**Destination Renewable Energy (DRE)**

Software Verification and Validation Report

Ref.: DestinE\_ESA\_DRE\_ SVVR \_v5.1



**Author's Table**

|  |  |  |
| --- | --- | --- |
| Written by: | George Koutalieris  Symeon Symeonidis  Vasillis Perifanis  Iphigenia Kapsomenaki  Rizos-Theodoros Chadoulis  Aggelos Georgakis  Christos Stathopoulos  Platon Patlakas | ENORA INNOVATION  ENORA INNOVATION  ENORA INNOVATION  ENORA INNOVATION  NOA  NOA  WeMET PC  WeMET PC |
| Verified by: | Theodora Papadopoulou | NOA |
| Approved by: | Haris Kontoes | NOA |

**Change Log**

|  |  |  |  |
| --- | --- | --- | --- |
| Issue | Date | Reason for change | Section(s) changed |
| 1.0 | 13 February 2024 | Creation of the document |  |
| 2.0 | 14 May 2024 | New Version | Edits in the entire document |
| 3.0 | 4 September 2024 | New Version | Edits in the entire document |
| 3.1 | 11 September 2024 | New Version | Comments integrated |
| 3.2 | 13 September 2024 | New Version | Appendix updated |
| 4.0 | 24 October 2024 | New Version | Edits in the entire document |
| 5.1 | 12 May 2025 | Fifth Version | Edits in the entire document |

**Table of Contents**

[1 Introduction 5](#_heading=h.30j0zll)

[1.1 Purpose and Scope 5](#_heading=h.1fob9te)

[2 Verification process & results 6](#_heading=h.3znysh7)

[3 Validation process & results 7](#_heading=h.2et92p0)

[4 Validation of the accuracy of forecasting 8](#_heading=h.oxtl4trsi8uv)

[4.1 Solar forecasting validation 8](#_heading=h.j3pd7g78shf)

[4.2 4.2 Wind forecasting validation 13](#_heading=h.uajv4tgpf6w0)

[5 Issue Tracking and Resolution 18](#_heading=h.tyjcwt)

[6 Resolved UAT Issues and New Testing Phase 19](#_heading=)

[7 Conclusion and Recommendations 20](#_heading=h.3dy6vkm)

[8 Appendix 21](#_heading=h.y4tr4fl37mxo)

[A. Assessment Testing Scenarios of User Acceptance Testing - Templates 22](#_heading=)

[B. Forecasting Testing Scenarios of User Acceptance Testing Templates 29](#_heading=)

[C. Assessment Users’ Testing Results 40](#_heading=h.kd4yftf01i7j)

[Solar Assessment 40](#_heading=h.hqq6zjfhvc2b)

[Wind Assessment 44](#_heading=h.2i1cihhhbqpv)

[D. Forecasting Users' Testing Results 49](#_heading=)

[Solar Forecasting 49](#_heading=h.8n3xlsguqrw0)

[Wind Forecasting 54](#_heading=h.y4ywxuhfbhiy)

# Introduction

This Software Verification and Validation (V&V) Report version 5.0 analyses the V&V activities conducted for the Destination Renewable Energy (DRE) Use Case. The primary objective of these activities is to ensure that the developed application, called Hybrid Renewable Energy Forecasting System (HYREF), meets the specified requirements and standards and performs its intended functions safely, reliably, and efficiently. Therefore, this document uses the terms DRE Use Case and HYREF interchangeably.

## Purpose and Scope

The scope of this report, version 5.0, covers all V&V activities undertaken from the inception of the DRE Use Case through its completion. These activities encompass a range of tests, analyses, and evaluations designed to authenticate the software's functional and non-functional requirements, including its performance, security, and compatibility.

The methodology section of this report outlines the approaches and techniques employed in executing the V&V processes. This includes a description of the testing frameworks, tools, and environments used, the criteria for test case selection, and the metrics for evaluating software quality and performance.

This report is structured to provide a clear and detailed account of the V&V activities, findings, and recommendations. It is intended for a broad audience, including project stakeholders, development teams, and quality assurance professionals, to facilitate informed decision-making and continuous improvement in software development practices.

# Verification process & results

The Verification report encompasses a systematic, disciplined, and documented process of evaluating the DRE Use Case application at different stages of its development lifecycle to ensure it meets the specified requirements and standards.

Based on the evaluation of the Criticality and Complexity criteria described in the Software Verification and Validation Plan (SVVP), the analysis depends on the provided conditions. Both criticality results are Medium, and complexity and result are Medium, as shown in Table 1.

**Table 1: Risk Assessment Matrix.**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | | **COMPLEXITY** | | |
| **HIGH** | **MEDIUM** | **LOW** |
| **CRITICALITY** | **HIGH** | **HIGH** | **HIGH** | **MEDIUM** |
| **MEDIUM** | **HIGH** | **MEDIUM** | **LOW** |
| **LOW** | **MEDIUM** | **LOW** | **LOW** |

# Validation process & results

The validation results for the DRE Use Case application in this stage showcase its robustness and reliability in forecasting solar and wind energy production, adhering to the highest accuracy and user satisfaction standards.

Through a validation strategy involving end-users via WeMET & QUEST ENERGY, the application underwent testing phases, including requirement verification, performance analysis, reliability tests, and scalability assessments.

The User Acceptance Testing (UAT) for the HYREF application is crucial to ensure it meets users' needs by providing accurate solar, wind, and hybrid energy forecasts. The application depends on data from the DESP (Digital Energy Simulation Platform) for digital twin simulations.

During WP5, a wide UAT process was established to ensure the DRE Advanced Service's robustness and accuracy for assessing and forecasting solar and wind energy. This iterative process began with the creation of structured testing scenarios tailored to evaluate the functionality and reliability of the assessment modules. These scenarios covered key parameters for energy yield estimations.

A Quality Control (QC) review was conducted on the assessment sections following the initial setup to verify data accuracy, interface consistency, and compliance with defined requirements. Once verified, the User Acceptance Testing phase was initiated, where end-users and stakeholders tested the solar and wind assessment features in realistic conditions. Their feedback was carefully collected, analysed, and used to inform subsequent refinements, ensuring the system met user needs effectively.

This cycle was then repeated for the forecasting modules, involving similar stages of testing scenario creation, quality control review, and UAT for both solar and wind forecasts. The emphasis was on validating user time series data integration, evaluating forecast accuracy, and ensuring the usability of visualisation outputs such as dashboards and time series graphs. Through this structured, feedback-driven approach, the DRE partners successfully delivered a high-quality service aligned with operational expectations and user requirements through this structured, feedback-driven approach.

The UAT and test cases are provided in the deliverables' Appendix.

The resolution of previous UAT issues confirms that the system functions as expected with the available data. Additional enhancements will optimise performance once full data integration from DESP is restored. Feedback from stakeholders has also guided planned improvements to the UI/UX and the introduction of new indicators, which will be incorporated into this DRE release 5.

UAT results show that the system performs as expected with limited data. When complete DESP data is integrated, future enhancements will improve various functionalities.

Partner collaboration ensures thorough evaluation against real-world scenarios, focusing on data integrity, model accuracy, interface, and forecast precision. This provides valuable insights into the application's practicality. Feedback was used to refine the DRE Use Case application, meeting quality requirements.

# Validation of the accuracy of forecasting

## Solar forecasting validation

Based on a 4-month accumulation of Digital Twin forecasting variables, and the actual power production results from the Kinigos (solar) site, a comprehensive validation was carried out, a comparison of the expected and actual results was analyzed for validation purposes. The overall accuracy for this validation demonstrated robust predictive capabilities with an R² value of 0.845, indicating a strong correlation between the predicted and actual power generation values.

Validation metrics used:

The validation process employed multiple complementary metrics to provide a comprehensive assessment of model performance, each offering unique insights:

* **Mean Absolute Error (MAE)** measures the average magnitude of errors without considering direction, providing an intuitive understanding of prediction accuracy in kilowatts. Lower MAE values indicate better overall accuracy regardless of whether predictions are too high or too low.
* **Root Mean Square Error (RMSE)** penalizes large errors more heavily than small ones by squaring the errors before averaging, making it particularly valuable for identifying forecasts with occasional large deviations that might impact grid stability or energy trading decisions.
* **Coefficient of Determination (R²)** quantifies how well the model explains the variance in actual production, with values closer to 1.0 indicating better fit. This metric is particularly useful for assessing the model's ability to capture the overall production patterns.
* **Normalized MAE and RMSE** express errors as a percentage of maximum production capacity, enabling fair comparison across different days and seasons when production scales vary significantly.
* **Mean Bias Error (MBE)** reveals systematic over or underprediction tendencies, with negative values indicating underprediction. For grid operators, this metric is crucial for understanding prediction bias direction for operational planning.
* **Mean Absolute Percentage Error (MAPE)** provides error magnitude relative to actual values, though it can be problematic during low-production periods when small absolute errors result in large percentage errors.
* **Coefficient of Variation of RMSE (CV-RMSE)** normalizes RMSE by the mean of observed values, offering a standardized measure of error relative to the average production.
* **Correlation Coefficient** measures the linear relationship between predicted and actual values, indicating how well the model captures production patterns even when absolute values may differ.

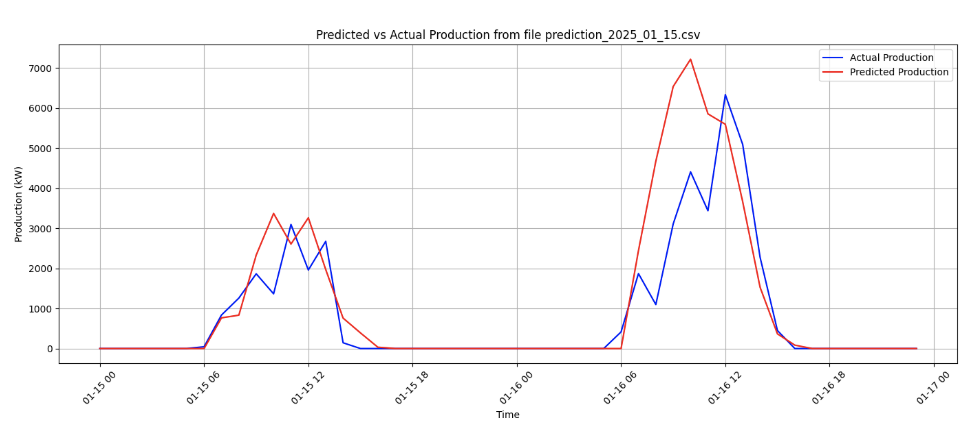
The validation methodology involved a rigorous comparison of hourly forecasted production values against the corresponding actual measurements from the site's monitoring system. For a total of 67 days with matching timestamps, the model achieved a Mean Absolute Error (MAE) of 437.37 kW and a Root Mean Square Error (RMSE) of 940.21 kW. When normalized against the maximum production capacity, these translate to normalized MAE and RMSE values of 4.59% and 9.87%, respectively, demonstrating good accuracy for solar forecasting applications. For the validation period, all the forecasts generated between 22/11/2024 and 22/04/2025 were considered, and selected the ones for which all the necessary DT variables had been successfully downloaded, thus letting us correctly inference the model, and compare its outputs. This resulted in 67 days.

Further analysis of the model's performance shows a mean bias error (MBE) of 27.39 kW, indicating a slight tendency toward overestimation. The Coefficient of Variation of RMSE (CV-RMSE) of 65.01% reflects the model's ability to handle variability across different weather conditions. Particularly noteworthy is the high correlation coefficient of 0.922, confirming the strong linear relationship between forecasted and actual production values.

The validation process incorporated several meteorological variables, including solar zenith angle, relative humidity, ambient temperature, and wind speed, all of which were processed through an XGBoost regression model previously trained on historical data. The model demonstrated particular strength in capturing daily production patterns and peak generation periods, though performance was slightly reduced during highly variable weather conditions.

For operational implementation, this level of accuracy provides reliable input for grid integration planning and energy trading decisions. The validation confirms that the Digital Twin approach effectively captures the complex interactions between meteorological conditions and solar power generation at the Kinigos site, making it a valuable tool for day-ahead production forecasting in real-world applications.

In the Plots 1 - 3 below, specific forecasts can be explored as well as the overall scatterplot of our model for all the validation dates:



**Plot 1**

Metrics for file prediction\_2025\_01\_15.csv:

- MAE: 476.39

- RMSE: 1014.41

- R²: 0.5479

- Normalized MAE: 0.0752

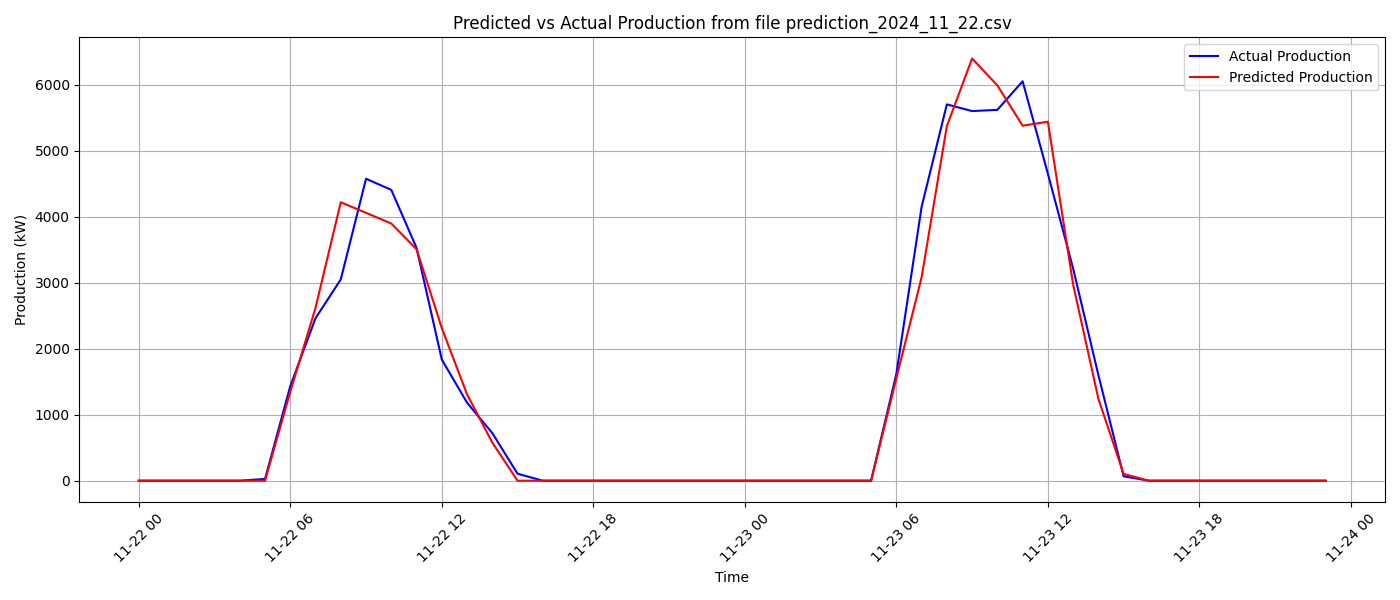
- Normalized RMSE: 0.1601

- MBE: 261.63 (negative = underprediction)

- MAPE: 10566840833367.61%

- CV-RMSE: 116.64%

- Correlation: 0.8704



**Plot 2**

Metrics for file prediction\_2024\_11\_22.csv:

- MAE: 167.97

- RMSE: 338.08

- R²: 0.9700

- Normalized MAE: 0.0278

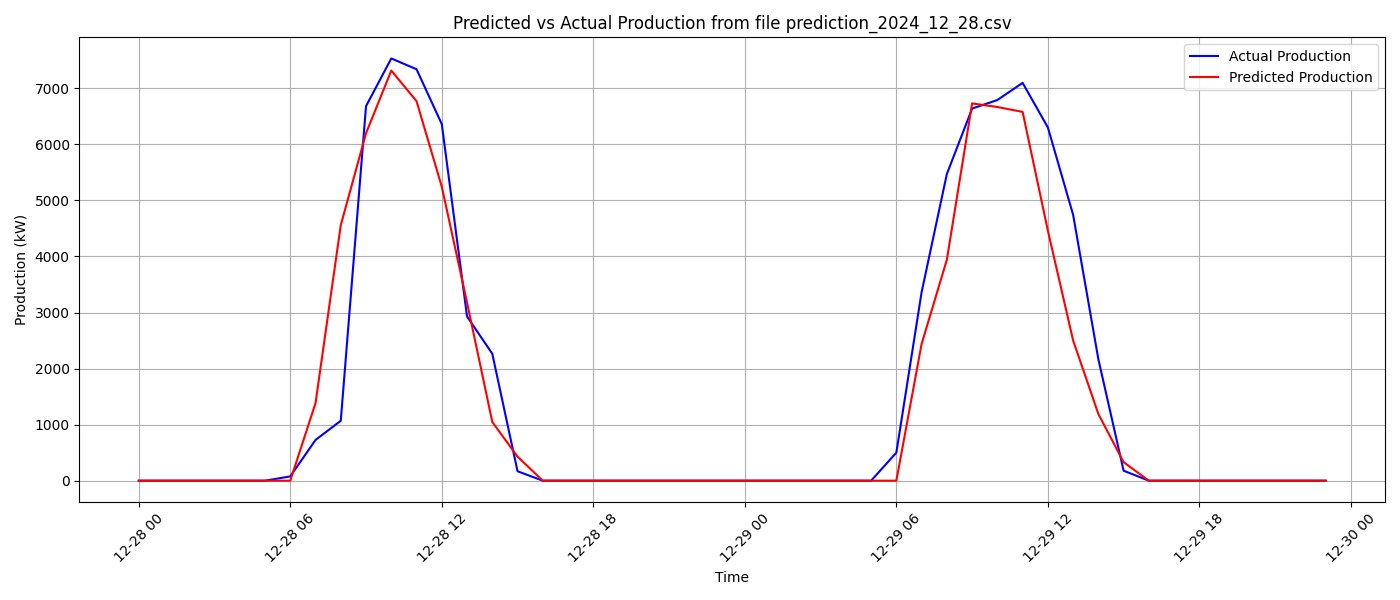
- Normalized RMSE: 0.0559

- MBE: -4.58 (negative = underprediction)

- MAPE: 10.38%

- CV-RMSE: 26.35%

- Correlation: 0.9851

**Plot 3**

Metrics for file prediction\_2024\_12\_28.csv:

- MAE: 358.98

- RMSE: 779.89

- R²: 0.9121

- Normalized MAE: 0.0477

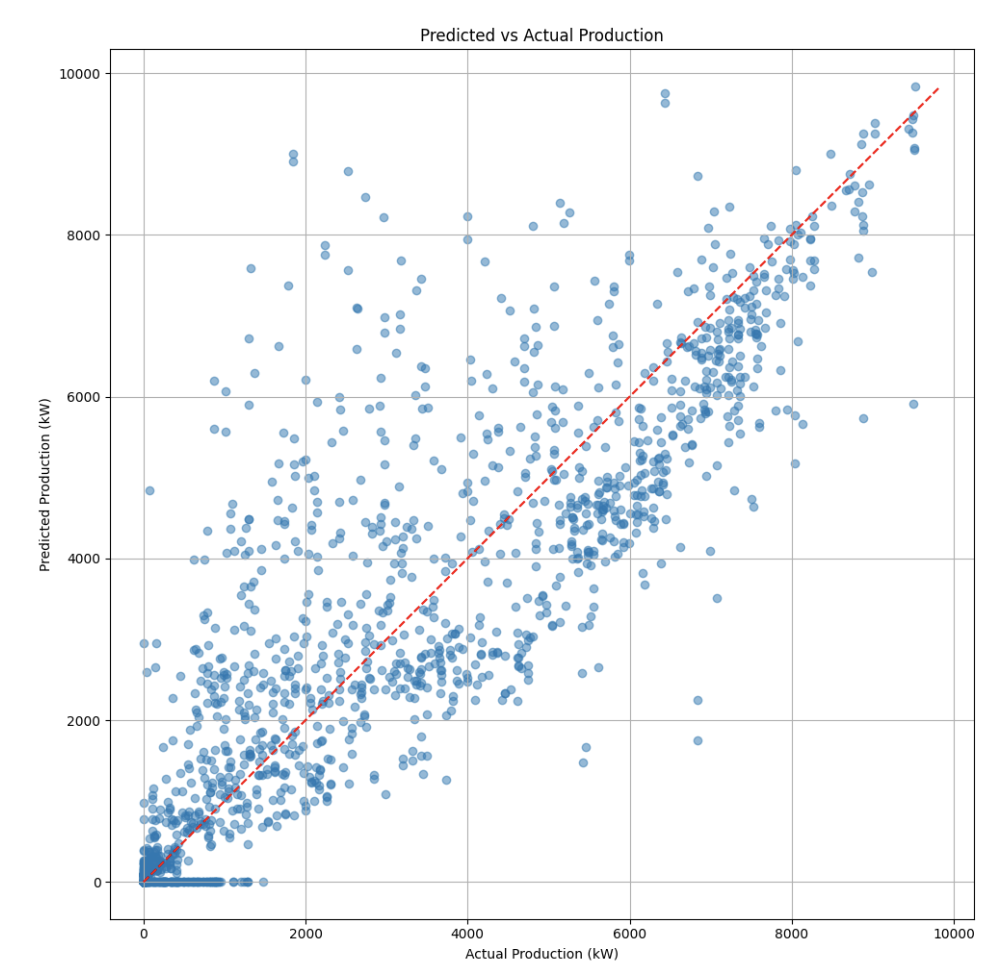
- Normalized RMSE: 0.1035

- MBE: -154.83 (negative = underprediction)

- MAPE: 23.79%

- CV-RMSE: 47.75%

- Correlation: 0.9579



## 4.2 Wind forecasting validation

In this part an evaluation process is carried out between observed and predicted wind power as derived from the use of the Digital Twin forecasting data combined with a machine learning inference. The evaluation is performed with the aid of corresponding power records in locations where wind parks operate. The evaluation covers five sites located in Greece and during the period from 2024-11-20 to 2025-02-06. It should be noted that the period is not completely continuous due to the fact that Digital Twin forecasting data was unavailable on several dates within this time interval. Moreover, part of this period was used for the training phase of the machine learning algorithm, reducing further the actual testing period.

The five sites are presented in Figure 1. One (Site 1) is located in northeastern Greece within a mountainous forest area. Another three sites are situated in central Greece, from which the one (Site 2) is in a mountain ridge near the coast, the second (Site 3) inland and the third one along the coastline (Site 4). The last site (Site 5) is located on an island in the Ionian Sea, at the top of a mountain ridge. Most selected locations are situated in complex terrain areas considering the geomorphological characteristics of the areas including dense forest areas, significant elevation changes as well as the land-sea contrasts. Figure 2 shows some visual examples of the sites under examination indicating the surroundings, topography and potential obstacles influencing the meteorological conditions and mainly the local wind flow at each location.

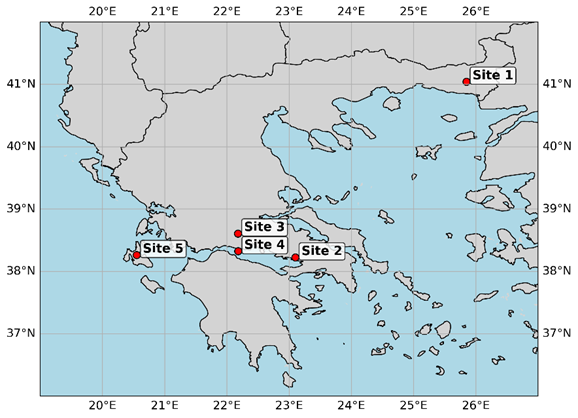


Figure 1 Geographical locations of the five study sites across Greece used for wind power prediction evaluation.

(a) (b)

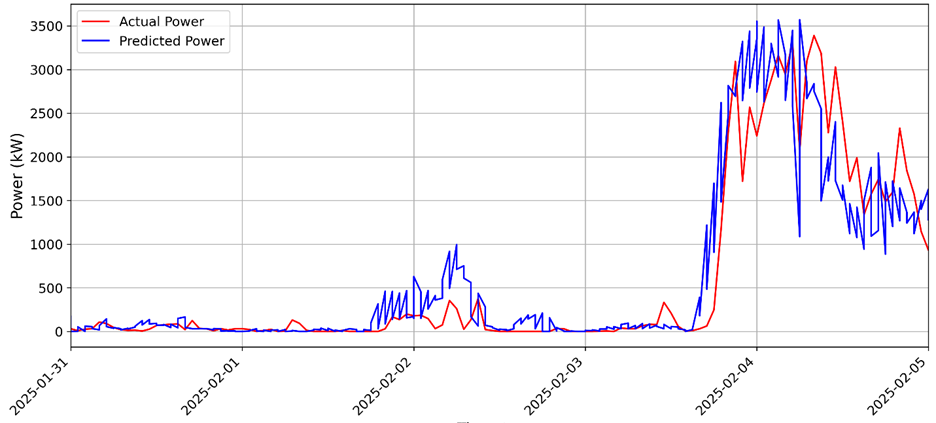


(c)

Figure 2: Visual representation of some of the sites used for evaluation obtained from Google Street View images, (a) Site 1, (b) Site 2 and (c) Site 5.

To estimate the energy yield, a machine learning adjustment was applied to refine the forecasts, with the aid of historical data, based on local weather and production conditions. More precisely, a regression version of the Support Vector Machine (SVM) algorithm was used, namely the Support Vector Regression (SVR) with RBF Kernel. The current scheme is designed to learn non-linear relationships by mapping the input features into a higher-dimensional space where a linear relationship can be found. The main advantage is the ability to model complex and non-linear patterns, using limited input data. The latter was the reason for the selection of the current methodology. More complicated and advanced methods could be used in case of long and continuous forecasting meteorological data. To achieve optimal performance, a training phase is required where specific hyperparameters have to be tuned. The ultimate result is the predicted wind power corresponding to the specific site. Representative examples for two sites are depicted in Figure 3, displaying both the actual and predicted power timeseries.

(a)



(b)

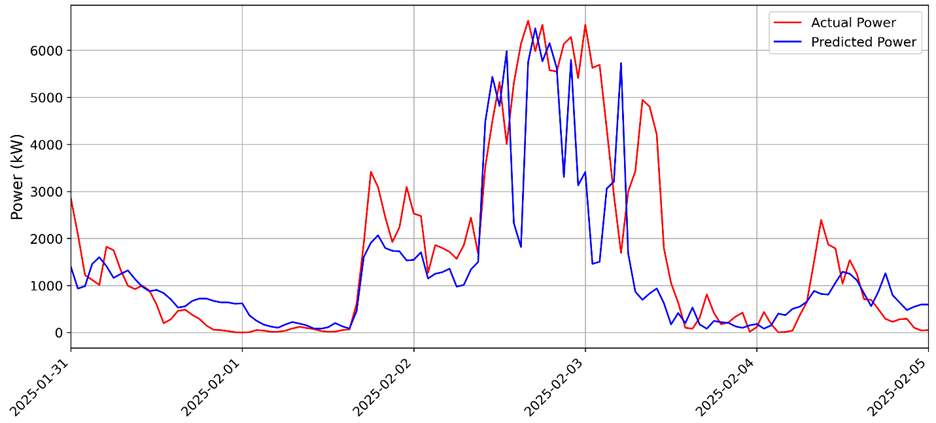


Figure 3 Time Series of forecasting and actual wind power for (a) Site 1 and (b) Site 5

The evaluation is based on the actual recorder power output. To perform the process, ensuring reliability, a quality control of the recorded data was carried out and data with an hourly temporal resolution was kept.

The statistical evaluation is based on the following statistical metrics: the Bias or Mean Bias Error, the Mean Absolute Error (MAE) and the Root Mean Square Error (RMSE). For compatibility across sites and better interpret the performance, the normalized forms of the aforementioned indexes were calculated, expressing the results as percentages, derived from the division with the nominal power corresponding to each site. For a forecasted value, and the corresponding observed one, the statistical indices are calculated based on the following the equations:

· *Normalized Bias Error of forecasted values:*

(1)

· *Normalized Mean absolute error:*

(2)

· *Normalized Root mean square error:*

  (3)

NBIAS indicates the difference between the mean values of the two compared parameters and informs whether and to what extent the prediction systematically overestimates (NBIAS>0) or underestimates (NBIAS<0) the observations. Ideal forecasts have bias values equal to zero. The NMAE measures the average magnitude of errors between predicted values and actual values. The NRMSE refers to the absolute value of the deviation between modeled and recorded, providing at the same time a quantification of the error variability. Ideal forecasts have NRMSE close or equal to zero.

The overall performance obtained from all the sites under examination is provided in Table 1.

Table 1 Results of statistical indicators of performance for the five sites

|  |  |  |  |
| --- | --- | --- | --- |
|  | ***NBIAS (%)*** | ***NMAE (%)*** | ***NRMSE (%)*** |
| **Site 1** | 3.32 | 8.51 | 13.80 |
| **Site 2** | 7.57 | 17.05 | 25.79 |
| **Site 3** | 1.64 | 12.67 | 18.68 |
| **Site 4** | -1.42 | 19.21 | 29.11 |
| **Site 5** | -1,10 | 16.02 | 23,37 |

Results suggest generally acceptable performance with some site-specific variability in prediction quality. Site 1 demonstrates the best accuracy, with low bias (3.32%) and the lowest NMAE and NRMSE, followed by Site 3 with minimal bias (1.64%) and moderate errors. Site 5 also performs fairly well, showing a slight underestimation bias (-1.10%) and NMAE of 16.02%. Site 4 exhibits underestimation (-1.42%) and the highest NMAE and NRMSE. On average, the wind power prediction shows a small overestimation tendency (mean NBIAS ≈ 2.00%), with NMAE around 14.7% and NRMSE near 22.8%.

The statistical analysis of performance indicates a generally satisfactory scheme for wind power prediction with the use of the Digital Twin data, showing low overall bias and normalized error metrics that reflect consistent predictive capability. The normalized mean absolute error (NMAE) averages around 15%, while the normalized root mean square error (NRMSE) remains below 25%, both within acceptable limits for wind prediction applications, especially for complex terrain cases. These values suggest that the wind power prediction method offers reasonably accurate and reliable forecasts across different locations. The slight overall overestimation bias is minor and does not significantly affect prediction quality. The results are sufficient to demonstrate the wind power model’s ability utilizing DT data, to generalize and capture key relationships for local adaptation. Greater improvements could be expected if longer-term, continuous datasets were available, as these would enhance the model’s learning capacity, reduce overfitting and support the application of more advanced techniques, such as deep learning or hybrid machine learning methods. Additionally, the inclusion of extended time series would allow the model to better learn seasonal patterns and temporal dependencies, leading to further refinement in predictive performance.

# Issue Tracking and Resolution

This section outlines the methodologies and practices employed to track, manage, and resolve issues uncovered during the DRE Use Case application verification and validation processes.

**DRE Use Case Tracking System**

Our approach ensures that all identified issues are systematically addressed, promoting the development of reliable and high-quality software. The process is designed to be iterative, ensuring continuous improvement and adherence to best practices in software development.

We employ a Jira issue-tracking system that records all identified issues, including bugs, defects, and dependencies with specified requirements.

**DRE Use Case JIRA Onboarding Project**

Serco has implemented a Jira issue-tracking system to record all identified issues, including bugs, defects, and dependencies, against specified requirements. This system supports onboarding DRE Use case services in DESP and resolves issues, contributing to software development. DRE's partners use this system to monitor onboarding progress, update the issue statuses appropriately, and utilise comments for support.

The DRE use case has encountered significant challenges due to recent changes in account status. Disabling accounts required for DESP access has impeded our ability to perform critical data collection, simulation, and integration tasks. Re-enabling these accounts is essential for resuming our integration trials with DESP, which are crucial for progressing with project deliverables and adhering to the planned timeline.

# Resolved UAT Issues and New Testing Phase

As the project concludes, all issues identified during the previous UAT sessions, including those related to WeMet results, have been successfully resolved. The HYREF software has completed its final User Acceptance Testing (UAT) phase, during which its functionality was validated against user requirements. The system has demonstrated its ability to deliver accurate solar and wind energy assessments and forecasts. HYREF is now fully operational and ready for deployment, meeting the performance standards established by the project.

# Conclusion and Recommendations

The Verification and Validation (V&V) process conducted during WP5 for the DRE Use Case application has thoroughly assessed the system’s compliance with its specifications and functional requirements, confirming its readiness for deployment. A combination of methodologies—including code reviews, system integration testing, and an iterative User Acceptance Testing (UAT) process—was employed to evaluate the software across key dimensions such as performance, stability, and security.

The results confirm that the application meets the predefined quality, reliability, and operational performance criteria, aligning with project goals and stakeholder expectations. Testing has validated the correct functioning of the application under a range of real-world conditions and use cases, including those related to solar, wind, and hybrid energy forecasting.

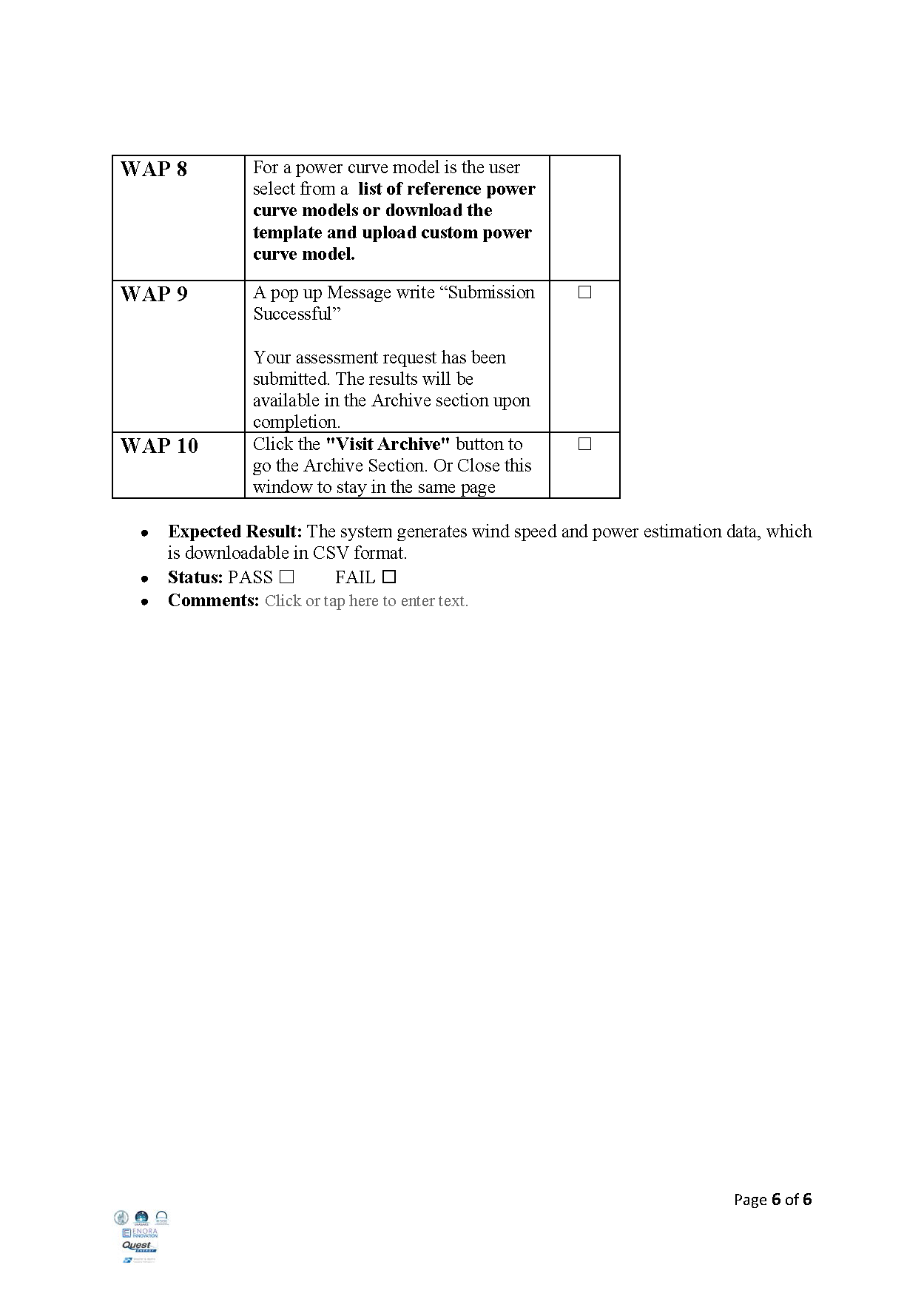
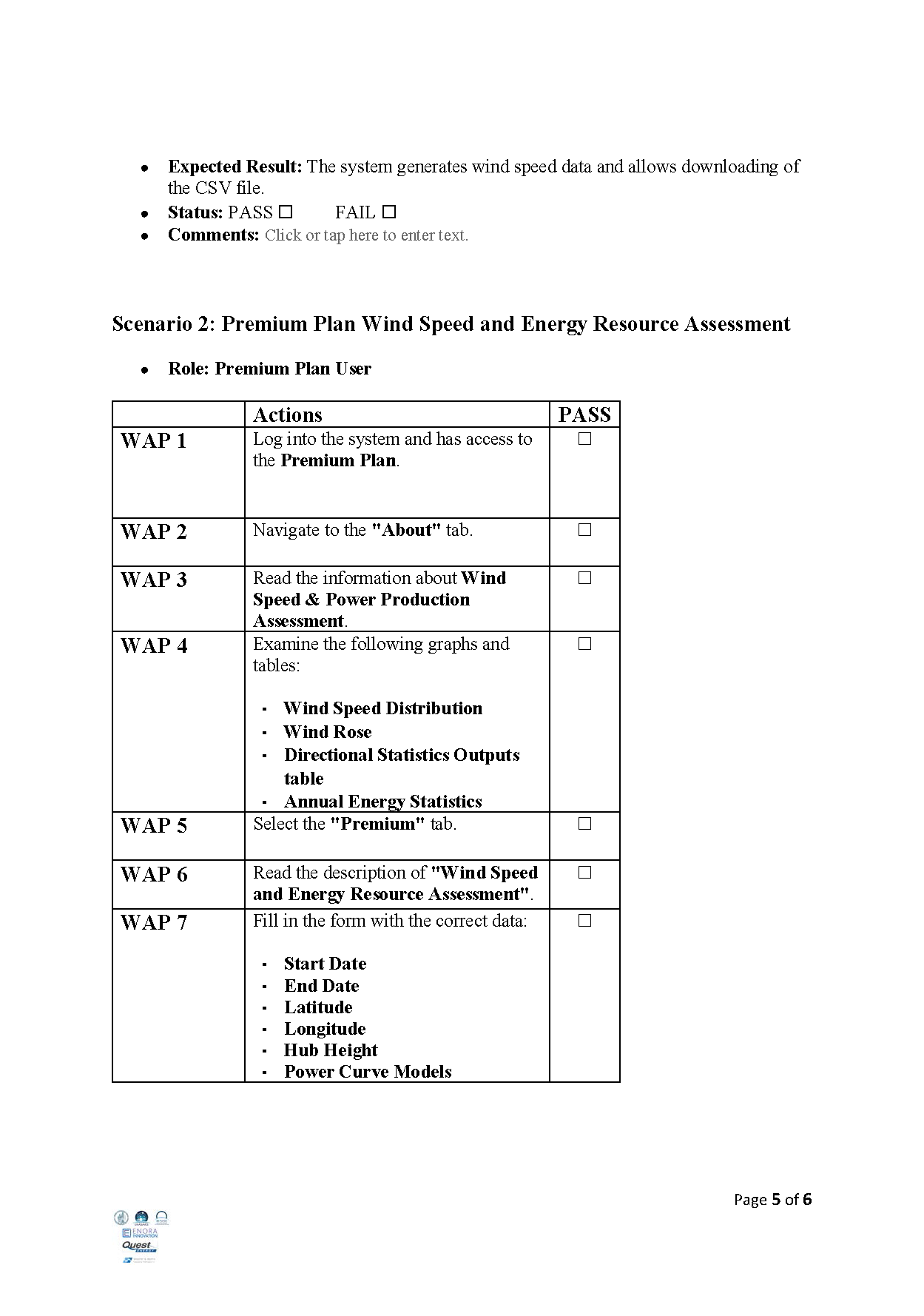
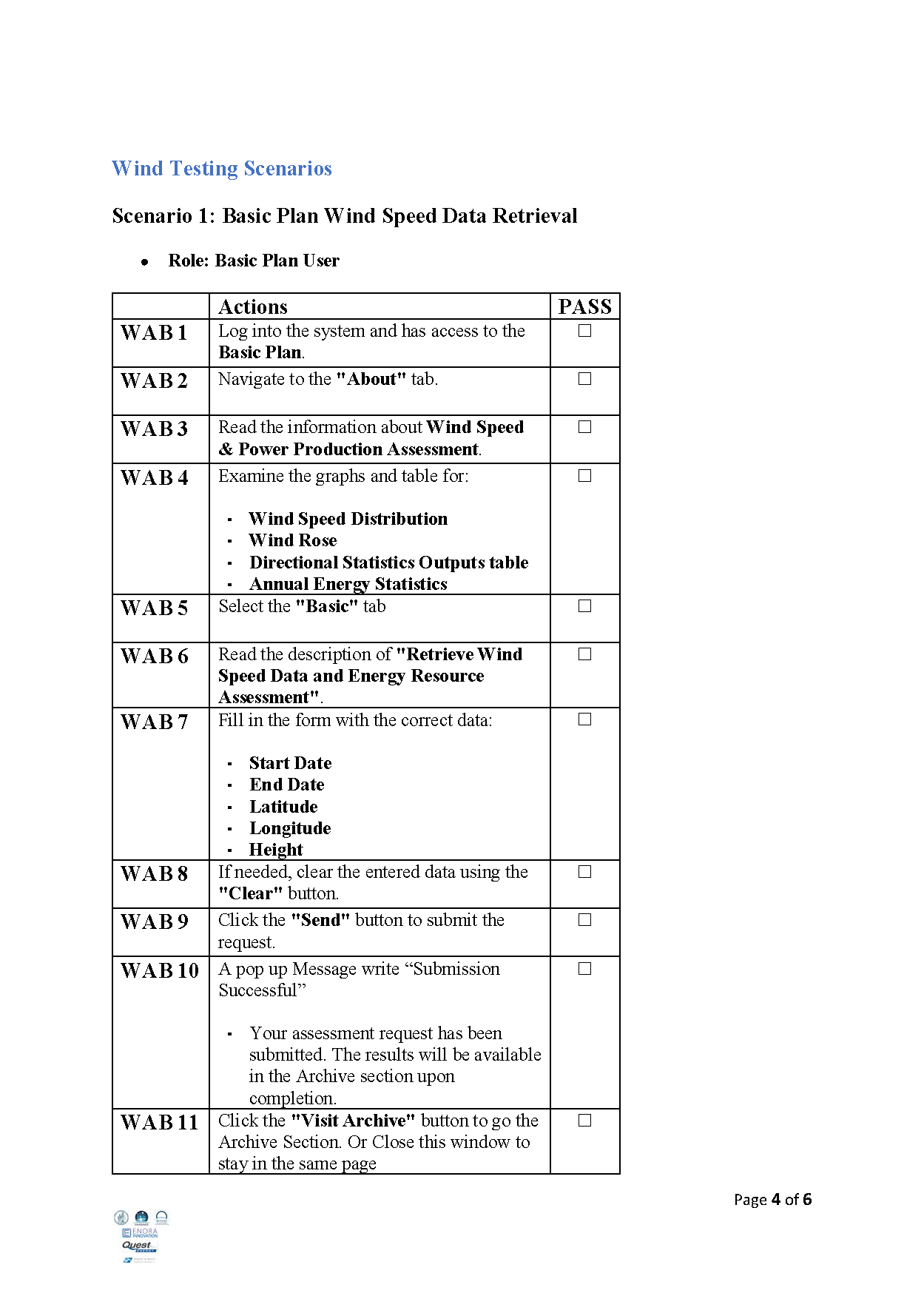
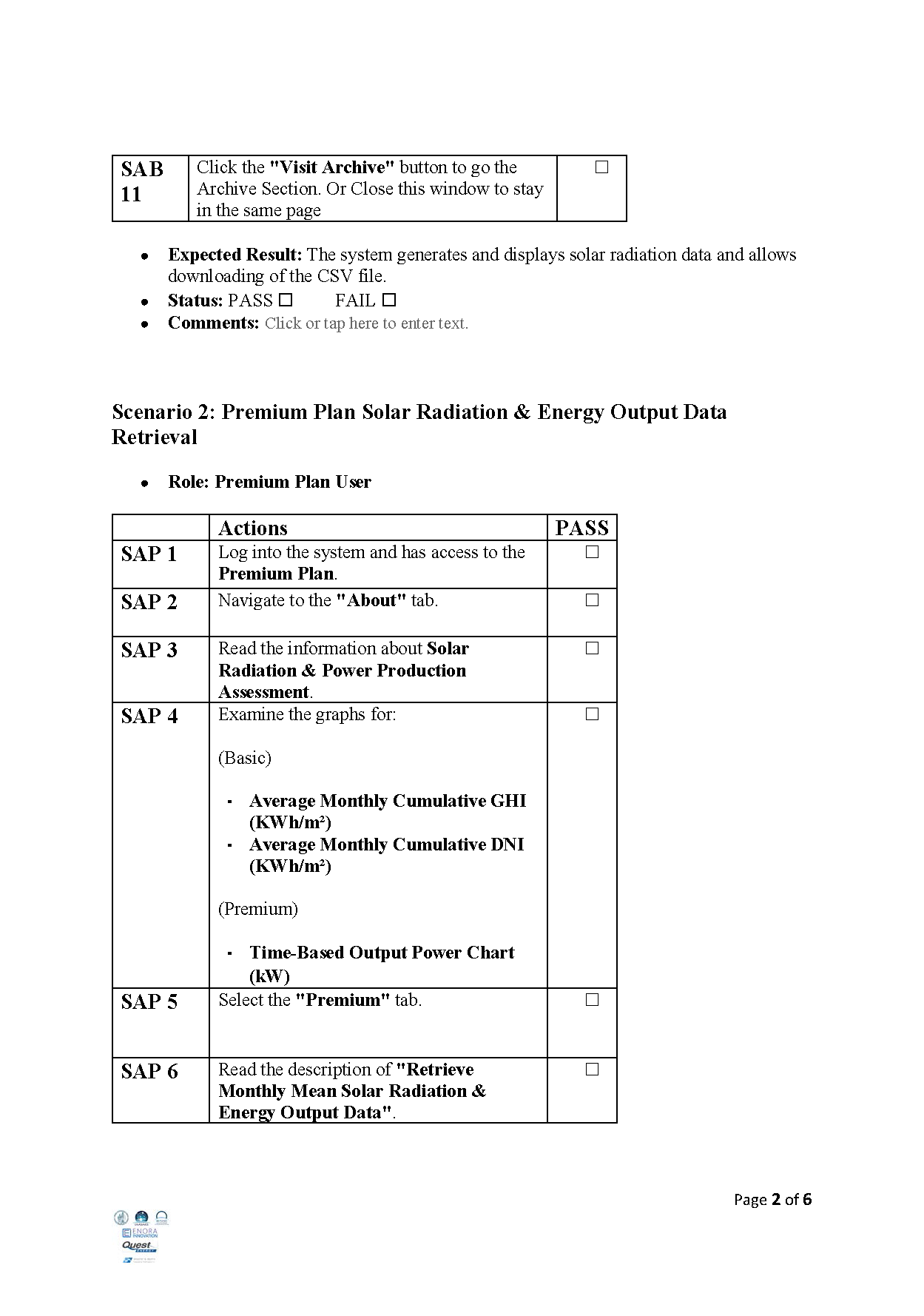
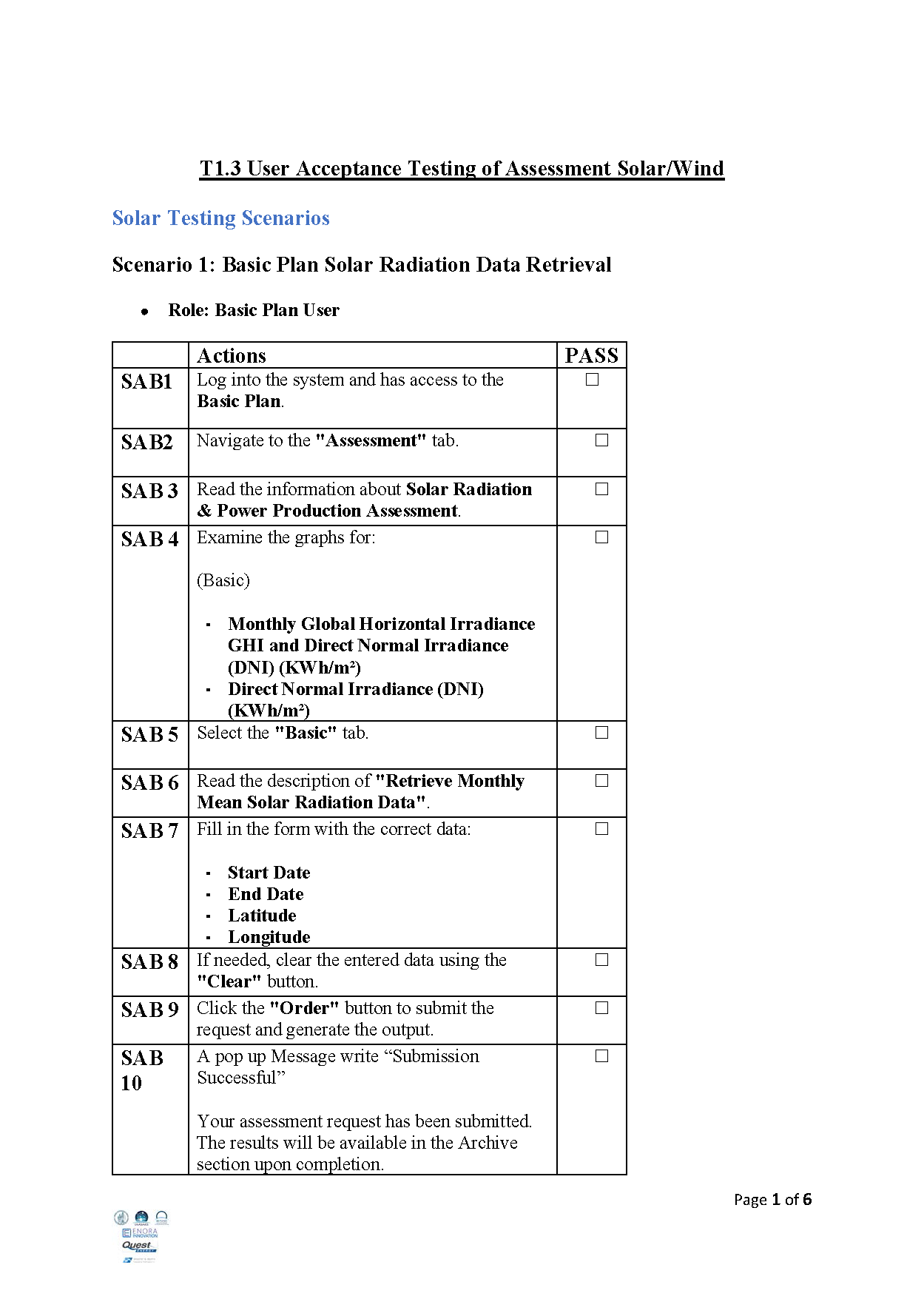
Notably, the UAT phase, conducted with end-users' involvement, highlighted the platform’s effectiveness in delivering accurate assessments and forecasts. Continuous feedback loops and quality control reviews contributed to refining the assessment and forecasting modules, resulting in a robust and user-aligned solution.

With these outcomes, the DRE Use Case application is now a validated, production-ready component of the HYREF system. This final deliverable reflects the completed status of the use case, concluding the iterative development and validation cycles carried out in WP5.

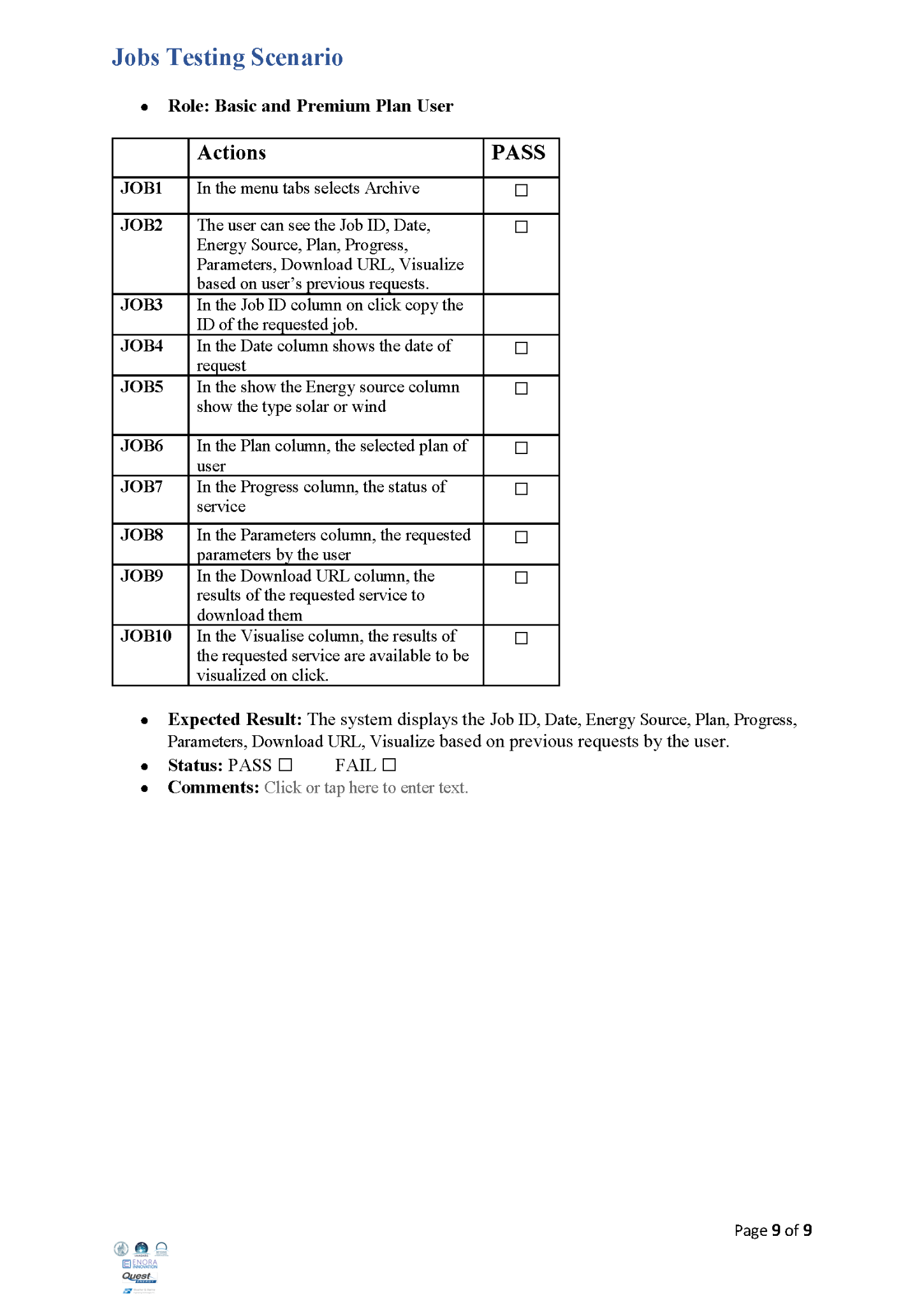
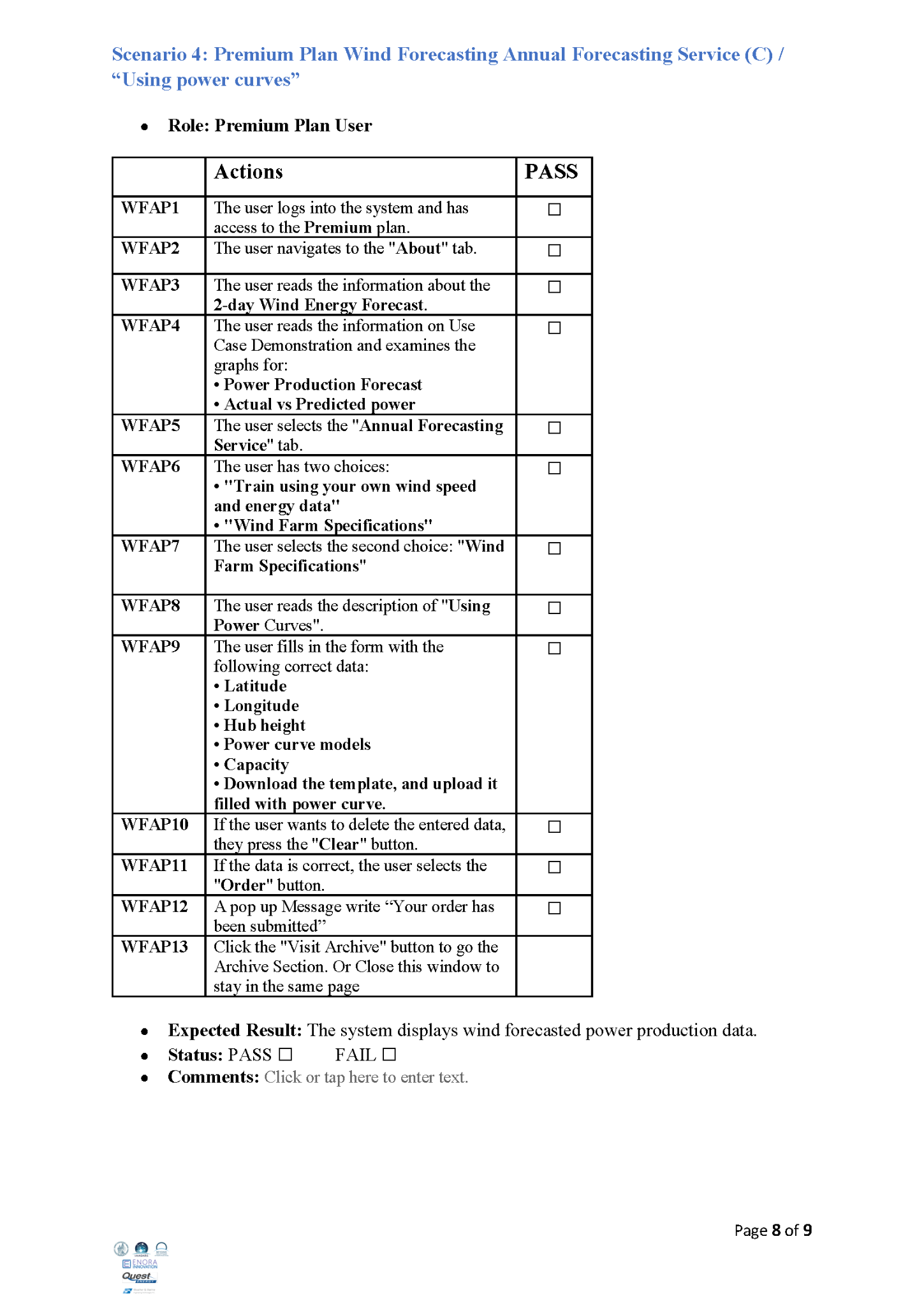
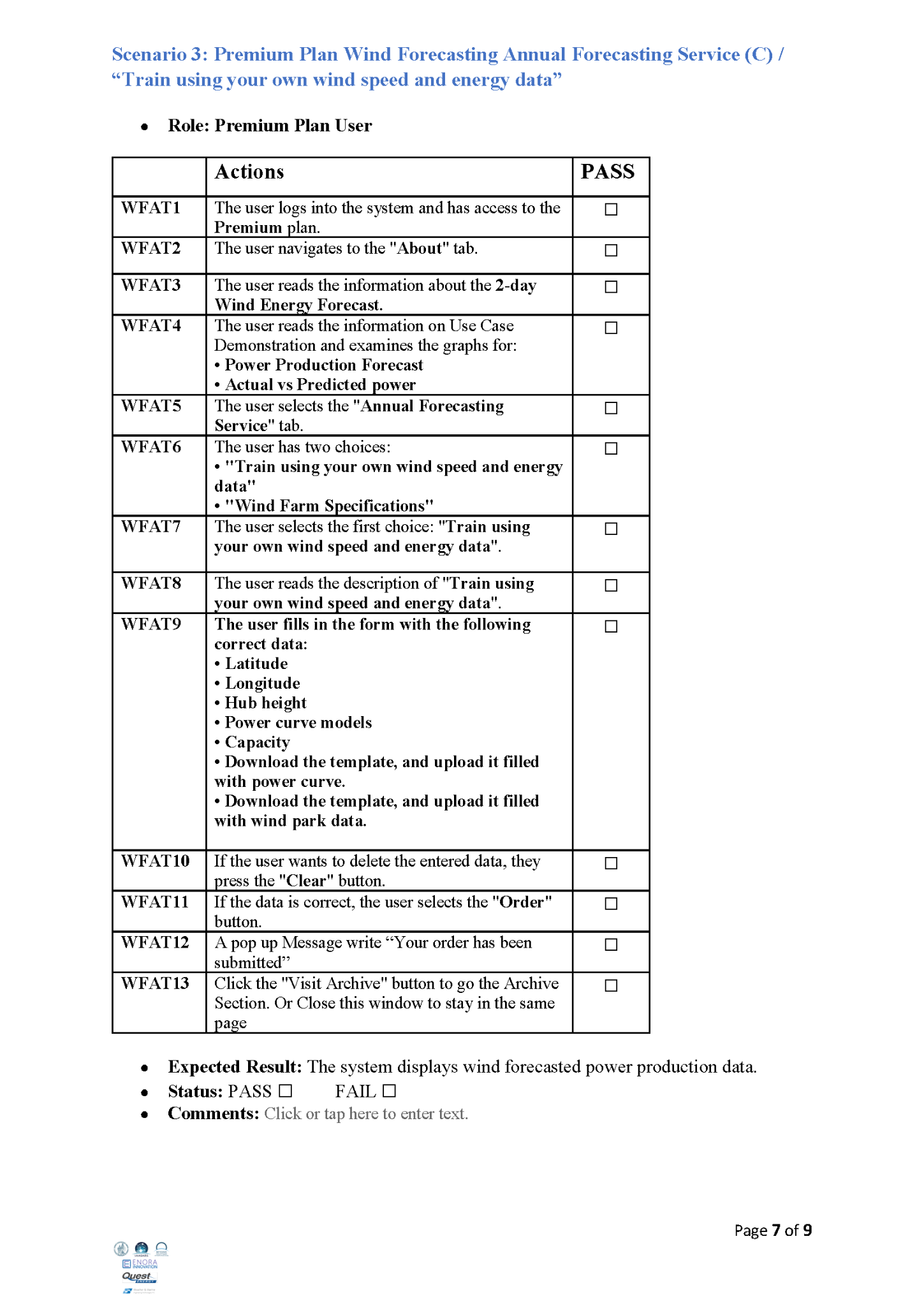
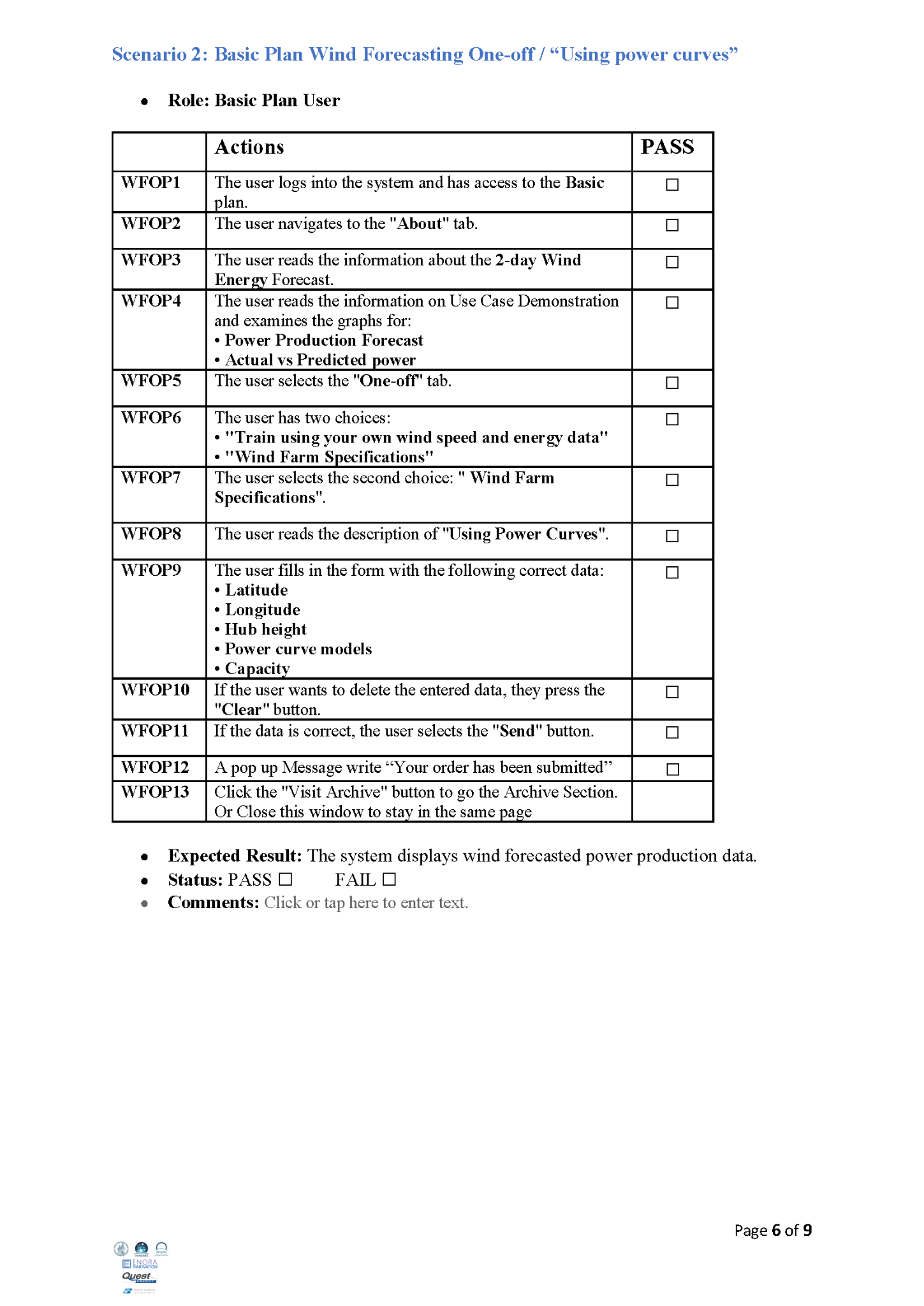
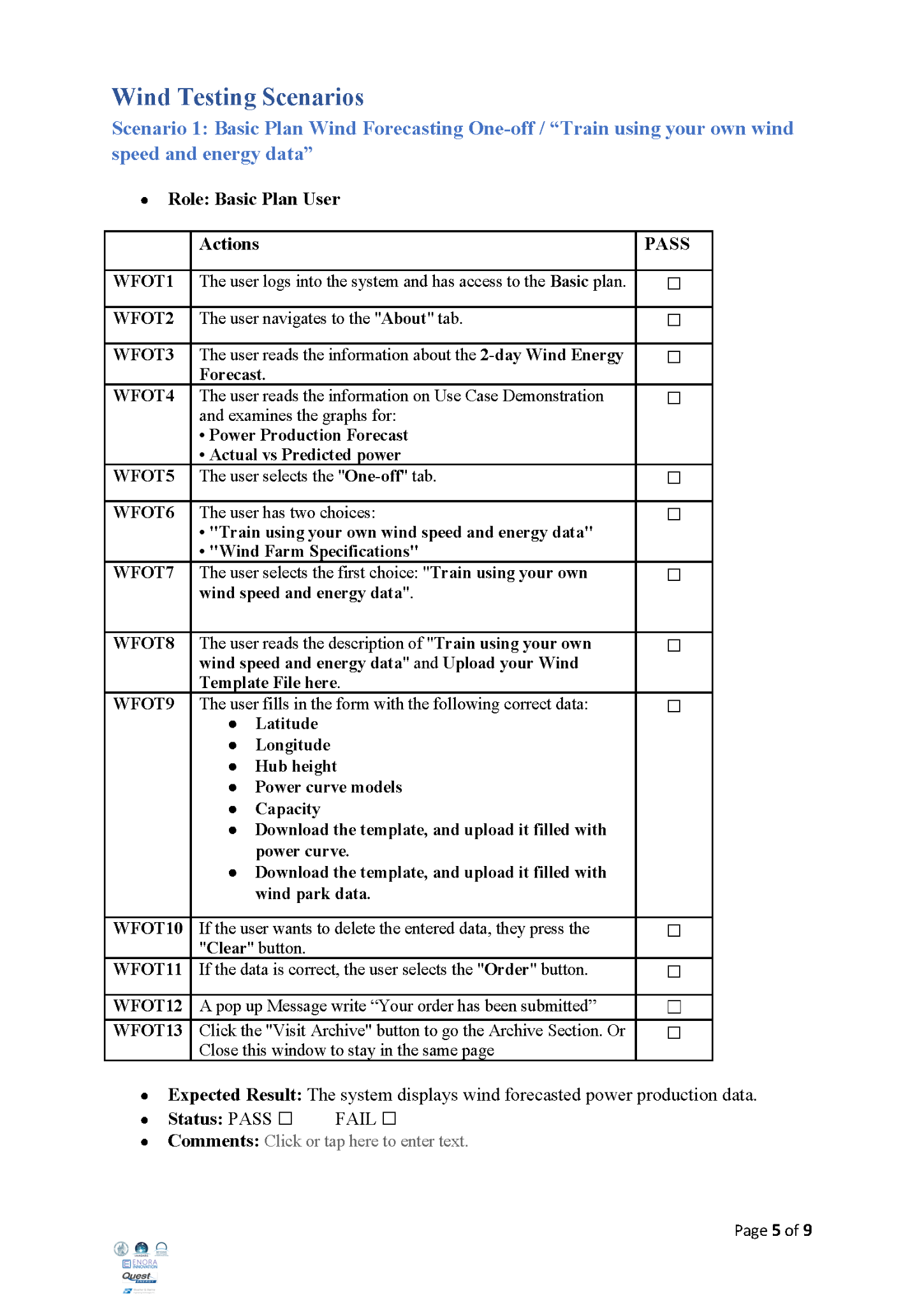
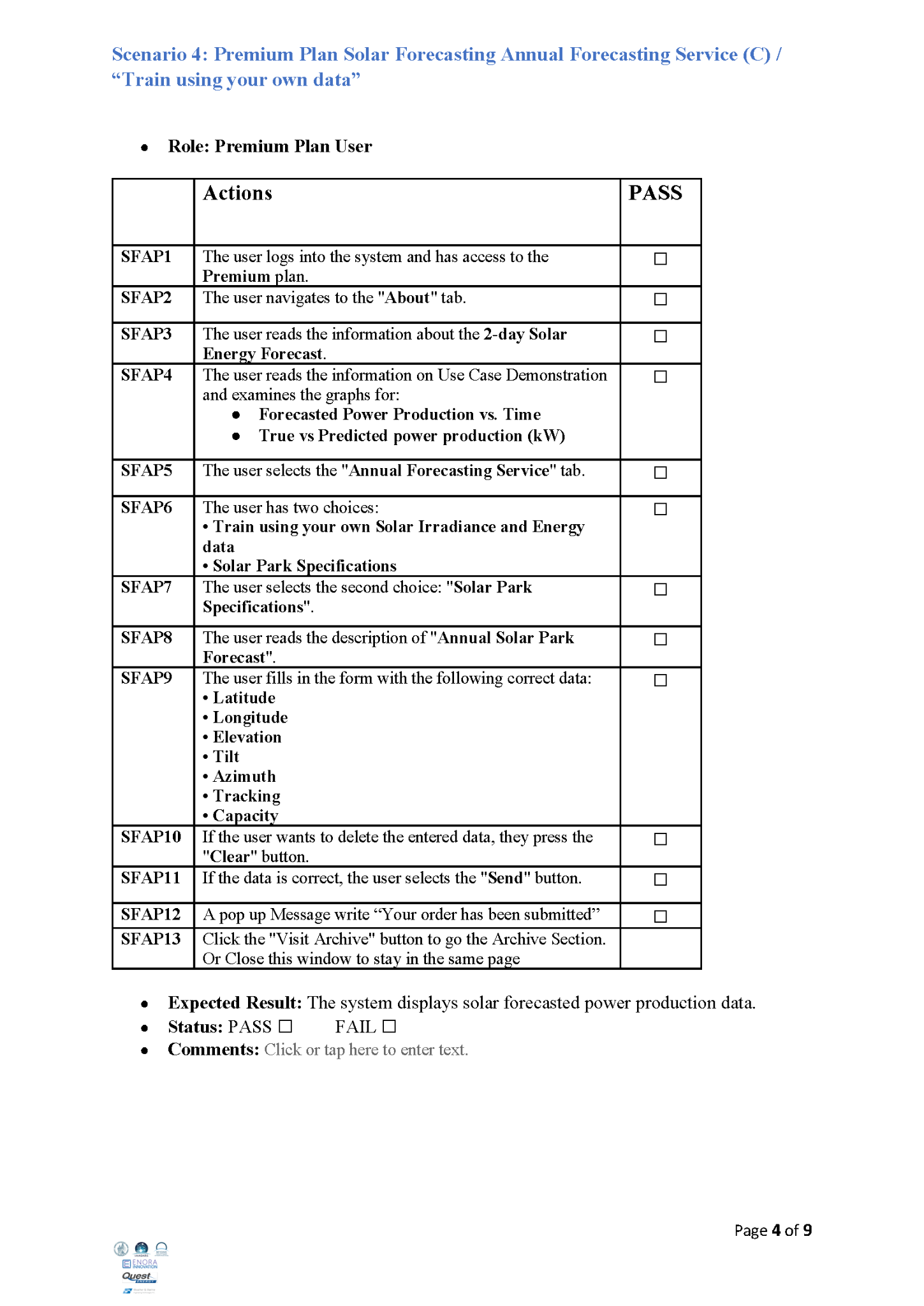
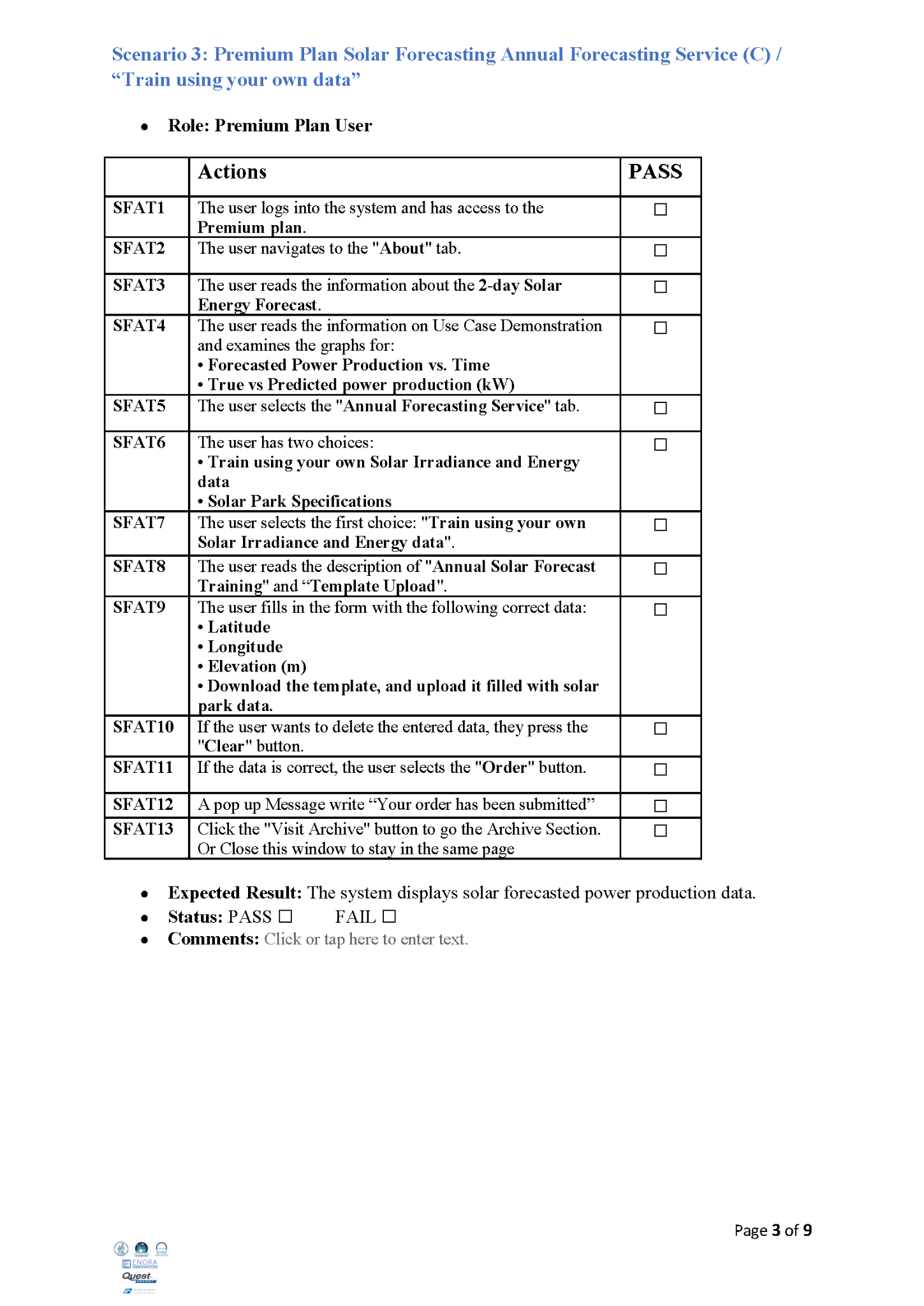
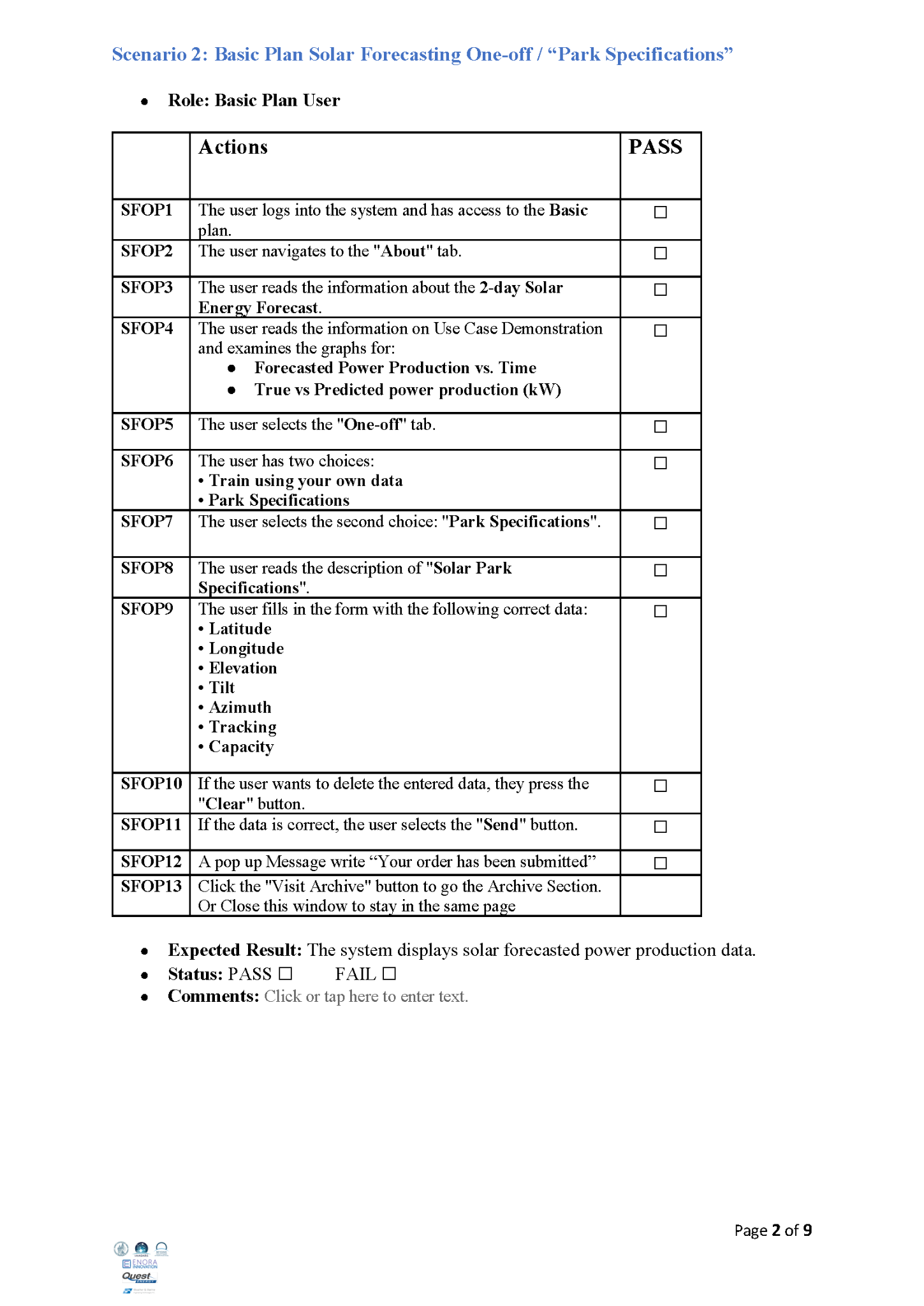
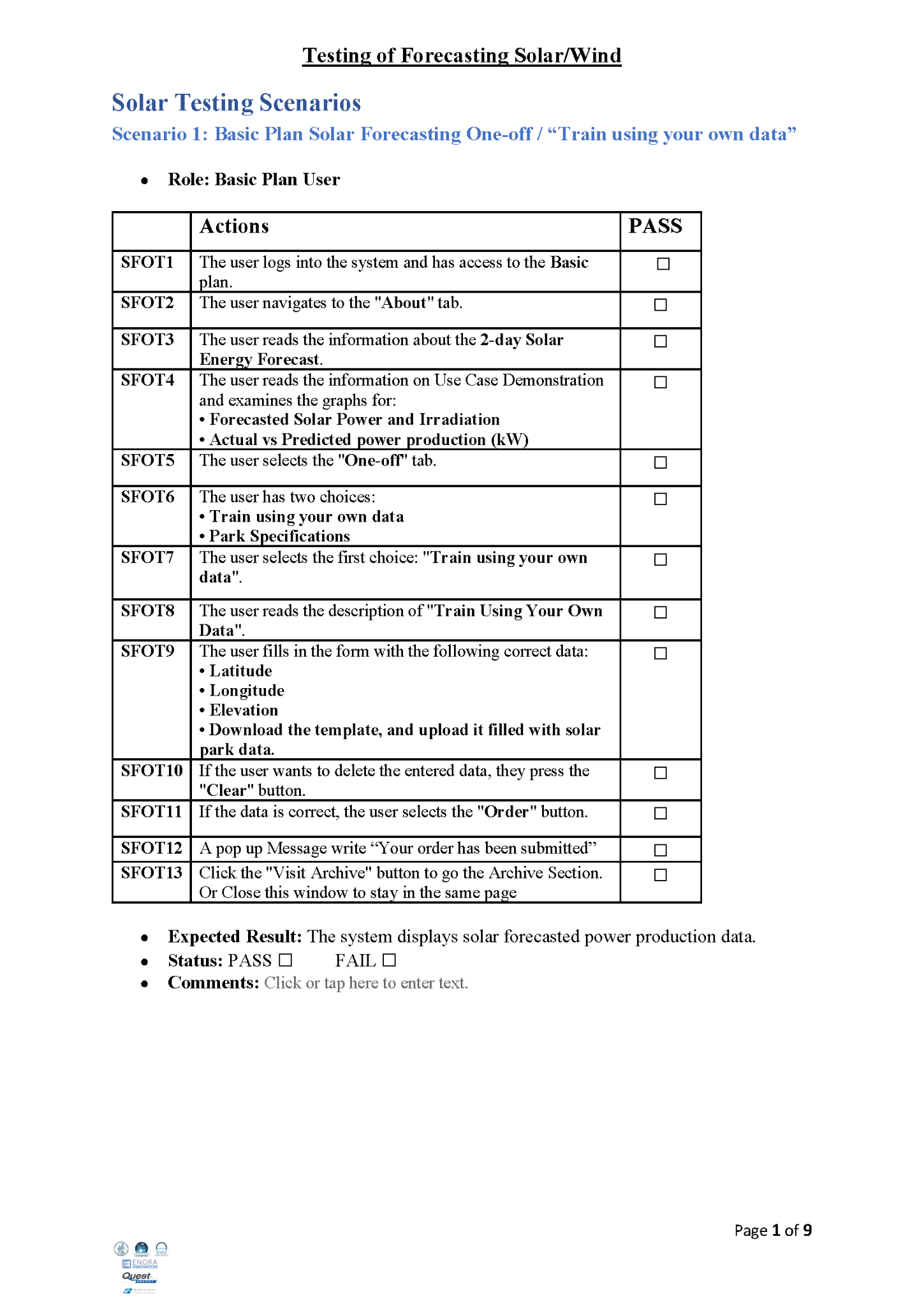
# Appendix

The User Testing for the HYREF application is crucial to ensure it meets users' needs by providing accurate solar and wind energy assessment and forecasts.

## Assessment Testing Scenarios of User Acceptance Testing **-** Templates

****

## Forecasting Testing Scenarios of User Acceptance Testing Templates



## Assessment Users’ Testing Results

### Solar Assessment

## 

## 

## 

### Wind Assessment

## **Εικόνα που περιέχει κείμενο, στιγμιότυπο οθόνης, γραμματοσειρά, έγγραφο Το περιεχόμενο που δημιουργείται από τεχνολογία AI ενδέχεται να είναι εσφαλμένο.**

## **Εικόνα που περιέχει κείμενο, γραμματοσειρά, στιγμιότυπο οθόνης, αριθμός Το περιεχόμενο που δημιουργείται από τεχνολογία AI ενδέχεται να είναι εσφαλμένο.**

## **Εικόνα που περιέχει κείμενο, στιγμιότυπο οθόνης, γραμματοσειρά, έγγραφο Το περιεχόμενο που δημιουργείται από τεχνολογία AI ενδέχεται να είναι εσφαλμένο.**

## **Εικόνα που περιέχει κείμενο, στιγμιότυπο οθόνης, γραμματοσειρά, αριθμός Το περιεχόμενο που δημιουργείται από τεχνολογία AI ενδέχεται να είναι εσφαλμένο.**

## Forecasting Users' Testing Results

### Solar Forecasting

## **Εικόνα που περιέχει κείμενο, στιγμιότυπο οθόνης, γραμματοσειρά, αριθμός Το περιεχόμενο που δημιουργείται από τεχνολογία AI ενδέχεται να είναι εσφαλμένο.**

## 

## 

## 

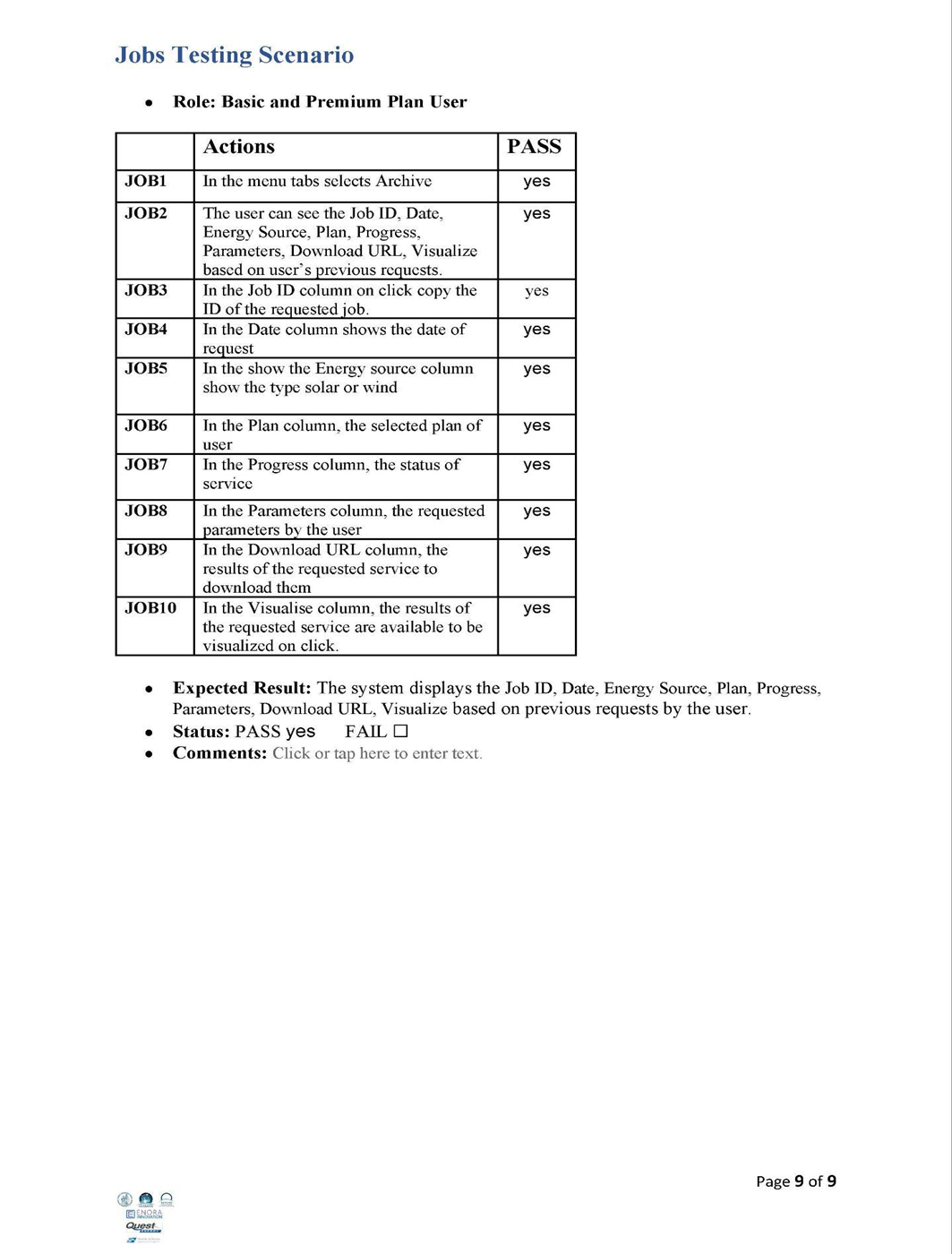
### Wind Forecasting

## **Εικόνα που περιέχει κείμενο, στιγμιότυπο οθόνης, γραμματοσειρά, αριθμός Το περιεχόμενο που δημιουργείται από τεχνολογία AI ενδέχεται να είναι εσφαλμένο.**

## **Εικόνα που περιέχει κείμενο, στιγμιότυπο οθόνης, γραμματοσειρά, αριθμός Το περιεχόμενο που δημιουργείται από τεχνολογία AI ενδέχεται να είναι εσφαλμένο.**

## **Εικόνα που περιέχει κείμενο, στιγμιότυπο οθόνης, γραμματοσειρά, αριθμός Το περιεχόμενο που δημιουργείται από τεχνολογία AI ενδέχεται να είναι εσφαλμένο.**

## **Εικόνα που περιέχει κείμενο, στιγμιότυπο οθόνης, γραμματοσειρά, αριθμός Το περιεχόμενο που δημιουργείται από τεχνολογία AI ενδέχεται να είναι εσφαλμένο.**

****