**Daily Log Project M.Sc. ECMM510**

**Student Name: James William Chamberlain**

**Project Name: Reusable tools for monitoring land use changes around UNESCO sites**

**Internal Supervisor: Dr. Milto Miltiadou**

**External Supervisor: Dr. Athos Agapiou**

**Daily Log of Activity**

|  |  |  |  |
| --- | --- | --- | --- |
| Date | Activity | Outcomes | Comments |
| *14/01/2025 (Pre)* | *MSc Research Project Introduction Seminar* | *Initial exploration and shortlisting of potential projects to go towards – Introduction to main project deliverables and deadlines* | *Noticed a few projects with the UNESCO project specifically standing out as the most interesting to pursue.* |
| *16/01/2025 (Pre)* | *Project Inquiry and Supervisor Approval Meeting* | *Approval to be on the project* | *Had an initial meeting with Milto, during which I received approval to join the project.* |
|  |  | Project Starts |  |
| 03/03/2025 -13/03/2025 (Week 1-2) | Initial Exploration into UNESCO world heritage sites and urbanism and its impacts | Explored related literature gaining a general understanding of a few potential directions this project can go in – with a primary outcome of UNESCO sites being divided into 3 categories that would majorly set the direction of the project: these being Natural Sites, Cultural Sites within Urban Areas, and Cultural Sites outside Urban Hubs.   More on related literature:  - Exploration into how monitoring is performed, with Landsat, Sentienl-1/2 were also compared - Exploration on Land Use Classification and Land Use Change were briefly explored. | This gave a good initial introduction towards the project and directions to explore; from reading the literature it was found that it would be best to look in the direction of exploring urban expansion towards sites outside of (or close to but not within) urban hubs – as this is where monitoring this threat would likely be most valuable.   Whereas sites within urban hubs a more pro-active approach towards designs and planning would be more affective – and in the case of this project being remote sensing it would only be post- or mid- construction that this project would notice any changes, as such would be less valuable. |
| 06/03/2025 (Week 1) | Initial Group Meeting *(Weekly Meeting)* | Initial Meeting officially kickstarting the project and started project proposal. | N/A – introduction to others, with useful brief discussions towards each of our projects. |
| 13/03/2025 (Week 2) | Initial Meeting with Athos | Further introduction to project, incredibly useful as this opened the direction the project can take. | Incredibly useful meeting for my understanding of the importance of project like this one – one key takeaway was the mention of how remote sensing for monitoring inaccessible sites such as those within warzones/similar. |
| Meeting for Project Proposal Outline *(Weekly Meeting)* | A draft outline for the project proposal was created, with recommendations on | N/A |
| 18/03/2025 (Week 3) | Google Earth Engine Lecture | Further introduction to Google Earth Engine// (GEE). | Very useful to understand the kind of that that I will be dealing with.   This gave me an initial understanding of direction for data collection. However, I don’t know JavaScript, and it’s also far more limited in terms of data manipulation, so the Python API would be a better path to pursue for greater versatility later. |
| 21/03/2025 (Week 3) | Discussion on Site Selection and later data collection *(Weekly Meeting)* | Discussion on site selection, and polygon selection around sites for PlotToSat and then ML (for project pipeline). | No site was selected this week; some exploration had taken place however this did not yield any results – as per looking for sites that had significant changes within the urban landscape in their surroundings. |
| Site Exploration | Focused on using easily assessable tools at this point, such as *Esri | World Imagery Wayback* and manually checking, however this was tedious an slow. | This was too tedious, so need to find a more elegant solution that allows for checking the UNESCO sites quicker. |

|  |  |  |  |
| --- | --- | --- | --- |
| 24/03/2025 (Week 4)  27/03/2025  28/03/2025 | Site Exploration with GEE and GEEMAP tools | Created a tool to aid in site exploration where I would take the site locations from the UNESCO  Explored sites initially around the Levant, 1 site in Crete and looked at pins within Romania, the UK and little bit within Spain – mainly stuck to looking at the Levant at first though however a site with large scale changes (excluding Palestine) were not really spotted.   Further exploration into the neighbouring countries – Egypt (briefly) and Iraq (heavily) identified a site in Smarra containing what I can assume to be older parts of the city and palaces from the Abbasid Caliphates (from approx. ~800s). | Would personally like to look more into the ancient cities within the area however, those appear to mostly be out of the way of Uran Areas or only have minor activity around them (excluding Ur. Where there is an airport right next to it. A day to perform further exploration would be nice – but might not be possible until the 7th of April as the project needs to move on)   A simple version of tiling was achieved but more as a coincidence while trying to load the images faster- was achieved by creating multiple tasks for GEE by tiling then requesting those segments as separate tasks then collating the image – no further exploration has been done as it was not useful to the aim of site exploration.   Initial exploration on ‘S2’ but moved to ‘S2\_HARMONIZED’ as there was less variety in image intensity (brightness) – used B4, B3 and B2 bands (RGB) |
| Group Meeting focusing on the Ethics and Risk Assessment for the project proposal.  *(Weekly Meeting Pt.1)* | Discussion on Risk assessments within the Project Proposal. A need for feasibility for parts of the project and where to move onto (contingencies) if parts do not work out as planned. | “FAIR” – Findable Accessible, Interoperable and Reusable.  Deliverables – actual deadlines for this project + parts of the project (that will be made open source)  Risk for project being noise in the data causing potential misclassification – contingency is to increase the resolution (larger polygon) so that the classifier becomes more accurate. |
| Codebase expansion – explore using polygons to collect region | Added ability to select time-span and date for selected polygons on `geemap` to collect and view for comparisons. This greatly expanding my ability to search around sites rather than it collecting just the surroundings around a UNESCO site point. | This allowed to further explore the surrounding areas of sites, Samarra is still the best choice up to this point.  This gives me a way greater |
| Meeting to discuss site findings and further discussion on creation of a project pipeline. *(Weekly Meeting Pt.2)* | Site at Smarra selected as primary site of interest. | Samarra Archaeological City, due to the scale of this site, only this site is going to be selected rather than multiple smaller sites, as it has a bunch of monuments in account of it being a mostly abandoned ancient city.   Project pipeline for the project proposal is now the next area the needs to be progressed. |
| 31/03/2025 – 06/04/2025 *(Week 4)* | Created initial project proposal document, including draft sections and an early pipeline. | Created the document overall structure with early content across all key sections | All major sections have been created; however, this is not filled in other than initial notes, but these give a good guideline for the direction to go in. The project pipeline style has been created (in latex) however its way too early to fill in, so this will be looked at closer to the end of the proposal. |
| 07/04/2025 – 13/04/2025 *(Week 5)* | Focus on exploring existing literature | Gained detailed understanding of Samarra’s background and significance to Islamic history as a well-preserved Abbasid city; identified key sources including an EAMENA damage study and two site-focused books. | Found no land-use change prediction studies for Samarra specifically, however it is brought up a lot as the main source of damage towards the site.  All other studies focusing on Samarra other than the EAMENA, were very limited, and focused specific aspects of the site. As such this is a very clear research gap for the site.  This leads the next direction to explore the literature for land-use change and it’s classification. |
| Focus on existing methods for land-use change detection and classification and satellite monitoring. | Reviewed LULC methods (RF, SVM, CNNs like U-Net); explored prediction models (Markov, Cellular Automata, LSTM); identified suitable satellite data (Landsat, Sentinel-1/2). | Work is currently in highly note-based form and has not yet been formally written up, as such this is the next direction to follow. However, this does give the direction for the methodology being to be Land-Use Change and Prediction – this also avoids the rehashing the recent EAMENA study and ensures that the direction this project follows fills a good research gap for the site. |
| 14/04/2025 -  20/04/2025 *(Week 6)* | Refinement of the Literature review and starting the methodology. | Organised previous week’s findings into a more structured literature section and began outlining and filling in the methodology focused on land-use change and prediction. | Next step is to complete and integrate the aims and objectives section (within the introduction), ensuring it all flows, does not overlap/repeat within the introduction, literature review and methodology for a cohesive project structure. |
| Focus on write-up flow across the document. | Refined aims and objectives to better reflect project direction. Then began to update the introduction, literature review and methodology to both better reflect these refined aims and objectives. | This was not fully completed, and the methodology is still a mix of first section drafts and notes for the section, so this needs to be refined, and the Discussion and Project Pipeline sections need to be completed.   The literature review and introduction have come a long way, but further refinement is still needed at this stage. |
| 21/04/2025 -  27/04/2025 *(Week 7)* | Refinement of the document and a focus on Methodology, Project Pipeline and Discussion sections (timeframe) | Refined the document further with rearrangements of the literature review and introduction due to there being too much of an overlap when it came the historical background for the site. The UNESCO section was rewritten as well due to the overlap with the Introduction + Literature Review to be more concise.  Methodology was refined into phases, with sub goals for each phase, going from data collection and pre-processing, classification and creation of land-use change maps for then use in creating predictive models (also discussion within these phase sections on measures of how to assess these classifiers and predictive models).  Made a first draft version of the project pipeline using the same style from earlier on. | The introduction and literature review at this point are getting close to ready for a first draft full review.  The methodology section has come a long way, the section on site selection is taking up too much space and repeats too much from, as such that needs to be refined to save space and improve readability.  The project pipeline has had some work done to it, to improve readability, and further discussions on what the deliverables are going to be on it needed to be added, alongside each subphase. |
| 28/04/2025 -  04/05/2025 *(Week 8)*  *Project Proposal Deadline* | Finalisation of the Project Proposal | Reduced the section on site selection down massively, whilst keeping its core information still intact.  Rewrite the literature review section for LULC classification and prediction, and this section reads closer to a discussion on what exists and what options there are to explore. The mirrored phase section in the methodology was also written to better reflect how classification will be used to create maps, and these maps will be analysed on their own, and fed into predictive models.  Finalised the Discussion section (very short)  Finalised and made major adjustments to the Project Pipeline to better reflect deadlines and finalised phases.  Draft was also reviewed by both Milto and Athos – making a note of this here as there has been correspondence primarily between me and Milto throughout the whole project proposal creation process, but this was the final and first full draft review, and last minor changes were made before submission. | This indicates the start of the project implementation and a shift away from writing for the time being.  Phase 1 being data collection and any necessary preprocessing, so that is the next step to go towards. |
| Handed in the Project Proposal | |
| 05/05/2025 – 11/05/2025 *(Week 9)* | Label the Core Site Boundaries | Created Polygons for each of the 10 major site boundaries.  UNESCO does not have exact polygons available online so had to adapt and estimate based on the 2007 maps and plans document. | This will be useable at a later stage of the project, for analysis and masking. |
| 12/05/2025 – 18/05/2025 *(Week 10)* | Focus on adapting existing code using gee and geemap to support visualisation and testing of tiling/polygon creation. | Enabled visualisation of site polygons and added functionality to export them as local shapefiles. | This now allows for easy testing, as it utilises geemap and the draw functionality to allow for custom polygons to tile. The next step requires creating tiles from these polygons. |
| Focus on creating tile generation for use with PlotToSat. | Created a basic tiling system that produces consistent tiles across an input polygon. | Tiles do not align with Sentinel-2 projection, and tile edges are clipped to the input polygon.  This creates some key concerns, tiles do not algin with the Sentienl-2 Projection, which may affect downstream as if the projection of the polygons is off, it may not be a consistent 5x5px when reduced down to the 50m resolution in PlotToSat. As such aligning the generated tiles is necessary. |
| 19/05/2025 – 25/05/2025 *(Week 11)* | Focus on refining tile generation by aligning to the s2 projection. | Came across a useful function within GEE called ReduceToVector that allows for reducing polygons to the same projection, and this means that edges also align to the projection rather than being clipped (however this is to the nearest pixel).  Chunking was implemented using the same functionality to create the tiles, and this allowed for easier visualisation and exporting to shapefiles. | This was nearly ready for exporting to GEE, however, upon the final tests an issue was discovered where along a certain line just the right of the city, (within the UNESCO zone but not toughing anything but zone 3 of the core zones). This requires fixing as entire chunks fail to generate along that strip. So before data collection can start this needs to be addressed. |
| Focus on Chunking | Chunking was implemented using the same functionality to create the tiles, and this allowed for easier visualisation and exporting to shapefiles.  *(this was worked on in conjunction with the previous cell)* |
| GEE Export for tiles | Created a function to export all chunks as separate shapefiles to GEE | This allows for using exported files with PlotToSat more easily, as the shapefiles are then already on GEE rather than having to be uploaded. |
| 26/05/2025 – 01/06/2025 *(Week 12)* | Split failed and generated tiles within failed chunks | Adjusted the code so that it splits the failed chunk section polygon, this is then exported as its own chunk.  These will be treated as their own chunks. | Through testing this strip fails no matter what, so just requesting GEE to re-attempt does not work. This splitting of the failed chunk is useful, as now there is a way to go forward and a specified region/polygon where it failed to generate. Next will need to look into filling this failed partial-chunk in. |
| Fix failed partial chunk | Attempted to regenerate the chunks, however it was very easy to create tiles that are the same size but getting the same alignment has caused some issues. | Further work on this in the following week is necessary. Issues include:   * Misalignment * Too many tiles causing overlapping with some of the neighbouring chunks. |
| 02/06/2025 – 08/06/2025 *(Week 13)* | Work on fixing and aligning generated tiles within failed chunk | Added functionality to calculate and select points that give the rotation of the tiles. | This is close, however if chunk sizes are large the minor error will build up, and all subsequent chunks have a 1x5 chunk at the start for some reason (found that out by further inspection – occurred on the oldest version as well but somehow made it past inspection due to smaller test areas rather than full site testing).   However, this is close enough that it means moving onto the next stage is possible, but it is a limitation. If going forward, larger chunk sizes are going to be used, to minimise the error within the strip, and subsequent only have on set of 1x5 polygons (which overlap with newly generated tiles is an option at a later date). |
| 09/06/2025 – 15/06/2025 *(Week 14)* | Create a script to make PlotToSat calls to collect the data for each shapefile/chunk. | Created `[pts\_samarra.ipynb](https://github.com/JamesChamberlainw/COMM514/blob/main/PlotToSat/pts_samarra.ipynb)` and `[pts\_runner.py](https://github.com/JamesChamberlainw/COMM514/blob/main/PlotToSat/pts_runner.py)`, where the notebook is an adapted version of a PlotToSat file for Samarra specifically, and pts\_runner.py makes the calls to execute the notebook multiple times making adjustments to year and chunk shapefile. | Running this there are some minior issues where its not the most stable, and either will fail within Google Earth Engine itself due to, what I can assume is too many calls being made at once or a weird interaction within PlotToSat where multiple instances being called to GEE, are causing errors causing it to fail.   Separate issue is that it keeps coming back with GEE auth randomly, so will need to keep an eye on the notebooks to ensure that it runs and does not stop making calls due to an auth issue. |
| 16/06/2025 – 22/06/2025 *(Week 15)* | Create a way to monitor and reduce the number of GEE calls from PlotToSat | Created `[pts\_check4usage.py](https://github.com/JamesChamberlainw/COMM514/blob/main/PlotToSat/pts_check4usage.py)` and integrated it into `[pts\_runner.py](https://github.com/JamesChamberlainw/COMM514/blob/main/PlotToSat/pts_runner.py)`, so that it’s a util function that checks the activity on GEE, and does not let `[pts\_runner.py](https://github.com/JamesChamberlainw/COMM514/blob/main/PlotToSat/pts_runner.py)` call the next notebook until the previous notebook as finished executing on GEE.  Additionally edited the configs for `[pts\_runner.py](https://github.com/JamesChamberlainw/COMM514/blob/main/PlotToSat/pts_runner.py)` and `[pts\_samarra.ipynb](https://github.com/JamesChamberlainw/COMM514/blob/main/PlotToSat/pts_samarra.ipynb)` which seems to have increased its stability, however time will tell, but the GEE auth issue does not seem to be popping up anymore, and no failures on GEE have been spotted for the tiles – as of so far.   ML4EO – made adaptions to the code to allow for stopping midway through, for a given year and chunk. | Due to ML4EO and not being able to keep the laptop running with 100% uptime, as I need to bring the laptop with me, additional edits were necessary to allow for stopping and resuming making PlotToSat calls.  ML4EO – incredibly useful got to speak to many people, this was deeply useful for further familiarising me with EO concepts and terminology.   IBM did a introduction to foundation models which may be incredibly useful as its powerful in the sense that it allows for adaption of a pre-trained model to fit and achieve good results for classification with a much smaller dataset (transfer learning\*).  As of the 22nd Data Collection Started on main PC which will have 100% uptime unless an auth issue occurs crashing the program – so will need to keep an eye on how this performs. |
| Machine Learning for Earth Observation conference (**ML4EO)** 3 Days: 18th to the 20th | Got to speak to many people within the earth observation field, and got to discuss their projects, and this one. Generally very helpful as it gives valuable ideas on directions this project can go down – and introduced ideas such as using transfer learning in classification by using foundation models that likely would yield better results due to the sheer quantity of data its trained on. |
| 23/06/2025 – 29/06/2025 *(Week 16)* | Continued Data Collection | Data collection appears to be taking a long time also did some minor bugfixes with `[pts\_runner.py](https://github.com/JamesChamberlainw/COMM514/blob/main/PlotToSat/pts_runner.py)`. | Forced to look into alternative approaches to data collection as this is taking too long, still letting this run however. |
| Collecting Raw Level 2 Data | Data Collection via just monthly median raster for bands 1 to 13. | Accidentally ran initially on level 1, but this collects dramatically faster (2-3 days).  There is not a lot of clouds in the region, so this likely will work fine, this however is at a much lower resolution compared to the 50m PlotToSat data and will require additional pre-processing. |
| Labelling of Land Use | Land use labelling for 2016, and 2025 using QGIS. | This is taking a very long time to create accurately labelled polygons as labels. |
| 30/06/2025 – 06/07/2025 *(Week 17)*  Minor overlap with following week by one day\* -> | Project Goals Adaption *(Weekly Meeting + 2 others to discuss and build a new plan to go forward with)* | During these meetings, clustering became the next way to use to forgo how long it’s taking to create labels. This opens up analysis on the bands (1 to 13 excluding 12), over the site as a whole and the generated clusters.  Decided to reuse the code created for collecting PlotToSat data, the collection of the indices was likely the main reason for why this took so long to collect, and a reduction in resolution from 50m to 100m (5->10px in raw) was decided on to make the data collection feasible (started running this\*).  The 100m resolution also removes all of the issues with having to regenerate tiles so, no error going forward within the data. | The next steps would be to wait for the new PlotToSat Data to finish collecting on GEE.  During the meantime, looking into how the raw level-2 data can still be used towards the project. The next step will be to combine the files, then stack them. |
| Combining the tif files | The raw level-2 s2 data (10m resolution) was clipped to the same region size takes up just shy of 100gbs. | Even without finishing generating the PlotToSat data, its clear this takes up dramatically more storage space, requires paying for space on google drive, and raises processing concerns on a single device – and have already experienced issues where rasterio crashes due to RAM filling up (32gbs). |
| 07/07/2025 – 13/07/2025 *(Week 18)* | Explored the idea of using both original resolution 100GBs of data vs PlotToSat Data | Was explored this data takes up a lot of space, and has been found to be infeasible so sticking to PlotToSat generated data  Functionality for margining tif files, and a lot of manipulation were explored. However as this is infeasible due to the quantity of data as such this does not really bring any value towards the project. | Back to purely PlotToSat data. After having a meeting his led to going back to using PlotToSat data, reducing the resolution form 50 to 100m and adjust the inputs and removed indices calculations as that’s what was causing it to take too long |
|  | Re-Generated PlotToSat Data merge csv files | Created function using pandas to merge files | Data is now in a single csv file per year |
| 14/07/2025 – 20/07/2025 (Week 19) | Implement Visualization functions for band data for a given set of tiles. | Implemented a function that creates a surface plot and individual plots for visualising the summary mean band data across a year for a data frame (each row is a single tile). | This will be very useful for comparing the bands, and their usefulness and can be used to compare clusters. |
| Point based Cluster labelling - planning | Planning for the final stage of development, using clustering and labelled points to create an interactive way of creating labelled land-use maps – and this reuses a lot of functionality written so far for the project, from the tiling, unused code created during attempts to fix tiling, and the clustering performed so far.   With the aim of creating 1. Land use maps for each given year which then can be compared at the end + still do the comparison with the clusters later but can compare grouped clusters but expected land use. | This is the final little adaption of the project that brings a lot of the project together. The next stage for this is implementing it. |
| Weekly Meeting on Report Structure and its making scheme | Following the meeting, started laying out the document further, brining across reusable text from the project proposal. | Some of this is directly reuseable, other bits will require adaption. |
| 21/07/2025 – 27/07/2025 *(Week 20)* | Bug Fixes and Land-Use Map generation testing. | A lot of bug fixing, and now maps are generatable as a tif based on the label inputs. | Minor bug fixes might be necessary, but the next step is to move into the class for multiple years to act as a wrapper on top of this.  However this works on its own for dealing with a single years’ worth of data so needs expansion to deal with more years, a secondary class to build on top of this would be the best next step as it then keep the functionality of this and expands into a way where this can be extracted and the wrapper just needs to build them with time-frame based labelling.  Ready to attempt to create a full year’s worth of data |
| 27/07/2025 - 03/08/2025  *(Week 21)* | Begin Presentation Preparation  Continued report writing, specifically the Introduction and Study Area sections  Bug fixing Clustering | Started work on presentation focusing on the Aims and Objectives and Narrative flow into the site.  Continued work on the introduction which is now close to ready for review, and generated new figures ready for Study Area section.   Fixed key issues with the clustering class, and significantly overhauled the ts\_cluster manager tool for whole project timeframe. Created a basic   Started Process of generating full Land-Use Maps + Change Maps   -=-=- (by End of Friday or Early Saturday)  Started and Completed the Process of generating full land-use maps, and difference generation to create land-use change maps.   Started Analysis of results. | Next Step is to continue working on the presentation, generate figures and incorporate generated imagery and analysis into the presentation. |
| 04/08/2025 – *10/08/2025 (Week 22)* | Working on Presentation and Figures for Report/Presentation   Label creation and map generation | Presentation Completed and Recorded, and figures and imagery created for report and  Best result achieved 581->900 labels on clustering  Clustering was taking a while so improvements were made to loading labels that dramatically increased speed (using geopandas functionality for comparing geometrise). | Maps generated, best being with 900 labels, more need to be tested as more labels could increase consistency.  Presentation complete and many reuseable bits for the report – some bits already done in the report just mainly notes and a plan. |
| Deadline for Presentation – Presentation complied and recorded and created alongside figures for the presentation and report. | | Moving fully into report writing |
| 11/08/2025 – *17/08/2025 (Week 22)* | Working Report | Working on Report to completion and uploaded to submission points (code OneDrive url, this log book and Report)  Attempted to add more labels but the clusters ended up being a bit of a limitation making the 900 label run the best one (as such used in the report). | Complete |
| Deadline for Report, Code and Logbook | | Complete |