

Piano di Qualifica

versione 1.0.0



7commits@gmail.com

Progetto di Ingegneria del Software

A.A. 2024/2025

Destinatari	Responsabile	Redattori	Verificatori
Prof. Tullio Vardanega	Marco Cola	Stefano Dal Poz	Giulia Hu
Prof. Riccardo Cardin		Marco Cola	Ruize Lin
Gruppo 🇮🇹 7Commits		Ruize Lin	Michele Ogniben
		Mattia Piva	Giada Rossi

Registro delle modifiche

Versione	Data	Autori	Verificatori	Descrizione
v1.0.0	2025-06-12	Marco Cola	Ruize Lin	Approvazione per RTB
v0.1.1	2025-06-04	Marco Cola	Ruize Lin	Aggiunta sezione 5 «Considerazioni finali»
v0.1.0	2025-06-03	Marco Cola	Ruize Lin	Corretti grafici e struttura dei grafici, con unionie grafici CV/SV, ETC/AC, EV/PV. Aggiunti NCR e RSI + Gulpease
v0.0.9	2025-05-30	Giada Rossi	Ruize Lin, Giulia Hu	Aggiunti grafici ETC, EAC
v0.0.8	2025-05-29	Giada Rossi	Ruize Lin, Giulia Hu	Aggiunti grafici AC, PV, EV
v0.0.7	2025-05-19	Stefano Dal Poz	Ruize Lin	Aggiunti riferimenti al Glossario
v0.0.6	2025-05-18	Mattia Piva	Ruize Lin	Ultimata sezione test di sistema e aggiunti valori sugli obiettivi metrici di qualità
v0.0.5	2025-05-16	Giada Rossi	Ruize Lin	Prima bozza test di sistema
v0.0.4	2025-05-12	Giada Rossi	Michele Ogniben, Ruize Lin	Inizio paragrafo 3 Metodologia di testing
v0.0.3	2025-05-09	Giada Rossi	Michele Ogniben, Ruize Lin	Obiettivi metrici di qualità
v0.0.2	2025-05-06	Giada Rossi	Michele Ogniben	Capitolo Introduzione
v0.0.1	2025-04-14	Marco Cola	Michele Ogniben	Prima Bozza

Indice

1. Introduzione	4
1.1. Scopo del documento	4
1.2. Scopo del prodotto	4
1.3. Glossario	4
1.4. Riferimenti	4
1.4.1. Riferimenti normativi	4
1.4.2. Riferimenti informativi	4
2. Obiettivi metrici di qualità	6
2.1. Qualità di processo	6
2.1.1. Processi primari	6
2.1.1.0.1. Fornitura	6
2.1.1.1. Sviluppo	6
2.1.2. Processi di supporto	6
2.1.2.1. Documentazione	6
2.1.2.2. Gestione della qualità	6
2.2. Qualità di prodotto	7
2.2.1. Funzionalità	7
2.2.2. Affidabilità	7
2.2.3. Manutenibilità	7
2.2.4. Efficienza	7
3. Metodologie di testing	8
3.1. Codifica dei test	8
3.2. Test di sistema	8
3.3. Test di accettazione	9
4. Cruscotto di valutazione	10
4.1. EAC (Estimated At Completion)	10
4.1.1. Descrizione	10
4.2. Earned Value (EV) e Planned Value (PV)	11
4.2.1. Descrizione	11
4.3. Actual Cost (AC) e Estimate to Complete (ETC)	12
4.3.1. Descrizione	12
4.4. Cost Variance (CV) e Schedule Variance (SV)	13
4.4.1. Descrizione	13
4.5. Requirements stability index (RSI)	14
4.5.1. Descrizione	14
4.6. Indice Gulpease	15
4.6.1. Descrizione	15
4.7. Non-Calculated Risk	16
4.7.1. Descrizione	16
5. Considerazioni finali in vista della revisione RTB	17
5.1. Analisi delle Pratiche	17
5.2. Valutazioni Generali	17
5.3. Miglioramenti Specifici nei Processi	17
5.4. Conclusione	17

1. Introduzione

1.1. Scopo del documento

Il presente documento ha l'obiettivo di delineare le procedure di verifica e validazione adottate per assicurare la *qualità_G* sia dei processi che del prodotto software. Essendo il miglioramento continuo un principio fondamentale dei sistemi di *qualità_G*, il contenuto del documento sarà aggiornato e ampliato progressivamente, con l'intento di riflettere i risultati delle verifiche eseguite e di garantire il costante miglioramento del progetto nel tempo.

1.2. Scopo del prodotto

Lo scopo principale del *prodotto_G* è quello di fornire all'azienda *proponente_G*, Zucchetti S.p.A., una piattaforma *software_G* denominata Artificial QI, pensata per supportare la valutazione delle prestazioni di sistemi basati su Large Language Models (LLM). Il *sistema_G* nasce dall'esigenza di disporre di uno strumento centralizzato, automatizzato e configurabile che consenta di testare, analizzare e confrontare le risposte generate da diversi modelli *LLM_G* rispetto a risposte attese predefinite.

La piattaforma dovrà permettere all'utente sviluppatore di inserire e gestire domande e risposte attese, interfacciarsi con *LLM_G* esterni tramite chiamate *API_G* conformi allo standard *OpenAPI 3.1_G*, ed eseguire valutazioni automatiche della coerenza tra input e output tramite metriche personalizzabili (come *BM25_G* o modelli neurali). Inoltre, dovrà offrire *funzionalità_G* avanzate di visualizzazione, ordinamento e confronto dei risultati ottenuti, favorendo anche l'archiviazione e l'analisi storica delle esecuzioni.

Il *prodotto_G* finale sarà un'unica *applicazione_G* integrata, che racchiude tutte le *funzionalità_G* necessarie in un flusso di lavoro continuo e organico. Tale soluzione dovrà garantire semplicità d'uso, flessibilità nella configurazione e supporto per l'analisi evolutiva delle prestazioni dei modelli, diventando così un valido supporto nelle *attività_G* di test e selezione degli *LLM_G* più adatti per specifici scenari applicativi.

1.3. Glossario

Per evitare ambiguità relative alla terminologia utilizzata, il documento fa riferimento a un Glossario, che raccoglie definizioni chiare e univoche dei termini specifici del dominio d'uso. I termini inclusi nel glossario sono contrassegnati nel testo con una G a pedice, in *questo modo_G*.

1.4. Riferimenti

1.4.1. Riferimenti normativi

- Norme di Progetto
- Capitolato d'appalto: <https://www.math.unipd.it/~tullio/IS-1/2024/Progetto/C1.pdf>
- Regolamento del progetto didattico: <https://www.math.unipd.it/~tullio/IS-1/2024/Dispense/PD1.pdf>

1.4.2. Riferimenti informativi

- Glossario
- Qualità del software: <https://www.math.unipd.it/~tullio/IS-1/2024/Dispense/T07.pdf>
- Qualità del processo: <https://www.math.unipd.it/~tullio/IS-1/2024/Dispense/T08.pdf>
- Verifica e validazione - Introduzione: <https://www.math.unipd.it/~tullio/IS-1/2024/Dispense/T09.pdf>
- Verifica e validazione - Analisi dinamica: <https://www.math.unipd.it/~tullio/IS-1/2024/Dispense/T11.pdf>

- Verifica e validazione - Analisi statica: <https://www.math.unipd.it/~tullio/IS-1/2024/Dispense/T10.pdf>

2. Obiettivi metrici di qualità

Ogni *processo_G* viene valutato mediante l'applicazione di metriche specifiche, le cui specifiche dettagliate sono riportate nel documento Norme di Progetto v1.0.0. Questa sezione delinea i criteri che le metriche devono rispettare per essere valutate come accettabili o eccellenti.

2.1. Qualità di processo

La qualità di processo ha come obiettivo quello di garantire che gli standard definiti nel documento Norme di Progetto vengano rispettati sempre in modo da svolgere al meglio l'intera gamma di *attività_G*. Questo serve per monitorare costantemente l'*efficacia_G* e l'*efficienza_G* dei processi.

2.1.1. Processi primari

2.1.1.0.1. Fornitura

Metrica	Descrizione	Valore accettabile	Valore ideale
CV	Cost Variance	$\geq -10\%$	≥ 0
PV	Planned Value	≥ 0	$\leq \text{BAC}$
EV	Earned Value	≥ 0	$\leq \text{EAC}$
AC	Actual Cost	≥ 0	$\leq \text{EAC}$
CPI	Cost Performance Index	$\pm 10\%$	0%
SPI	Schedule Performance Index	≥ 0.9	≥ 1
BAC	Budget At Completion	-	-
EAC	Estimated At Completion	$\leq 1.1 \times \text{BAC}$	$= \text{BAC}$
VAC	Variance At Completion	$\geq -10\%$	≥ 0
ETC	Estimated To Completion	≥ 0	$\leq \text{EAC}$
SV	Schedule Variance	$\geq -10\%$	≥ 0
BV	Budget Variance	$\pm 10\%$	0%

2.1.1.1. Sviluppo

Metrica	Descrizione	Valore accettabile	Valore ideale
SC	Statement Coverage	$\geq 70\%$	100%

2.1.2. Processi di supporto

2.1.2.1. Documentazione

Metrica	Descrizione	Valore accettabile	Valore ideale
IG	Indice Gulpease	$\geq 60\%$	$\geq 80\%$

2.1.2.2. Gestione della qualità

Metrica	Descrizione	Valore accettabile	Valore ideale
MNS	Metriche Non Soddisfatte	≤ 3	0

2.2. Qualità di prodotto

La qualità del prodotto indica quanto un *software_G* soddisfa requisiti ed esigenze, sia esplicite che implicite. Garantirla significa offrire un *prodotto_G* stabile, efficiente e usabile. L'uso di metriche specifiche consente di valutare aspetti chiave come *funzionalità_G*, affidabilità e manutenibilità, permettendo interventi rapidi in caso di scostamenti dagli standard attesi.

2.2.1. Funzionalità

Metrica	Descrizione	Valore accettabile	Valore ideale
ROS	Requisiti Obbligatori Soddisfatti	100%	100%
RDS	Requisiti Desiderabili Soddisfatti	≥ 50%	100%
RPS	Requisiti Opzionali Soddisfatti	≥ 0%	≥ 50%

2.2.2. Affidabilità

Metrica	Descrizione	Valore accettabile	Valore ideale
PTCP	Passed Test Cases Percentage	≥ 75%	100%
CC	Code Coverage	≥ 80%	100%

2.2.3. Manutenibilità

Metrica	Descrizione	Valore accettabile	Valore ideale
SFIN	Structure Fan IN	-	Va massimizzato
SFOUT	Structure Fan OUT	-	Va minimizzato

2.2.4. Efficienza

Metrica	Descrizione	Valore accettabile	Valore ideale
TDE	Tempo di Elaborazione	≤ 15 minuti	≤ 5 minuti

3. Metodologie di testing

In questa sezione vengono illustrate le modalità con cui sono stati individuati e classificati i *test_G* da applicare al progetto per verificare il rispetto delle metriche previste e le specifiche del documento *Analisi dei Requisiti_G*.

I *test_G* sono suddivisi nelle seguenti categorie:

- **Test di Unità:** riguardano le singole unità *software_G* sviluppate dai programmatori del team ed eseguite in modo automatizzato;
- **Test di Integrazione:** verificano l'integrazione corretta tra le varie parti del *sistema_G*, evidenziando eventuali difetti a livello architetturale, per questo sono svolti successivamente ai precedenti *test_G*
- **Test di Sistema:** assicurano che l'intero prodotto software funzioni correttamente nel suo insieme, verificando il rispetto del documento *Analisi dei Requisiti_G*
- **Test di Accettazione:** verificano che il *prodotto_G* finale sia conforme a quanto richiesto nel *capitolato_G* per questo vengono fatti alla presenza del *committente_G*.

3.1. Codifica dei test

Ogni *test_G* viene associato ad un codice univoco nel formato **T[Tipologia]-[Numero]**, dove Tipologia indica la tipologia del *test_G*:

- **U:** di unità;
- **I:** di integrazione;
- **S:** di sistema;
- **A:** di accettazione.

Inoltre, ogni *test_G* appartiene ad uno Stato, che può essere:

- **V:** verificato, cioè con esito positivo;
- **E:** non verificato, cioè con esito negativo;
- **NI:** non implementato.

3.2. Test di sistema

I test di sistema rappresentano una *fase_G* fondamentale del *processo_G* di verifica del *software_G*, in cui viene valutato il comportamento dell'intero applicativo una volta che tutte le sue componenti sono state integrate. L'obiettivo è accertarsi che il *sistema_G*, nel suo complesso, soddisfi i requisiti funzionali e non funzionali definiti nell'*Analisi dei Requisiti_G*. Questa *attività_G* viene svolta in un ambiente che riproduce il più fedelmente possibile le condizioni d'uso reali, simulando l'interazione dell'utente finale con il *sistema_G*.

Codice	Descrizione	Requisito	Stato
TS-01	Verificare che l'utente sviluppatore possa archiviare una lista di domande e risposte attese	RFO01	NI
TS-02	Verificare che sia possibile organizzare le domande in set riutilizzabili	RFO02	NI
TS-03	Verificare che sia possibile configurare l'endpoint API di un LLM esterno	RFO03	NI
TS-04	Verificare l'invio di domande a un LLM esterno tramite API e la ricezione delle risposte.	RFO04	NI
TS-05	Verificare il confronto automatico risposte attese e risposte ricevute	RFO05	NI

TS-06	Verificare il salvataggio e l'archiviazione dei risultati delle valutazioni	RFO06	NI
TS-07	Verificare che sia possibile visualizzare e consultare i risultati	RFO07	NI
TS-08	Verificare che sia possibile il confronto tra diverse esecuzioni di test	RFO08	NI
TS-09	Verificare che sia possibile filtrare lo storico dei test per diversi criteri	RFO09	NI
TS-10	Verificare che sia possibile aggiungere, modificare e gestire i modelli LLM multipli	RFO10	NI

3.3. Test di accettazione

Codice	Descrizione	Fonte	Stato
TA-01	Verificare che l'utente possa creare, visualizzare, modificare e cancellare set di domande	UC1, UC5, UC6, UC7. UC8	NI
TA-02	Verificare che l'utente possa gestire i modelli LLM, visualizzando, aggiungendo, selezionando e rimuovendo un LLM	UC3, UC20, UC21, UC22	NI
TA-03	Verificare che il test venga avviato correttamente e che il sistema registri le risposte	UC4, UC5, UC6, UC9	NI
TA-04	Verificare che l'utente possa visualizzare e confrontare test effettuati in precedenza	UC10, UC11, UC12, UC14, UC15, UC16, UC17, UC18, UC19	NI
TA-05	Verificare che l'utente possa scaricare i risultati	UC13	NI

4. Cruscotto di valutazione

4.1. EAC (Estimated At Completion)

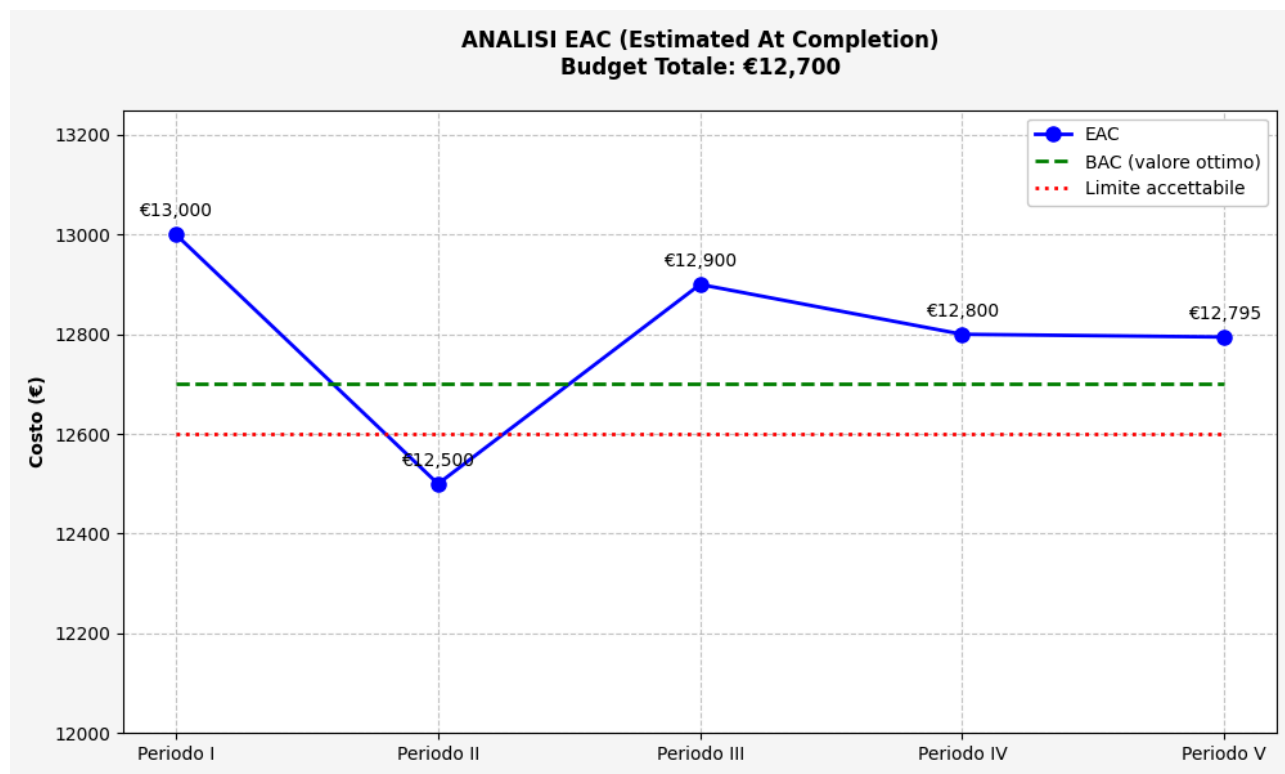


Figura 1: Proiezione della stima del costo totale nei vari periodi di progetto

4.1.1. Descrizione

Il grafico confronta il budget totale (€12,700) con l'EAC nei diversi periodi. Mentre il BAC (valore ottimo) rappresenta l'obiettivo ideale, l'EAC si avvicina progressivamente al limite accettabile, indicando una gestione dei costi efficace. I dati mostrano un lieve sforamento iniziale, corretto nei periodi successivi grazie a revisioni delle stime. Ciò riflette l'adattamento del gruppo alle dinamiche di progetto, mantenendo i costi sotto controllo.

4.2. Earned Value (EV) e Planned Value (PV)

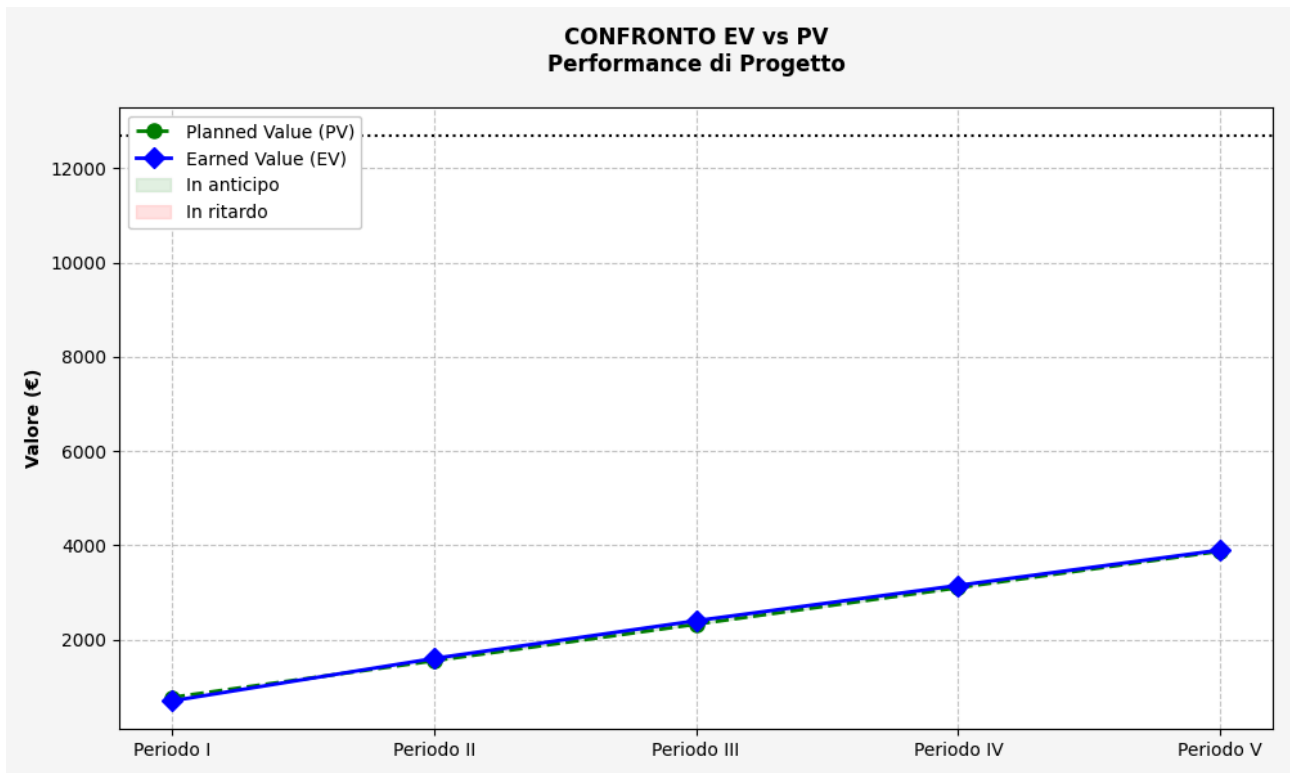


Figura 2: Proiezione dell'EV e del PV nei vari periodi di progetto.

4.2.1. Descrizione

Il grafico EV vs PV mostra un sostanziale allineamento tra Earned Value e Planned Value, confermando che l'avanzamento reale del progetto rispecchia fedelmente la pianificazione iniziale. Questa sovrapposizione delle curve indica una gestione efficace, con il team che ha mantenuto una produttività costante senza significativi scostamenti temporali o quantitativi.

4.3. Actual Cost (AC) e Estimate to Complete (ETC)

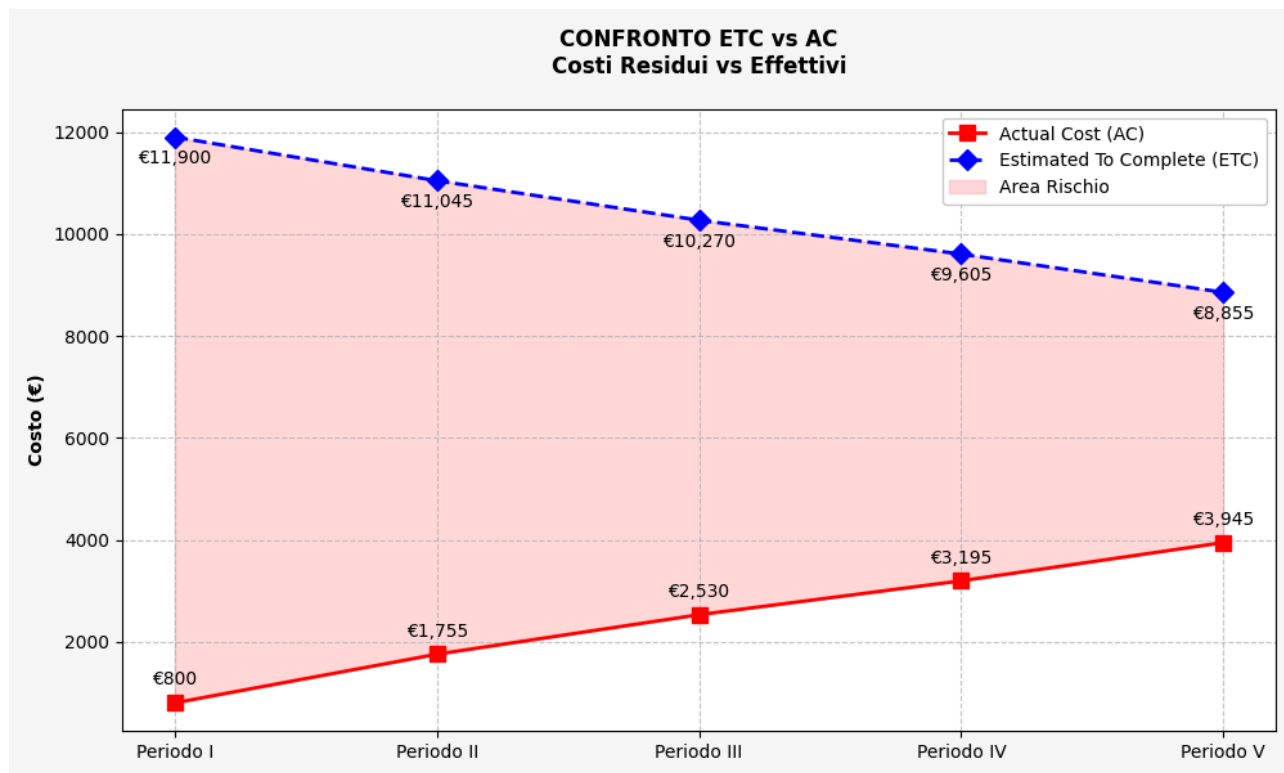


Figura 3: Proiezione dell'AC e dell'ETC nei vari periodi di progetto.

4.3.1. Descrizione

Il grafico mostra l'andamento dell'Estimate to Complete (ETC) e dell'Actual Cost (AC). L'ETC diminuisce progressivamente, indicando la riduzione del costo residuo previsto, mentre l'AC cresce in modo proporzionale, riflettendo i costi effettivi sostenuti periodo dopo periodo. Questo andamento contrapposto conferma un avanzamento del progetto coerente con le previsioni.

4.4. Cost Variance (CV) e Schedule Variance (SV)

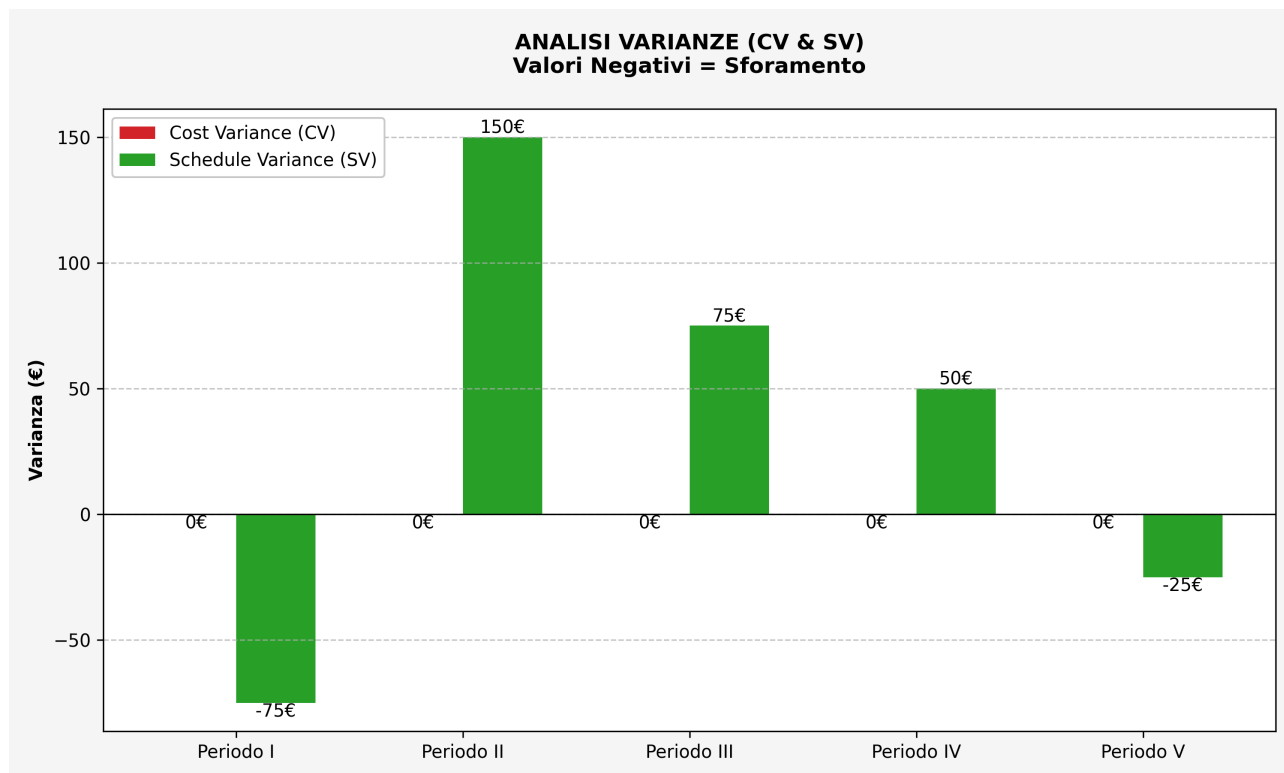


Figura 4: Proiezione della CV e della SV nei vari periodi di progetto.

4.4.1. Descrizione

Le varianze di costo (CV) e schedulazione (SV) rivelano andamenti distinti. Il CV costante a 0€ indica un perfetto allineamento tra costi e avanzamento. Lo SV, invece, mostra fluttuazioni: negativo nei Periodi I e V (ritardi per incomprensioni e revisioni del PoC), positivo nei Periodi II-III-IV. Questa analisi aiuta a identificare le fasi critiche del progetto.

4.5. Requirements stability index (RSI)

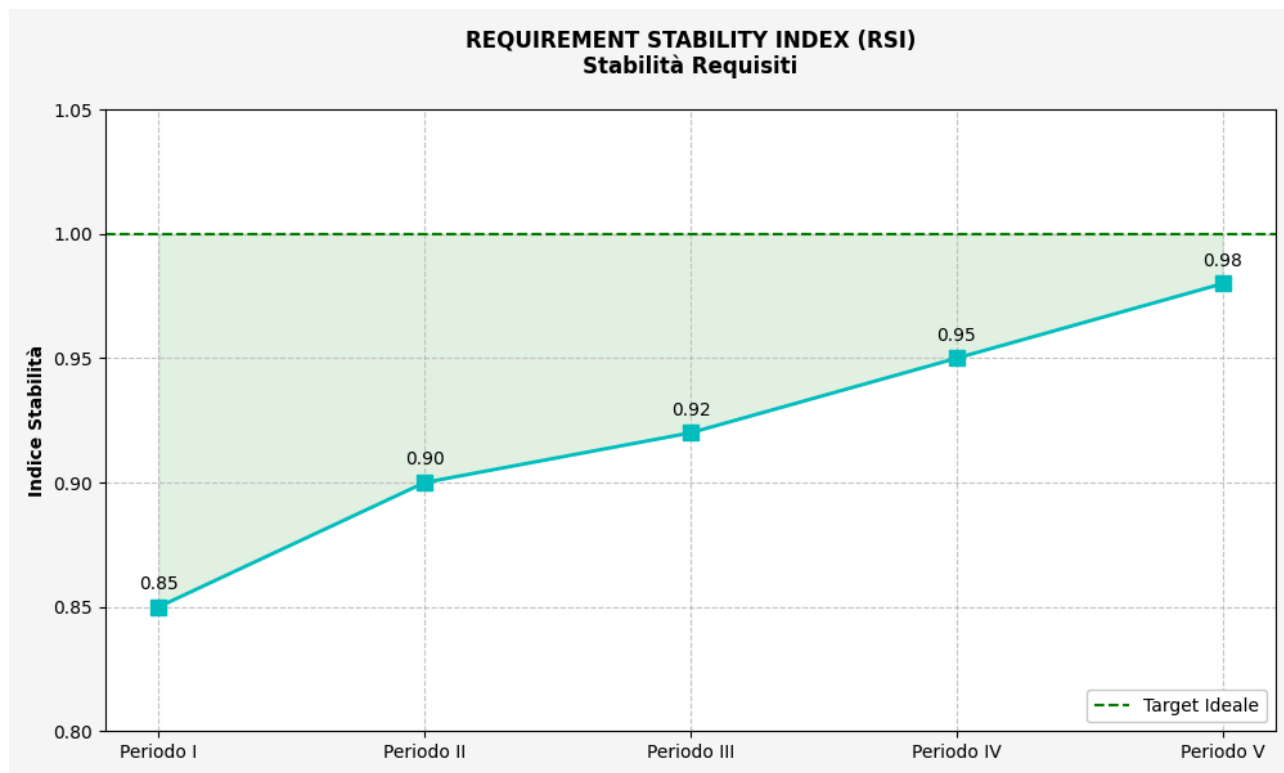


Figura 5: Proiezione del RSI nei vari periodi di progetto.

4.5.1. Descrizione

L'RSI misura la stabilità dei requisiti nel tempo. Il picco veloce iniziale (Periodo I) riflette l'attività intensiva di analisi, seguito da cali (Periodi II-III) per aggiornamenti e maggior dettaglio. Il raggiungimento del 98% nel Periodo V indica la maturità dei requisiti, senza ulteriori modifiche.

4.6. Indice Gulpease

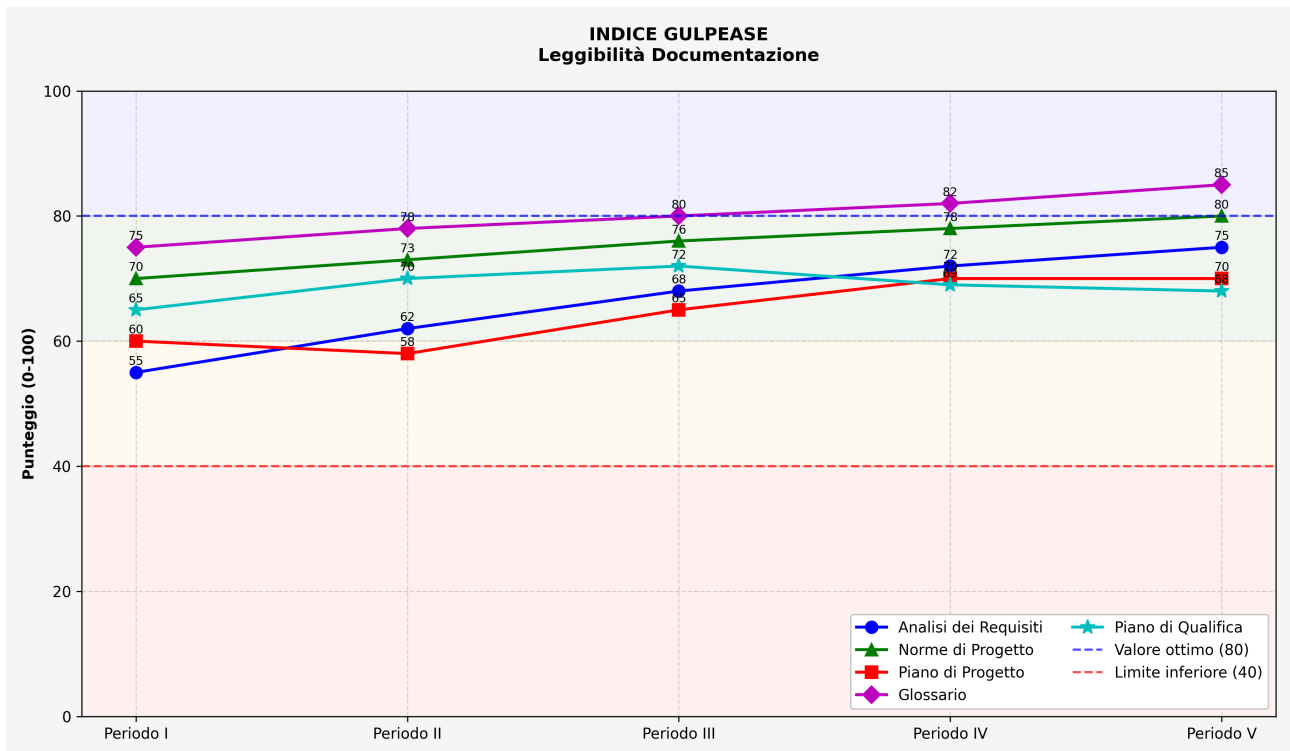


Figura 6: Proiezione dell'indice Gulpease per ogni documento (RTB) nei vari periodi di progetto.

4.6.1. Descrizione

L'indice valuta la leggibilità dei documenti. La tendenza generale è positiva, con tutti i documenti sopra la soglia di accettabilità.

4.7. Non-Calculated Risk

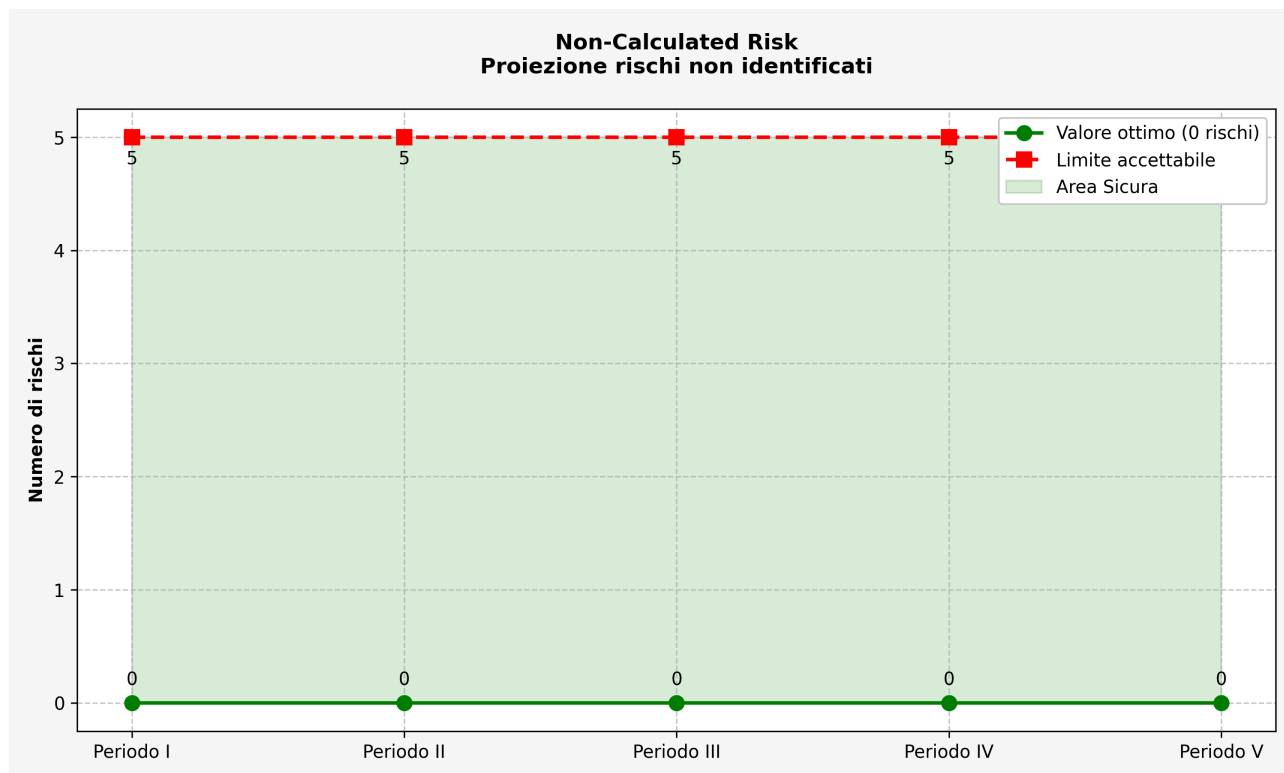


Figura 7: Proiezione dell'indice Gulpease per ogni documento (RTB) nei vari periodi di progetto.

4.7.1. Descrizione

Il grafico illustra la proiezione dei rischi non identificati durante i vari periodi del progetto. La linea del Valore ottimo (0 rischi) e il Limite accettabile superiore mostrano un'area sicura (verde) in cui il gruppo ha mantenuto un controllo efficace. L'assenza di picchi evidenzia una previsione accurata dei rischi.

5. Considerazioni finali in vista della revisione RTB

Questa sezione riassume i risultati dell'analisi retrospettiva condotta sul percorso progettuale, evidenziando l'evoluzione del way of working e gli impatti sul team. L'obiettivo è delineare le trasformazioni operative che hanno aumentato efficacia ed efficienza, identificando sia i progressi raggiunti sia le aree di ulteriore ottimizzazione.

5.1. Analisi delle Pratiche

Inizialmente, il team ha adottato un approccio reattivo, correggendo gli errori man mano che emergevano, a causa della scarsa esperienza con il ciclo PDCA (Plan-Do-Check-Act). Successivamente, l'adozione di obiettivi quantificabili e selettivi ha permesso di passare a un miglioramento proattivo, ottimizzando risorse e prevenendo criticità.

5.2. Valutazioni Generali

Nonostante l'assenza di obiettivi formali iniziali, l'esperienza acquisita ha portato a un affinamento progressivo dei processi. Le autocorrezioni sono diminuite solo quando il way of working si è stabilizzato, allineandosi alle buone pratiche apprese.

5.3. Miglioramenti Specifici nei Processi

Comunicazioni e incontri:

- Riduzione della durata degli incontri (max 1 ora) e introduzione di sondaggi per ottimizzare la presenza.
- Ogni membro condivide pubblicamente lo stato del lavoro, garantendo partecipazione attiva.
- Gli assenti ricevono un riassunto vocale dei punti chiave.

Pianificazione:

- Rotazione dei ruoli gestita attraverso discussioni collettive sulla struttura del progetto, evitando duplicazioni o confusioni.
- Domande specifiche affrontate in chat private per ottimizzare il tempo dei meeting.

5.4. Conclusione

Il percorso ha dimostrato come l'esperienza diretta e l'adattamento iterativo siano fondamentali per trasformare un approccio inizialmente caotico in un metodo strutturato e sostenibile. Le soluzioni implementate, seppur semplici, hanno ridotto sprechi di tempo e migliorato la coesione del team.