Testing & Evaluation

When developing any form of software, even if everything appears to be working, there will be uncaught errors and issues in the program. It is therefore essential to ensure that rigorous testing is undertaken to discover as many of these errors as possible.

For this project two forms of testing were used to evaluate the program, and decide whether it meets its goals. To test the system’s functionality a series of tests, which correlated to specific user stories, were run manually. User experience was tested with a short survey, where users were given tasks to follow and provide their answers, as well as providing general feedback about the user experience, successfully testing both the system itself and showcasing the views of potential users.

System Testing

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Test ID | Description | Stories Tested | Postman | Website |
| 1 | Run forecast to display energy production statistics | 1 | Pass | Fail |
| 2 | Run simulation to see energy production statistics | 2 | Pass | Fail |
| 3 | Input a forecast with hourly set, and compare to one not set hourly | 1,3 | Pass | Fail |
| 4 | Run two simulations with all same settings other than timescale and compare results | 2,6 | Pass | Fail |
| 5 | Run simulations with different solar/offshore/onshore settings and with different input values | 2,4,5 | Pass | Fail |
| 6 | Update/delete a previously run forecast and simulation and check stored values | 8 | Pass | Pass |
| 7 | Check lists of forecasts and simulations on webpage against database | 7 | Pass | Pass |
| 8 | Save a forecast or simulation to database | 9 | Pass | Pass |
| 9 | Rerun a stored simulation and check data is valid | 2,7,10 | Pass | Fail |
| 10 | Search for specific known plants from database in online display | 11 | Pass | Pass |
| 11 | Add a new theoretical plant to a specific region | 12 | Pass | Pass |
| 12 | Add a new user to system | 15 | Pass | Pass |
| 13 | Try to login with a created user | 15 | Fail | Fail |
| 14 | Call API independently of a forecast | 14 | Pass | Pass |
| 15 | Check that you can’t add user with existing email or username | 15 | Pass | Pass |

The following table outlines all the primary technical tests run to analyse the product. For each test, two methods were used to test functionality (where applicable).

Table :List of System Tests Undertaken

The two methods of testing shown in the table were

* Test the program using the website to check for full functionality and meeting of user story requirements and checking for bugs
* Testing by sending appropriate HTTP requests using Postman to check that if there is an issue whether it is relating to how the front end is receiving and handling the data or the back-end logic. Some of these tests can be seen here: . Link also in Appendix III.

User Testing

A simple questionnaire was designed to allow users to provide feedback on the product. Several of the questions in the questionnaire pertain to simple tasks set for the users, as well as several questions giving the users to provide general views on the usability, aesthetic and general feel of the program. The general questions were, where possible, kept to neutral language as to prevent any unintentional bias towards a specific view of the program. Where users were allowed to input their own choice of values, they were instructed to mention these choices in their response section. The reason for allowing users to input their own choices rather than giving a set of inputs in the task description was to ensure that varied testing was undertaken to check for any bugs and prevent only testing the system in a way that didn’t only test parts of the system known to be fully functional. In addition to this, allowing the user to input their own values more accurately matches how a non-technical user such as a student may approach using the application, the users were set up with the website and a brief guide to the question set, and then were left on their own to complete the tasks and questionnaire. “Stories Tested” is only shown here for clarity, this was not shown to the users.

|  |  |  |  |
| --- | --- | --- | --- |
|  | Task | Response | Story |
| 1 | Find the forecast page, run and save a new forecast with your choice of settings. Answer with Forecast ID |  | 1,3,4,6,9 |
| 2 | Try to add a new onshore windfarm in the location of your choice. Or update an existing one. Give plant ID |  | 11,12 |
| 3 | Navigate to the info page, read the documentation and run an API call, did you receive a 24hr forecast? |  | 14 |
| 4 | From the simulation page, access the list of previously run simulations, then rerun a chosen simulation.  Write down your chosen simulation and its first value |  | 2,7,10 |
| 5 | Register an account for yourself, were you successful |  | 15 |
| 6 | Rerun a forecast of your choosing and answer as for Q4 |  | 7,8 |
| 7 | Run a new simulation using your choice of settings |  | 1,2,4,5,6 |
| 8 | Navigate to all the pages, were you successful? |  | General |
| 9 | Delete a forecast of your choice, were you able to delete it? |  | 7,8 |
| 10 | Update a forecast, what happened? |  | 7,8 |
|  | General Questions | Response |  |
| 11 | How did you find the navigation around the website? |  |  |
| 12 | What do you think of website colour scheme? |  |  |
| 13 | What opinions do you have on the page layout if any? |  |  |
| 14 | Did you understand what the information you were looking at represented? |  |  |
| 15 | What features would you like to see in a future version of the program? |  |  |
| 16 | Were you able to complete most of the tasks from the instructions? |  |  |

Table : User Testing Survey \*Full results of user questionnaires can be found in the appendices.

Evaluation

System Testing

Tests 1-5 were used to test the Must Have user stories outlined in Chapter 2 and Appendix I. All of these essential functions work in full on the backend, with only small adjustments from initial design principles, however these tests were unsuccessful on the frontend. The primary cause of tests 1-5 not being passed in website testing is due to the lack of a graphical or even numerical output of the energy forecasting/simulation, this was due to lengthy database restructuring during the project which delayed development, issues with getting Chart.js graphs to display on pages and difficulty with parsing data sent from the backend to the front end. While it is predominantly issues with the displaying of data that have caused the front-end failure of these functionalities and they are fully functional on the back end, it is far from ideal that these Must Have stories were not completed during the initial development stage. Future work on the program should first and foremost focus on bringing these stories to completion before any other less important features are worked on.

The other tests focused on the lower priority user stories, all of these bar one worked on the back end with only two not also working on the front end. The only system test to fail in every aspect was test 13, this was an expected failure as only the very basic structure of an authentication system was attempted to be implemented in the initial design period. The other test of tests 6-15 that failed on the front end was test 9, this was again due to the issues with being able to access a readout of data from the backend, and not issues with the backend calculations itself. This shows that most of the lower priority user stories have been met in completion, while this means the product has most of its functions working, clearly it would be more ideal to have resolved the issues with the higher priority functions over having these lower priority functions fully fledge for an initial release of a minimum viable product.

It however should be noted that several of the functions that were lower priority needed to be implemented on the backend in order to allow for high priority functions to work across the full stack and were therefore easy to implement on the front-end ahead of other, more essential, functions. For example, while displaying a list on the front end of previously run forecasts was less important than being able to rerun a previous forecasts, the backend logic of saving and retrieving forecasts had to be implemented, and thus it was an easy task to implement the display of the list prior to the rerun function on the front end. This means that, provided that a single short future development can resolve the issues related to the Must Have user stories, the product will be able to be released in the near future with not just essential features, but also many supplementary desirable features.

In summary, the system testing shows that the vast majority of the main functionalities work on the backend. However, several fail tests when using the web application, this highlights the difficulties that were encountered during the build of the front end, and shows points that could be improved in future development cycles.

Validity of System Testing

With any software product it is important not only to test the programs functionality and usability, but it is also important to evaluate the validity of the testing methods used.

The system tests were chosen to cover the whole scope of the project, from the Must Have down to the Would Like to Have stories outlined in Chapter 2 and Appendix III, this means that to some extent every function of the system was test solely by the System Testing, with supplementary system testing performed during the User Testing of the system.

There is always the possibility of making testing more thorough, and if further work were to be undertaken with this project, a wider range of testing methods and individual tests would improve the rigour and accuracy of the testing. Implementing some form of automated system testing in future iterations of the program would be the top priority for improving system testing. In addition to this using a wider range of scenarios to test primary features would be desirable, in order to ensure that the main use cases of the program are fully functional.

User Testing

The responses of the user testing survey were analysed, and common themes identified, these were collated and then used to evaluate the systems usability and functionality. The users were informed prior to participating in the survey, that the displaying of forecast and simulation data wasn’t working, this was done in order to explain what the users should expect to see when undertaking the tasks, but that they should still comment on this aspect of the UI in the general comments section.

In the Table 3 below common responses from the users surveyed are shown.

|  |  |  |
| --- | --- | --- |
| Topic No. | Description of Response | Users Referencing Topic |
| 1 | A graphic display would be desirable in future | 3,4,6 |
| 2 | Navigation was easy | 1,2,3,4,5,6 |
| 3 | Simple Layout (seen as a positive) | 3,4,5,6 |
| 4 | Simple Layout (seen as a negative) | 1,2 |
| 5 | Would like to see more colour to highlight important information | 3,4,6 |
| 6 | All pages accessible from navigation bar, reordered on navbar | 1,3 |
| 7 | Questions (particularly questions 6) unclear | 1,3,6 |
| 8 | Questions were clear | 5,6 |
| 9 | More/more informative pop-ups upon submitting information | 2,3 |
| 10 | More information about what each function does | 1,2,3,4,6 |
| 11 | Colour scheme clear and makes information easy to read | 2,5,6 |
| 12 | Colour scheme gives professional look to program | 1,6 |
| 13 | Improve page layouts, and individual element layouts | 1,2 |
| 14 | Make login feature functional | 1 |

Several users complained of a lack of clarity of what the information represented, or asked for a tutorial explaining more thoroughly how the software works. This is a valid concern as in some parts of the website the guidance given was slightly broad, this choice was made to minimise clutter on the pages, with supporting information grouped together on the Info page of the website, however the info page could have more description or instructions as to how all the systems worked. The lack of clarity in some places may be an issue for non-experienced users (i.e students), so this issue must be addressed, although the only response that was common to all users was that the website was easily navigable, which is an important element of user satisfaction. In future the addition of a tutorial video, or a walkthrough such as in Appendix III, could be added to the info page to give greater clarity.

While there were some complaints that the colour scheme was too simple, overall users found it to aid with the readability of text and gave a professional look to the website. These were the intended goals of the colour scheme chosen in Chapter 3, of mainly muted tones with some colour for highlighting, this was done to give the website a professional feel, well suited to the target users groups of academics and energy industry workers. The complaints about lack of colour would be addressed in future developments, as it was always planned to use more colour in graphic displays of the website and to avoid using too much colour elsewhere, which would distract users as they tried to analyse the data shown in graphic displays.

Some of the common responses related to the desire of features which were already planned for the website, such as a graphic interface or making the login feature fully functional, this shows that the initial design of the project is valid, as users are requesting features that were intended to be implemented at some point. This further confirms that the top priority in any future developments would be to get the forecast and simulation pages to display all data produced by the back end in a graphical way.

Overall, users mainly seemed to comment on the desire for more documentation to clarify what each feature of the program does, this may be in part be caused by the ambiguous instructions that some respondents found unhelpful, but regardless is easily implementable in future, and for a graphic and numerical display of forecast and simulation data, which will be the top priority in any future development. This gives clearly defined goals for improvements of the user interface and functionality of the system. Other common responses related either positively or negatively to the simple page layout and colour scheme. The goal was always to try and appeal to both professional and casual users of the program, this requires finding a middle ground between the desired specifications of both user groups, and inherently when finding a compromise between starkly different options some users from each group will not like the final decision made. Despite the opposing opinions, both the positive and negative responses highlighted the fact that the layout and colour scheme were simple, clear and professional, which were the main criteria when choosing these design elements.

User Testing Validity

For the project evaluation a small group of users were set tasks in order to gain feedback on the system. Ideally a larger group of users would have been surveyed, but as the system is presently only hosted on a single machine, and with the ongoing issues with face-to-face interaction stemming from the Covid pandemic, only a small number of users were able to be given direct access to the program. Before a full scale release a wider scope of user testing would be essential for testing the product.

All members of the testing group were aged between 18-30, while this maps well to one of the target user groups of the software, students, it did not include many of the other two primary target user groups, only one academic was surveyed, and no energy industry professionals were surveyed. In future testing the group should be widened to include more people, and from a wider range of ages, and include more academics and energy industry workers, in order to ensure that the respondents views more accurately match those of the target user base as a whole. Gender breakdown of users was fairly even, with three female and four male respondents to the survey, while this could be improved on, it is an acceptable ratio at this stage of testing.

Some of the tasks were noted by users to be unclear. In attempting to avoid any bias, or too strong guidance, some of the language used in the survey may have been confusing, this should be reworked before future testing.

It should be noted that the answers in two of the responses, answers had to be changed as the users wrote down the name that they had registered an account with, any information that suggested who the user may be was removed and a substitute generic phrase of similar meaning was inserted, this was to ensure no user information was disseminated through this report and fully anonymise the responses, but maintain as close as possible the exact answers given by the users.