Mission Space Lab Phase 4 report

Team name: Parsec

Chosen theme: Life on Earth

Organisation name: Liceo Leonardo Da Vinci, Gallarate

Country: Italy

1. Introduction

The abnormal floods verified in our country in May remind us that climate change is endangering several areas of the world. Adapting the territory to face these issues will be essential to prevent damage to the population and to natural resources, and predicting the evolution of the affected sites will be essential to take successful measures.
Remote sensing is an effective way of tracking the changes of water basins, such as rivers or lakes, as it allows direct monitoring of wide areas of land.

Our experiment consists in verifying if the AstroPi camera is suitable for analysing the evolution of water basins in relation to the vegetation around them.
With the aim of creating a model able to predict the evolution of those areas, we took advantage of satellite imagery datasets to gather pictures of the same sites over which the ISS passed when our program was running to compare the data and identify any changes that took place over the years.

2. Method

We used the AstroPi camera to take pictures of the Earth and the Sense HAT to collect data. The code was optimised to take as many pictures as possible without running out of space.
To choose only the relevant pictures, an algorithm assessed the presence of water (NDWI) and land covered by vegetation (NDVI). The interval at which the pictures were taken was regulated based on the remaining storage space and the average size of the previous pictures.

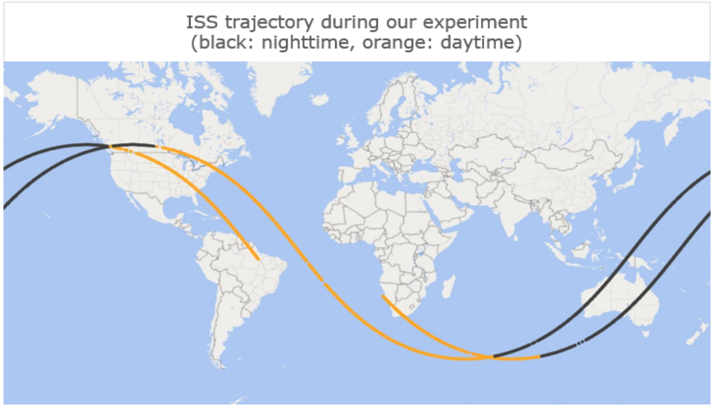
For phase 4, we made a program that uses Google Earth Engine API to get photos of the most relevant areas from phase 3 from the MODIS dataset between 2000 and 2023 and then generates a CSV file that contains parameters like the area covered by water and vegetation and the mean NDVI and NDWI values throughout each image. This file is then used by a SARIMA statistical model to make predictions of the evolution of such parameters.

The grid search technique proved computationally too expensive, thus the SARIMA model was optimised by running another program that employs a genetic algorithm to find the best hyperparameters for each specific dataset.

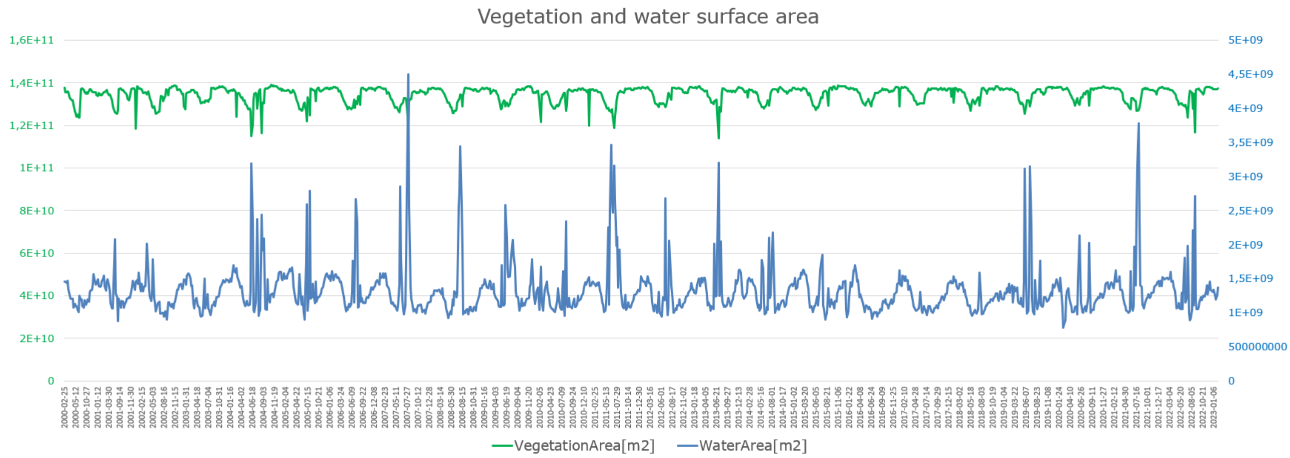
Our Github repository: https://github.com/Parsec2k23/Parsec_AstroPi_2022-23/

3. Experiment results

Many of the pictures taken on the ISS were cloudy, but some of them were workable. The ISS passed mainly over North America, so we focused on the Charleston area, which contains several lakes and its sea level rise has been assessed by [other sources](#).



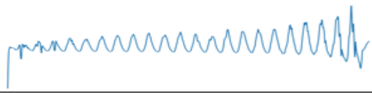
The charts derived from the dataset showed a clear seasonality:



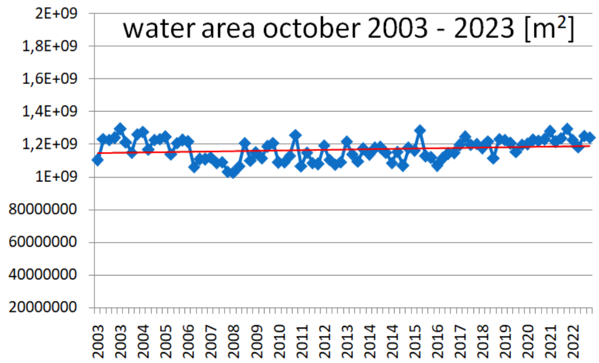
Therefore, we opted for the SARIMA model instead of a LSTM neural network, which gave us good predictions, with a root-mean-square error as low as *1.8%* of the correct values.

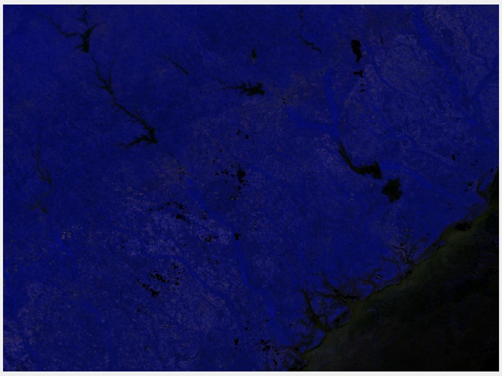
The spikes in the graph are a consequence of an imperfection in the MODIS dataset: some images that were returned by the API presented inexplicable black patches in seemingly random areas. We could have composited more than one image to remove these flaws, but since this did not have a significant impact on the results, we chose not to spend much time on it.

Another program calculated the cross correlation and Kendall's tau of each two columns of the original dataset. On the NDWI and the vegetation area, this returned a Kendall's τ of about *0.33* with a *p-value* of $4.04 * 10^{-67}$, which indicates a weak but statistically significant correlation: correlation does not mean causation, but this means that a higher water index corresponds to more vegetation.

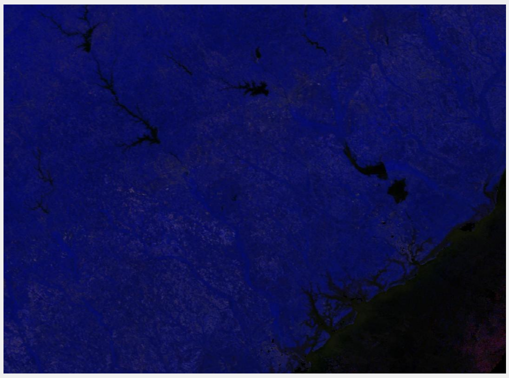
	Kendall's τ	p-value	Cross correlation
Vegetation area - NDWI	0,3338717	4,04E-67	
NDWI - Vegetation%	0,3338717	4,04E-67	

After creating separate charts for each month, we observed the trend of the surface covered by water and vegetation, determined by counting the pixels with high NDWI or NDVI values, and multiplying by the area of each pixel.
A regression analysis with Visual Basic for Excel indicated that the vegetation area exhibits an increasing trend in 11/12 months, while the water area did not change significantly. The pictures confirm this, being brighter as time passes:



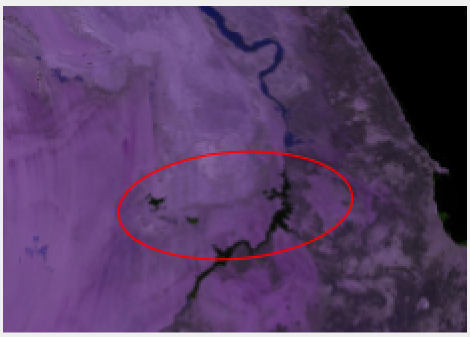


The coast of South Carolina in 2010 in IR

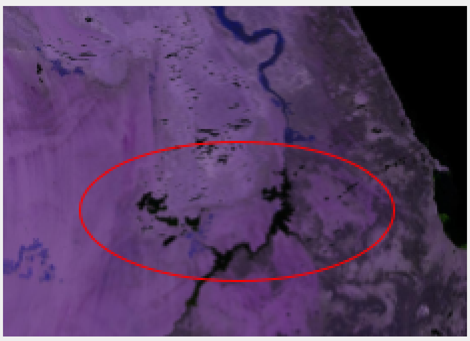


The coast of South Carolina in 2022 in IR

In other areas the variation is clearly visible:



Nile and Toshka lake in 2010 in IR



Nile and Toshka lake in 2022 in IR

4. Learnings

For most of us, working in AstroPi was the first time coding in a team, so we got to know how to use platforms to work together on the same program, such as Replit.

Platforms like GitHub played a crucial role in facilitating seamless sharing of individual results and saving valuable time. Additionally, we utilised communication tools like Discord and WhatsApp to keep in touch.

One of the major challenges was the constant race against time while maintaining organisation. However, by implementing effective project management techniques, we were able to overcome this challenge successfully.

We also acquired many new skills. Firstly, we significantly improved our Python skills, which served as the primary programming language for our project. We gained expertise in working collaboratively on GitHub, enabling efficient version control and seamless integration of everyone's contributions. We also acquired a substantial understanding of advanced techniques such as computer vision and machine learning. Moreover, we explored the potential of utilising APIs, allowing us to incorporate the functionality of various web services into our scripts.

5. Conclusion

In the studied region there has not been a significant change of the water coverage index over the years, only seasonal fluctuations. We expected it to vary due to climate change. On the other hand, the vegetation coverage increased.

A possible explanation is the fact that some of the water basins that we considered are artificially regulated and are not affected by changes detectable by the AstroPi camera at a resolution of 130 m/px. With such small variations to find, any imperfection of the images might compromise the results of the analysis. Hence, the AstroPi is only suitable to study the evolution of water basins with a high variability (hundreds of metres). If we had the chance to study geographical sites such as the Nile and Toshka Lake or the Aral Sea, we would have found significant changes over the years. Sensors with a higher resolution than the AstroPi camera, along with techniques able to distinguish imperfections of the images from significant elements, would make this type of study more accurate.

Despite these issues, we were able to accurately predict the evolution of the considered areas using the SARIMA model and we found a possible correlation between the NDWI and the area covered by vegetation.