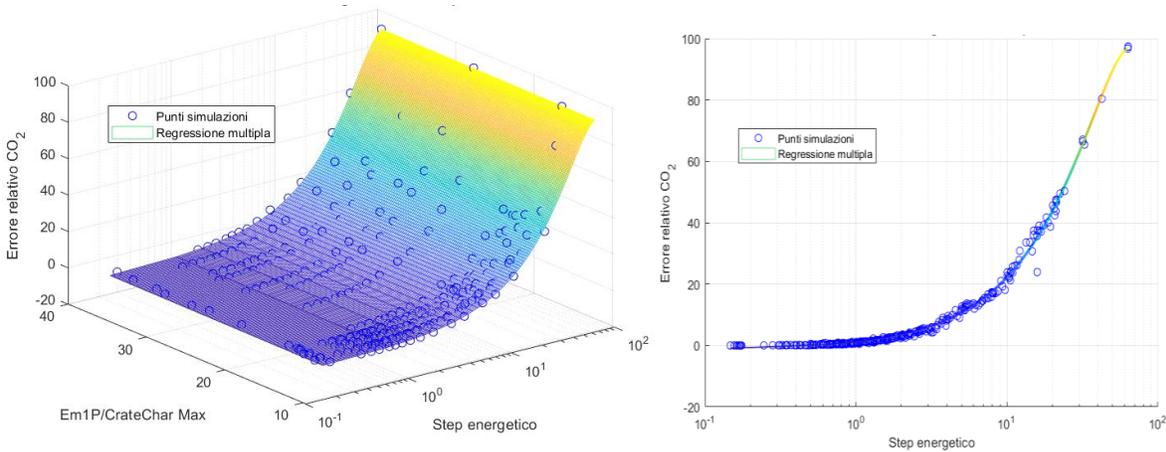


Figura 8.3 – Regressione multipla, viste in tre dimensioni e viste del piano $y-x_1$.

Tabella 8.3 – Coefficienti stimati

Termine	Valore	pValue
Intercetta	-1,4214	0,0000
X_1	2,3598	0,0000
X_2	0,0064	0,5438
X_1^2	-0,0054	0,0223
$X_1 X_2$	0,0032	0,1020
X_1^3	-0,0001	0,0000
$X_1^2 X_2$	-0,0000	0,2171

La regressione multipla ha permesso di visualizzare la superficie su cui giacciono i punti di errore relativo rispetto a step energetico e potenza del motore elettrico rapportata a Crate Char. Si nota come i punti vengano rappresentati fedelmente dalla superficie. Si nota che tenendo conto anche dei parametri Crate e non solo della potenza del motore elettrico si ottiene una superficie più vicina ai punti risultato delle simulazioni, la regressione è più accurata. Anche in questo caso i termini con maggiore significatività statistica sono quelli legati allo step energetico.

Regressione multipla: taglia della batteria

In questo paragrafo è stata eseguita la regressione multipla per descrivere l'errore relativo in base a step energetico e taglia della batteria. L'equazione utilizzata per la regressione è espressa nell'equazione 8.1, in questo caso x_1 rappresenta lo step energetico ed x_2 rappresenta la variabile taglia della batteria. In figura 8.4 si possono visualizzare le superfici ottenute tramite una vista tridimensionale ed una vista secondo il piano $y-x_1$. Entrambe le viste sono state riportate utilizzando l'asse logaritmico per lo step energetico. La tabella 8.4 riporta i coefficienti ricavati insieme ai rispettivi valori p. Si riporta inoltre per questo tipo di regressione il valore p per il test F, il coefficiente di determinazione, il coefficiente di determinazione corretto ed il valore di RMSE.

Root Mean Squared Error	1,18
R-squared	0,995
Adjusted R-Squared	0,995
p-value	0,000

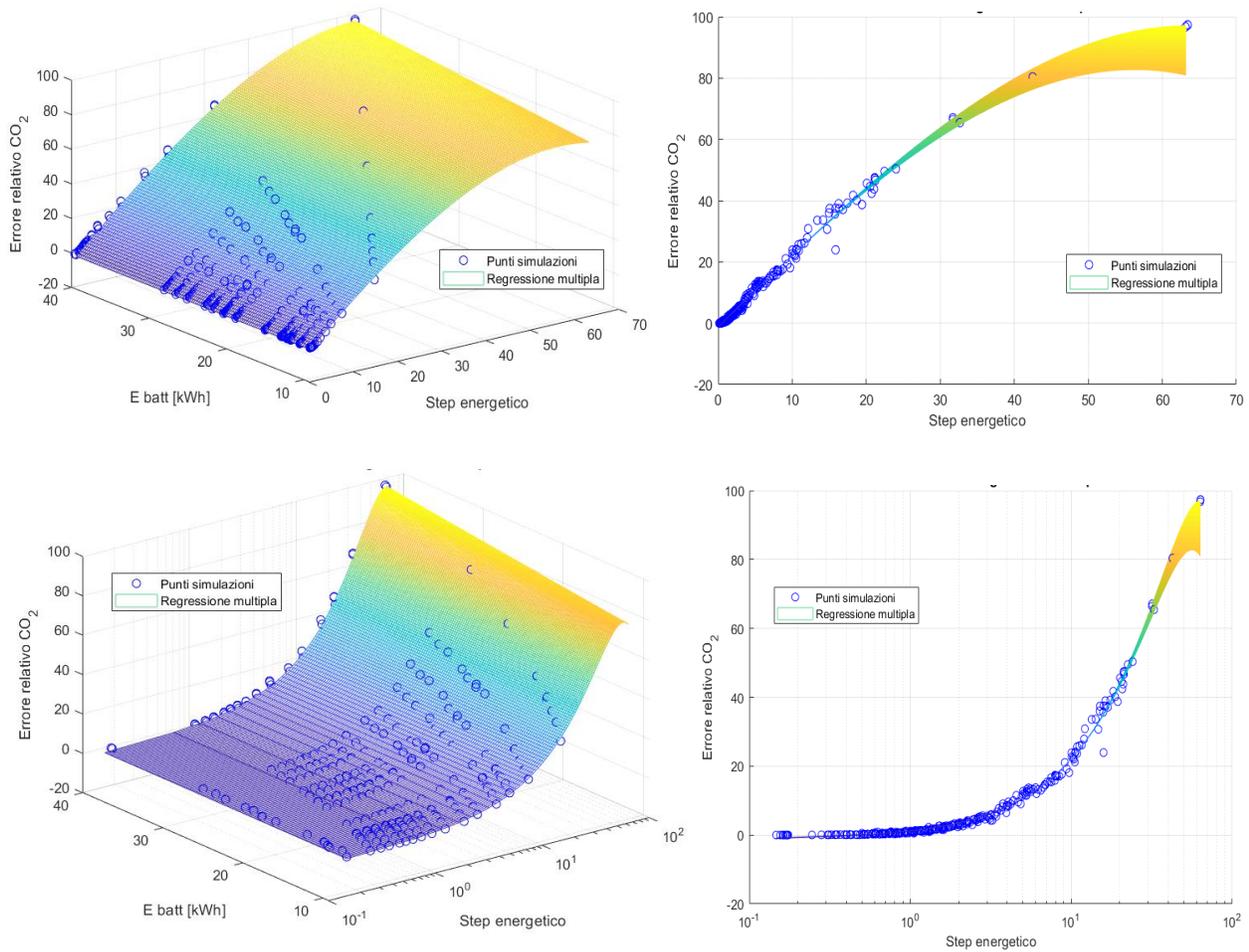


Figura 8.4 – Regressione multipla, viste in tre dimensioni e viste del piano $y-x_1$.

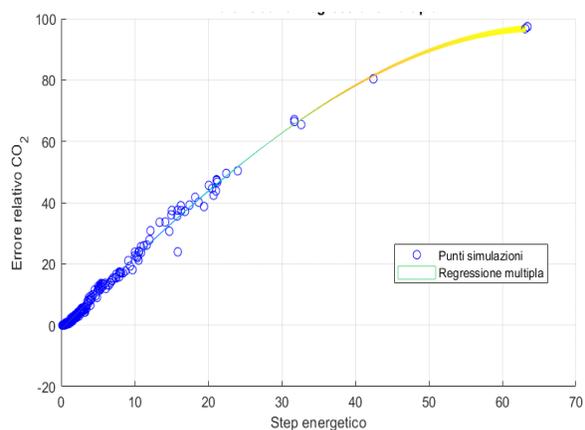
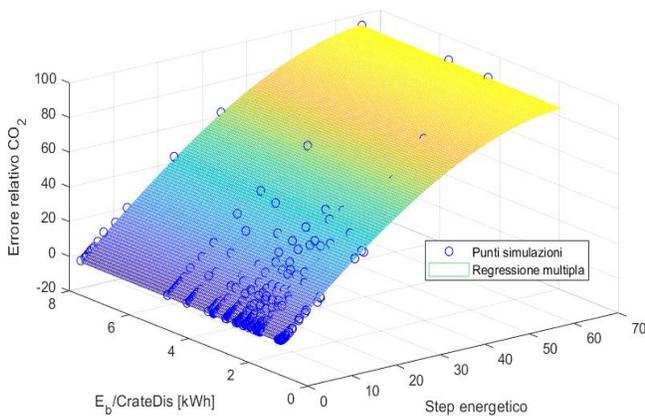
Termine	Valore	pValue
Intercetta	-1,1945	0,0000
X ₁	2,4324	0,0000
X ₂	-0,0050	0,6650
X ₁ ²	-0,0079	0,0020
X ₁ X ₂	-0,0009	0,7600
X ₁ ³	-0,0002	0,0007
X ₁ ² X ₂	0,0001	0,2124

La regressione multipla ha permesso di visualizzare la superficie su cui giacciono i punti di errore relativo rispetto a step energetico e taglia della batteria. Si nota come i punti risultato delle simulazioni vengano ben rappresentati dal piano della regressione soprattutto per valori di step energetico inferiori a 10. Anche in questo caso la variabile più significativa statisticamente è lo step energetico. Si vede che la superficie di regressione va comunque a rappresentare bene i dati, la non significatività statistica può derivare dal fatto che step energetico e taglia della batteria sono due variabili altamente correlate, per cui si ottengono pvalue maggiori di 0,05.

Regressione multipla: taglia della batteria rapportata a C rate di scarica

In questo paragrafo è stata eseguita la regressione multipla per descrivere l'errore relativo in base a step energetico e taglia della batteria rapportata a C rate di scarica. L'equazione utilizzata per la regressione è espressa nell'equazione 8.1, in questo caso x_1 rappresenta lo step energetico ed x_2 rappresenta la variabile taglia della batteria rapportata al C rate di scarica. In figura 8.5 si possono visualizzare le superfici ottenute tramite una vista tridimensionale ed una vista secondo il piano $y-x_1$. Entrambe le viste sono state riportate utilizzando l'asse logaritmico per lo step energetico. La tabella 8.5 riporta i coefficienti ricavati insieme ai rispettivi valori p. Si riporta inoltre per questo tipo di regressione il valore p per il test F, il coefficiente di determinazione, il coefficiente di determinazione corretto ed il valore di RMSE.

Root Mean Squared Error	1,19
R-squared	0,995
Adjusted R-Squared	0,995
p-value	0,000



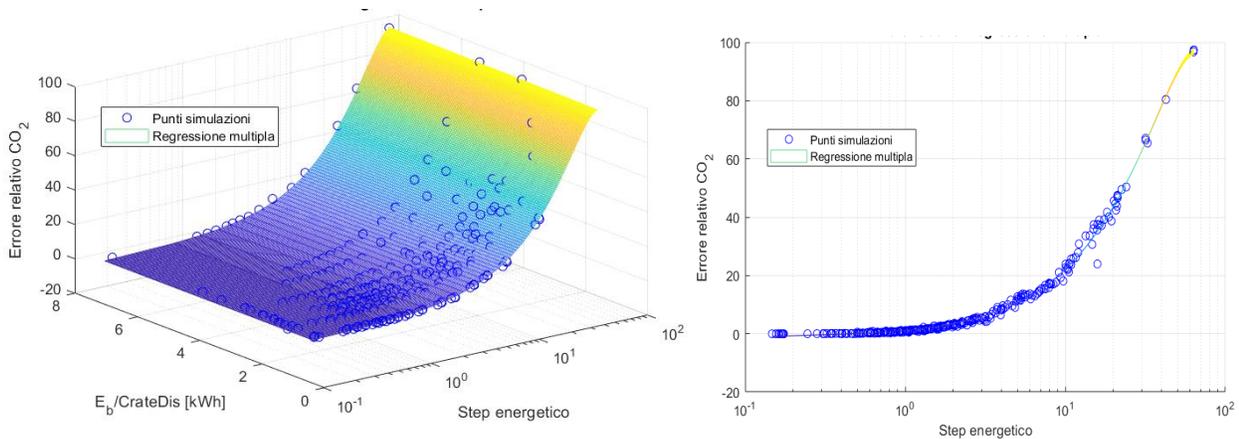


Figura 8.5 – Regressione multipla, viste in tre dimensioni e viste del piano $y-x_1$.

Tabella 8.5 – Coefficienti stimati

Termine	Valore	pValue
Intercetta	-1,3262	0,0000
X ₁	2,4279	0,0000
X ₂	0,0104	0,8732
X ₁ ²	-0,0059	0,0121
X ₁ X ₂	-0,0043	0,6915
X ₁ ³	-0,0001	0,0000
X ₁ ² X ₂	0,0001	0,50941

La regressione multipla ha permesso di visualizzare la superficie su cui giacciono i punti di errore relativo rispetto a step energetico e taglia della batteria rapportata a C rate di scarica. Si nota come i punti risultato delle simulazioni vengano ben rappresentati dal piano della regressione soprattutto per valori di step energetico inferiori a 10. Anche in questo caso i termini con maggiore significatività statistica sono quelli legati allo step energetico, si ricorda inoltre che la variabile taglia della batteria è legata allo step energetico.

8.2 Regressioni multiple senza outliers

Regressione multipla: taglia della macchina elettrica

Sono state ripetute le regressioni dei paragrafi precedenti andando ad escludere gli outliers dai dati. Come visto al capitolo 6, sono stati esclusi i dati che si collocano prima del ginocchio della curva CO₂-N SOC. Viene quindi eseguita una regressione multipla tenendo conto di step energetico e potenza della macchina elettrica in modo da rappresentare tramite una superficie l'andamento dell'errore relativo. La superficie viene descritta tramite l'equazione 8.1, in questo caso x_1 rappresenta lo step energetico ed x_2 rappresenta la variabile potenza della macchina elettrica. In figura 8.6 si possono visualizzare le superfici ottenute tramite una vista tridimensionale ed una vista secondo il piano $y-x_1$. Entrambe le viste sono state riportate utilizzando l'asse logaritmico per lo step energetico. La tabella 8.6 riporta i coefficienti ricavati insieme

ai rispettivi valori p. Si riporta inoltre per questo tipo di regressione il valore p per il test F, il coefficiente di determinazione, il coefficiente di determinazione corretto ed il valore di RMSE.

Root Mean Squared Error	0,348
R-squared	0,982
Adjusted R-Squared	0,981
p-value	0,000

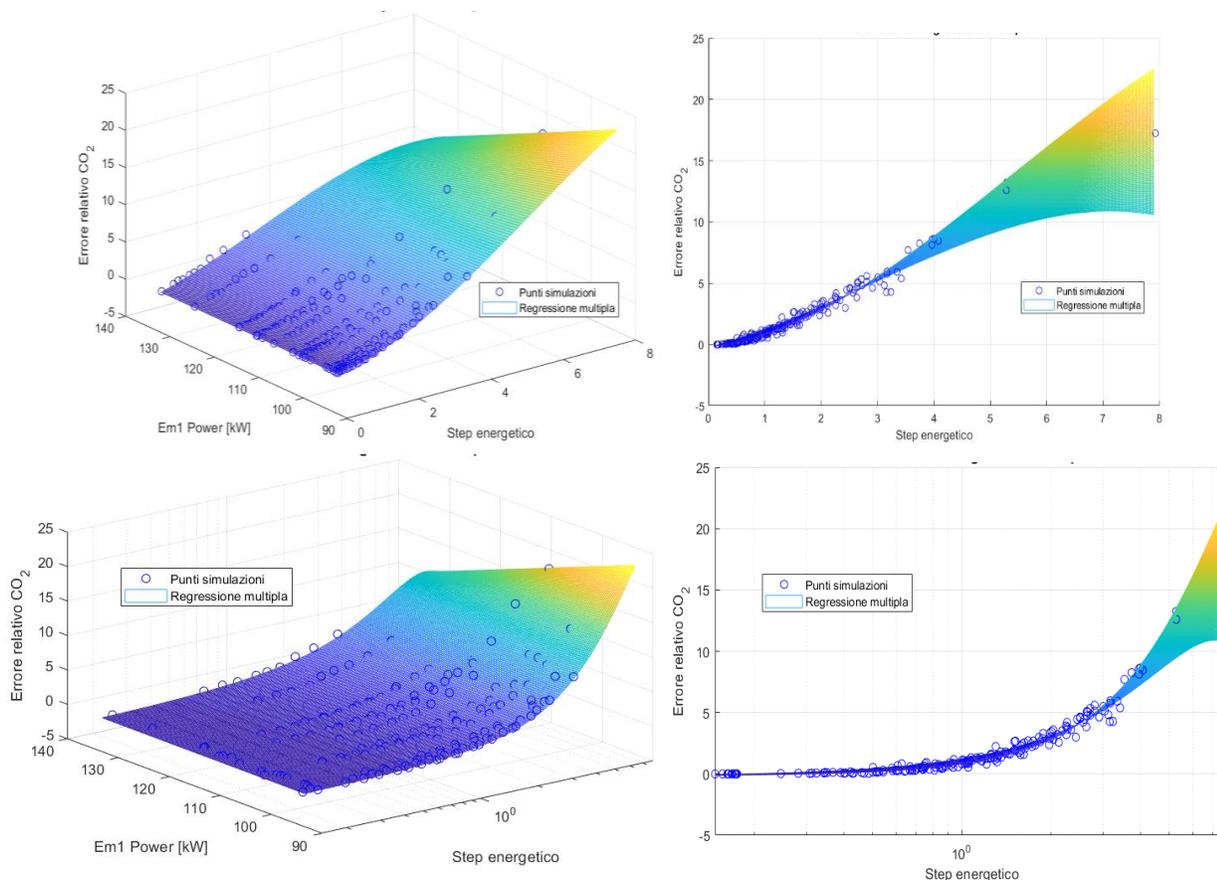


Figura 8.6 – Regressione multipla, viste in tre dimensioni e viste del piano y-x₁.

Tabella 8.6 – Coefficienti stimati

Termine	Valore	pValue
Intercetta	-0,0912	0,8818
X ₁	-1,329	0,1019
X ₂	-0,0011	0,8532
X ₁ ²	1,2788	0,0000
X ₁ X ₂	0,0189	0,0169
X ₁ ³	-0,0381	0,0000
X ₁ ² X ₂	-0,0071	0,0010

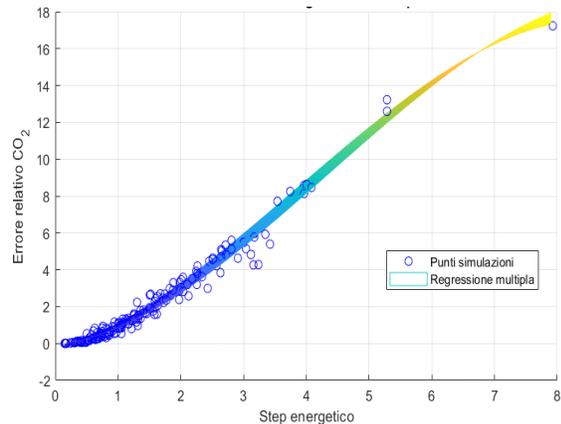
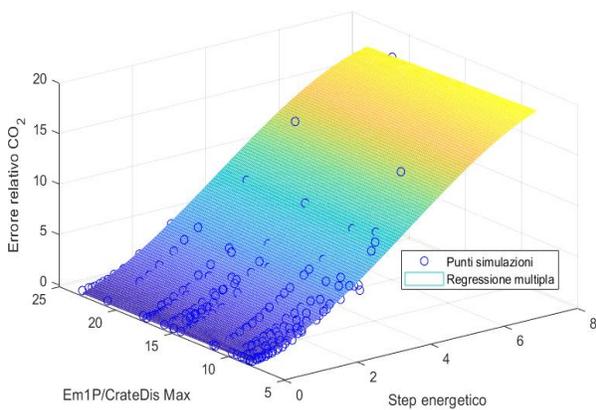
La superficie diventa più irregolare rispetto alla superficie ottenuta per questo tipo di regressione considerando i dati completi ed inoltre si ottengono valori di R² sono inferiori. La superficie di

regressione va comunque a rappresentare in modo fedele i dati. Anche in questo caso si ottiene come variabile statisticamente significativa lo step energetico.

Regressione multipla: taglia della macchina elettrica rapportata a C rate di scarica

Viene eseguita una regressione multipla tenendo conto di step energetico e potenza della macchina elettrica rapportata a C rate di scarica, in modo da rappresentare tramite una superficie l'andamento dell'errore relativo in funzione di step energetico e della variabile appena introdotta. La superficie, i cui coefficienti sono il risultato della regressione, presenta l'equazione indicata dalla 8.1. In questo caso x_1 rappresenta lo step energetico ed x_2 rappresenta la variabile potenza della macchina elettrica rapportata a C rate di scarica. In figura 8.7 si possono visualizzare le superfici ottenute tramite una vista tridimensionale ed una vista secondo il piano $y-x_1$. Entrambe le viste sono state riportate utilizzando l'asse logaritmico per lo step energetico. La tabella 8.7 riporta i coefficienti ricavati insieme ai rispettivi valori p. Si riporta inoltre per questo tipo di regressione il valore p per il test F, il coefficiente di determinazione, il coefficiente di determinazione corretto ed il valore di RMSE.

Root Mean Squared Error	0,35
R-squared	0,982
Adjusted R-Squared	0,981
p-value	0,000



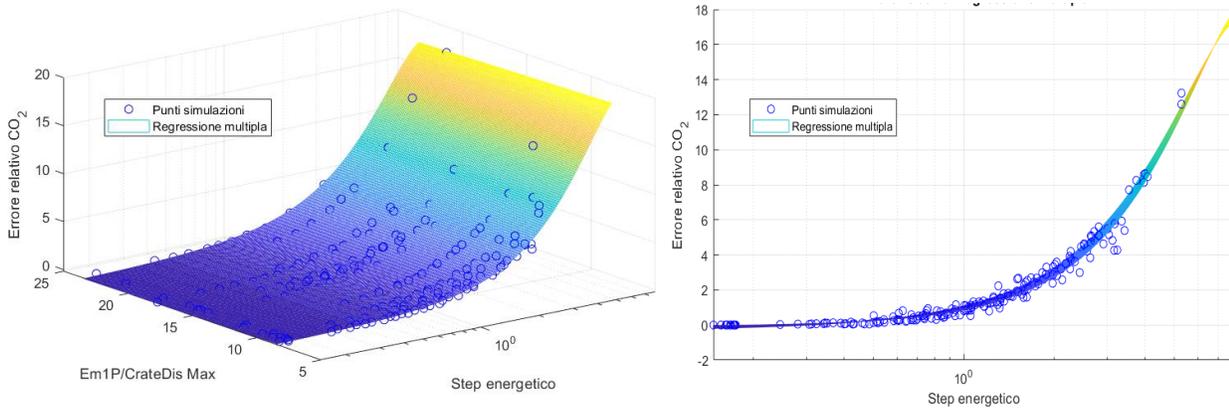


Figura 8.7 – Regressione multipla, viste in tre dimensioni e viste del piano y-x1.

Tabella 8.7 – Coefficienti stimati

Termine	Valore	pValue
Intercetta	0,0145	0,9339
X ₁	0,1546	0,4435
X ₂	-0,0152	0,2252
X ₁ ²	0,6230	0,0000
X ₁ X ₂	0,0348	0,0231
X ₁ ³	-0,0443	0,0000
X ₁ ² X ₂	-0,0048	0,1849

In questa regressione si ottengono valori di R² inferiori rispetto alla corrispondente regressione con i dati completi. La superficie di regressione va comunque a rappresentare in modo fedele i dati. Anche in questo caso si ottiene come variabile statisticamente significativa lo step energetico.

Regressione multipla: taglia della macchina elettrica rapportata a C rate di carica

In questo caso la regressione eseguita tiene conto dello step energetico e della potenza della macchina elettrica rapportata al valore C rate di carica. La superficie, i cui coefficienti sono il risultato della regressione, presenta l'equazione indicata dalla 8.1. In questo caso x₁ rappresenta lo step energetico ed x₂ rappresenta la variabile potenza della macchina elettrica rapportata a C rate di carica. In figura 8.8 si possono visualizzare le superfici ottenute tramite una vista tridimensionale ed una vista secondo il piano y-x₁. Entrambe le viste sono state riportate utilizzando l'asse logaritmico per lo step energetico. La tabella 8.8 riporta i coefficienti ricavati insieme ai rispettivi valori p. Si riporta inoltre per questo tipo di regressione il valore p per il test F, il coefficiente di determinazione, il coefficiente di determinazione corretto ed il valore di RMSE.

Root Mean Squared Error	0,328
R-squared	0,984
Adjusted R-Squared	0,983
p-value	0,000

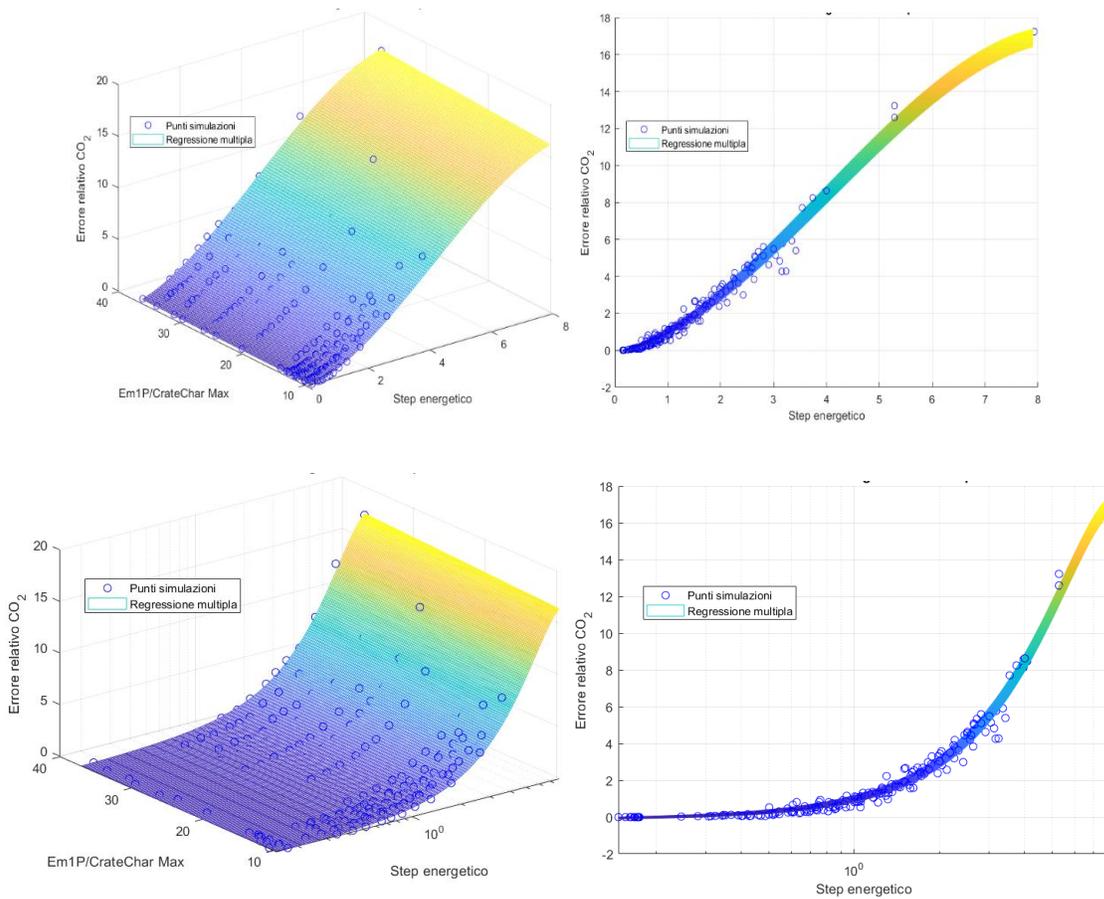


Figura 8.8 – Regressione multipla, viste in tre dimensioni e viste del piano $y-x_1$.

Tabella 8.8 – Coefficienti stimati

Termine	Valore	pValue
Intercetta	-0,1984	0,1184
X_1	0,3459	0,0322
X_2	0,0021	0,7607
X_1^2	0,5896	0,0000
$X_1 X_2$	0,0124	0,1384
X_1^3	-0,0472	0,0000
$X_1^2 X_2$	-0,0010	0,6190

La regressione ha permesso di visualizzare la superficie che lega errore relativo a step energetico e taglia della macchina elettrica rapportata a C rate di carica. La variabile statisticamente più significativa risulta essere lo step energetico.

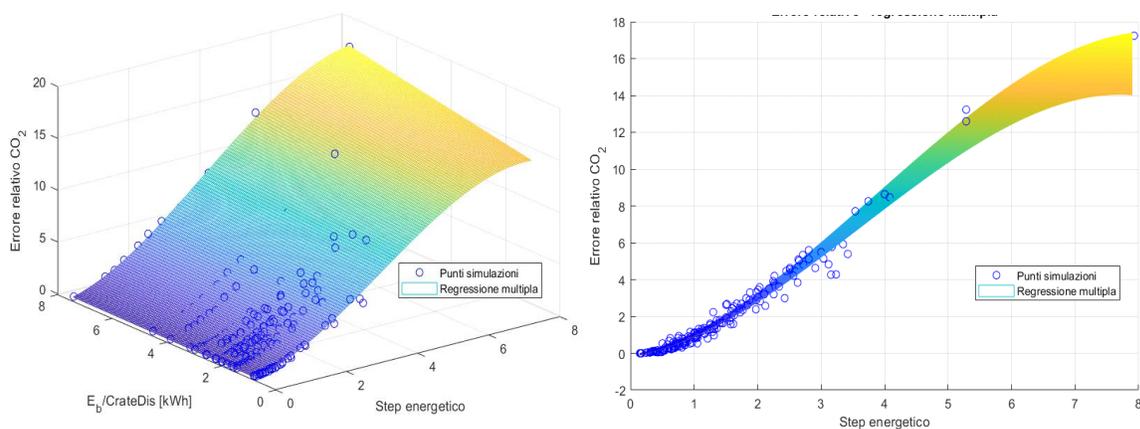
Termine	Valore	pValue
Intercetta	-0,4029	0,0063
X ₁	0,8905	0,0000
X ₂	0,0117	0,1076
X ₁ ²	0,4949	0,0000
X ₁ X ₂	-0,0156	0,0702
X ₁ ³	-0,0588	0,0000
X ₁ ² X ₂	0,0053	0,0129

La regressione ha permesso di visualizzare la superficie che lega errore relativo a step energetico e taglia della batteria. Anche in questo caso si ottengono valori di R² inferiori rispetto al caso in cui si considerano tutti i dati. La variabile statisticamente più significativa risulta essere lo step energetico.

Regressione multipla: taglia della batteria rapportata a C rate di scarica

In questo caso la regressione eseguita tiene conto dello step energetico e della taglia della batteria rapportata al valore di C rate di scarica. La superficie presenta l'equazione indicata dalla 8.1. In questo caso x₁ rappresenta lo step energetico ed x₂ rappresenta la variabile taglia della batteria rapportata a C rate di scarica. In figura 8.10 si possono visualizzare le superfici ottenute tramite una vista tridimensionale ed una vista secondo il piano y-x₁. Entrambe le viste sono state riportate utilizzando l'asse logaritmico per lo step energetico. La tabella 8.10 riporta i coefficienti ricavati insieme ai rispettivi valori p. Si riporta inoltre per questo tipo di regressione il valore p per il test F, il coefficiente di determinazione, il coefficiente di determinazione corretto ed il valore di RMSE.

Root Mean Squared Error	0,347
R-squared	0,982
Adjusted R-Squared	0,981
p-value	0,000



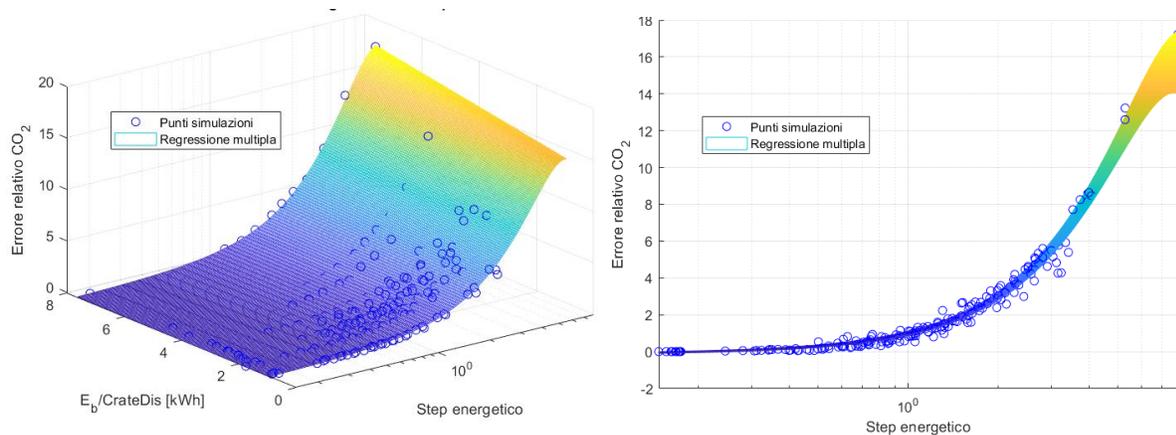


Figura 8.10 – Regressione multipla, viste in tre dimensioni e viste del piano y-x1.

Tabella 8.10 – Coefficienti stimati

Termine	Valore	pValue
Intercetta	-0,1929	0,0839
X1	0,5239	0,0001
X2	0,0102	0,8164
X1 ²	0,5740	0,0000
X1X2	0,0197	0,6727
X1 ³	-0,0532	0,0000
X1 ² X2	0,0054	0,6105

La regressione ha permesso di visualizzare la superficie che lega errore relativo a step energetico e taglia della batteria rapportata a C rate di scarica. La variabile statisticamente più significativa risulta essere lo step energetico.

8.3 Regressioni multiple a tre variabili

Regressione multipla: taglia della macchina elettrica e taglia della batteria

Poiché ai paragrafi precedenti sono state evidenziate le dipendenze dell'errore relativo da parametri quali taglia della macchina elettrica e taglia della batteria potrebbe essere interessante ricercare un'unica equazione in grado di descrivere l'andamento dell'errore relativo a partire da tre variabili esplicative: step energetico, potenza della macchina elettrica, taglia della batteria. Per tale tipo di regressione viene utilizzata l'equazione 8.2:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_1^2 + \beta_5 x_1 x_2 + \beta_6 x_1 x_3 + \beta_7 x_2 x_3 + \beta_8 x_1^3 + \beta_9 x_1^2 x_2 + \beta_{10} x_1^2 x_3 + \beta_{11} x_1 x_2 x_3 \quad (8.2)$$

In cui x_1 rappresenta lo step energetico, x_2 rappresenta la potenza della macchina elettrica ed x_3 rappresenta la taglia della batteria. Si riportano per questo tipo di regressione il valore p per il test F, il coefficiente di determinazione, il coefficiente di determinazione corretto ed il valore di RMSE.

Root Mean Squared Error	1,15
R-squared	0,995
Adjusted R-Squared	0,995
p-value	0,000

In figura 8.11 vengono riportate le viste in tre dimensioni della superficie di regressione ottenuta. La tabella 8.11 riporta il valore dei coefficienti ottenuti insieme ai pValue.

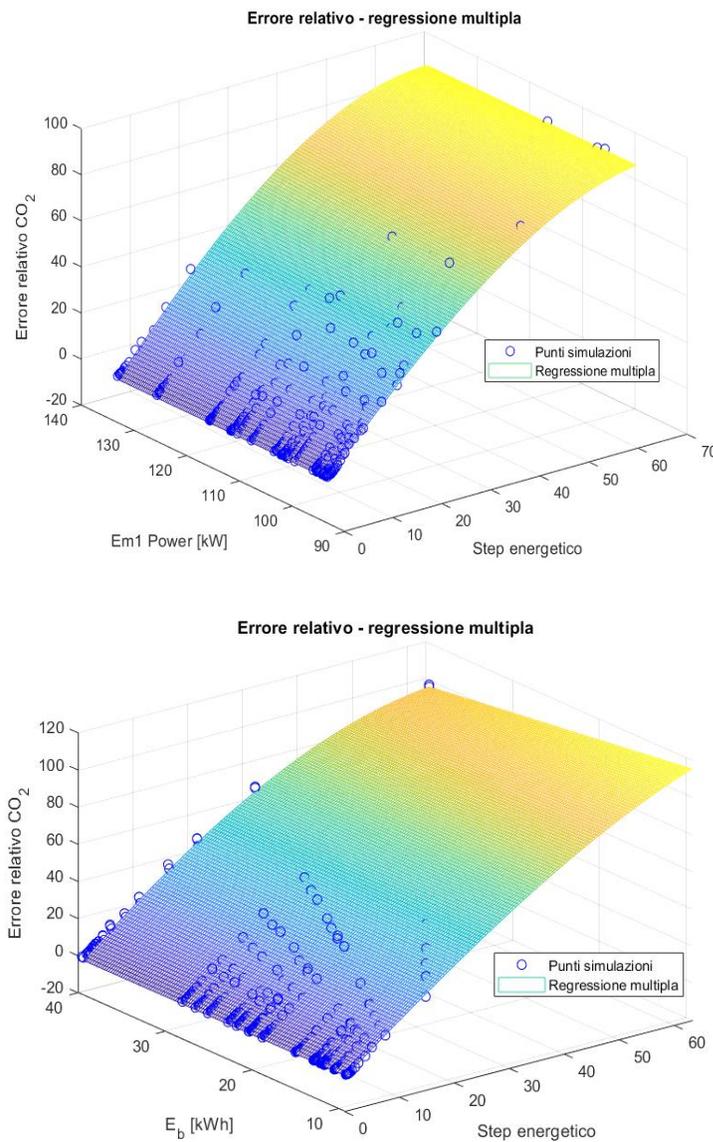


Figura 8.11 – Regressione multipla con tre variabili, viste in tre dimensioni

<i>Tabella 8.11 – Coefficienti stimati</i>		
<i>Termine</i>	<i>Valore</i>	<i>pValue</i>
Intercetta	3,1889	0,2501
X ₁	1,1204	0,0017
X ₂	-0,0426	0,1162
X ₃	-0,2617	0,0660
X ₁ ²	-0,0241	0,0059
X ₁ X ₂	0,0124	0,0002
X ₁ X ₃	0,0637	0,0009
X ₂ X ₃	0,0025	0,0722
X ₁ ³	-0,0001	0,0002
X ₁ ² X ₂	0,0001	0,0349
X ₁ ² X ₃	0,0001	0,2875
X ₁ X ₂ X ₃	-0,0006	0,0006

Questo tipo di regressione mostra il legame tra taglia della batteria, potenza del motore elettrico ed errore relativo di CO₂. Andando ad osservare i pValue dei coefficienti ottenuti si può notare come abbiano significatività statistica i termini di interazione oltre ai termini legati puramente allo step energetico. Si ottiene un valore di R² alto per cui si può dire che la superficie di regressione si adatta ai dati.

Lo stesso tipo di regressione è stata ripetuta andando a considerare i dati con l'esclusione degli outliers. Si riporta il valore p per il test F, il coefficiente di determinazione, il coefficiente di determinazione corretto ed il valore di RMSE.

Root Mean Squared Error	0,306
R-squared	0,986
Adjusted R-Squared	0,985
p-value	0,000

In figura 8.12 vengono riportate le viste in tre dimensioni della superficie di regressione ottenuta. La tabella 8.12 riporta il valore dei coefficienti ottenuti insieme ai pValue.

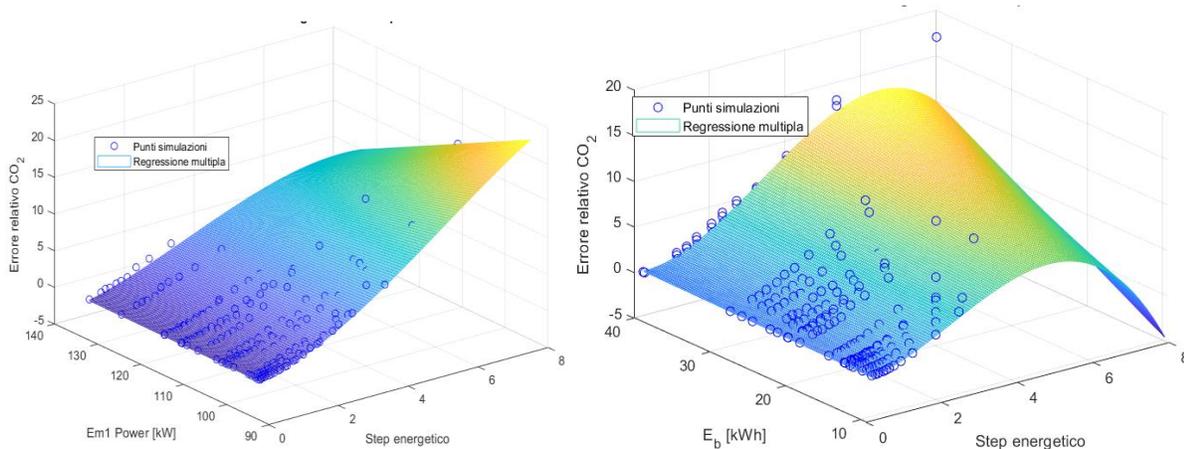


Figura 8.12 – Regressione multipla con tre variabili, viste in tre dimensioni

<i>Tabella 8.12 – Coefficienti stimati</i>		
<i>Termine</i>	<i>Valore</i>	<i>pValue</i>
Intercetta	-1,16	0,3203
X ₁	0,7903	0,3283
X ₂	0,0080	0,4891
X ₃	0,0917	0,1597
X ₁ ²	1,8484	0,0000
X ₁ X ₂	-0,0005	0,9512
X ₁ X ₃	-0,1829	0,0000
X ₂ X ₃	-0,0008	0,1998
X ₁ ³	-0,0434	0,0000
X ₁ ² X ₂	-0,0130	0,0000
X ₁ ² X ₃	0,0024	0,2020
X ₁ X ₂ X ₃	0,0017	0,0000

Differentemente dal caso in cui si considerano tutti i dati, in questa regressione si evidenzia significatività statistica solamente nei termini legati allo step energetico. I valori di R² sono più alti per la regressione con i dati completi.

8.4 Osservazioni

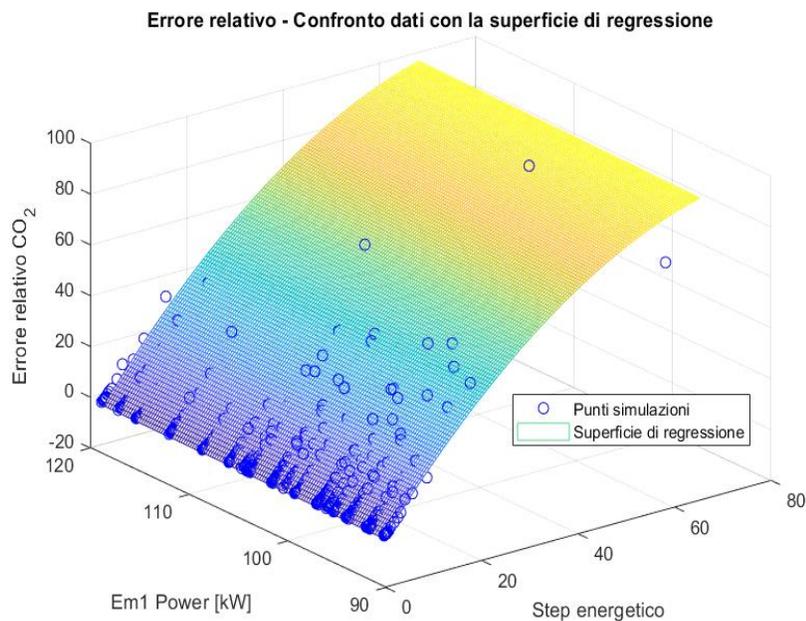
Osservando i modelli ottenuti confrontati ai dati risultato delle simulazioni si può affermare che effettivamente esistono dei legami tra l'errore relativo ed ulteriori variabili oltre allo step energetico. Sono emerse quindi relazioni tra errore relativo e parametri quali taglia della macchina elettrica, taglia della batteria e C rate sia di carica che di scarica. È da notare però che spesso i valori p dei relativi coefficienti non sembrano essere significativi: questo potrebbe essere dovuto al fatto che i dati utilizzati per le regressioni sono stati ricavati andando a modificare tutti i parametri di design e non solo i parametri coinvolti nella regressione, per cui diventa più difficile andare a valutare l'effettiva influenza di un parametro. Inoltre, i valori p vengono calcolati per ogni singolo coefficiente andando a mantenere costanti tutti gli altri coefficienti, da qui, si nota l'influenza di tutti i parametri del modello sulla significatività del singolo coefficiente. I termini di interazione tra le variabili infatti vanno ad influire sui risultati del test t sul singolo coefficiente. Per tale motivo, andare ad aumentare il grado dell'equazione polinomiale può essere da un lato utile a creare un modello più preciso per lo step energetico, dall'altro può essere svantaggioso poiché va ad aumentare i termini di interazione e quindi, può influire sulla significatività statistica dei singoli parametri. Si può pensare inoltre che per creare modelli più precisi sia necessario andare a considerare oltre a queste variabili, ulteriori variabili esplicative ovvero i parametri di design che vanno ad influenzare l'errore relativo in modo da andare a creare modelli più precisi.

9. Test delle superfici di regressione

Obiettivo di questa sezione è andare a verificare se le superfici di regressione ottenute per ogni variabile di design finora considerata si adattano bene a dei dati differenti che sono stati ricavati e non utilizzati per le regressioni stesse. Lo scopo principale è quello di andare a determinare la bontà della superficie di regressione che è stata ricavata con un numero di dati ridotto e quindi verificare se si adatta bene a dei dati ulteriori.

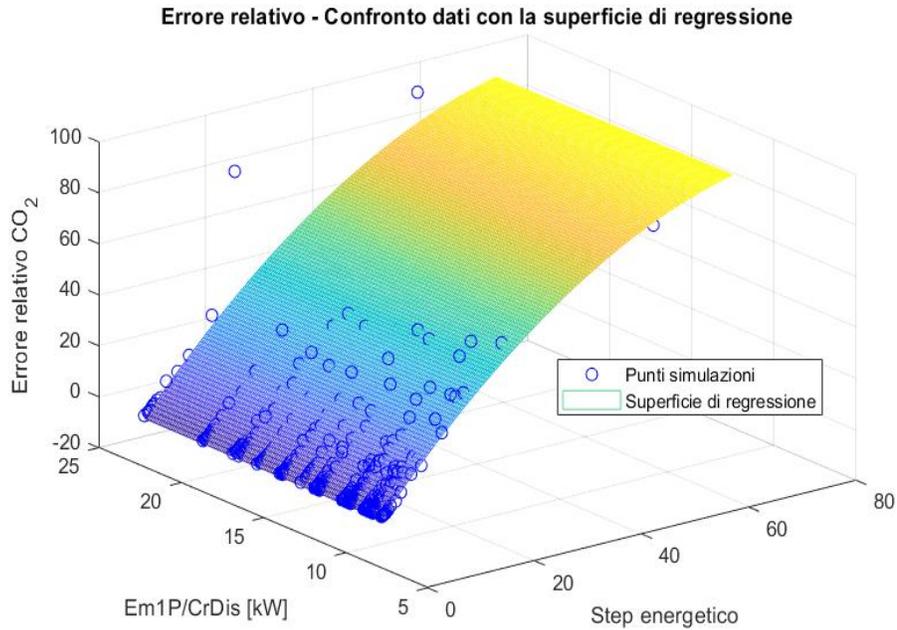
9.1 Superfici di secondo grado

Per ogni regressione svolta al capitolo 7 si riportano nelle immagini da 9.1 a 9.5 le regressioni testate su dati ulteriori insieme al valore di R^2 calcolato. Per tutte le regressioni il valore di R^2 calcolato è dell'ordine di 0,93, inferiore ai valori di R^2 calcolati in ogni regressione.



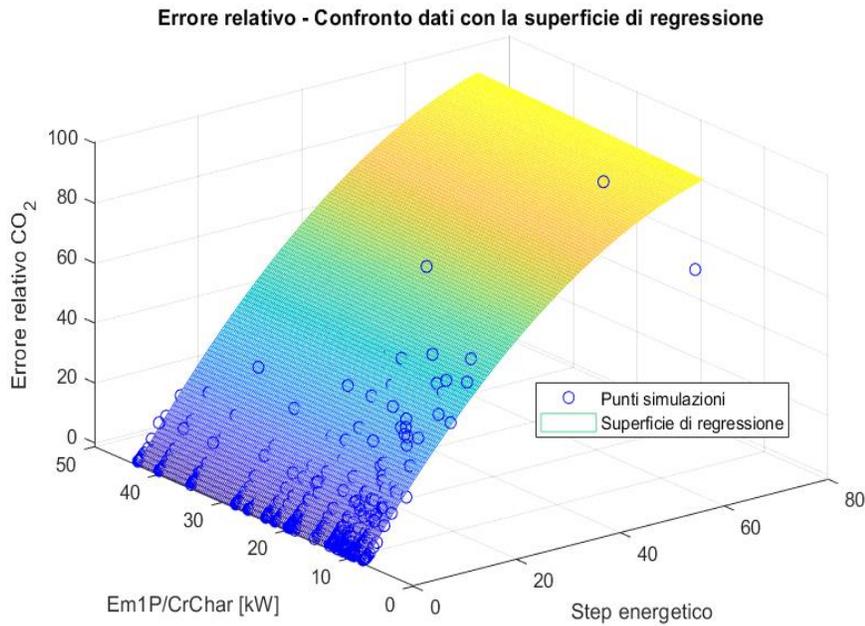
$$R^2 = 0,9305$$

Figura 9.1 – Errore relativo, step energetico, potenza della macchina elettrica



$$R^2 = 0,9300$$

Figura 9.2 – Errore relativo, step energetico, potenza della macchina elettrica rapportata a C rate di scarica



$$R^2 = 0,9308$$

Figura 9.3 – Errore relativo, step energetico, potenza della macchina elettrica rapportata a C rate di carica

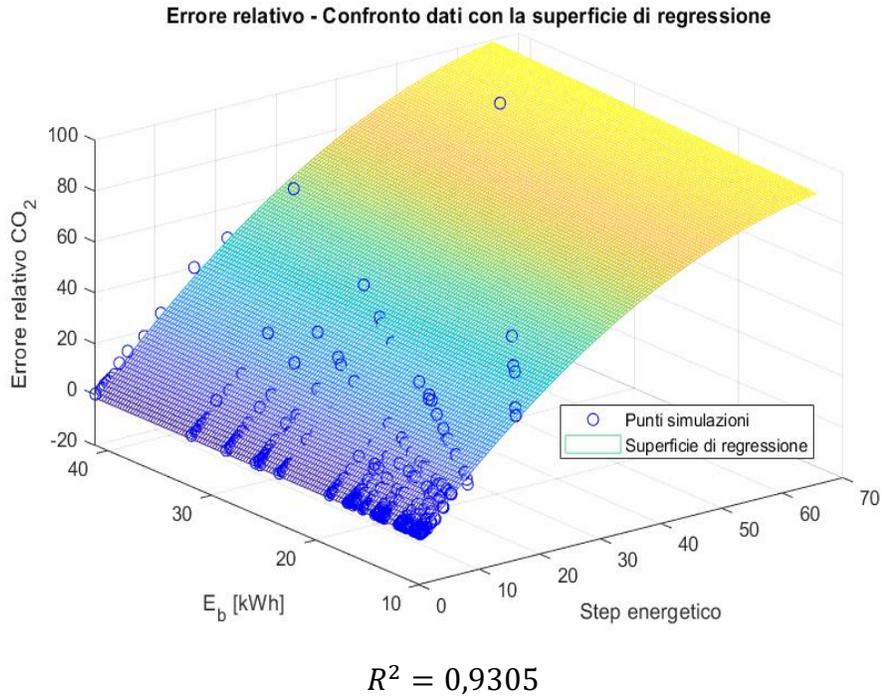


Figura 9.4 – Errore relativo, step energetico, taglia della batteria

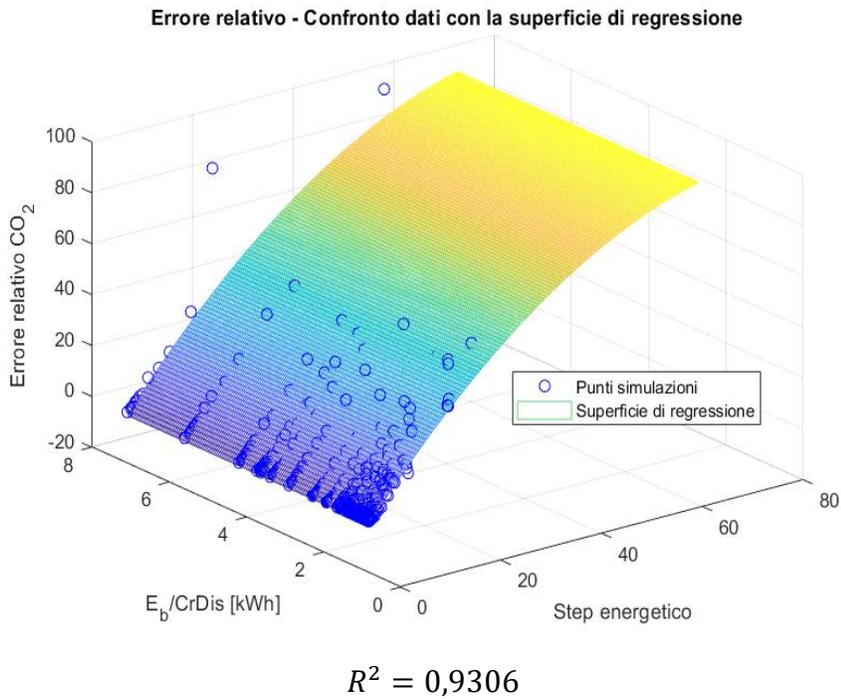


Figura 9.5 – Errore relativo, step energetico, taglia della batteria rapportata a C rate di scarica

9.2 Superfici di terzo grado

Per ogni regressione svolta al capitolo 8 si riportano nelle immagini da 9.6 a 9.10 le regressioni testate su dati ulteriori insieme al valore di R^2 calcolato.

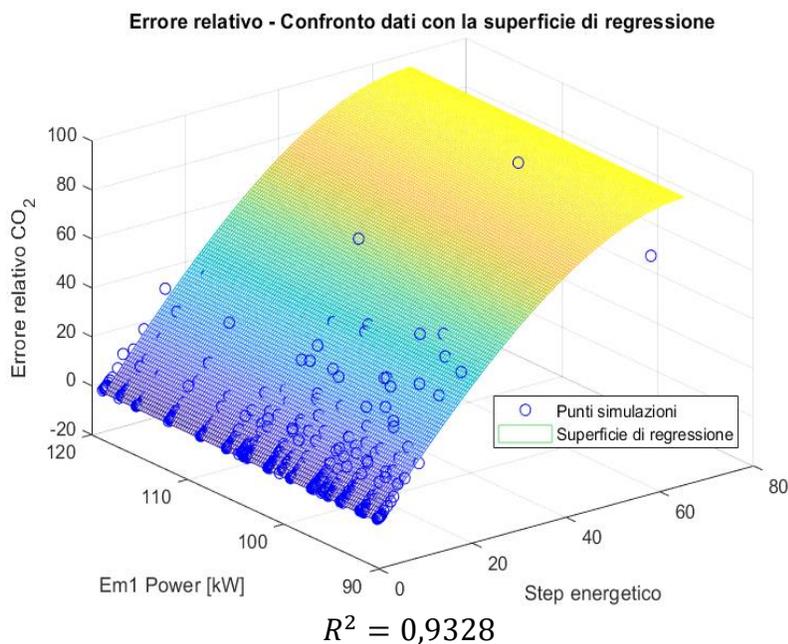
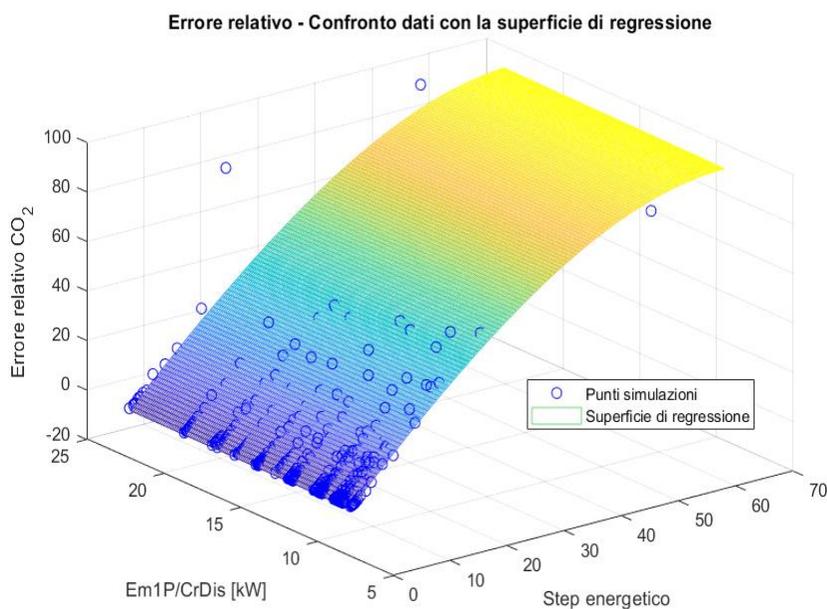


Figura 9.6 – Errore relativo, step energetico, potenza della macchina elettrica



$$R^2 = 0,9328$$

Figura 9.7 – Errore relativo, step energetico, potenza della macchina elettrica rapportata a C rate di scarica

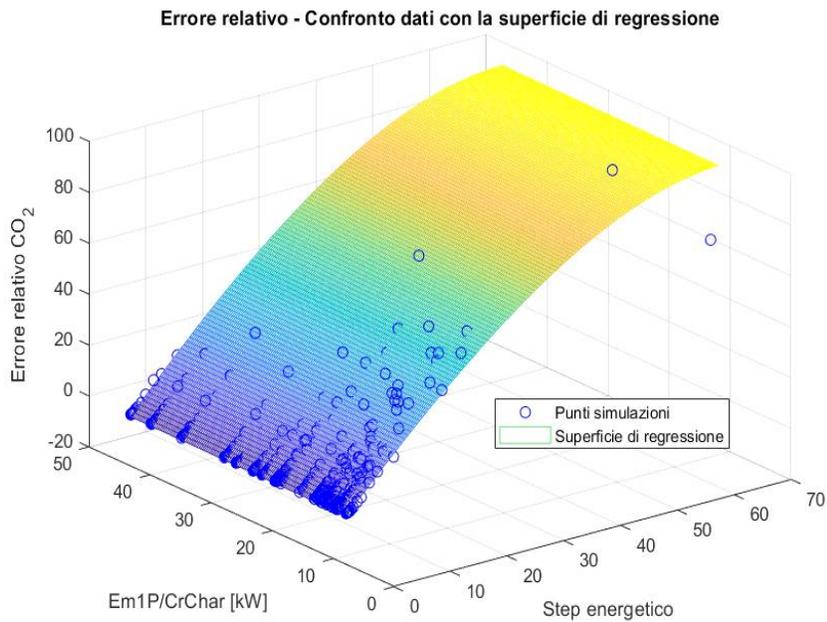


Figura 9.8 – Errore relativo, step energetico, potenza della macchina elettrica rapportata a C rate di carica

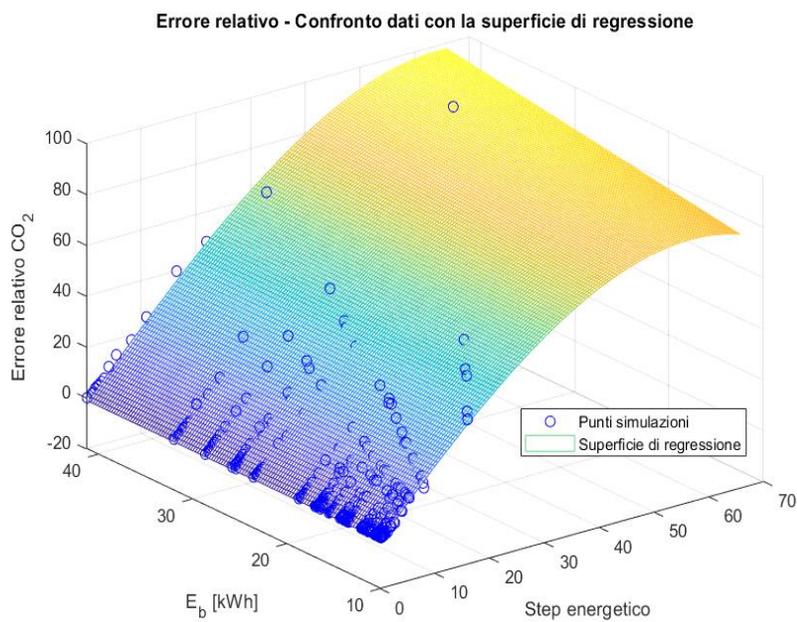
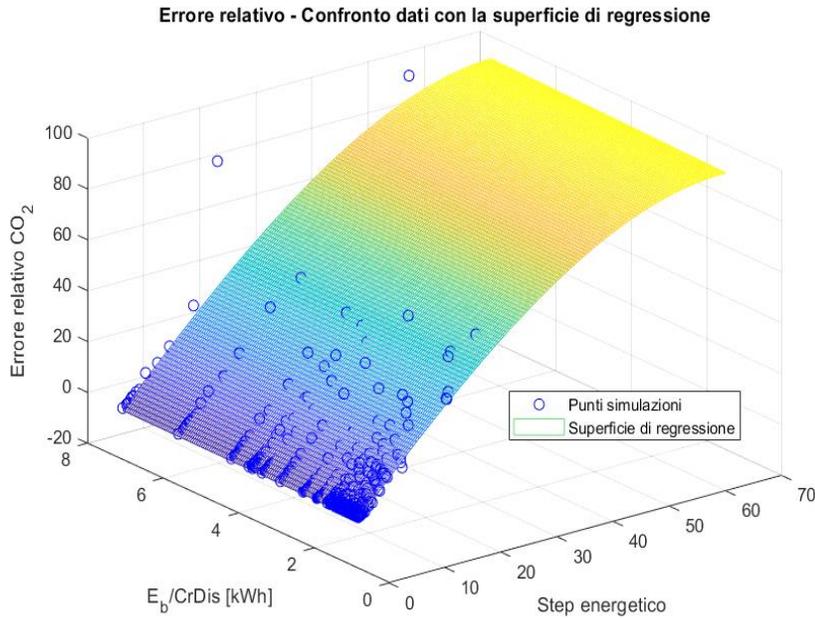


Figura 9.9 – Errore relativo, step energetico, taglia della batteria



$$R^2 = 0,9332$$

Figura 9.10 – Errore relativo, step energetico, taglia della batteria rapportata a C rate di scarica

9.3 Osservazioni

Osservando i valori di R^2 ottenuti si può affermare che questi sono decisamente inferiori rispetto ai valori di R^2 ottenuti durante lo svolgimento delle regressioni. Questo significa che andando a modificare i layout presi in considerazione si vanno a modificare le superfici di regressione. Potrebbe essere utile ricavare nuove superfici con un data set più ampio in modo da ottenere superfici meno sensibili al tipo di dati utilizzato.

10. Regressioni multiple per l'architettura p3

Le analisi svolte ai capitoli precedenti riguardavano i dati raccolti per differenti layout di un'architettura p2. L'interesse verso altri tipi di architetture ha evidenziato la necessità di estendere l'analisi all'architettura p3, in modo da evidenziare eventuali analogie o differenze tra i dati raccolti ma soprattutto tra le regressioni eseguite per le due diverse architetture.

In primo luogo è stata svolta la raccolta dei dati tramite le simulazioni. Come descritto al paragrafo 6.1, sono stati raccolti i valori di CO₂ e di fc per 10 layout differenti con architettura p3 e sono stati calcolati errore relativo e step energetico. In figura 10.1 sono raccolti gli andamenti di CO₂ ed fc in funzione del numero di livelli di SOC utilizzato, viene inoltre mostrato l'andamento dell'errore relativo di CO₂ rispetto allo step energetico.

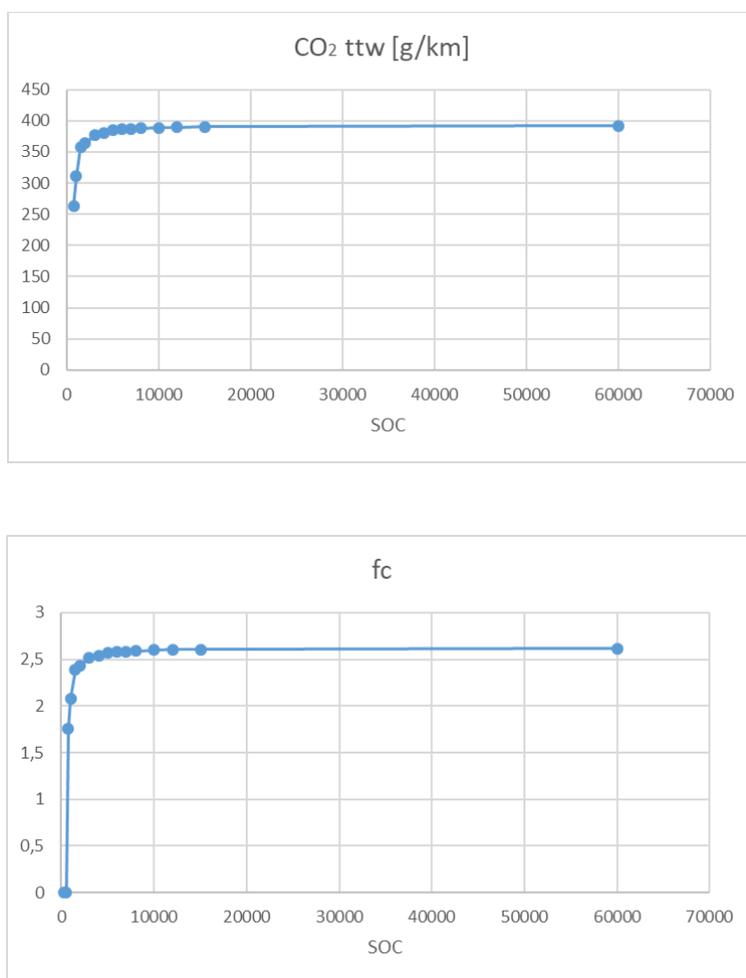


Figura 10.1 – Andamenti di CO₂, fc ed errore relativo per il layout 1

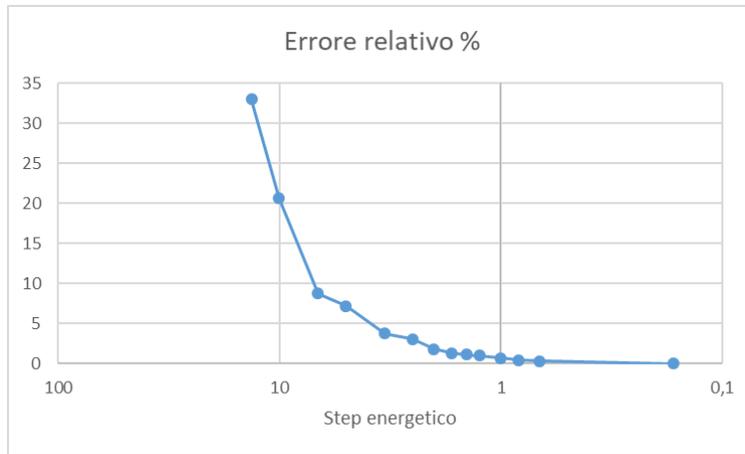


Figura 10.1 – Andamenti di CO₂, fc ed errore relativo per il layout 1

Una volta ottenuti i dati iniziali vengono svolte le regressioni multiple in modo da ottenere superfici che descrivano l’andamento dell’errore relativo in base alle variabili di design: taglia della macchina elettrica, taglia della batteria e valori di C rate di carica e scarica. In questo capitolo vengono svolte le regressioni multiple andando ad assegnare termini di secondo grado alla variabile step energetico.

10.1 Regressioni multiple dei dati completi

Regressione multipla: taglia della macchina elettrica

Il primo tentativo di regressione multipla svolto considera come variabili esplicative lo step energetico e la potenza della macchina elettrica. Tramite Matlab è stata eseguita una regressione multipla in modo da correlare errore relativo a step energetico e potenza della macchina elettrica. Si ottiene una superficie su cui sono distribuiti i punti ottenuti tramite simulazione. In particolare la superficie è espressa come nell’equazione 10.1:

$$y = \beta_0 + \beta_1x_1 + \beta_2x_2 + \beta_3x_1^2 + \beta_4x_1x_2 \quad (10.1)$$

In cui x1 rappresenta lo step energetico ed x2 rappresenta la potenza del motore elettrico. La figura 10.2 riporta la superficie ottenuta. La tabella 10.1 riporta i coefficienti ricavati insieme ai rispettivi valori p. Si riporta inoltre per questo tipo di regressione il valore p per il test F, il coefficiente di determinazione, il coefficiente di determinazione corretto ed il valore di RMSE.

Root Mean Squared Error	2,38
R-squared	0,975
Adjusted R-Squared	0,974
p-value	0,000

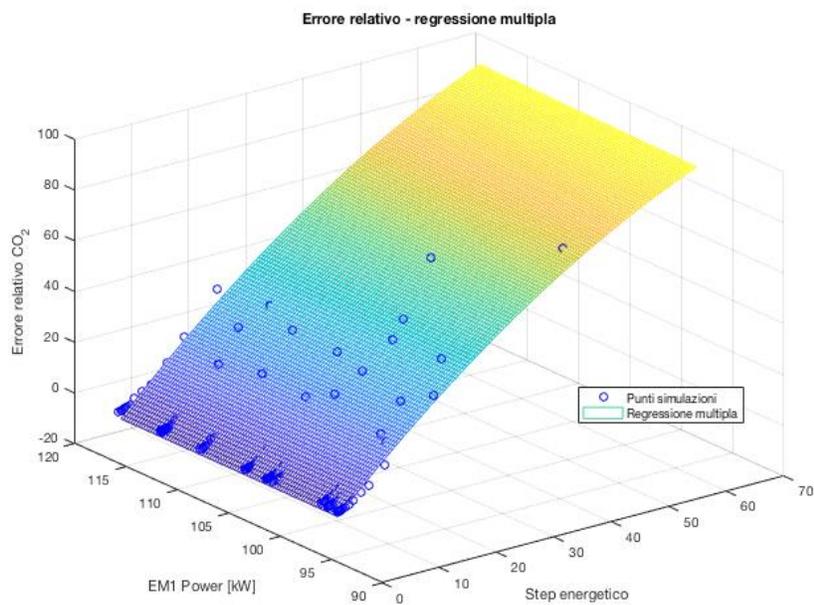


Figura 10.2 – Regressione multipla

Tabella 10.1 – Coefficienti stimati

Termine	Valore	pValue
Intercetta	-1,9218	0,5780
X ₁	2,16	0,0001
X ₂	-0,0052	0,8759
X ₁ ²	-0,0108	0,0000
X ₁ X ₂	0,0013	0,8007

Come verificato per l'architettura p2 per questo tipo di regressione la variabile più significativa risulta essere lo step energetico.

Regressione multipla: taglia della macchina elettrica rapportata a C rate di scarica

È stata eseguita la regressione multipla per descrivere l'errore relativo in base a step energetico ed potenza della macchina elettrica rapportata a C rate di scarica. L'equazione utilizzata per la regressione è la medesima espressa nell'equazione 10.1, in questo caso x_1 rappresenta lo step energetico ed x_2 rappresenta la variabile potenza della macchina elettrica rapportata al C Rate di scarica. In figura 10.3 è riportata la superficie ottenuta tramite la regressione. La tabella 10.2 riporta i coefficienti ricavati insieme ai rispettivi valori p. Si riporta inoltre per questo tipo di regressione il valore p per il test F, il coefficiente di determinazione, il coefficiente di determinazione corretto ed il valore di RMSE.

Root Mean Squared Error	2,37
R-squared	0,975
Adjusted R-Squared	0,974
p-value	0,000

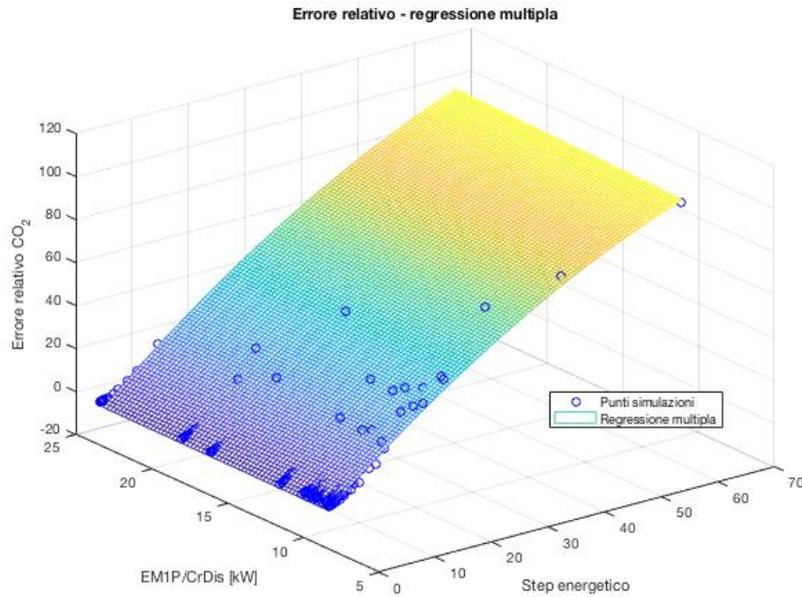


Figura 10.3 – Regressione multipla

Tabella 10.2 – Coefficienti stimati

Termine	Valore	pValue
Intercetta	-2,7696	0,0000
X ₁	2,2595	0,0000
X ₂	0,0240	0,6358
X ₁ ²	-0,0107	0,0000
X ₁ X ₂	0,0032	0,7451

Il termine più significativo per questa regressione risulta essere lo step energetico. I valori di R² dimostrano un buon adattamento del modello ai dati.

Regressione multipla: taglia della macchina elettrica rapportata a C rate di carica

È stata eseguita la regressione multipla per descrivere l'errore relativo in base a step energetico ed potenza della macchina elettrica rapportata a C rate di carica. L'equazione utilizzata per la regressione è la medesima espressa nell'equazione 10.1, in questo caso x₁ rappresenta lo step energetico ed x₂ rappresenta la variabile potenza della macchina elettrica rapportata al C Rate di carica. In figura 10.4 è riportata la superficie ottenuta tramite la regressione. La tabella 10.3 riporta i coefficienti ricavati insieme ai rispettivi valori p. Si riporta inoltre per questo tipo di regressione il valore p per il test F, il coefficiente di determinazione, il coefficiente di determinazione corretto ed il valore di RMSE.

Root Mean Squared Error	2,37
R-squared	0,975
Adjusted R-Squared	0,974
p-value	0,000

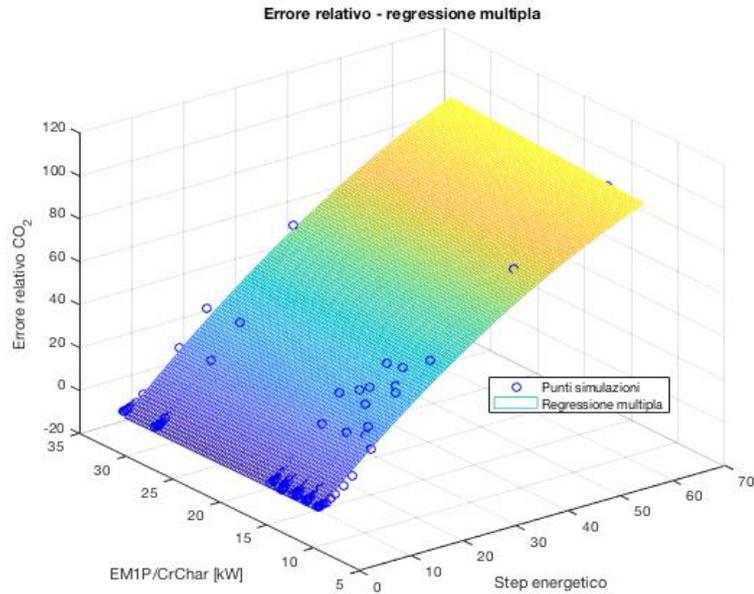


Figura 10.4 – Regressione multipla

Tabella 10.3 – Coefficienti stimati

Termine	Valore	pValue
Intercetta	-1,8619	0,0016
X ₁	2,2194	0,0000
X ₂	-0,0379	0,2621
X ₁ ²	-0,0107	0,0000
X ₁ X ₂	0,0045	0,2658

In questo tipo di regressione la variabile statisticamente più significativa è lo step energetico. I valori di R² mostrano la bontà del modello.

Regressione multipla: taglia della batteria

È stata eseguita la regressione multipla per descrivere l'errore relativo in base a step energetico e taglia della batteria. L'equazione utilizzata per la regressione è la medesima espressa nell'equazione 10.1, in questo caso x₁ rappresenta lo step energetico ed x₂ rappresenta la variabile taglia della batteria. In figura 10.5 è riportata la superficie ottenuta tramite la regressione. La tabella 10.4 riporta i coefficienti ricavati insieme ai rispettivi valori p. Si riporta inoltre per questo tipo di regressione il valore p per il test F, il coefficiente di determinazione, il coefficiente di determinazione corretto ed il valore di RMSE.

Root Mean Squared Error	2,37
R-squared	0,975
Adjusted R-Squared	0,974
p-value	0,000

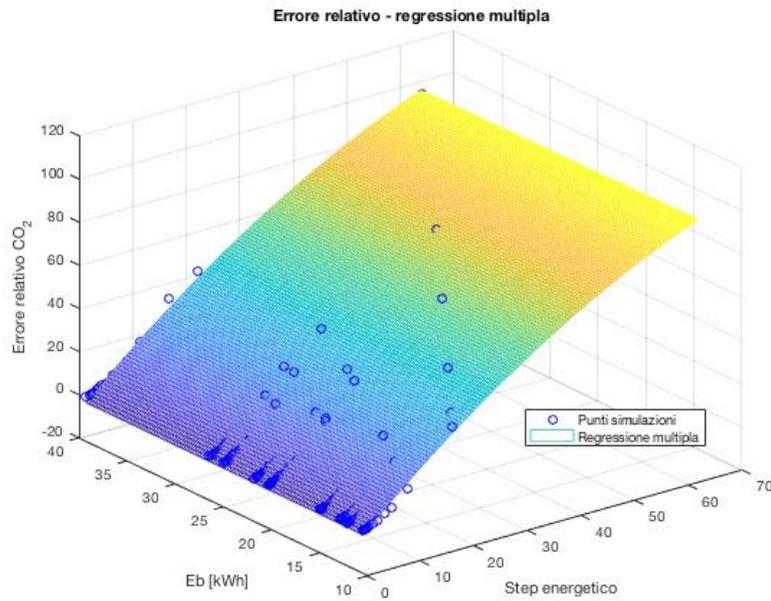


Figura 10.5 – Regressione multipla

Tabella 10.4 – Coefficienti stimati

Termine	Valore	pValue
Intercetta	-2,1712	0,0006
X ₁	2,3101	0,0000
X ₂	-0,0145	0,6341
X ₁ ²	-0,0107	0,0000
X ₁ X ₂	-0,0007	0,8562

Regressione multipla: taglia della batteria rapportata a C rate di scarica

È stata eseguita la regressione multipla per descrivere l'errore relativo in base a step energetico e taglia della batteria rapportata al valore di C rate di scarica. L'equazione utilizzata per la regressione è la medesima espressa nell'equazione 10.1, in questo caso x_1 rappresenta lo step energetico ed x_2 rappresenta la variabile taglia della batteria rapportata a C rate di scarica. In figura 10.6 è riportata la superficie ottenuta tramite la regressione. La tabella 10.5 riporta i coefficienti ricavati insieme ai rispettivi valori p. Si riporta inoltre per questo tipo di regressione il valore p per il test F, il coefficiente di determinazione, il coefficiente di determinazione corretto ed il valore di RMSE.

Root Mean Squared Error	2,37
R-squared	0,975
Adjusted R-Squared	0,974
p-value	0,000

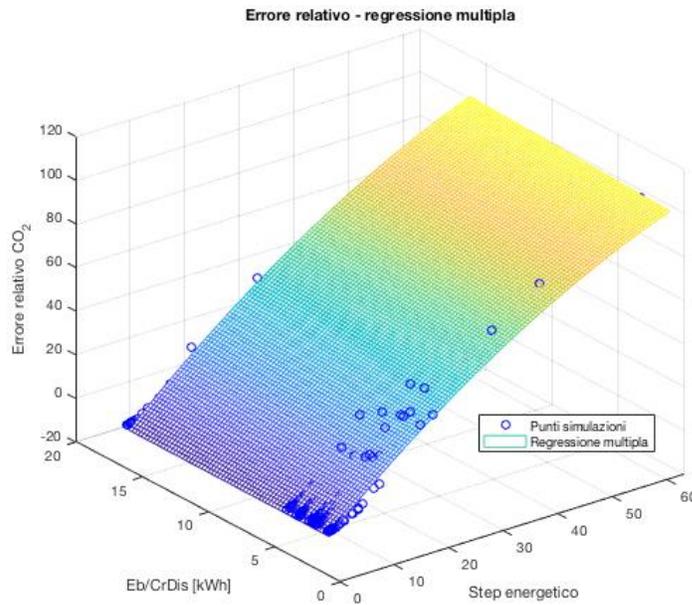


Figura 10.6 – Regressione multipla

Tabella 10.5 – Coefficienti stimati

Termine	Valore	pValue
Intercetta	-2,4963	0,0000
X ₁	2,2798	0,0000
X ₂	0,0110	0,8374
X ₁ ²	-0,0109	0,0000
X ₁ X ₂	0,0040	0,587

10.2 Regressioni multiple a tre variabili

Regressione multipla: potenza della macchina elettrica e taglia della batteria

Si è inoltre pensato di svolgere una regressione multipla tenendo conto di tre variabili invece di due come fatto ai capitoli precedenti, in particolare l'obiettivo è di descrivere l'andamento dell'errore relativo di CO₂ in funzione di differenti parametri:

- Step energetico
- Potenza della macchina elettrica [kW]
- Taglia della batteria [kWh]

Il risultato della regressione è una superficie di equazione:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_1^2 + \beta_5 x_1 x_2 + \beta_6 x_1 x_3 + \beta_7 x_2 x_3 \quad (10.2)$$

In cui x_1 rappresenta lo step energetico, x_2 la potenza del motore elettrico ed infine x_3 rappresenta la taglia della batteria. In figura 10.7 è riportata la superficie ottenuta tramite la regressione. La tabella 10.6 riporta i coefficienti ricavati insieme ai rispettivi valori p. Si riporta inoltre per questo tipo di regressione il valore p per il test F, il coefficiente di determinazione, il coefficiente di determinazione corretto ed il valore di RMSE.

Root Mean Squared Error	2,39
R-squared	0,975
Adjusted R-Squared	0,974
p-value	0,000

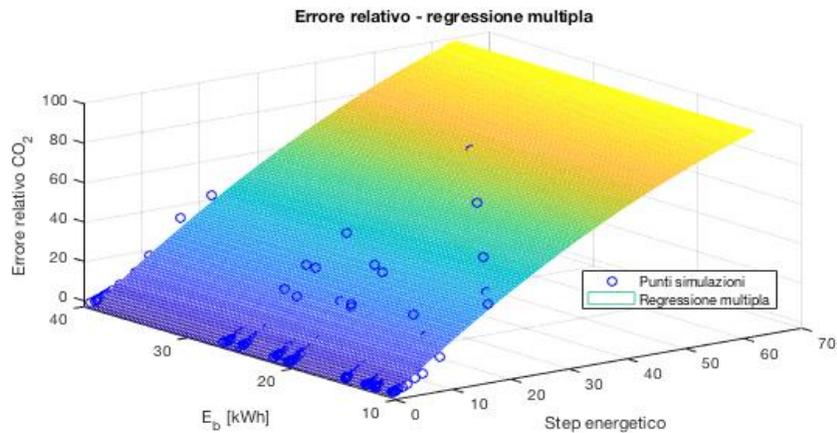
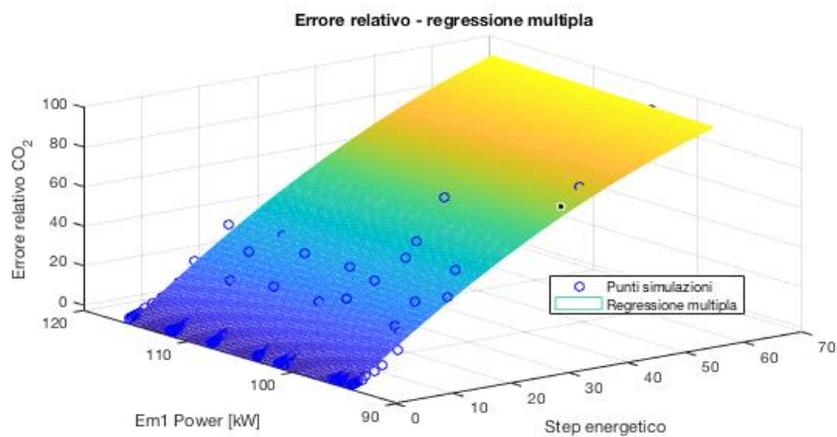


Figura 10.7 – Regressione multipla a tre variabili

Termine	Valore	pValue
Intercetta	7,4275	0,4342
X ₁	2,2743	0,0000
X ₂	-0,0957	0,311

X3	-0,5293	0,2799
X1 ²	-0,0106	0,0000
X1X2	0,0004	0,9430
X1X3	-0,0008	0,8444
X2X3	0,0051	0,2919

Osservazioni

In generale si può osservare che i legami evidenziati tra i parametri di design e l'errore relativo per l'architettura p2 si ripresentano anche per l'architettura p3. È infatti possibile anche per questa seconda architettura andare a creare superfici di regressione che tengano conto di potenza della macchina elettrica, taglia della batteria e C rate. Come espresso dai test F infatti, i modelli sono adatti alla rappresentazione dei dati anche per l'architettura p3. Osservando inoltre l'andamento dei dati risultato delle simulazioni ed osservando le regressioni ottenute si è cercato di determinare quale dei parametri presi in considerazione va ad influenzare maggiormente i valori di errore relativo di CO2. Tenendo in considerazione infatti le regressioni ottenute per l'architettura p2 si nota l'influenza marcata di un parametro sull'andamento dell'errore relativo, in particolare si tratta della taglia della batteria. Estendendo l'analisi all'architettura p3 mantenendo la stessa missione si è potuto constatare che il legame evidenziato per la p2 è presente anche per la nuova architettura analizzata.

La tabella 10.7 raccoglie tutti i coefficienti ottenuti per l'architettura p2 mentre la tabella 10.8 raccoglie tutti i coefficienti ottenuti per l'architettura p3. In particolare si può notare che per tutte le regressioni, ad eccezione della regressione che utilizza come variabile la potenza del motore elettrico, le superfici vengono traslate verso il basso, ovvero si ha un valore di intercetta inferiore.

Tabella 10.7 – Coefficienti delle regressioni con architettura p2

Regressione	β_0	β_1	β_2	β_3	β_4
Em1	-2,1742	2,4848	0,0059	-0,0158	0,0008
Em1/CrDis	-1,8028	2,5916	0,0192	-0,0158	-0,0016
Em1/CrC	-1,7587	2,5495	0,0116	-0,0159	0,0010
Ebatt	-1,2602	2,5277	-0,0152	-0,0166	0,0025
Ebatt/CrDis	-1,5059	2,5675	-0,0179	-0,0159	0,0016

Tabella 10.8 – Coefficienti delle regressioni con architettura p3

Regressione	β_0	β_1	β_2	β_3	β_4
Em1	-1,9218	2,16	-0,0052	-0,0108	0,0013
Em1/CrDis	-2,7696	2,2595	-0,0240	-0,0107	0,0033
Em1/CrC	-1,8619	2,2194	-0,0379	-0,0107	0,0045
Ebatt	-2,1712	2,3101	-0,0145	-0,0106	-0,0007
Ebatt/CrDis	-2,4963	2,2798	0,0109	-0,0109	0,0040

In figura 10.8 sono riportati i coefficienti per le differenti regressioni in modo da poter verificare il loro range di variabilità. Si nota che i coefficienti dei termini di primo e secondo grado hanno un range di

variabilità molto simile passando dall'architettura p2 all'architettura p3 e si notano inoltre come detto precedentemente sostanziali differenze a riguardo del termine intercetta (β_0).

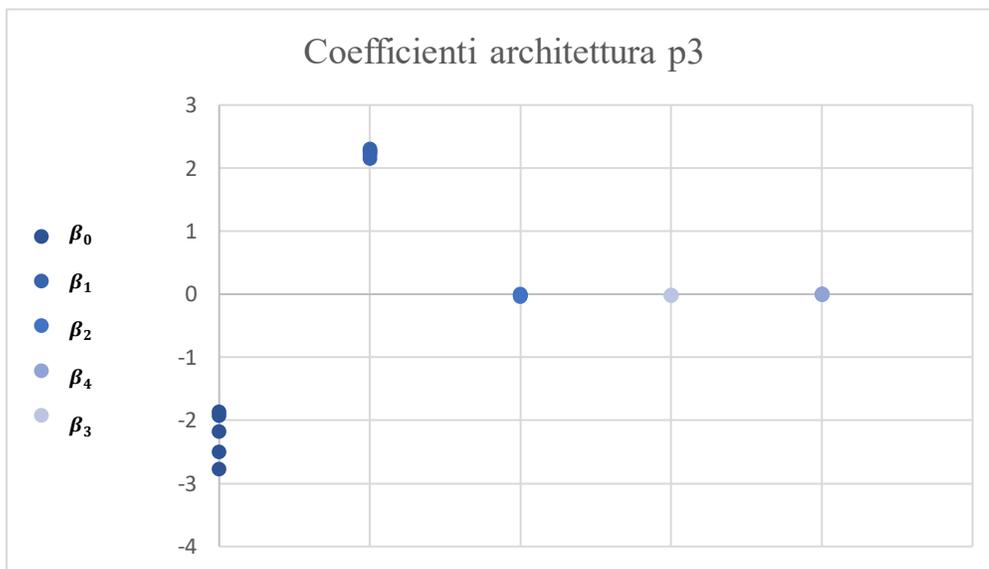
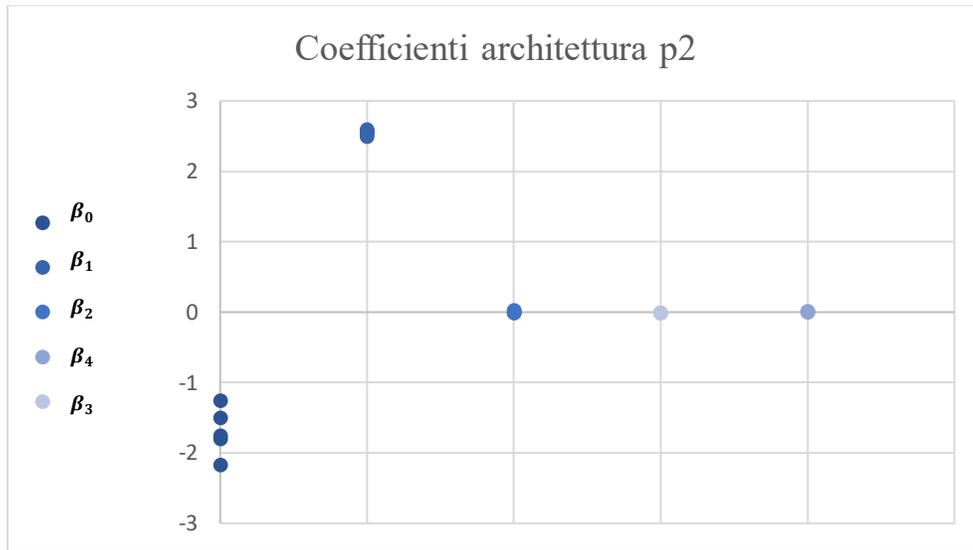


Figura 10.8 – Coefficienti per le regressioni svolte

11. Influenza degli altri parametri di design sull'errore relativo

Ai capitoli precedenti sono state svolte regressioni multilineari in modo da evidenziare relazioni tra l'errore relativo e alcuni dei parametri di design ovvero taglia della batteria, taglia della macchina elettrica e C rate sia di carica che di scarica. Questo capitolo ha invece lo scopo di andare ad investigare l'influenza che assumono i rimanenti parametri di design utili a definire il layout del veicolo. Visto l'alto numero di predittori che possono essere presi in considerazione per la descrizione dell'andamento dell'errore relativo si è deciso di provare in primo luogo a descrivere tale andamento andando ad aggiungere o rimuovere le variabili statisticamente più significative, insieme all'utilizzo di tecniche statistiche e di machine learning per validare le regressioni svolte e per scegliere le variabili di regressione migliori da utilizzare. Le regressioni svolte ai capitoli precedenti hanno permesso di determinare il parametro con maggior influenza sull'errore relativo ovvero step energetico e dunque la taglia della batteria. Scopo di questa sezione è determinare l'effettiva influenza di questo parametro e successivamente l'interazione e l'influenza che hanno gli altri parametri di design sull'errore relativo di CO₂. Si è scelta quindi una strategia per determinare l'influenza degli altri parametri svolgendo una serie di regressioni multiple in cui vengono aggiunte o eliminate le variabili indipendenti. In particolare si è scelto di verificare l'influenza della taglia del motore e della trasmissione che nelle regressioni precedenti non sono state prese in considerazione.

Le regressioni svolte in questo capitolo sono state svolte con l'utilizzo del toolbox di Matlab per statistica e machine learning, in particolare è stata usata l'applicazione "regression learner", la quale ha permesso innanzi tutto di poter svolgere le regressioni utilizzando la tecnica di "cross validation" per evitare il fenomeno dell'overfitting, descritta al paragrafo 5.4.

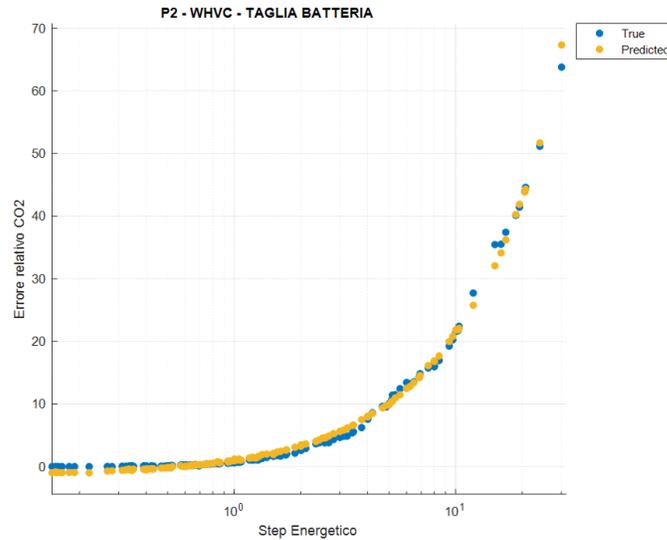
11.1 Regressioni multilineari

Taglia della batteria

È stata individuata la taglia della batteria come parametro probabilmente più influente sull'errore relativo oltre allo step energetico. Per poter effettivamente verificare il legame tra i due parametri sono stati raccolti ulteriori dati tramite simulazioni con il codice HEVbox. Sono stati scelti 8 layout differenti in cui tutti i parametri di design vengono tenuti fissi ad eccezione della taglia della batteria che è variabile da layout a layout. Una volta raccolti i valori di errore relativo e step energetico, calcolati come illustrato al capitolo 6, è stata svolta la prima regressione, in cui le variabili esplicative sono solamente step energetico e taglia della batteria. Per questa regressione sono stati utilizzati i dati ottenuti dalle simulazioni con missione WHVC, architettura p2, parametri di design fissi ad eccezione della taglia della batteria, variabile per ogni layout. L'equazione utilizzata per ricavare la superficie di regressione è espressa in seguito:

$$ErrRelCO_2 = \beta_0 + \beta_1 StepEn + \beta_2 EBatt + \beta_3 StepEn * EBatt + \beta_4 StepEn^2 + \beta_5 Ebatt^2 \quad (11.1)$$

In figura 11.1 è riportato il risultato della regressione insieme ai valori di R². Poiché si tratta di una regressione multipla è stato riportato per chiarezza il piano errore relativo – step energetico in modo da poter confrontare i dati sperimentali con i punti calcolati tramite la regressione.



$$R^2 = 0,9973$$

$$R^2_{adj} = 0,9971$$

$$MSE = 0,55499$$

Figura 11.1 – Regressione multipla: piano Errore relativo-step energetico

In tabella 11.1 sono riportati i coefficienti ottenuti tramite la regressione insieme ai valori p.

<i>Tabella 11.1 – Coefficienti</i>		
<i>Coefficiente</i>	<i>Stima</i>	<i>P-value</i>
<i>Intercetta</i>	-0,4829	0,7595
<i>Step En</i>	2,2464	0,0000
<i>EBatt</i>	-0,0990	0,6611
<i>StepEn*Ebatt</i>	0,0061	0,1480
<i>Step En ^2</i>	-0,0053	0,0000
<i>E batt ^2</i>	0,0024	0,7586

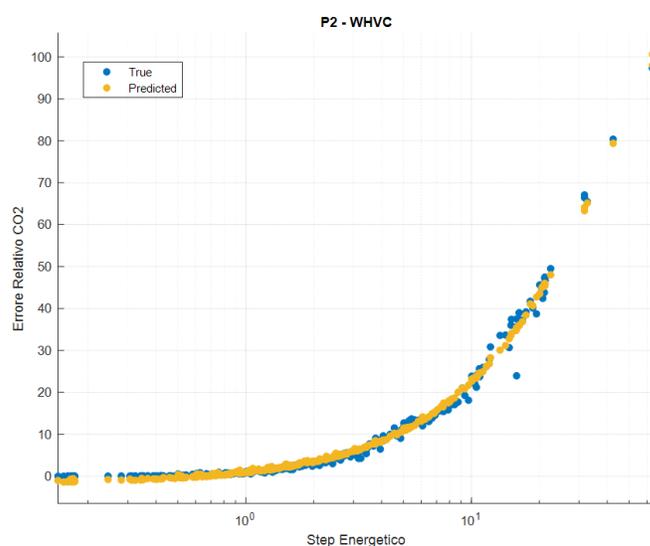
Il modello va a riprodurre in modo fedele i dati visti i valori di R^2 . Bisogna però tenere conto del valore p riguardante il coefficiente della taglia della batteria il quale assume valore superiore a 0.05. Per tale motivo la variabile sembra essere non significativa. Potrebbe esserci un'ulteriore ragione alla base del p-value di questo coefficiente: la taglia della batteria è significativamente legata allo step energetico per cui può non essere utile considerare entrambe le variabili all'interno della medesima regressione.

Cilindrata

Un altro parametro di layout interessante che potrebbe andare ad influire sull'errore relativo è la cilindrata del motore termico. Per verificare questa eventuale dipendenza è stata fatta una regressione che considera come variabili esplicative lo step energetico e la cilindrata (engine displacement). La regressione è stata svolta considerando i dati ricavati per missione WHVC, con architettura p2 e parametri di design tutti variabili. L'equazione utilizzata per la regressione è la seguente:

$$ErrRelCO_2 = \beta_0 + \beta_1 StepEn + \beta_2 EngDisp + \beta_3 StepEn * EngDisp + \beta_4 StepEn^2 + \beta_5 EngDisp^2 \quad (11.2)$$

In figura 11.2 viene riportato il risultato della regressione in particolare si riporta la vista dal piano errore relativo-step energetico. Sono riportati inoltre i valori di R^2 .



$$R^2 = 0,9942$$

$$R^2_{adj} = 0,9941$$

$$MSE = 1,55$$

Figura 11.2 – Regressione multipla: piano errore relativo-step energetico

In tabella 11.2 sono riportati i coefficienti ottenuti tramite la regressione.

Tabella 11.2 - Coefficienti		
Coefficiente	Stima	P-value
Intercetta	6,5625	0,00127
Step En	2,6572	0,0000
EngDisp	-4,8350	0,0023
StepEn*EngDisp	-0,0253	0,2570

$Step\ En^2$	-0,0155	0,0000
$EngDisp^2$	0,7023	0,0028

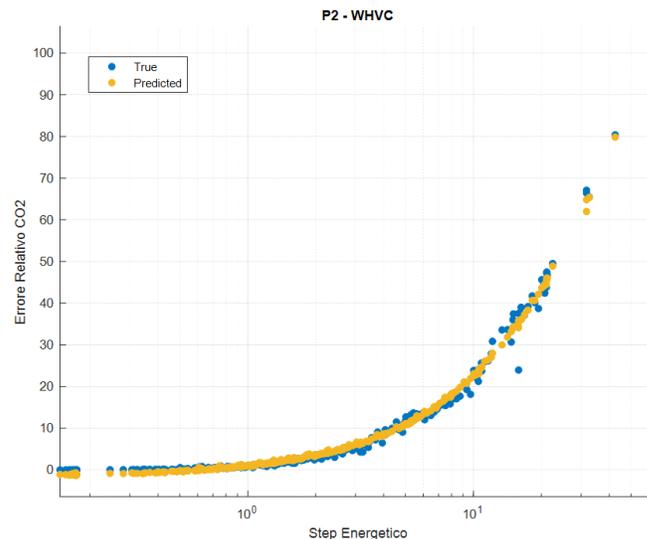
Questa regressione è stata svolta per determinare l'influenza del parametro Eng Disp ovvero la cilindrata del motore termico. Il parametro sembra essere significativo andando ad osservare il valore p associato al coefficiente. Si può notare che il termine di interazione in questo caso sembra non essere statisticamente significativo.

Trasmissione

Un ulteriore parametro di interesse è la trasmissione, in particolare il rapporto di trasmissione del riduttore. Per poter definire l'influenza della trasmissione sull'errore relativo è stata svolta una regressione che considera step energetico e trasmissione come variabili esplicative. Questa regressione è stata svolta andando ad utilizzare i dati ricavati dal tool utilizzando con missione la missione WHVC, con architettura p2 e con i parametri di design variabili. L'equazione utilizzata per la regressione è la seguente:

$$ErrRelCO_2 = \beta_0 + \beta_1 StepEn + \beta_2 Em1SpR + \beta_3 StepEn * EM1SpR + \beta_4 StepEn^2 + \beta_5 Em1SpR^2 \quad (11.3)$$

In figura 11.3 è riportata la regressione ottenuta, in particolare viene riportato il piano errore relativo-step energetico. La tabella 11.3 riporta i coefficienti ottenuti tramite la regressione. Il parametro Em1SpRatio sembra essere non significativo per l'errore relativo dato il valore del p-value associato al coefficiente corrispondente.



$$R^2 = 0,9942$$

$$R^2_{adj} = 0,9941$$

$$MSE = 2,8533$$

Figura 11.3 - Regressione multipla: piano errore relativo-step energetico

Tabella 11.3 – Coefficienti

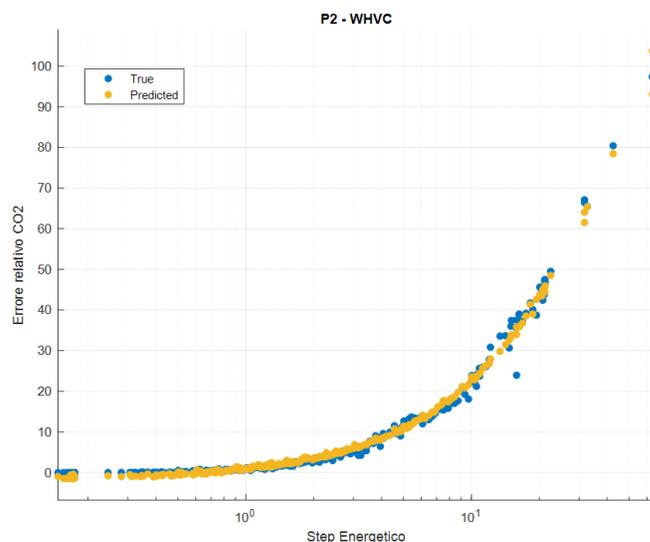
<i>Coefficiente</i>	<i>Stima</i>	<i>P-value</i>
<i>Intercetta</i>	0,9495	0,4683
<i>Step En</i>	2,6115	0,0000
<i>Em1SpR</i>	-1,1630	0,0927
<i>StepEn*Em1SpR</i>	-0,0127	0,2531
<i>Step En^2</i>	-0,0156	0,0000
<i>Em1SpR^2</i>	0,1254	0,1595

Cilindrata e trasmissione

Per poter valutare l'effetto combinato delle due variabili analizzate singolarmente con le due precedenti regressioni è stata condotta una regressione multipla in cui vengono considerati oltre allo step energetico, la cilindrata e la trasmissione. Vengono utilizzati i dati ricavati per la missione WHVC con architettura p2 e tutti i parametri di design variabili. L'equazione utilizzata per la regressione è la seguente:

$$\begin{aligned}
 ErrRelCO_2 = & \beta_0 + \beta_1 StepEn + \beta_2 EngDisp + \beta_3 Em1SpR + \beta_4 StepEn * EngDisp + \\
 & + \beta_5 StepEn * Em1SpR + \beta_6 EngDisp * Em1SpR + \beta_7 StepEn^2 + \\
 & + \beta_8 SngDisp^2 + \beta_9 Em1SpR^2 \quad (11.4)
 \end{aligned}$$

Il risultato della regressione viene rappresentato in figura 11.4 in cui è riportata la vista del piano errore relativo- step energetico, in tabella 11.4 sono riportati i coefficienti con i rispettivi p-value.



$$R^2 = 0,9943$$

$$R^2_{adj} = 0,9941$$

$$MSE = 1,7421$$

Figura 11.4 - Regressione multipla: piano errore relativo-step energetico

Tabella 11.4 - Coefficienti		
Coefficiente	Stima	P-value
Intercetta	6,4787	0,0254
Step En	2,6595	0,0000
EngDisp	-3,6450	0,0487
Em1SpR	-0,9466	0,4305
Stepen*EngDisp	-0,0129	0,5915
StepEn*Em1SpR	-0,0133	0,2720
EngDisp*Em1SpR	0,0372	0,8970
Step En ²	-0,0155	0,0000
Eng Disp ²	0,5126	0,1388
Em1SpR ²	0,0915	0,3154

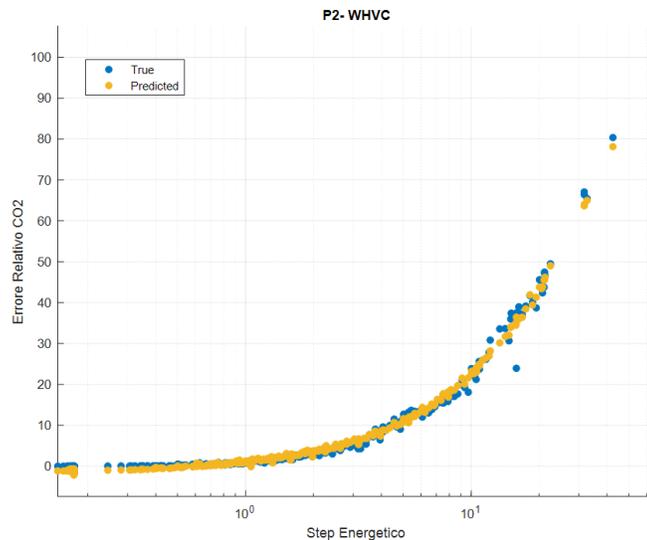
In questo caso si osserva che sia la cilindrata che lo step energetico hanno un'influenza significativa sulla variabile step energetico. Non viene invece considerata significativa la variabile EM1Speed Ratio per questo tipo di regressione, visto il corrispettivo p-value. C'è da notare che i termini di interazione hanno poca significatività statistica, insieme ai termini di secondo grado riguardanti cilindrata e trasmissione.

Cilindrata, taglia della batteria e trasmissione

In questo paragrafo l'obiettivo è quello di svolgere una regressione che tenga conto dei parametri analizzati singolarmente ai paragrafi precedenti. Viene quindi svolta una regressione multipla che tiene conto di step energetico, cilindrata, taglia della batteria e trasmissione come variabili indipendenti. Anche in questo caso la regressione è stata svolta considerando i dati ricavati per la missione WHVC, architettura p2 e parametri di design variabili. L'equazione utilizzata è la seguente:

$$\begin{aligned} ErrRelCO_2 = & \beta_0 + \beta_1 StepEn + \beta_2 EngDisp + \beta_3 Em1SpR + \beta_4 EBatt + \beta_5 StepEnEngDisp + \\ & + \beta_6 StepEnEBatt + \beta_7 StepEnEm1SpR + \beta_8 EngDispEBatt + \beta_9 EngDispEm1SpR + \\ & + \beta_{10} Ebatt * Em1SpR + \beta_{11} StepEn^2 + \beta_{12} EngDisp^2 + \beta_{13} Ebatt^2 + \beta_{14} Em1SpR^2 \quad (11.5) \end{aligned}$$

Viene riportata la vista step energetico- errore relativo in figura 11.5 in modo da poter visualizzare i risultati della regressione. In tabella 11.5 sono riportati i coefficienti risultato della regressione. Analizzando i risultati si può dire che il modello va a riprodurre in modo fedele i dati. Bisogna necessariamente tenere conto dei valori p dei differenti coefficienti: solamente il coefficiente legato allo step energetico sembra essere statisticamente significativo. Per quanto riguarda la taglia della batteria, è necessario ricordare la correlazione tra la variabile e lo step energetico. La variabile Em1SpRatio sembra non essere significativa per andare a descrivere l'errore relativo come anche la cilindrata del motore termico. Poiché è stato dimostrato precedentemente che la variabile cilindrata può avere significatività statistica si deduce che in questo caso non si verifica significatività a causa dell'interazione con gli altri termini del modello.



$$\begin{aligned} R^2 &= 0,9947 \\ R^2_{adj} &= 0,9944 \\ MSE &= 1,7519 \end{aligned}$$

Figura 11.5 - Regressione multipla: piano errore relativo-step energetico

Tabella 11.5 - Coefficienti

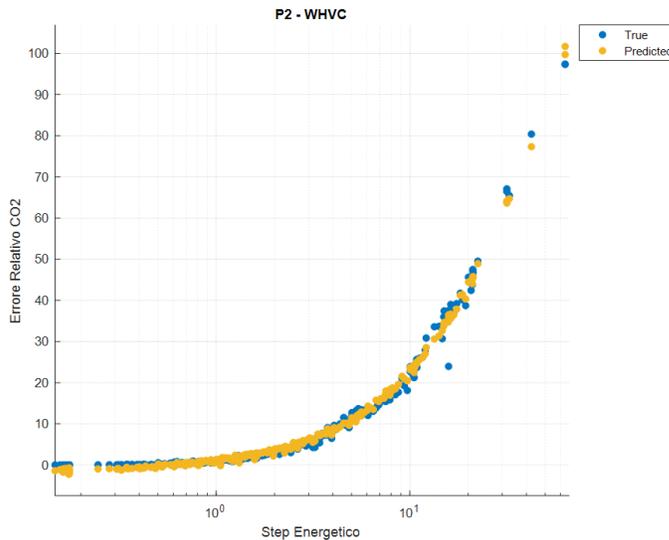
<i>Coefficiente</i>	<i>Stima</i>	<i>P-value</i>
<i>Intercetta</i>	8,0371	0,0247
<i>Step En</i>	2,6480	0,0000
<i>EngDisp</i>	-2,2723	0,2827
<i>Ebatt</i>	-0,1420	0,2186
<i>Em1SpR</i>	-2,1341	0,1215
<i>Stepen*EngDisp</i>	-0,0212	0,3737
<i>StepEn*Ebatt</i>	0,0032	0,0082
<i>StepEn*Em1SpR</i>	-0,0181	0,1369
<i>EngDisp*Ebatt</i>	0,0411	0,0961
<i>EngDisp*Em1SpR</i>	0,1133	0,7033
<i>Ebatt*Em1SpR</i>	0,0204	0,1219
<i>Step En^2</i>	-0,0164	0,0000
<i>Eng Disp^2</i>	0,1683	0,6447
<i>E batt^2</i>	-0,0022	0,0409
<i>Em1SpR^2</i>	0,1566	0,1777

Cilindrata, taglia della batteria e trasmissione

In questo paragrafo si vuole determinare l'influenza del parametro final drive sull'errore relativo. Viene quindi svolta una regressione multipla che tiene conto di step energetico, cilindrata, taglia della batteria e trasmissione (final drive speed ratio) come variabili indipendenti. Anche in questo caso la regressione è stata svolta considerando i dati ricavati per la missione WHVC, architettura p2 e parametri di design variabili. L'equazione utilizzata è la seguente:

$$\begin{aligned}
 ErrRelCO_2 = & \beta_0 + \beta_1 StepEn + \beta_2 EngDisp + \beta_3 FdSpR + \beta_4 EBatt + \beta_5 StepEnEngDisp + \\
 & + \beta_6 StepEnEBatt + \beta_7 StepEnFDSpR + \beta_8 EngDispEBatt + \beta_9 EngDispFDSpR + \\
 & + \beta_{10} Ebatt * FDSpR + \beta_{11} StepEn^2 + \beta_{12} EngDisp^2 + \beta_{13} Ebatt^2 + \beta_{14} FDSpR^2 \quad (11.6)
 \end{aligned}$$

Viene riportata la vista step energetico- errore relativo in figura 11.6 in modo da poter visualizzare i risultati della regressione. In tabella 11.6 sono riportati i coefficienti risultato della regressione.



$$R^2 = 0,9949$$

$$R^2_{adj} = 0,9946$$

$$MSE = 1,8247$$

Figura 11.6 - Regressione multipla: piano errore relativo-step energetico

Tabella 11.6 - Coefficienti		
Coefficiente	Stima	P-value
Intercetta	-4,8839	0,4280
Step En	2,8348	0
EngDisp	-0,7070	0,7640
Ebatt	0,0324	0,7432
FDSpR	0,7532	0,1197
Stepen*EngDisp	-0,0549	0,0212
StepEn*Ebatt	0,0030	0,0124
StepEn*FDSpR	-0,0385	0,0116
EngDisp*Ebatt	0,0211	0,5156
EngDisp*FDSpR	-0,3062	0,2543
Ebatt*FDSpR	-0,0029	0,8106
Step En^2	-0,0166	0

$Eng\ Disp^2$	0,1855	0,5040
$E\ batt^2$	-0,0024	0,0277
$FDSpR^2$	-0,2535	0,1048

Analizzando i risultati si può dire che il modello va a riprodurre in modo fedele i dati. Bisogna necessariamente tenere conto dei valori p dei differenti coefficienti: solamente il coefficiente legato allo step energetico sembra essere statisticamente significativo. Per quanto riguarda la taglia della batteria, è necessario ricordare la correlazione tra la variabile e lo step energetico. La variabile FDSpRatio sembra non essere significativa per andare a descrivere l'errore relativo come anche la cilindrata del motore termico. Poiché è stato dimostrato precedentemente che la variabile cilindrata può avere significatività statistica si deduce che in questo caso non si verifica significatività a causa dell'interazione con gli altri termini del modello.

11.2 Stepwise selection

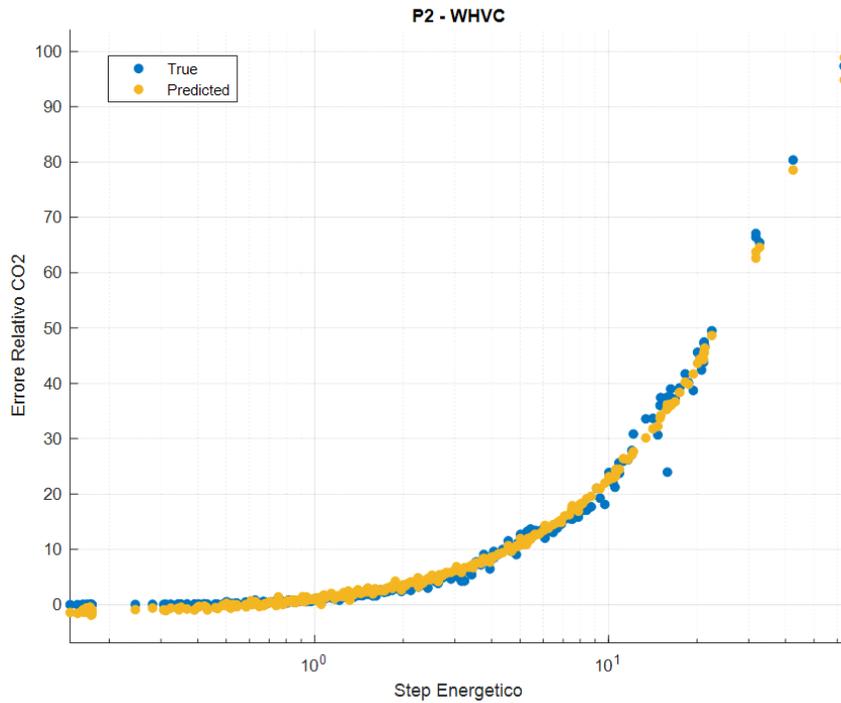
Visti i risultati delle regressioni svolte, non è semplice determinare quali siano i predittori effettivamente significativi e quindi non è facile andare a determinare il modello migliore per descrivere il fenomeno. Per andare a decretare quali sono effettivamente i predittori migliori si è scelto di utilizzare l'algoritmo Stepwise. La regressione stepwise è infatti un metodo di regressione automatico, consigliabile in studi esplorativi in presenza di molteplici variabili per la stima di modelli lineari multivariati. Tale tipo di regressione nasce dalla necessità di selezionare un sottoinsieme ottimo tra un numero di variabili esplicative per la costruzione di un modello efficiente. Qualora si debba effettivamente andare a scegliere un modello di regressione per un insieme limitato di variabili predittive si gestisce manualmente l'operazione andando a considerare tutte le possibili combinazioni di variabili esplicative. Se il numero delle variabili esplicative aumenta, il metodo stepwise è decisamente più efficiente per la risoluzione di questo tipo di problema.

Stepwise linear to quadratic

Il primo tentativo di regressione stepwise svolto considera come modello di partenza un modello lineare e come modello finale un modello quadratico. La regressione è stata svolta utilizzando i dati risultato delle simulazioni per l'architettura p2 con la missione WHVC, i parametri di design sono variabili. Questo tipo di regressione mette effettivamente alla luce il legame tra le variabili di design e l'errore relativo. Si nota anche che vengono eliminati i termini di interazione tra le variabili i quali risultavano essere ridondanti. L'equazione utilizzata è la seguente:

$$ErrRelCO_2 = \beta_0 + \beta_1 StepEn + \beta_2 EngDisp + \beta_3 EBatt + \beta_4 Em1SpR + \beta_5 StepEn * Ebatt + \\ + \beta_6 EngDisp * Ebatt + \beta_7 StepEn^2 \quad (11.7)$$

Vengono riportati i risultati ottenuti in figura 11.7 e in tabella 11.7.



$$R^2 = 0,9944$$

$$R^2_{adj} = 0,9943$$

Figura 11.7 – Regressione multipla: piano errore relativo-step energetico

Tabella 11.7

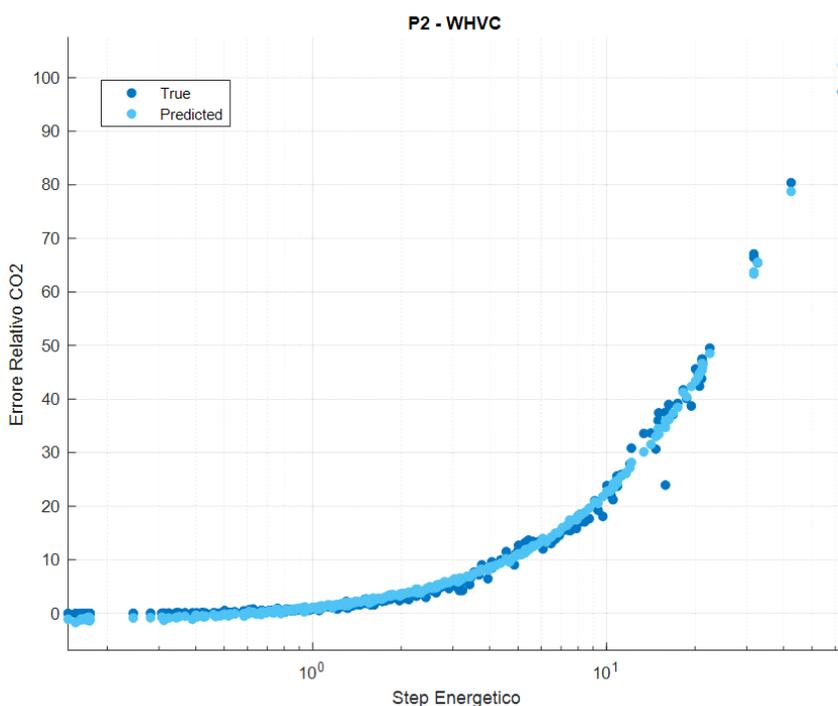
Coefficiente	Stima	P-value
Intercetta	5,2570	0,0058
Step En	2,5168	0,0000
EngDisp	-1,4024	0,004
Ebatt	-0,2503	0,003
EmISpR	-0,3267	0,0000
StepEn*Ebatt	0,0029	0,0156
EngDisp*Ebatt	0,0617	0,0061
Step En^2	-0,0166	0,0000

Stepwise constant to quadratic

Il secondo tentativo di regressione stepwise svolto considera come modello di partenza un modello costante e come modello finale un modello quadratico. La regressione è stata svolta utilizzando i dati risultato delle simulazioni per l'architettura p2 con la missione WHVC, i parametri di design sono variabili. Questo tipo di regressione permette di determinare il parametro di design con più influenza sull'errore relativo. Si nota anche che vengono eliminati i termini di interazione tra le variabili i quali risultavano essere ridondanti. L'equazione utilizzata è la seguente:

$$ErrRelCO_2 = \beta_0 + \beta_1 StepEn + \beta_2 Em1SpR + \beta_3 StepEn^2 \quad (11.8)$$

In figura 11.8 è riportato il risultato della regressione ed in tabella 11.8 sono riportati i coefficienti ottenuti. Sembra che il parametro più influente sull'errore relativo sia la trasmissione.



$$R^2 = 0,9941$$

$$R^2_{adj} = 0,9940$$

Figura 11.8 – Regressione multipla: piano errore relativo-step energetico

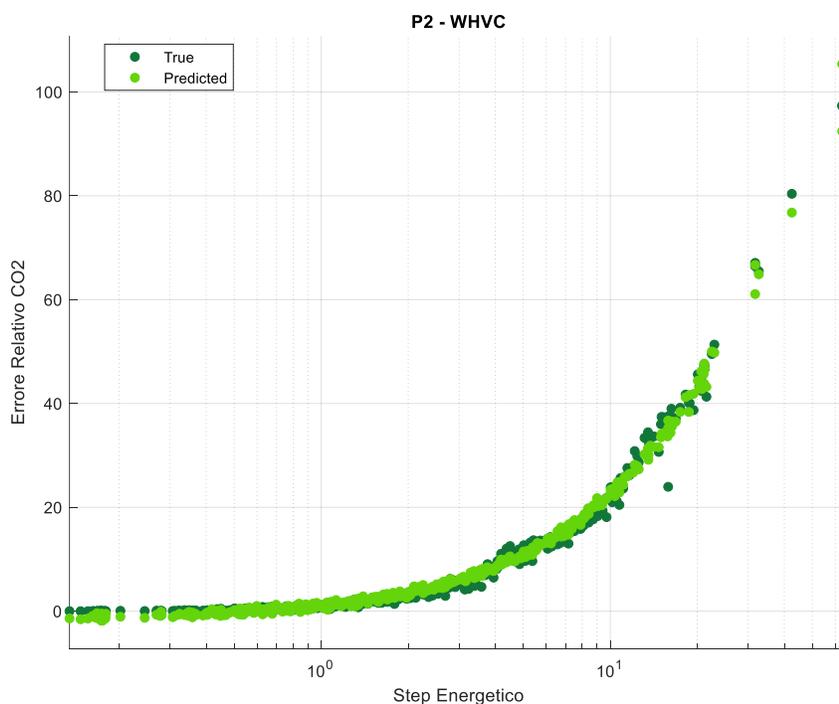
Tabella 11.8 - Coefficienti		
Coefficiente	Stima	P-value
Intercetta	-0,5651	0,1019
Step En	2,5668	0,0000
Em1SpR	-0,2692	0,0033
Step En^2	-0,0156	0,0000

Stepwise linear to quadratic con dati ulteriori

Si è deciso di svolgere la regressione stepwise partendo da un modello lineare per arrivare ad un modello quadratico considerando più dati per la missione WHVC, con architettura p2 e layout variabili ricavati. Sono stati infatti uniti ai dati delle precedenti regressioni i dati ricavati da altri studenti. Si ottiene un modello espresso dall'equazione:

$$\begin{aligned}
 ErrRelCO_2 = & \beta_0 + \beta_1 StepEn + \beta_2 EngDisp + \beta_3 Ebatt + \beta_4 Em1SpR + \\
 & + \beta_5 StepEn * EngDisp + \beta_6 StepEn * Ebatt + \beta_7 StepEn * Em1SpR + \\
 & + \beta_8 EngDisp * Ebatt + \beta_9 StepEn^2 \quad (11.9)
 \end{aligned}$$

In figura 11.9 è riportato il risultato della regressione ed in tabella 11.9 sono riportati i coefficienti ottenuti. Sembra che i parametri di design abbiano significatività statistica. Rispetto alle regressioni normali svolte nei capitoli precedenti, in questo caso l'algoritmo stepwise è andato ad eliminare molti dei termini di interazione tra le variabili di design. Si vede inoltre che è rimasto come unico parametro di secondo grado lo step energetico.



$$\begin{aligned}
 R^2 &= 0,9937 \\
 R^2_{adj} &= 0,9936
 \end{aligned}$$

Figura 11.9 – Regressione multipla: piano errore relativo-step energetico

<i>Tabella 11.9 - Coefficienti</i>		
<i>Coefficiente</i>	<i>Stima</i>	<i>P-value</i>
<i>Intercetta</i>	4,4298	0,0001
<i>Step En</i>	2,7460	0,0000

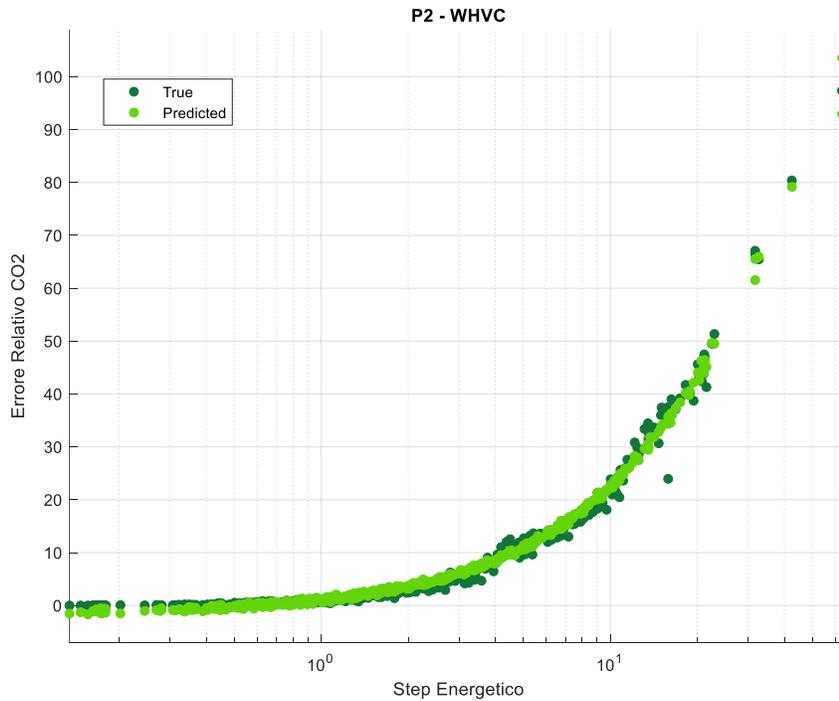
<i>EngDisp</i>	-1,3463	0,0000
<i>Ebatt</i>	-0,2610	0,0000
<i>Em1SpR</i>	-0,1857	0,0038
<i>StepEn*EngDisp</i>	-0,0459	0,0090
<i>StepEn*Ebatt</i>	0,0032	0,0018
<i>StepEN*Em1SpR</i>	-0,0237	0,0055
<i>EngDisp*Ebatt</i>	0,0665	0,00001
<i>Step En^2</i>	-0,0162	0,0000

Stepwise constant to quadratic con dati ulteriori

Il secondo tentativo di regressione stepwise svolto considera come modello di partenza un modello costante e come modello finale un modello quadratico. La regressione è stata svolta utilizzando ulteriori dati risultato delle simulazioni per l'architettura p2 con la missione WHVC, i parametri di design sono variabili. Questo tipo di regressione permette di determinare il parametro di design con più influenza sull'errore relativo. Si nota anche che vengono eliminati i termini di interazione tra le variabili i quali risultavano essere ridondanti. L'equazione utilizzata è la seguente:

$$\begin{aligned}
 ErrRelCO_2 = & \beta_0 + \beta_1 StepEn + \beta_2 EngDisp + \beta_3 Em1SpR + \\
 & + \beta_4 StepEn * Em1SpR + \beta_5 StepEn^2 \quad (11.10)
 \end{aligned}$$

In figura 11.10 è riportato il risultato della regressione ed in tabella 11.10 sono riportati i coefficienti ottenuti. Sembra che i parametri di design con maggiore significatività statistica siano la trasmissione e la cilindrata del motore termico. Rispetto alle regressioni normali svolte nei capitoli precedenti, in questo caso l'algoritmo stepwise è andato ad eliminare molti dei termini di interazione tra le variabili di design. Si vede inoltre che è rimasto come unico parametro di secondo grado lo step energetico.



$$R^2 = 0,9932$$

$$R^2_{adj} = 0,9932$$

Figura 11.10 – Regressione multipla: piano errore relativo-step energetico

Tabella 11.10 - Coefficienti		
Coefficiente	Stima	P-value
Intercetta	-0,0621	0,8661
Step En	2,6537	0,0000
EngDisp	-0,2158	0,0264
Em1SpR	-0,1882	0,0037
StepEn*Em1SpR	-0,0275	0,0011
Step En^2	-0,0153	0,0000

11.3. Osservazioni

Non è facile determinare quali siano effettivamente le variabili più significative per la descrizione dell'andamento dell'errore relativo. Per tale motivo sono state svolte più regressioni andando a modificare ogni volta il numero di predittori considerati in modo da cercare effettive correlazioni tra le variabili di design e l'errore relativo. La bontà delle regressioni è stata valutata osservando i valori di R^2 mentre l'efficacia di ogni variabile è stata determinata osservando i p-value dei coefficienti corrispettivi.

È stata utilizzata la tecnica di cross validation per la validazione dei modelli, in modo che questi fossero protetti dal fenomeno dell'overfitting. Disponendo effettivamente di una quantità non esorbitante di dati potrebbe infatti verificarsi il fenomeno per il quale si ottengono ottime superfici di regressione con valori

di R^2 molto alti che però potrebbero non essere significative qualora si volessero fare previsioni su dati differenti.

È stato tentato il metodo stepwise per migliorare la regressione ottenuta per i dati appartenenti alla medesima missione ed alla medesima architettura. Grazie a tale regressione sono stati eliminati alcuni dei termini di interazione ed è stato possibile ottenere una superficie di regressione che tenesse conto dei dati di design. La regressione stepwise può non funzionare in differenti casi: innanzitutto quando si hanno molti predittori potenziali ma pochi dati per stimare coefficienti significativi, inoltre, quando due variabili in un modello sono altamente correlate, una delle due può essere esclusa dal modello (per esempio step energetico e taglia della batteria). È una buona regola solitamente avere almeno cinque osservazioni per ogni variabile considerata per la selezione.

12. L'influenza della missione sull'errore relativo

Le analisi precedenti sono state svolte tenendo in considerazione i dati ottenuti per la missione WHVC. Scopo di questo capitolo è verificare come vanno a modificarsi le superfici di regressione per missioni differenti, in particolare è stato fatto il confronto tra la missione WHVC e il ciclo ARTEMIS, composto da tre tratti: AUDC, ARDC, AMDC.

Il ciclo WHVC è suddiviso in tre tratti: urbano, rurale ed autostrada. La durata totale del ciclo è di 1800 secondi, ogni tratto di strada ha le proprie caratteristiche. Il tratto di guida urbano dura 900 secondi e tiene conto di fermate, partenza e velocità di minimo regime. Il tratto urbano ha una velocità media di 21,3 km/h Il tratto rurale ha una durata di 481 secondi con velocità media di 43,6 km/h. Il tratto di guida in autostrada dura 419 secondi, con velocità media di 76,6 km/h.

Il ciclo guida ARTEMIS è invece suddiviso nei seguenti tratti:

- AUDC: tratto urbano, durata 993 secondi, velocità media 17,7 km/h;
- ARDC: tratto rurale, durata 1082 secondi, velocità media 57,5 km/h;
- AMDC: tratto autostradale, durata 1068 secondi, velocità media 99,6 km/h.

In figura 12.1 sono visibili i tre tratti del ciclo ARTEMIS.

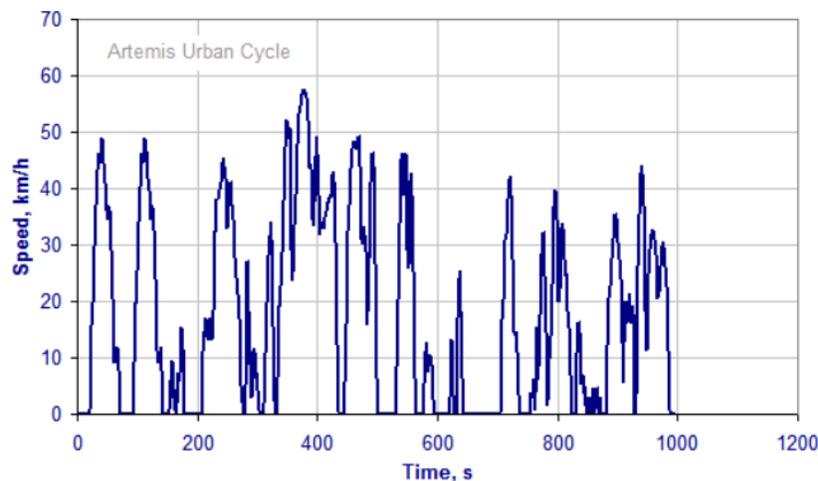


Figura 12.1 – Ciclo Artemis, tratto urbano

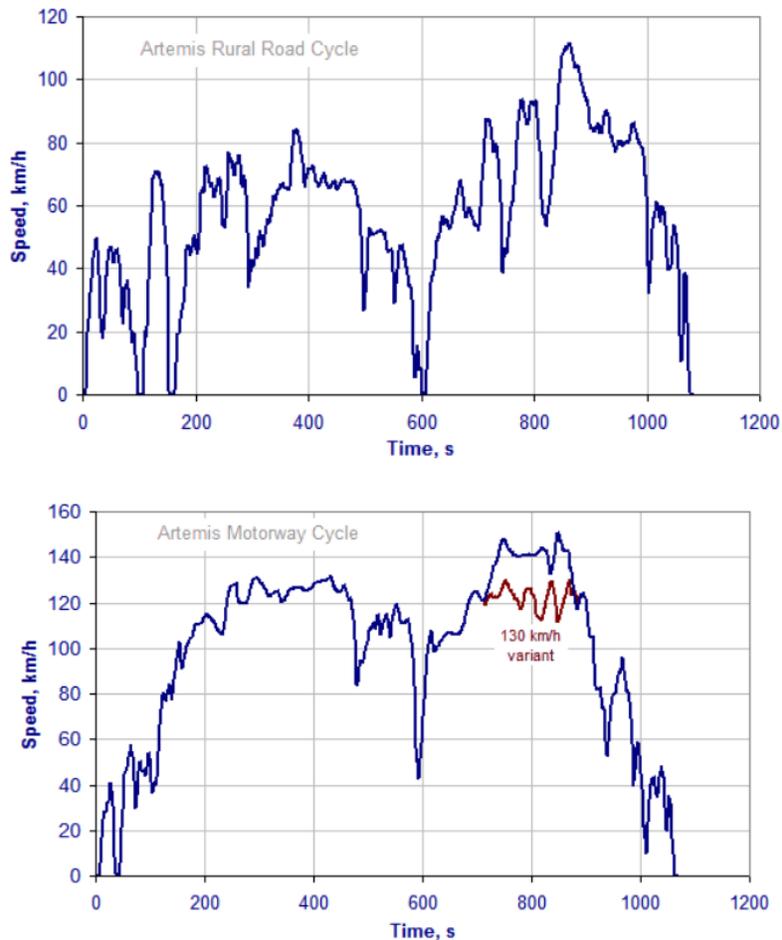


Figura 12.1 – Ciclo ARTEMIS, tratto rurale ed autostradale

Lo scopo di questo capitolo è quello di andare a ripetere la regressione svolta al capitolo precedente in cui si considera errore relativo, step energetico ed errore relativo per la missione WHVC sui dati ottenuti per le altre missioni, in particolare il confronto viene fatto tenendo conto della velocità media lungo la missione e del numero di accelerazioni nel tratto considerato.

12.1 Calcolo della velocità media e del numero di accelerazioni

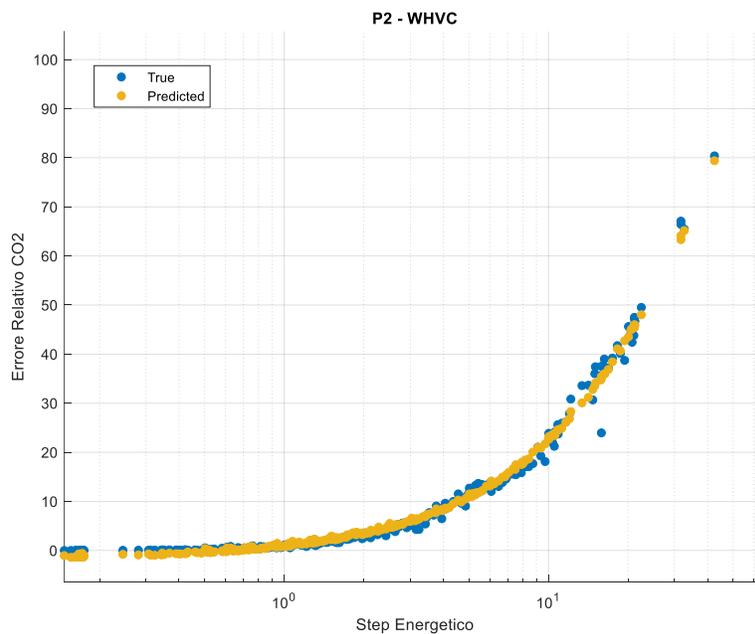
Si è pensato quindi di caratterizzare ogni missione tramite il parametro di velocità. Ogni tratto della missione è definito dalla velocità istantanea che il veicolo deve mantenere per un determinato periodo di tempo, per tale motivo è stata definita una variabile per ogni tratto delle missioni che tenesse conto della velocità media in quel tratto. Potrebbe inoltre essere ragionevole per descrivere una missione andare ad aggiungere un secondo parametro che tenga conto del numero di accelerazioni effettuate in un determinato tratto. Gli step seguiti per definire il numero di accelerazioni sulla missione sono i seguenti:

- Si calcola la velocità istantanea per ogni istante di tempo della missione

- Si calcola l'accelerazione in ogni intervallo di tempo
- Si impone una soglia per poter definire il numero di accelerazioni ($0,3 \text{ m/s}^2$)
- Per ogni missione si contano il numero di accelerazioni sopra la soglia prestabilita.
- Il numero di accelerazioni così ottenuto viene pesato sul tempo del tratto rispetto al tempo totale della missione considerata.

12.2 Confronto tra WHVC e ciclo ARTEMIS

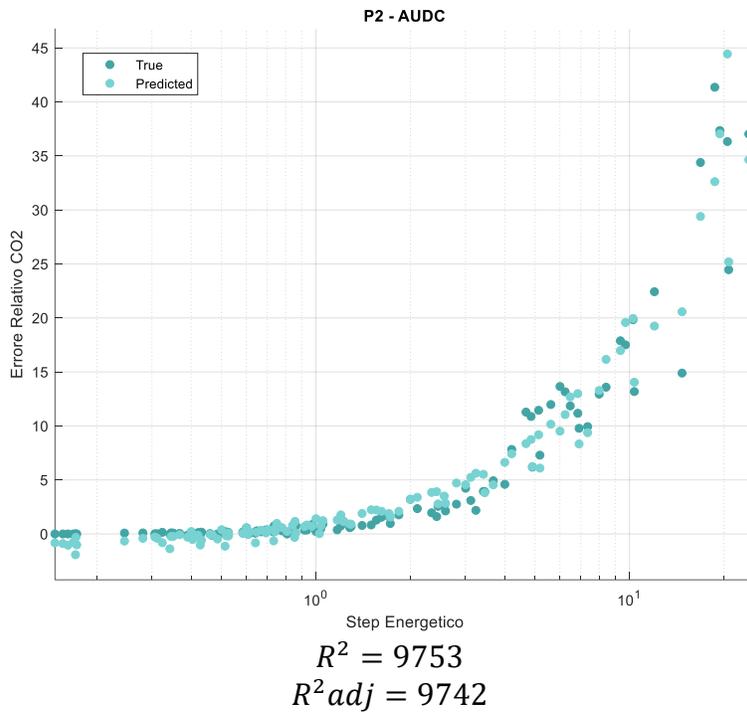
La regressione considerata per comparare i differenti cicli è quella svolta al paragrafo 11.1 in cui si tiene conto di cilindrata e di step energetico come variabili esplicative. La superficie di regressione ottenuta è descritta dall'equazione 11.2. Tale regressione è stata svolta anche sui dati ottenuti per i cicli AUDC, ARDC ed AMDC. In figura 12.2 sono riportate le regressioni ottenute.



Velocità media: 47,1693 km/h

Numero di accelerazioni: 299

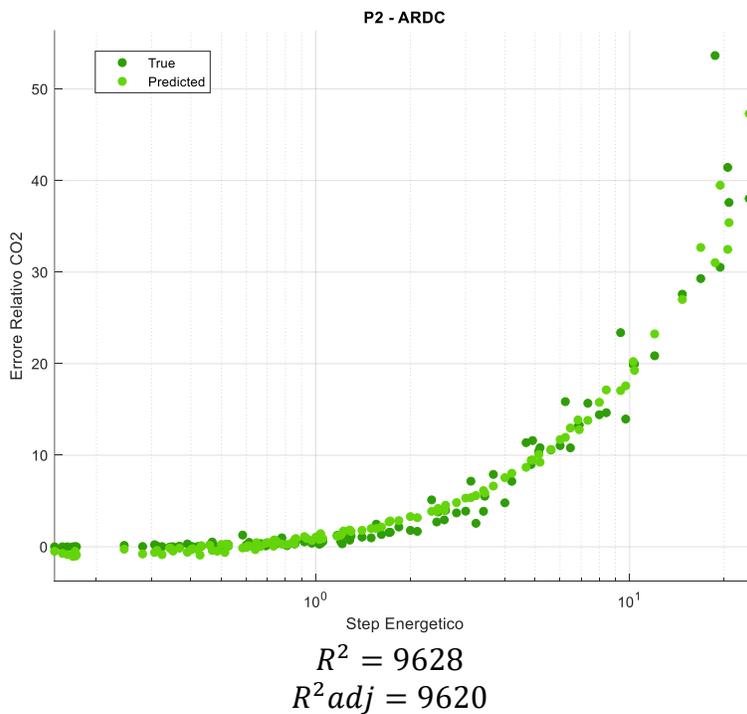
Numero di accelerazioni pesate sul tempo del tratto rispetto al tempo totale della missione: 124



Velocità media: 18,65 km/h

Numero di accelerazioni: 258

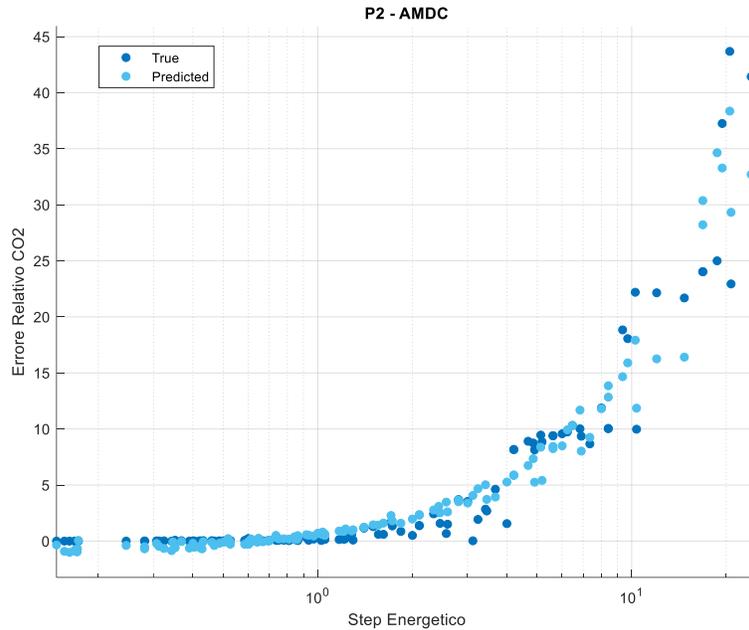
Numero di accelerazioni pesate sul tempo del tratto rispetto al tempo totale della missione: 81



Velocità media: 57,41 km/h

Numero di accelerazioni: 252

Numero di accelerazioni pesate sul tempo del tratto rispetto al tempo totale della missione: 87



$$R^2 = 9628$$

$$R^2_{adj} = 9620$$

Velocità media: 57,41 km/h

Numero di accelerazioni: 252

Numero di accelerazioni pesate sul tempo del tratto rispetto al tempo totale della missione: 87

Figura 12.2 – Regressioni step energetico, cilindrata, errore relativo per le differenti missioni

Nelle seguenti tabelle, dalla 12.1 alla 12.4 sono riportati i risultati ottenuti per le differenti missioni.

Tabella 12.1 – Risultati WHVC

Coefficiente	Stima	P-value
Intercetta	6,5625	0,00127
Step En	2,6572	0,0000
EngDisp	-4,8350	0,0023
StepEn*EngDisp	-0,0253	0,2570
Step En^2	-0,0155	0,0000
EngDisp^2	0,7023	0,0028

Tabella 12.2 – Risultati AUDC

Coefficiente	Stima	P-value
Intercetta	-9,5258	0,0044
Step En	5,5662	0,0000
EngDisp	40,2729	0,0061
StepEn*EngDisp	-0,09765	0,0000
Step En^2	-0,0089	0,0388
EngDisp^2	-5,1247	0,0077

Rispetto alla missione WHVC si può notare nella missione AUDC una diminuzione del numero di accelerazioni (pesate rispetto al tempo), nei coefficienti si può notare in particolare una diminuzione del valore di intercetta. Rispetto alla missione WHVC in questo caso si ha inoltre una diminuzione della velocità media.

<i>Coefficiente</i>	<i>Stima</i>	<i>P-value</i>
<i>Intercetta</i>	-1,8827	0,9661
<i>Step En</i>	3,1123	0,0000
<i>EngDisp</i>	-0,4487	0,9847
<i>StepEn*EngDisp</i>	-0,2422	0,09009
<i>Step En^2</i>	-0,0133	0,0578
<i>EngDisp^2</i>	0,1643	0,9571

Anche in questo caso, rispetto alla missione WHVC si può notare una diminuzione del numero di accelerazioni, si nota inoltre una diminuzione del valore di intercetta. Il numero di accelerazioni pesate rispetto al tempo è maggiore rispetto a quelle per la missione AUDC, si nota che si ha un valore di intercetta maggiore rispetto alla superficie di regressione per la missione AUDC. Rispetto alla missione AUDC in questo caso si ha inoltre un aumento della velocità media.

<i>Coefficiente</i>	<i>Stima</i>	<i>P-value</i>
<i>Intercetta</i>	3,3168	0,9151
<i>Step En</i>	4,5404	0,0000
<i>EngDisp</i>	-3,2100	0,8445
<i>StepEn*EngDisp</i>	-0,8044	0,0000
<i>Step En^2</i>	0,0070	0,1520
<i>EngDisp^2</i>	0,5499	0,7970

Anche in questo caso, rispetto alla missione WHVC si può notare una diminuzione del numero di accelerazioni ed anche in questo caso si va a notare una diminuzione del valore di intercetta. Rispetto alla missione WHVC si ha inoltre un aumento della velocità media e si nota un aumento di velocità media anche rispetto ai cicli AUDC ed ARDC: il valore di intercetta è superiore rispetto a queste ultime missioni.

Dalle regressioni ottenute con differenti missioni si può osservare che i piani vengano spostati verso il basso se diminuisce il numero di accelerazioni. Si può anche osservare che l'aumento della velocità media contribuisce a spostare le superfici di regressione ottenute verso l'alto, facendo aumentare i valori di intercetta. Poiché la missione viene descritta tramite entrambi i due parametri citati, le superfici di regressione vanno a modificarsi tenendo conto della combinazione tra questi due parametri. È possibile notare che il termine che va a variare maggiormente è il termine legato alla cilindrata, si osserva inoltre che non è statisticamente significativo per tutte le regressioni. Questo può essere dovuto al numero di dati ridotto con cui sono state eseguite le regressioni per i cicli ARTEMIS. Si è comunque evidenziato un legame importante tra velocità media, numero di accelerazioni, forma e posizione delle superfici di regressione ottenute. L'analisi svolta è un'analisi di base che necessita approfondimenti, è necessario in futuro raccogliere un numero superiore di dati su cui svolgere le regressioni

13. Possibili sviluppi futuri

Il lavoro svolto in questa tesi ha permesso di analizzare i risultati forniti dal tool di ottimizzazione HEVbox creato precedentemente in modo da determinare l'effetto di differenti discretizzazioni della batteria sull'errore che il codice fornisce sul risultato finale. L'obiettivo iniziale era quello di analizzare l'andamento dell'errore relativo in funzione dello step energetico per l'architettura p2 in modo da verificare l'eventuale esistenza di un'unica equazione per tutti i layout dell'architettura. L'analisi svolta ha fatto emergere la probabile dipendenza dell'errore relativo dai parametri di design che descrivono il veicolo ibrido.

Il lavoro svolto ha quindi permesso di evidenziare relazioni tra l'errore relativo, lo step energetico ed i parametri che vanno a definire il layout del veicolo ibrido. Tali relazioni sono state messe in luce dapprima tramite l'utilizzo della regressione semplice e successivamente tramite l'utilizzo della regressione multipla con l'aiuto di strumenti statistici ulteriori quali la convalida incrociata e l'algoritmo stepwise selection.

Sono stati ricavati ed analizzati diversi modelli per l'architettura p2 considerando la missione WHVC. Sono state eseguite le stesse regressioni anche per l'architettura p3: le regressioni sono state eseguite però andando a considerare un numero inferiore di dati, per cui queste ultime sono state utilizzate per confrontare la differenza tra le superfici ottenute per l'architettura p2 e per l'architettura p3. In futuro, si potrebbe approfondire questa indagine andando a ricavare un numero maggiore di dati per l'architettura p3 in modo da svolgere analisi più accurate. Potrebbe essere inoltre interessante indagare su ulteriori tipi di architettura non considerati in questa tesi.

Come ulteriore sviluppo è da considerare dopo le analisi svolte l'utilizzo di dati ricavati per cicli guida differenti, in modo da verificare la differenziazione delle equazioni ottenute quando si vanno a considerare missioni ulteriori. In questa tesi infatti sono stati infatti considerati solamente i cicli WHVC ed ARTEMIS.

Si può inoltre pensare di svolgere in futuro differenti analisi statistiche rispetto a quelle presentate in questo lavoro. Per esempio, sarebbe interessante creare un modello in cui vengono coinvolti tutti i parametri che descrivono il layout del veicolo ed in cui si utilizzano tipi differenti di regressione, per esempio una Lasso regression che permette di scegliere i parametri più adatti per costruire un modello.

Ringraziamenti

Al termine di questo percorso devo innanzitutto ringraziare mamma, papà e Beatrice per l'aiuto, per la forza e per il conforto che mi hanno dato ogni giorno.

Un ringraziamento particolare va alla mia relatrice, la professoressa Daniela Misul, che mi ha dato la possibilità di mettermi alla prova, di maturare e di svolgere un lavoro interessante, complesso e gratificante. Ringrazio anche Alessia Musa, che mi ha sempre seguita durante il periodo di lavoro ed è stata sempre molto precisa e attenta nelle indicazioni che mi ha fornito.

Non posso dimenticare Luca e Camilla a cui va un ringraziamento sentito. Grazie anche a tutti i compagni che ho incontrato in questo percorso, soprattutto Francesco, Alberto e Francesca con cui ho condiviso ogni giorno al Politecnico.

Bibliografia

- [1] *HEV introduction and architectures* – Ezio Spessa - 2018
- [2] <https://climate.nasa.gov/> - 2019
- [3] https://ec.europa.eu/clima/policies/transport_en - 2019
- [4] *Zero Emission Vehicles, An optimal control for hybrid vehicles* – Gianmarco Passiatore, 2018
- [5] *Misurare per decidere* - G. Barbato, A. Germak, G. Genta - 2014
- [6] *Statistica* - David M. Levine, Timothy C. Krehbiel, Mark L. Berenson -2000
- [7] *An introduction to statistical learning: with application in R* – Gareth James, Daniela Witten, Trevor Hastie, Robert Tibshirani - 2013
- [8] *HEVBox user manual* – revisione 2/10/2019