Monitoring Property Parcels Using Semantic EO Data Cubes

Course: Analysis & Modelling | Topic: Parcel-based analysis

Rama Kamala Rajeswari Parasa

August, 2023

Contents

Abstract

Monitoring the development status of real-estate properties is in the monetary interest of banks and other financing institutions. Further, assessing the environmental impact and climate risk is increasingly becoming an important part of the calculation of creditworthiness of the properties due to push from regulations and policies for climate resilience. Open and freely available timeseries satellite data have proved to be a critical resource in retrieving latest land cover information and also temporal changes. This study explores the potential of utilising open and free Copernicus Sentinel-2 data specifically, in the context of property parcels for – (a) retrieving the semantic parcel history and, (b) measuring a proposed 'green score' for vegetation on the parcel. These two components partially contribute towards understanding the developmental changes and the green cover on the parcel respectively. In order to fulfil its objectives, the study fundamentally uses Semantique python library to query an Austrian semantic Earth observation (EO) data cubes infrastructure (sen2cube.at).

1. Introduction & Literature

1.1. Background

Valuation and credit ratings of real-estate properties are important tools for assessing their monetary value and their creditworthiness. They provide an insight into the financial health of the property, its risk profile, ability to recover costs and generate income. The ratings inform investors and lending institutions to make critical financial decisions such as identifying the most profitable locations to invest, estimating projected rent income from a property, ascertaining collateral value and potential losses from a property, and so on (Plazzi et al., 2008; Ivanov & Faulkner, 2020).

The real estate assessments consider data inputs of the parcel such as location, size, age, surrounding neighbourhood, travel accessibility, land development status, sale price, zoning, utilities, etc (Pagourtzi et al., 2003; Kokot & Gnat, 2019). Conventionally, much of this information is gathered from sources such as public records maintained by government agencies, physical property inspections, online real estate databases, local government data and building permits and records (European Valuation Practice: Theory and Techniques, 1996). Some of these data sources and data collection practices can be prohibitively expensive, error prone or infrequent. For instance, conducting multiple site visits for physical inspections can be highly expensive for the evaluators in the long run (Nightingale & Rossman, 1994). This can hinder the monitoring of construction activity or other land development changes. Similarly, public government records may fail to provide up-to date information due to reporting delays (Gee, 2008).

Additionally, monitoring of environmental aspects of the real-estate properties is becoming increasingly important. This is because, international commitments like the Paris Agreement, 2015, and legislation like the EU's Sustainable Finance Taxonomy are now attempting to establish concepts such as 'sustainable finance' and align cash flows with sustainable economic activities (OBSERVER, 2023). For real-estate sector, this could potentially translate into policies such as evaluating the green cover on the property parcels or identifying solar roofs within them resulting in a favourable rating of those properties.

In this context, open and freely available satellite data is proposed as a viable resource to be integrated with the conventional practices of real-estate assessment (Wei et al., 2022). Openly available satellite data has been established in literature as a valuable data source for studying several land related phenomena. Examples of such phenomena include, land cover changes, urban expansions, neighbourhood characteristics, urban heat islands, crop activity and so on (Weng et al., 2004, Munafò & Congedo, 2017; Deng et al., 2017; Guerri et al., 2022, Segara et al., 2020). Characteristics such as high temporal resolution and considerable spatial resolution allow for EO data to be used for such rigorous analysis of land over time. Hence, these properties of EO data are of great value in real-estate. Consider, for instance, Sentinel-2 that has a spatial resolution of 10m (also 20m and 60m) and a revisit time of 5 days at equators with both operational satellites and more frequent at high latitudes (at different viewing angles). Contrasting the temporal frequency of Sentinel-2 against that of the government records that are updated and published either quarterly or less frequently, paints a clearer picture of the potential of EO data in providing regular, up-to date and verifiable information about the land. When integrated with the current assessment processes, it can also substantially bring down the costs for site visits.

1.2. Parcel-based analysis

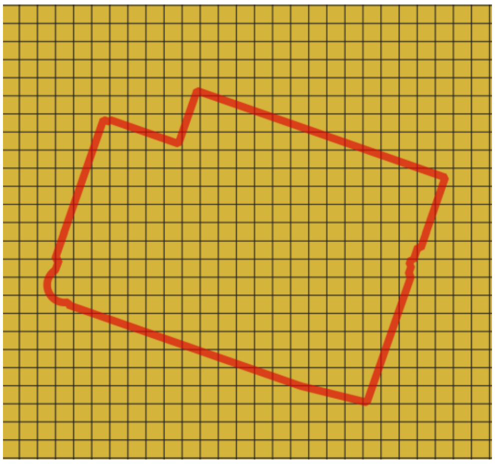

Figure 1: Illustration for parcel-based approach (Source: Author's own)

Parcel-based analysis techniques analyse satellite data at the level of individual parcels or land units (Blaschke et al., 2008). Figure 1 shows an illustration of such an approach. The grid (yellow) represents a raster from which information is extracted for the parcel (in red). These land units are generally property boundaries defined in cadastral maps or administrative units. Parcel-based methods can be contrasted with pixel-based approaches that perform analysis at the level of individual pixels, considering each pixel as discrete unit. Further, compared to the image-object-based approaches, wherein image objects are created strictly based on internal homogeneity principles, parcel-based approaches differ in the sense that each parcel represents a real-world unit of land that is under a unified ownership (*Ibid*.).

There are some notable advantages to parcel-based approaches compared to other approaches discussed above. Firstly, they can incorporate semantic information associated with the parcel, such as ownership information, policy or regulation related information, etc. Secondly, they capture internal heterogeneity within the individual units. This is useful when the spatial context and spectral variation within the parcel needs to be analysed. Finally, since the parcels hold an inherent meaning, i.e., each unit is under one ownership, the results produced are potentially transferable to other use cases that use the same parcels as the unit of analysis. One main limitation to parcel-based approaches is that, when the size of the parcel is much smaller than a pixel of the imagery, then it is highly complex to interpret the results. This complexity is resulted from mixed pixel problems that occur due to presence of overlapping spectral information, within the same pixel, received from two different but adjacent physical features, such as vegetation and built-up (Sozzi et al., 2020; Vélez et al., 2020; Qarallah et al., 2022). In such scenarios, obtaining high-resolution imagery, such as from UAV platforms or commercial satellites, is recommended (Sozzi et al., 2020 ; Vélez et al., 2020).

Extensive literature is available showing use of parcel-based approaches in diverse agricultural studies involving study of cropping patterns, soil mapping, grassland monitoring, and more. Further, the techniques used to extract information from the EO data for the parcels depend on the type of EO data used, such as multispectral or Synthetic Aperture Radar (SAR), etc; and also, on the specific use case. Tamm et al. (2016) use SAR images from Sentinel-1 to correlate the interferometric coherence in image pairs to mowing events in grassland parcels. Hartmann et al. (2023) also look at mowing events in grasslands. They demonstrate the use of Sentinel-2 multispectral imagery to differentiate between the temporal heterogeneity of mowing events of hay milk grasslands to that of conventionally managed grasslands in Austria. On the other hand, Reinermann et al. (2022) combines Sentinel-1 and Sentinel-2 time series for detection of grassland mowing events in Germany. Dusseux et al. (2014) also combine optical images from SPOT-5 and Landsat sensors, and SAR images from RADARSAT-2 to train a support vector machine (SVM) to classify grassland parcels and crop parcels in France. Due to difference in spatial resolution of the optical and SAR images, they aggregate the 15 types of variables calculated at the pixel level from both sources, to the parcel or field level using the mean statistic. Such studies provide valuable information for monitoring of agricultural land parcels that can be useful for implementing crop policies. In fact, ESA launched the Sentinels for Common Agricultural Policy - Sen4CAP as a project dedicated to provide 'validated algorithms, products, workflows and best practices' derived from Sentinel missions that are relevant for the implementation of the CAP (Sen4CAP, n.d.). The algorithms are generated to check farmers' compliance with sustainable farming practices on their farming parcels against specific subsidy scheme applications.

This study couldn't find literature on use of open and free satellite datasets to study temporal changes in urban real-estate property parcels such as residential or industrial. Urban applications that happen to use parcel-based approaches, generally use high resolution or very high-resolution imagery obtained from privately deployed UAVs or commercial satellite missions (Zhou & Troy, 2007; Zhang et al., 2017) and not free and open EO data. This is also the case with studies analysing change in parcels over time (Bin et al., 2013). This striking discrepancy in exploration of parcel-based approaches for agricultural applications vis-à-vis built real-estate applications, using open and free satellite data, can be majorly due to the difference in the parcel sizes between agricultural applications and real-estate applications. In the latter, parcel sizes are much smaller causing a highly pronounced mixed pixel problem.

1.3. Research questions

In the context of parcel-based analysis for real-estate assessment and sustainable finance, this paper asks two main research questions. Can open and freely available Copernicus Sentinel-2 data potentially be used for an insight into –

1. the semantic history of a parcel?
2. the intensity and extent of vegetation on the parcel?

2. Materials & Methods

2.1. Study Parcels

I accessed the data on cadastre parcels from the open data portal – (https://data.bev.gv.at/geonetwork, accessed on 10/07/2023) maintained by the Federal Office of Metrology and Surveying of Austria (BEV). Digital Cadastre Map (DKM KAT_DKM_GST_epsg31287_20221001.gpkg) is the official spatial database of the cadastre parcels in Austria.

I selected three arbitrary parcels from the dataset located in the city of Vienna for the analysis. The parcels were finalised upon identifying some kind of construction activity by visual inspection in the historical imagery from Google Earth Imagery. These are shown in Figure 2. New buildings were observed in all three parcels in 2020. For parcel A and B, the Google Earth Imagery reveals two important stages before the appearance of new buildings – clearing of land (2018) followed by the construction activity (2019). For parcel C, these two stages cannot be spotted due to lack of adequately frequent imagery before the appearance of the new building.

(A) (B) (C)

Figure 2: Study Parcels. Data source: Digital Cadastre Map (DKM) Basemap: Google Earth Imagery

The study parcels are of varying spatial extent and as a result, each of them subsumes a different number of Sentinel-2 pixels. Parcel A derives information from 32 pixels, parcel B from 21 pixels and parcel C from 4 pixels.

2.2. Data Access

2.2.1. Semantic EO data cube & Semantique

In this study, I utilised data from the Sentinel-2 satellite mission of the European Copernicus Earth Observation Programme. The mission consists of two operational satellites, Sentinel-2A and Sentinel-2B. Together, they provide a revisit cycle of about 5 days at the equator (2-3 days at high latitudes with different viewing angles). The satellites carry a Multi Spectral Instrument (MSI) that captures reflectance values across 13 spectral bands at different resolutions. In this study, I use the Red, Blue, Green, Near Infrared (at 10m resolution) and Shortwave Infrared bands (at 20m resolution) (ESA, n.d.).

I retrieved the Sentinel-2 data, semantic information and greenness-index layers used in this analysis from the spatiotemporal Austrian Sentinel-2 semantic EO data cube, also called as Sen2Cube.at (demo access). Sen2Cube.at offers semantic information, which is a categorical explanation of each pixel, as well as Sentinel-2, Level 1C image data, a brightness layer, a greenness-index layer, and semantic information (Augustin et al., 2019). The SIAM (Satellite Image Automated Mapper) software, which automatically constructs spectral categories from multispectral picture data calibrated to at least top-of-atmosphere (TOA) reflectance (Baraldi et al., 2010) is used to obtain the semantic information. The categorisation offered by SIAM is based on predefined rulesets encoded in a decision tree that takes the physical and spectral information as input and applies to each pixel. Figure 3 shows the legend for SIAM colour names grouped by related semantic associations for a granularity of 33 spectral categories. In the context of our application, the vegetation spectral categories can be viewed as varying intensities of vegetation. The other relevant colour categories are bare soil or built-up. Sen2Cube.at provides a graphical interface for allowing querying of the data at the parcel level. However, for more control over the workflow, better transferability and reproducibility, a programmatic access was desirable. This was overcome by Semantique python library.

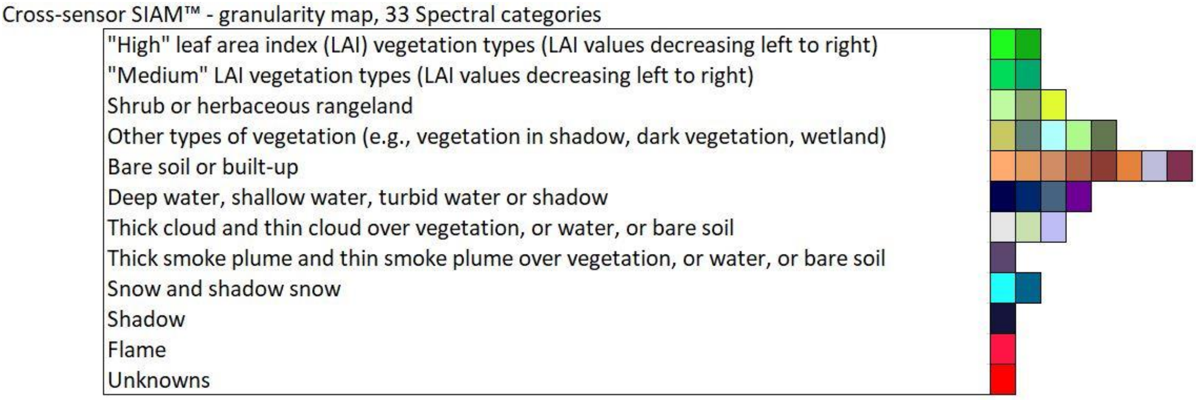

Figure 3: Legend for SIAM colour names grouped by related semantic associations and represented by related colours for visualisation. Source: Sen2Cube.at, accessed 27/07/2023

Semantique is a python package that allows for an ontology-based querying of the EO data cube (van der Meer et al., 2022). The ontology codifies representations of real-world entities. And through a mapping component constructed by the user (an EO expert) these representations are then mapped to data values in the data cube (*Ibid*.). This essentially cuts off the need for general users to directly interface with the data values. I created workflows as Jupyter Notebooks that are configured on the Sen2Cube.at server with all the necessary environment requirements including the Semantique library.

2.2.2. Cloud and snow filters

For each parcel, I generated cloud filter for each parcel using the semantic layer in the EO data cube. The semantic categories provided in the layer for each pixel produce adequate results for cloud cover estimation (Tiede et al., 2021). I used the semantic categories 'smoke plume' and 'cloud' in this study to create a cloud filter. For the generation of snow filter for each parcel, categories 'snow' and 'snow shadow' were used. For both filters, the threshold was set at 5%. This means that all the acquisition date timestamps (referred to as timestamps from here) where the cloud coverage is greater than 5% were removed from the analysis. Similarly for the snow coverage. Then, an intersection of the filters is applied to obtain data of timestamps of acceptable snow and cloud cover. The resultant set of timestamps are used in this study to filter the semantic parcel history and layers used in 'green score' calculation.

Parcel	Total scenes available	Scenes with acceptable snow cover	Scenes with acceptable cloud cover	Intersection (Acceptable snow & cloud)	Scenes lost
Parcel A	1003	900	726	632	371 (~37%)
Parcel B	1003	895	759	661	342 (~34%)
Parcel C	1003	903	759	665	338 (~34%)

Analysis time period: 2015-07-10 to 2021-04-30

2.2.3. Vegetation mask

In order to retrieve greenness-index information strictly from vegetated pixels in the parcel, or 'intelligent' greenness-index (Baraldi et al., 2010), I generated a vegetation mask for the parcels for each timestamp. A vegetation mask is useful because, greenness-index values can be generated for non-vegetation pixels as well, due to reasons such as spectral information from neighbouring pixels. In order to generate this mask, the semantic information layer is accessed and a Boolean array is returned for each timestamp with a True value for pixels where vegetation is observed and a False for vegetation absence.

2.3. Semantic Parcel History

In order to address the first research question of this study, I generated the semantic parcel history for the three study parcels. It gives an insight into the temporal occurrence of different spectral categories in the parcel. The hypothesis behind studying the semantic parcel history is that it can potentially hint at the major changes across spectral categories that the parcel may have undergone during the time period of interest. For example, an increase in the proportion of built surface extent within the parcel will result in a decrease in the vegetation extent observed in it, at least in the short term. If so, the semantic parcel history should be able to indicate such a phenomenon. For this, I queried the EO data cube for the colour types information for each timestamp in the time period. I then applied the snow and the cloud filters to retrieve scenes of timestamps with acceptable cloud and snow cover. Then the count of pixels within each parcel was retrieved for each spectral category and stored in a tabular format supported by Python's data structures. Each row in the table corresponds to the semantic categories information of the parcel associated to one timestamp as seen in the Figure 5. The workflow for the semantic parcel history is shown in Figure 4.

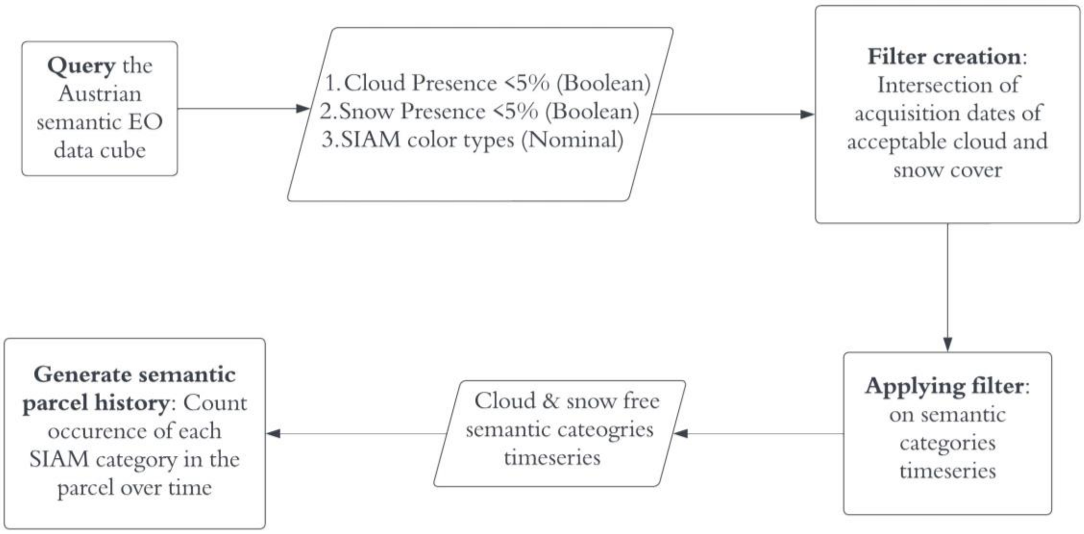

Figure 4: Workflow for generating semantic parcel history

time	33.0	31.0	30.0	28.0	27.0	25.0	23.0	22.0	19.0	17.0	...	13.0	11.0	10.0	8.0	7.0	6.0	5.0	4.0	3.0	1.0
2015-07-11 10:00:08.459000	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	1.0	5.0	4.0	0.0	0.0	0.0	11.0	11.0
2015-07-31 10:00:09.460000	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	3.0	3.0	4.0	0.0	0.0	0.0	17.0	4.0
2015-08-07 09:50:07.458000	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	3.0	3.0	13.0	0.0	0.0	0.0	12.0	0.0
2015-08-10 10:00:10.462000	0.0	0.0	0.0	0.0	5.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	2.0	9.0	2.0	0.0	0.0	4.0	10.0	0.0
2015-08-27 09:54:42.047000	0.0	0.0	0.0	0.0	8.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	1.0	2.0	1.0	0.0	0.0	2.0	18.0	0.0
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
2021-03-31 10:07:16.906165	0.0	0.0	0.0	5.0	7.0	0.0	3.0	0.0	0.0	5.0	...	0.0	0.0	5.0	1.0	0.0	0.0	0.0	3.0	0.0	0.0
2021-04-05 10:07:17.205533	0.0	0.0	0.0	0.0	10.0	0.0	3.0	0.0	0.0	0.0	...	0.0	0.0	15.0	1.0	0.0	0.0	0.0	1.0	2.0	0.0
2021-04-10 10:07:14.743547	0.0	0.0	0.0	2.0	30.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
2021-04-12 09:57:18.501935	0.0	0.0	0.0	0.0	32.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
2021-04-25 10:07:14.335154	0.0	0.0	0.0	2.0	29.0	0.0	1.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

302 rows × 23 columns

Figure 5: Sample table with results for semantic parcel history. Each value represents the count of pixels observed in a semantic category at a given acquisition date. Rows represent acquisition dates. Columns represents semantic categories observed in the parcel.

2.4. Proposed ‘Green Score’

In order to address the second research question asked by this paper, I propose a ‘green score’ which is conceived as a product of aggregated (mean, maximum and minimum) greenness-index at the parcel level and the proportion of spatial extent of vegetation in the parcel. Baraldi et al. (2010) introduced the greenness-index and it uses the red, near infrared and shortwave infrared bands. They found that greenness-index correlates better with the biophysical variable, Leaf Area Index (LAI) compared to the Normalised Difference Vegetation Index (NDVI). The proportion of vegetation extent is derived based on the semantic information available for each pixel from the data cube.

The rationale behind defining the score is that, neither greenness-index nor proportion of area under vegetation alone comprehensively defines the nature of vegetation observed within the parcel. To unpack this further, a relatively high aggregated greenness-index value of a parcel doesn’t necessarily mean that a considerable part of the parcel is covered with intense vegetation. The high greenness value could be resulting from a very few pixels in the parcel with intense vegetation and it doesn’t reveal much about the spatial extent of such high intensity vegetation in the parcel. Similarly, proportion of vegetation in the parcel only gives a measure of the extent of vegetation and doesn’t reveal anything about the quality or intensity of such vegetation. Combining these two measures into a product gives a more comprehensive insight into the vegetation in the parcel. Further, the timeseries curve of such a product is more representative of the combined temporal changes taking place in the spatial extent and intensity of vegetation on the parcel on ground.

In order to calculate the proposed green score, I queried the data cube for greenness-index layer and for presence of vegetation in the parcel. I used the latter to create a vegetation mask and applied it on the greenness-index layer. Snow and cloud filters generated for the parcel were also applied. Then, for each parcel, the product of the aggregated greenness index and percentage of area under vegetation was calculated. The results were stored in a table such that each row corresponds to the green score information of the parcel associated to one timestamp as seen in Figure 7. The workflow for the green score profile of a parcel is shown in Figure 6.

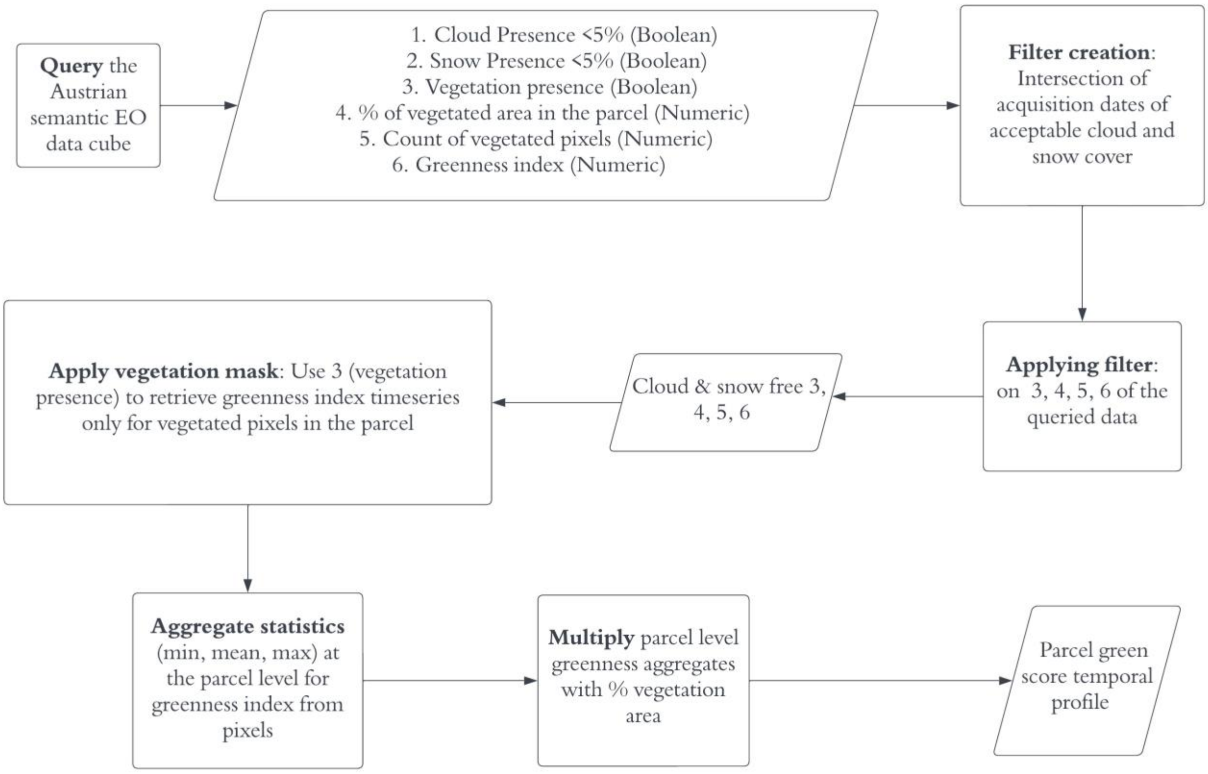

Figure 6: Workflow for generating green score profile

time	min_green	mean_green	max_green	veg_perc	veg_count_outof_4	min_score	mean_score	max_score
2017-03-29 09:50:24.458000	2.327586	2.702540	3.000000	100.0	4.0	232.758617	270.254040	300.000000
2017-04-01 10:00:22.464000	2.409722	2.933961	3.306324	100.0	4.0	240.972233	293.396109	330.632401
2017-05-01 10:00:29.460000	3.241539	3.562556	3.959184	100.0	4.0	324.153852	356.255645	395.918369
2017-05-11 10:05:39.727000	3.358627	3.358627	3.358627	25.0	1.0	83.965671	83.965671	83.965671
2017-05-18 09:50:32.459000	3.638814	3.811466	3.980769	75.0	3.0	272.911048	285.859948	298.557687

Figure 7: Sample table with results for green score profile of a parcel.

3. Results

3.1. Semantic Parcel History

Construction activity over previously vegetated land causes a loss in vegetation. Consider Parcel A and the corresponding timeseries in Figure 9. As expected, we observe a drop in the number of pixels of vegetation semantic categories (labelled from 1 to 9) from 2019 onwards. This is accompanied by an increase in the number of pixels of weak rangeland/built/barren semantic categories. In the year 2020, the building construction appears to be complete with a faint regrowth of some vegetation. This is reflected in the timeseries with a rise in vegetation category pixels with a continued large number of pixels under rangeland /barren/built categories. Figure 10 shows the timeseries of grouped categories based on their semantic association resulting in two broad groups green (SIAM categories 1 to 9) and barren/built-up (SIAM categories 10 to 20). Other SIAM categories (SIAM categories >20) are ignored from the timeseries analysis to limit the scope and keep the analysis relevant for construction activity and resultant changes in vegetation in the parcel. However, other categories can also be studied within the parcel through this technique, say, loss of any previously present water bodies in the parcel.

Parcel B exhibits a similar trend, wherein, the number of vegetation pixels drop and that of barren/built pixels rise at the start of clearing of land (2018) through the construction period (2019) as seen in Figures 12 & 13. It is pertinent to note that throughout the timeseries, the canopy of a neighbouring tree is observed to be covering a portion of the parcel. Hence, the canopy could be contributing to the count of vegetation pixels observed in the parcel since the neighbourhood is not cleared during the construction activity. The analysis of the effect on the parcel due its neighbourhood features is not within the scope of this study, however, crucial for decision-making.

The smallest parcel of the three, parcel C spans across only four Sentinel-2 pixels. This causes no visible difference to the count of pixels under vegetation and built/barren due to the construction activity. Hence the timeseries charts in Figure 15 and Figure 16 do not reveal any major trends and do not lend themselves to a meaningful interpretation.

Charts: Semantic Parcel History

Semantic parcel history for parcel A (32 Sentinel-2 pixels)

Figure 8: Temporal high-resolution images for parcel A obtained from Google Earth Imagery

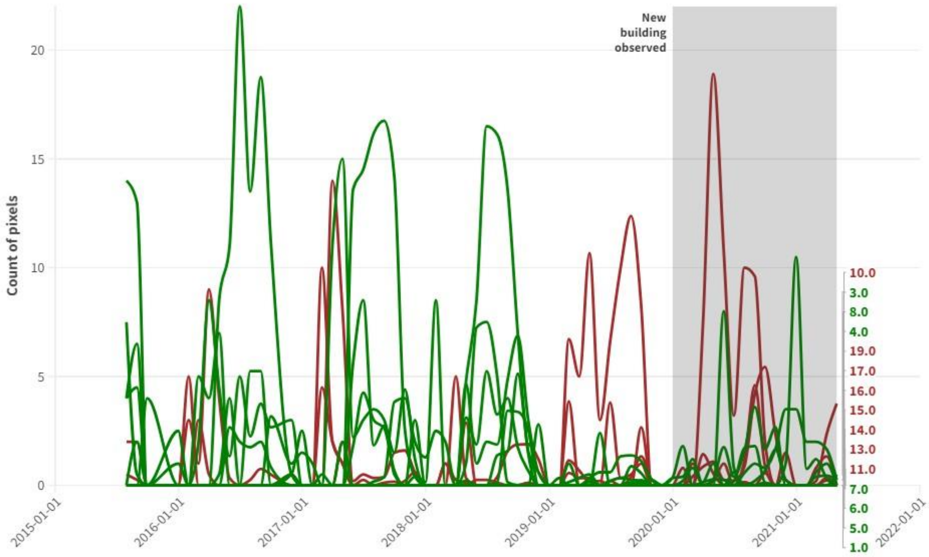

Figure 9: Timeseries curves of count of pixels for semantic categories (SIAM 1 to 20) for parcel A as a monthly mean aggregate

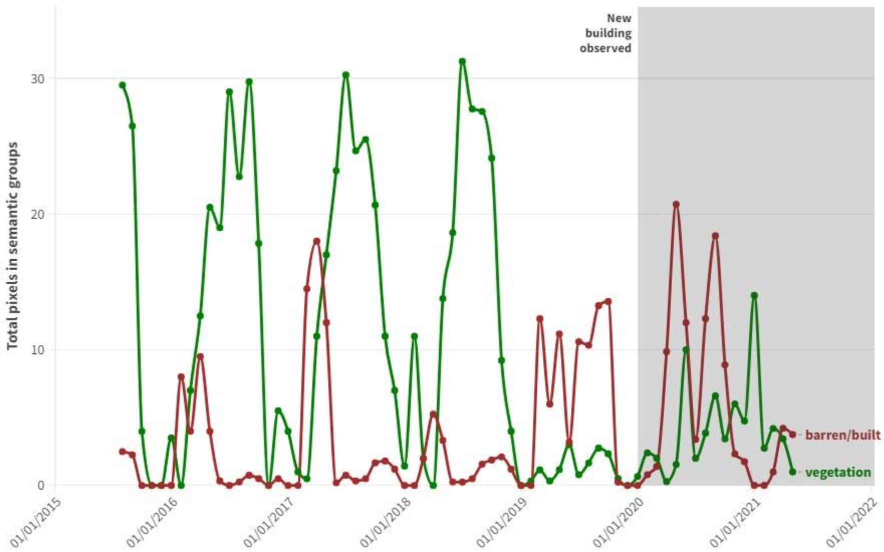

Figure 10: Timeseries curves of count of pixels for semantic groups, vegetation (SIAM 1 to 9) and barren/built (SIAM 10 to 20), for parcel A as a monthly mean aggregate

Semantic parcel history for parcel B (21 Sentinel-2 pixels)

Figure 11: Temporal high-resolution images for parcel B obtained from Google Earth Imagery

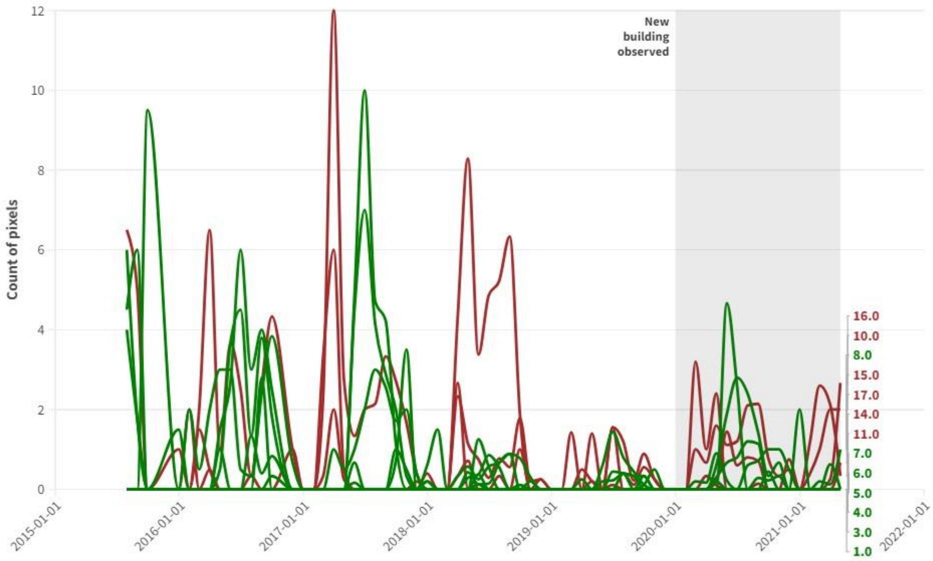

Figure 12: Timeseries curves of count of pixels for semantic categories (SIAM 1 to 20) for parcel B as a monthly mean aggregate

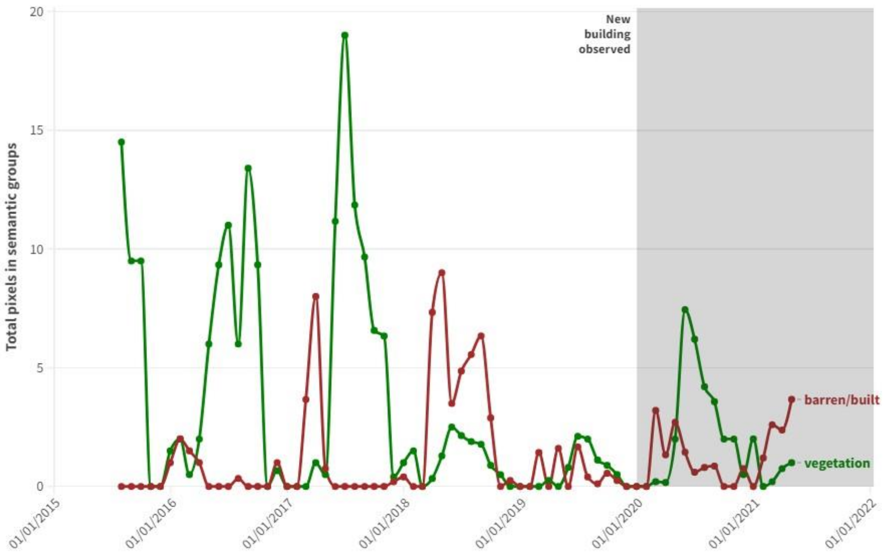

Figure 13: Timeseries curves of count of pixels for semantic groups, vegetation (SIAM 1 to 9) and barren/built (SIAM 10 to 20), for parcel B as a monthly mean aggregate

Semantic parcel history for parcel C (4 Sentinel-2 pixels)

Figure 14: Temporal high-resolution images for parcel C obtained from Google Earth Imagery

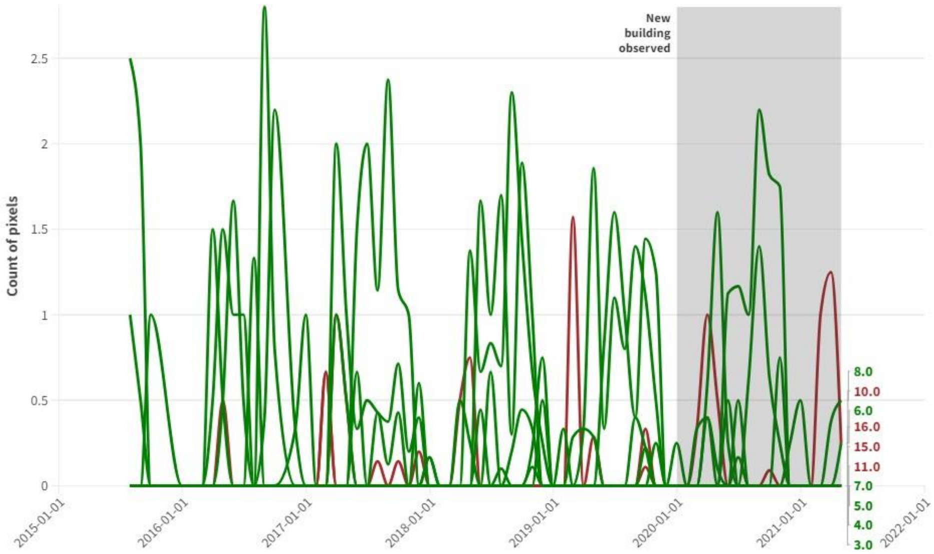

Figure 15: Timeseries curves of count of pixels for semantic categories (SIAM 1 to 20) for parcel C as a monthly mean aggregate

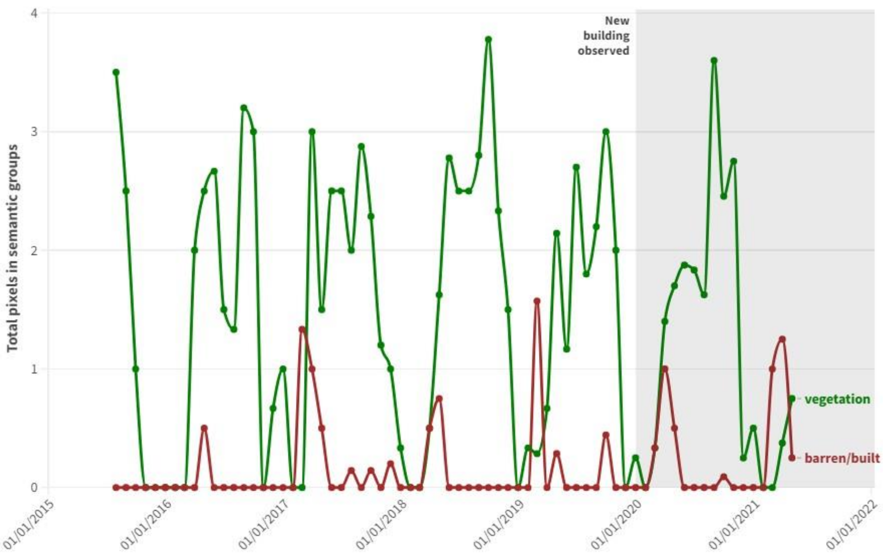

Figure 16: Timeseries curves of count of pixels for semantic groups, vegetation (SIAM 1 to 9) and barren/built (SIAM 10 to 20), for parcel C as a monthly mean aggregate

3.2. Green Score

We expect green score timeseries to be representative of the changes in green cover both in terms of its intensity and spatial extent. The timeseries charts of parcels A (Figure 19) and B (Figure 22) for the green score show a combined view of the trends from timeseries of both the percentage area of vegetation and greenness index (Figures 18 and 21 for parcels A and B respectively). As expected, there is a sudden drop in green score at the start of the construction activity. Then, there is a gradual rise in the green score after the construction activity is completed indicating regrowing vegetation. The green score profile of parcel C (Figure 25) doesn't reveal much information because of the challenges mentioned earlier. The small size of the parcel (four Sentinel-2 pixels), makes it difficult for the parcel analysis to provide any meaningful results.

(Charts follow from next page)

Charts: Derived Green Score

Green score profile of parcel A (32 Sentinel-2 pixels)

Figure 17: Temporal high-resolution images for parcel A obtained from Google Earth Imagery

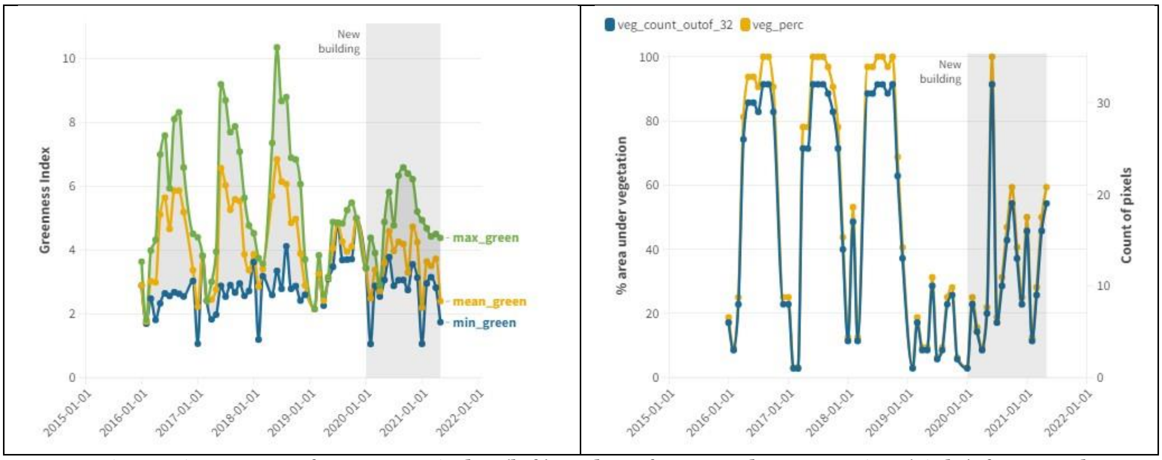

Figure 18: Timeseries curves of greenness-index (left) and % of area under vegetation (right) for parcel A as monthly aggregates (maximum)

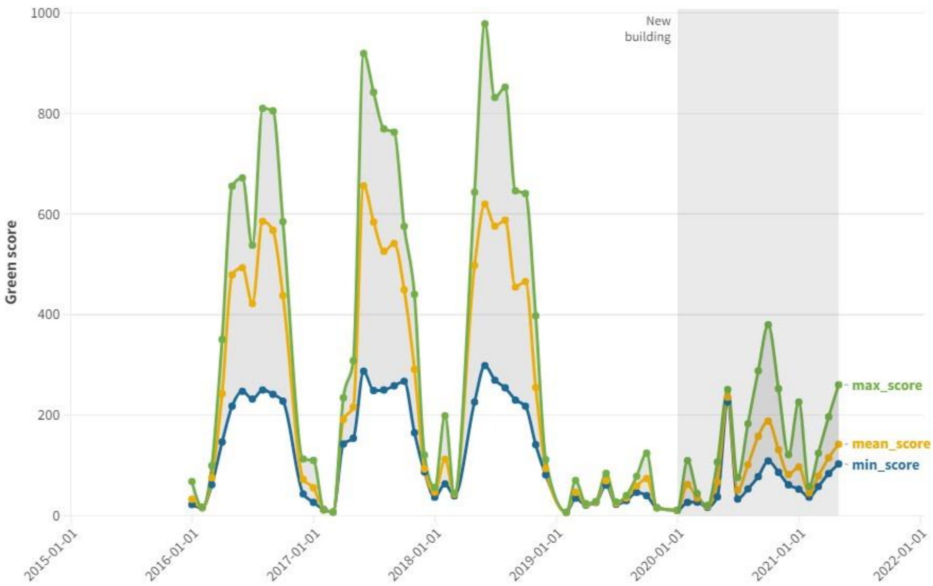

Figure 19: Timeseries curve of the proposed green score for parcel A as a monthly aggregate (maximum)

Green score profile of parcel B (21 Sentinel-2 pixels)

Figure 20: Temporal high-resolution images for parcel B obtained from Google Earth Imagery

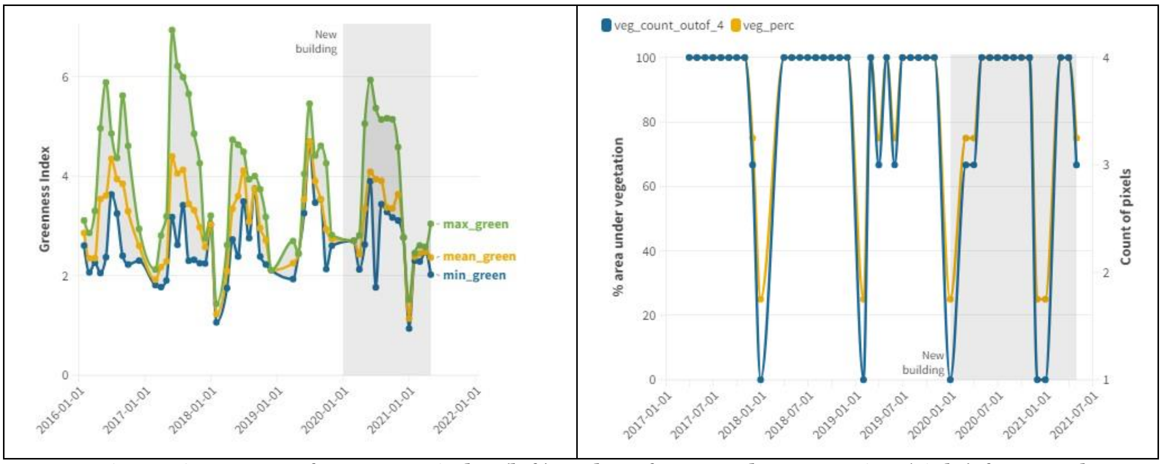

Figure 21: Timeseries curves of greenness-index (left) and % of area under vegetation (right) for parcel B as monthly aggregates (maximum)

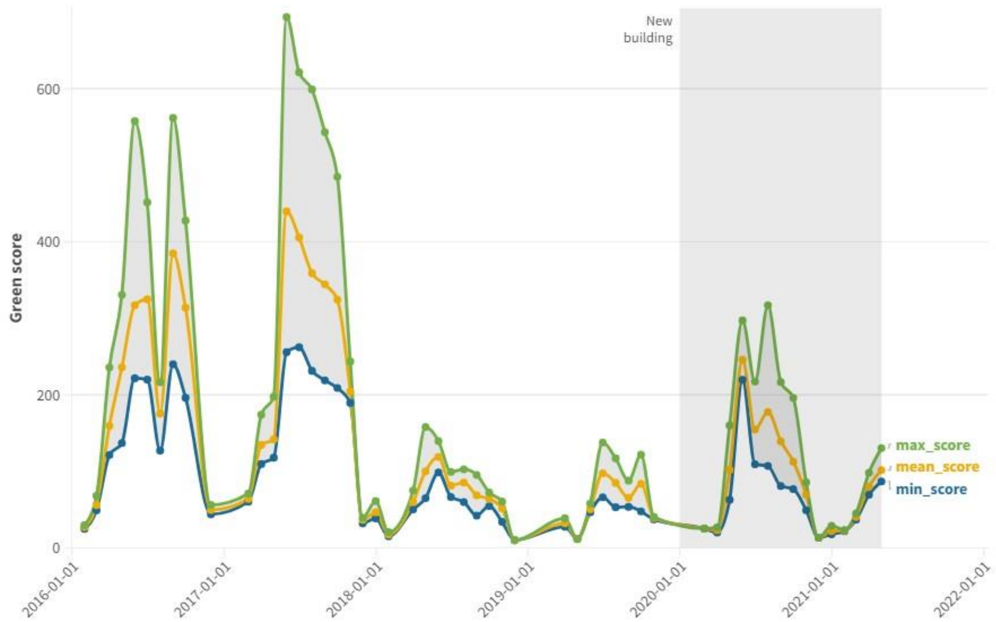

Figure 22: Timeseries curve of the proposed green score for parcel B as a monthly aggregate (maximum)

Green score profile of parcel C (4 Sentinel-2 pixels)

Figure 23: Temporal high-resolution images for parcel C obtained from Google Earth Imagery

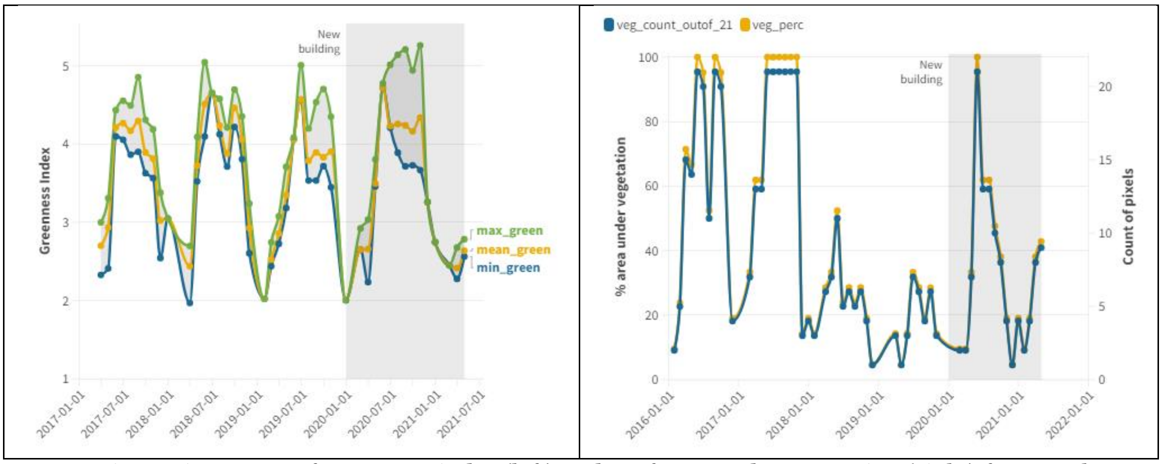

Figure 24: Timeseries curves of greenness-index (left) and % of area under vegetation (right) for parcel C as monthly aggregates (maximum)

Figure 25: Timeseries curve of the proposed green score for parcel C as a monthly aggregate (maximum)

4. Discussion & Conclusions

This study shows that the freely available Sentinel-2 imagery can be used to not only retrieve latest information about a real-estate parcel but also in 'looking back into its past' and understanding the changes that took place on the parcel over time. This can be a critical resource in producing environmentally-aware credit ratings of properties. For example, a parcel that is built in a previously fully vegetated parcel is perhaps less environmentally friendly vis-à-vis a parcel that is redeveloped (i.e., previously non-vegetated). In such scenarios, generating a *semantic parcel history* can be a useful input to the decision-making. Further, such a history can also help flag other kinds of changes over the parcel such as a water body being converted into built-up.

The *green score*, proposed as a multiplicative product of greenness-index aggregated at the parcel level and % area of parcel under vegetation, has the potential to provide a comprehensive picture at the vegetation in a parcel. It can allow for comparisons among different parcels in terms of their relative vegetation. In buildings where some regrowth of vegetation is observed post construction, such as through greening of the roofs, it allows us to hypothesize the following: *Even when there is some recovery of vegetation on the parcel post the construction, the new vegetation can be of a decreased intensity or spatial extent or both compared to the pre-construction vegetation*. Future studies can test this hypothesis and explore the potential of this new indicator at setting post-construction greening targets for property owners relative to the pre-construction state of vegetation on the parcel.

I found availability of limited literature on use of Sentinel-2 data and parcel-based analysis for real-estate assessment as a major challenge in carrying out this study. This calls for a need to explore the Sentinel-2 data for applications in the real-estate domain, alongside usage of high and very high-resolution imagery. Further, for each parcel, cloud and snow cover together resulted in a loss of about 34%-37% scenes from analysis. This can be a challenge when imagery is not available for a prolonged period causing a challenge to decision-making. Finally, the adequacy of the spatial resolution of Sentinel-2 imagery for the analysis of real-estate property parcels is an important aspect to consider. This study indicates that the interpretation of temporal patterns might be more relevant for larger property parcels compared to smaller ones. Hence, while workflows built on top of Sentinel-2 data can be fruitful, they cannot fully replace the conventional monitoring systems such as site visits, and hence need to be meaningfully integrated with them. Further, they can also be integrated with monitoring systems built upon UAV-based high-resolution imagery.

The present study is only exploratory and indicative. Hence, the limitations of this study include lack of statistical analysis on a significant number of parcels. Future studies can look at extending this analysis to several parcels at neighbourhood or city level. Further, issues arising out of the mixed pixel problem and potential pathways to handle it, need to be explored more in the context of real-estate parcels. Finally, studies focused on setting a minimum threshold for the size of a real-estate parcel that can produce statistically significant results for the analyses workflows presented in this study, would also be valuable.

5. Acknowledgements

This work is greatly benefitted by the knowledge and research developed by Thomas Strasser, Lorena Abad Crespo and Hannah Lucile Augustin, researchers at the EO Analytics, especially as part of the ESG Pro project. I am also grateful to Prof Dirk Tiede and Dr Martin Sudmanns, conversations with whom have informed and guided the research.

6. References

Augustin, H., Sudmanns, M., Tiede, D., Lang, S., & Baraldi, A. (2019). Semantic Earth Observation Data Cubes. Data, 4(3), 102. https://doi.org/10.3390/data4030102

Baraldi, A., Durieux, L., Simonetti, D., Conchedda, G., Holecz, F., & Blonda, P. (2010). Automatic Spectral-Rule-Based Preliminary Classification of Radiometrically Calibrated SPOT-4/-5/IRS, AVHRR/MSG, AATSR, IKONOS/QuickBird/ORBView/GeoEYE, and DMC/SPOT-1/-2 Imagery—Part I: System Design and Implementation. IEEE Transactions on Geoscience and Remote Sensing, 48(3), 1299–1325. https://doi.org/10.1109/tgrs.2009.2032457

Blaschke, T., Lang, S., & Hay, G. J. (2008). Object-Based image analysis. In Lecture notes in geoinformation and cartography. https://doi.org/10.1007/978-3-540-77058-9

Datenkatalog BEV. (n.d.). Accessed on 10/07/2023. https://data.bev.gv.at/geonetwork

Deng, J., Huang, Y., Chen, B., Tong, C., Liu, P., Wang, H., & Hong, Y. (2019). A methodology to monitor urban expansion and green space change using a time series of Multi-Sensor SPOT and Sentinel-2A images. Remote Sensing, 11(10), 1230. https://doi.org/10.3390/rs11101230

Dusseux, P., Corpetti, T., Hubert-Moy, L., & Corgne, S. (2014). Combined use of Multi-Temporal optical and radar satellite images for grassland monitoring. Remote Sensing, 6(7), 6163–6182. https://doi.org/10.3390/rs6076163

European Valuation Practice: Theory and techniques. (1969, December 31). Routledge & CRC Press. https://www.routledge.com/European-Valuation-Practice-Theory-and-Techniques/Adair-Downie-McGreal-Vos/p/book/9780419200406

Gee, H. (2010). Residential real estate data on the Internet: Benefits and limitations. Journal of Business & Finance Librarianship, 15(2), 104–122. https://doi.org/10.1080/08963560903554658

Guerri, G., Crisci, A., Cresci, I., Congedo, L., Munafò, M., & Morabito, M. (2022). Residential buildings' real estate values linked to summer surface thermal anomaly patterns and urban features: a Florence (Italy) case study. Sustainability, 14(14), 8412. https://doi.org/10.3390/su14148412

Hartmann, A., Sudmanns, M., Augustin, H., Baraldi, A., & Tiede, D. (2023). Estimating the temporal heterogeneity of mowing events on grassland for haymilk-production using Sentinel-2 and greenness-index. Smart Agricultural Technology, 4, 100157. https://doi.org/10.1016/j.atech.2022.100157

Ivanov, S. I., & Faulkner, M. (2020). Corporate real estate ownership, cash and credit ratings. Journal of Economic Studies. https://doi.org/10.1108/jes-05-2020-0231

Kokot, S., & Gnat, S. (2019). Identification and Classification of Real Estate Features for the Purpose of an Algorithm-Based Valuation– Case Study within Szczecin. Folia Oeconomica Stetinensia, 19(2), 38–55. https://doi.org/10.2478/foli-2019-0012

Munafò, M., & Congedo, L. (2017). Measuring and monitoring land cover. In Urban Expansion, Land Cover and Soil Ecosystem Services. https://doi.org/10.4324/9781315715674-2

Nightingale, D. S., & Rossman, S. B. (2015). Collecting Data in the Field. In Handbook of Practical Program Evaluation. https://doi.org/10.1002/9781119171386.ch17

OBSERVER: How Copernicus helps implement sustainable finance and ensure sustainable growth | Copernicus. (2023). https://www.copernicus.eu/en/news/news/observer-how-copernicus-helps-implement-sustainable-finance-and-ensure-sustainable-growth

Pagourtzi, E., Assimakopoulos, V., Hatzichristos, T., & French, N. (2003). Real estate appraisal: a review of valuation methods. Journal of Property Investment & Finance, 21(4), 383–401. https://doi.org/10.1108/14635780310483656

Plazzi, A., Torous, W. N., & Valkanov, R. (2008). The Cross-Sectional Dispersion of commercial real estate returns and rent growth: time variation and economic fluctuations. Real Estate Economics, 36(3), 403–439. https://doi.org/10.1111/j.1540-6229.2008.00218.x

Qarallah, B., Othman, Y. A., Al-Ajlouni, M. G., Alheyari, H. A., & Qoqazeh, B. A. (2022). Assessment of Small-Extent forest fires in Semi-Arid Environment in Jordan using Sentinel-2 and Landsat Sensors data. Forests, 14(1), 41. https://doi.org/10.3390/f14010041

Reinermann, S., Gessner, U., Asam, S., Ullmann, T., Schucknecht, A., & Kuenzer, C. (2022). Detection of grassland mowing events for Germany by combining Sentinel-1 and Sentinel-2 time series. Remote Sensing, 14(7), 1647. https://doi.org/10.3390/rs14071647

Segarra, J., Buchaillot, M. L., Araus, J. L., & Kefauver, S. C. (2020). Remote Sensing for Precision Agriculture: Sentinel-2 improved features and applications. Agronomy, 10(5), 641. https://doi.org/10.3390/agronomy10050641

Sen2Cube.at. (n.d.). Accessed on 10/07/2023. https://demo.sen2cube.at/

Sen4Cap. (n.d.). Accessed on 10/07/2023. http://esa-sen4cap.org/

Sentinel-2 - Missions - Sentinel Online - Sentinel Online. (n.d.). Sentinel Online. https://sentinel.esa.int/web/sentinel/missions/sentinel-2

Sozzi, M., Kayad, A., Marinello, F., Taylor, J. A., & Tisseyre, B. (2020). Comparing vineyard imagery acquired from Sentinel-2 and Unmanned Aerial Vehicle (UAV) platform. OENO One, 54(2), 189–197. https://doi.org/10.20870/oeno-one.2020.54.1.2557

Tamm, T., Zalite, K., Voormansik, K., & Talgre, L. (2016). Relating Sentinel-1 interferometric coherence to mowing events on grasslands. Remote Sensing, 8(10), 802. https://doi.org/10.3390/rs8100802

Tiede, D., Sudmanns, M., Augustin, H., & Baraldi, A. (2021). Investigating ESA Sentinel-2 products' systematic cloud cover overestimation in very high altitude areas. Remote Sensing of Environment, 252, 112163. https://doi.org/10.1016/j.rse.2020.112163

Van Der Meer, L., Sudmanns, M., Augustin, H., Baraldi, A., & Tiede, D. (2022). SEMANTIC QUERYING IN EARTH OBSERVATION DATA CUBES. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XLVIII-4/W1-2022, 503–510. https://doi.org/10.5194/isprs-archives-xlviii-4-w1-2022-503-2022

Vélez, S., Barajas, E., Rubio, J. A., Vacas, R., & Poblete-Echeverría, C. (2020). Effect of Missing Vines on Total Leaf Area Determined by NDVI Calculated from Sentinel Satellite Data: Progressive Vine Removal Experiments. Applied Sciences, 10(10), 3612. https://doi.org/10.3390/app10103612

Wei, C., Fu, M., Wang, L., Yang, H., Tang, F., & Xiong, Y. (2022). The research development of Hedonic Price Model-Based Real Estate Appraisal in the era of Big Data. Land, 11(3), 334. https://doi.org/10.3390/land11030334

Weng, Q., Lu, D., & Schubring, J. (2004). Estimation of land surface temperature–vegetation abundance relationship for urban heat island studies. Remote Sensing of Environment, 89(4), 467–483. https://doi.org/10.1016/j.rse.2003.11.005

Wu, B., Yang, J., Zhao, Z., Meng, Y., Yue, A., Chen, J., He, D., Liu, X., & Liu, S. (2014). Parcel-Based change detection in Land-Use maps by adopting the holistic feature. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing. https://doi.org/10.1109/jstars.2013.2291773

Zhang, W., Li, W., Zhang, C., Hanink, D. M., Li, X., & Wang, W. (2017). Parcel-based urban land use classification in megacity using airborne LiDAR, high resolution orthoimagery, and Google Street View. Computers, Environment and Urban Systems, 64, 215–228. https://doi.org/10.1016/j.compenvurbsys.2017.03.001

Zhou, W., & Troy, A. (2008). An object-oriented approach for analysing and characterizing urban landscape at the parcel level. International Journal of Remote Sensing, 29(11), 3119–3135. https://doi.org/10.1080/01431160701469065