EO Clinic project:

Estimating the Magnitude and Spatial Distribution of Informal Trade in Central Asia

Work Order Report

Support requested by:
World Bank Group (WBG)

Reference: EOC0014_WOR_v01
Date: 2021 June 29

TABLE OF CONTENTS

ABOUT THIS DOCUMENT

This publication was prepared in the framework of the EO Clinic (Earth Observation Clinic, see below), in partnership between ESA (European Space Agency), the World Bank Group (WBG) and team of service providers contracted by ESA: e-GEOS S.p.A. (Italy) as Prime with support from GAF (Germany) and PLANETEK Italia (Italy).

This Work Order Report (WOR) describes the context of the team activities on Estimating the Magnitude and Spatial Distribution of Informal Trade in Central Asia, the geoinformation requirements of the activities and finally, the EO products and services delivered by the EO Clinic service providers in support of those activities.

ABOUT THE EO CLINIC

The EO Clinic (Earth Observation Clinic) is an ESA (European Space Agency) initiative to create a rapid-response mechanism for small-scale and exploratory uses of satellite EO information in support of a wide range of International Development projects and activities. The EO Clinic consists of "on-call" technically pre-qualified teams of EO service suppliers and satellite remote sensing experts in ESA member states. These teams are ready to demonstrate the utility of satellite data for the development sector, using their wide range of geospatial data skills and experience with a large variety of satellite data types.

The support teams are ready to meet the short delivery timescales often required by the development sector, targeting a maximum of 3 months from request to solution.

The EO Clinic is also an opportunity to explore more innovative EO products related to developing or improving methodologies for deriving socio-economic and environmental parameters and indicators.

The EO Clinic was launched in March 2019 and is open to support requests by key development banks and agencies during the 3 years project duration.

AUTHORS

The present document was prepared and coordinated by Valeria Donzelli (project manager, e-GEOS) with support from the following contributors: Isabella Bodechtel and Andreas Lippert (GAF), Nicolò Taggio (Planetek Italia).

ACKNOWLEDGEMENTS

The following colleagues provided valuable inputs, insights and evaluation feedback on the work performed: Cordula Rastogi (Sr Economist, Trade and Regional Integration, WBG), Daniel Saslavsky (WBG), and Zoltan Bartalis (ESA Coordinator and Technical Officer).

For further information	Please contact Valeria Donzelli, project manager, e-GEOS S.p.A. (valeria.donzelli@e-geos.it) with copy to Zoltan.Bartalis@esa.int if you have questions or comments with respect to content or if you wish to obtain permission for using the material in this report.

	Visit the ESA EO Clinic: https://eo4society.esa.int/eo_clinic.

1 DEVELOPMENT CONTEXT AND BACKGROUND

Policy making should be based on thorough knowledge of the respective field. In the case of economic decisions, usually statistics and databases provided by authorities are used as a basis. However, in emerging countries a good part of the economies can comprise large informal sectors, either in the field of manufacturing or in the field of commerce. Statistics collected from local surveys to cover the informal sectors suffer from incompleteness and dependencies from trustful local government structures and authorities.
Using a completely independent approach to sample commerce activities would help to provide statistics that are not influenced by willingness or trustfulness of small local authorities and offices.

These project results would be incorporated to an ongoing 2-year analytical project implemented by the World Bank (Central Asia: Regional Trade Connectivity Linkages – project number P171131) aimed at improving regional trade integration in Central Asia. The World Bank Group is supporting client countries in Central Asia with the objective of enhancing regional trade, investment and connectivity. Informal trade is known to provide a large amount of employment in these countries, but this economic activity is generally not recorded in official statistics.

Extracting indicators from EO-data is a way to provide information that is easy to crosscheck as it is derived from imagery data that is easily (albeit commercially) accessible. The idea behind is to define simple and robust indicators that can be derived from satellite imagery without being too volatile or too complicated to retrieve. A robust approach can cope with satellite imagery from different satellite data providers and sensors, different spatial and spectral resolution.
At the beginning, preliminary information have been exchanged with WBG references on the magnitude of informal trade in both countries as well as the distribution across sectors/ products.

2 PROPOSED WORK LOGIC FOR EO-BASED SOLUTIONS

We proposed to test how to use remote sensing methods to observe the peri-urban landscape of inland bazaars as a conduit to estimate current and past informal trade (non-standard or bazaar commerce). The RFP listed four markets/bazaars located in four different central Asian countries and suggested conducting a study of at least one of them:
1. Dordoi Bazaar (Bishkek, Kyrgyz Republic)
2. Barakholka Bazaar (Almaty, Kazakhstan)
3. Abu Sahiy Bazaar (Tashkent, Uzbekistan)
4. Korvon Bazaar (Dushanbe, Tajikistan).

We proposed to use two of the bazaars in the study:
1. Dordoi Bazaar (Bishkek, Kyrgyz Republic)
2. Barakholka Bazaar (Almaty, Kazakhstan)

We had already suggested both bazaars in the very first version of our proposal before any change in the list with suggestions. The two suggested bazaars consist of over 40,000 (Bishkek) and 15,000 outlets (Almaty). They are the largest bazaars in the region.
Both bazaars had on freely available imagery the exact appearance of a group of dynamic and well frequented bazaars. They consist of warehouses, storage spaces, parking facilities and many sub-bazaars consisting of makeshift buildings, apparently constructed from stacked containers and with the alleys between the containers being roofed over with sheet metal constructions. Both bazaars awere apparently growing and intermingled already with surrounding residential areas.
Both bazaars are well recognizable in satellite imagery and are clearly identifiable as centres of commerce. The study concentrates on two markets to make efficient use of EO data.

The data extracted from the EO data besides the economic and other data are the input the a predictive analysis which try to estimate the amount of informal trades. Predictive analytics is a branch of advanced analytics and data mining. These methodologies are used in predicting future events analyzing current and historical data. To make predictions, these methods employ techniques from statistics, data mining, machine learning, and artificial intelligence. With these approaches it is possible to combine heterogeneous information to forecast future events.

The project is structured in 4 tasks or Work Packages (WP):

- WP1 covers all the management activities and lasts for the whole project duration.
- WP2A covers all activities regarding data procurement (WP is integrated into Service 1).
- WP2B generates Service 1: Past and Present Analysis of Markets and Bazaars.
- WP3 generates Service 2 Predictive Analysis.

3 DELIVERED EO-BASED PRODUCTS AND SERVICES

3.1 Work logic and results

WP2A has been integrated into WP2B of Service 1. So the research about the markets, their development and about any major events could be used for determining best image acquisition dates as well as background information for geospatial analysts.

During the phase of Service 1 preliminary data was provided to service 2 to prepare service 2 for the incoming data. Moreover, in this phase, it has been performend an extensive search of ancillary data to be used in the predictive analysis.

3.1.1 Service 1: Past and Present Analysis of Markets and Bazaars

3.1.1.1 Overview

The general approach is using satellite imagery and visual image interpretation to detect, analyse, measure and document certain bazaar-related features to extract indicators that are accessible to earth observation methodology and represent bazaar activity.

Prior to ordering satellite imagery, an open source analysis was conducted to collect ancillary information about the recent operations and the development of the relevant bazaars. The open source analysis helped in determining meaningful observation dates for the geospatial analysis and helped in understanding the overall functioning of the bazaars (important for geospatial analysts to interpret detected features and objects).

Satellite data was ordered with GAF-internal licences to save cost. One satellite scene for each bazaar will receive a licence uplift for usage in reports, display or further discussions. The additional licence is issued to ESA and to WBG.

All satellite scenes per bazaar underwent a bundle block adjustment process to be aligned geometrically. This process is not necessary for the sole purpose of indicator extraction (as the indicators are expressed in m^2 thematic surface area, the exact location is not relevant to the indicator). However, it simplifies the geospatial analysis process and supports any QC processes.

With background knowledge from ancillary data, the geospatial analysts check all satellite images, starting with the most recent and going backwards in the time series. Features and objects identified as indicators were delineated and GIS data was collected in one geodatabase per bazaar. Features and indicators were assigned the respective date/time codes, depending from which image they were extracted.

A 100% quality control was performed by an independent, experienced senior geospatial analyst and QC expert. Focus of QC was completeness, thematic and geometric accuracy. Semantic relationship was checked to the necessary level, as no indicator feature should lay outside the geometry of indicator “Total Bazaar Area”.

After QC, surface area and number of all features per market and observation date was calculated. The results are presented as tables and graphs in an Excel file. The geometries of all features are available for usage in maps.

3.1.1.2 Pre-Check on Bazaars

Dordoi Bazaar, Bishkek

- In fact different bazaars in one, mainly consumer goods, mainly Chinese, Turkish and Russian products

- Part of the "modern silk road"
- Re-Export is part of the operation
- Is regarded well organized under the leadership of commercial operators
- Started operation in the nineties, consisted of ~6 to 7,000 containers in 2005
- Open daily 05:00 – 16:00 local time, closed during new year holidays. New year holidays and public holidays were omitted in satellite data aquisition
- Events
    - 2015 Kirgizstan member of EEU (Eurasian Economic Union). Affected relationship to China, Re-export activities were at danger, closing was imminent
    - 2019 Bishkek was declared part of the Chinese "Belt and Road Initiative" and is part of the routes China-Kirgizstan-Uzbekistan and China-Kirgizstan-Iran/West Asia
    - Fires 2013
    - 2020 partly closed due to Covid-19. Albeit partly open, activities were affected by closed borders towards China

Preliminary discoveries: changes in bazaar appearance do not impact very much upon surface area. The decision was made to reduce observation periodicity for this bazaar to once per year. Effort was redirected to a more detailed indicator extraction regarding the building types. The indicator "Total roofed over Area" was divided into two separate indicators: "Container Building (including their support structures)" and "Fixed Buildings" for both bazaars.

Barakholka Bazaar, Almaty

- Different bazaars
- Operation is affected by proximity to China
- Goods from China come to Barakholka and are further transferred to Russia, Turkey and Korea.
- Open Tuesday to Sunday 10:00 – 17:00, closed Monday. Monday was subsequently omitted as day of satellite acquisitions for this bazaar (additional to public holidays).
- Events
    - Was from 2011 on subject of an "Initiative to eliminate chaotic circumstances in bazaars"
    - From 2014 on disassembling of container buildings and construction of fixed buildings and malls.
    - Partly closed in 2020

Extraction of two indicators regarding buildings (container buildings and fixed buildings) was also applied to this bazaar.

3.1.1.3 Acquisition of satellite imagery

Following VHR satellite images between 2012 and 2020 from spring/summer and autumn season were used for both bazaars to have the ability to detect extra activities like construction sites as these are expected to happen in summer instead of winter season. This reduction to the summer season helps also to avoid any seasonal effects in the derived indicators.

Table 1: VHR satellite images used for Barakholka, Almaty

Sensor	Acquisition Date	ID
WV2	11.04.2012 (Wednesday)	10300100123CF900
Pleiades	14.03.2013 (Thursday)	DS_PHR1B_201303140553185_SE1_PX_E076N43_1108_03978
Pleiades	03.05.2014 (Saturday)	DS_PHR1A_201405030600341_SE1_PX_E076N43_1117_09738
Pleiades	01.11.2014 (Saturday)	DS_PHR1A_201411010600506_SE1_PX_E076N43_1007_03223
Pleiades	28.05.2015 (Thursday)	DS_PHR1A_201505280600393_FR1_PX_E076N43_1008_03316

WV2	30.06.2015 (Tuesday)	10300100440D9100
WV2	03.09.2016 (Saturday)	103001005B0D6A00
Pleiades	28.03.2017 (Tuesday)	DS_PHR1B_201703280556536_FR1_PX_E076N43_1108_03466
Pleiades	08.10.2017 (Sunday)	DS_PHR1B_201710080604310_FR1_PX_E076N43_1207_02030
Pleiades	22.04.2018 (Sunday)	DS_PHR1B_201804220556023_FR1_PX_E076N43_1107_05280
WV2	01.08.2019 (Thursday)	103001009601F300
WV2	28.04.2020 (Tuesday)	10300100A583E000

Table 2: VHR satellite images used for Dordoi, Bishkek

Sensor	Acquisition Date	ID
WV2	30.04.2012 (Monday)	10300100181C0900
WV2	23.05.2013 (Thursday)	1030010022237B00
WV2	18.07.2014 (Friday)	1030010032087300
Pleiades	19.04.2015 (Sunday)	DS_PHR1B_201504190600144_FR1_PX_E074N42_0922_02646
WV2	24.03.2016 (Thursday)	1030010054284800
Pleiades	10.04.2017 (Monday)	DS_PHR1A_201704100556545_FR1_PX_E074N42_0722_02560
Pleiades	21.04.2018 (Saturday)	DS_PHR1A_201804210603441_FR1_PX_E074N43_0701_08226
Pleiades	24.06.2019 (Monday)	DS_PHR1B_201906240604281_FR1_PX_E074N43_0704_05036
GeoEye	25.09.2020 (Friday)	105001001FD85600

3.1.1.4 Geospatial Analysis Phase

On the strength of past experience we developed the following workflow for the feature extraction in the different mapping phases between 2012 and 2020:

Having the basis on the most recent situation in 2020, the progress and change mapping was performed on the historic satellite datasets. Both bazaars were analysed by an experienced interpreter with regard to changes compared to the dataset one year before.

The following indicators were extracted for each satellite scene / observation date:

- Construction Sites: clearly visible, larger construction sites (construction activity must be visible, no demolition activity, no small maintenance activities). Featureclass name in geodatabases: ConstructionSiteA
- Vehicle Lots: only permanent, well-marked parking lots, no roadside parking, not backyard parking, no car dealerships, not workshop or junkyard parking. Featureclass name in geodatabases: VehicleLotA
- Fixed Buildings: permanent, “classical”, buildings. Featureclass name in geodatabases: BuildingA
- Container Buildings: Buildings made from Containers: distinct containers including the covered aisles. Featureclass name in geodatabases: ContainerBuildingA
- Total Bazaar Area: Overall bazaar areas, including the surface area of the above mentioned indicators and including roads, open spaces, workshops. The indicator is shown in maps as “Other bazaar areas” as it fills the spaces between the other indicators. Featureclass name in geodatabases: BazaarA
- The indicator “Total Roofed Over Area” is the sum of Fixed Buldings plus Container Buildings.

All indicators are reported as areas in square meters per timestamp. The exact timestamp comes from the date at which a satellite image was acquired.

Resulting data are numeric values per indicator, observation date and bazaar (surface area in square meters and number of features (frequency)) and GIS data with the delineations of all indicator observations, correctly and completely attributed.

3.1.1.5 Quality Control and Validation

The technical production workflow contains two independent quality control steps (QC1 and QC2) for the mapped vector geometries:

QC1: Semi-automatic/automatic quality checks (QC1), performed by one geospatial analyst for each bazaar.

These are final automatic checks of the mapping result concerning the minimum object sizes, correct data attribution, overlap and gaps checks, multi-part polygons, etc. Further on, plausibility checks and database queries were performed. Homogeneity of analysis results were maintained by frequent crosschecks between analysts of both bazaars.

QC2: A 100% quality control on the extracted features and GIS geometries in the geodatabases was performed by an independent, experienced senior geospatial analyst and QC expert. QC was done against imagery. Focus of QC was completeness, thematic and geometric accuracy.

Calculation of surface areas was crosschecked by another operator. The overall results were checked for plausibility and consistence with prior knowledge and results of open source analysis by the project managers (FAT Factory acceptance test).

3.1.1.6 Intermediate Results and Input into Service 2

Barakholka Bazaar Results

Two satellite scenes per year were available for the years 2014, 2015 and 2017 in Barakholka. The in-year observations show changes in the indicator that are consistent with the overall trend. While the increase of the total Area (BazaarA) has abated, the built-up part of the bazaar (TotalRoofedA, Totally roofed over area) is still growing. A qualitative change is visible in the fact that Container Buildings are reduced while Fixed Building are increasing, accompanied by expanded Vehicle Lot Areas when compared with 2012/13. Construction activity has a maximum in 2017.

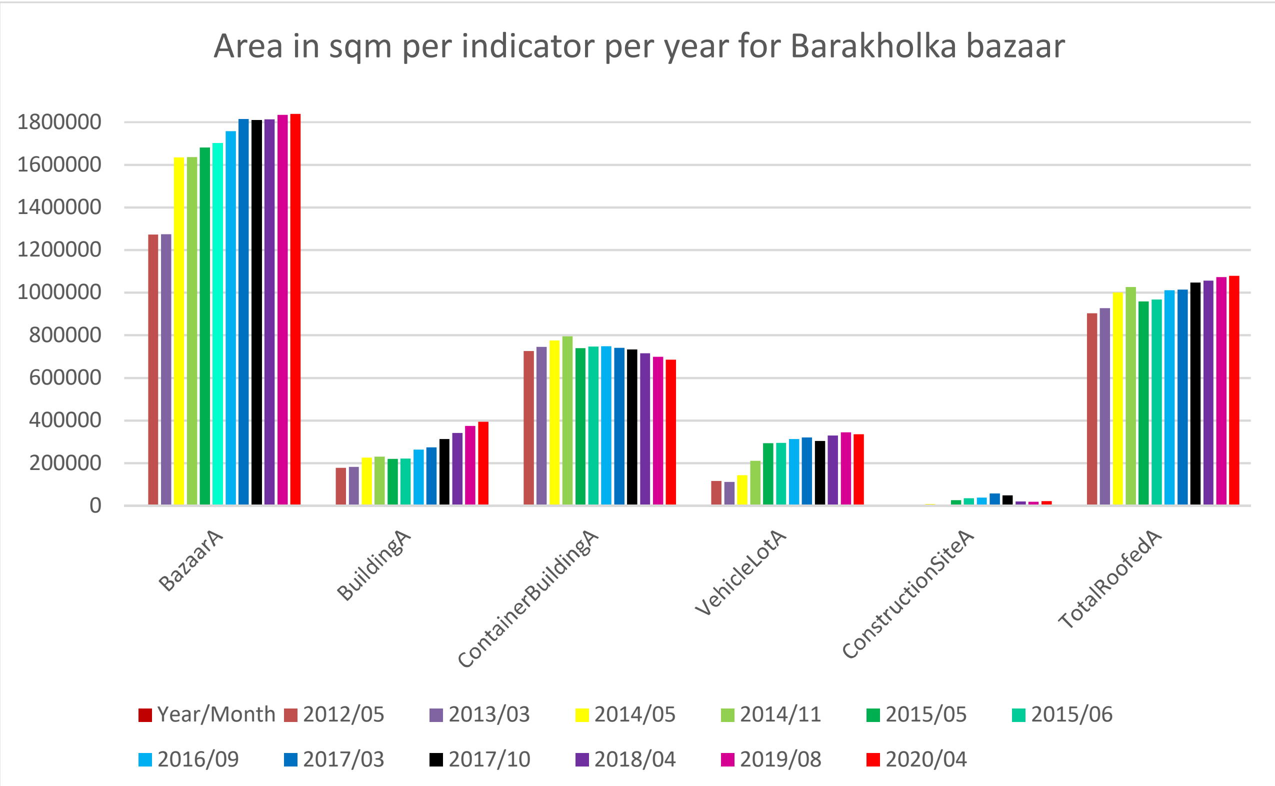

Figure 1: Area per indicator between 2012 and 2020 for Barakholka bazaar

Table 3: Surface areas indicated in square meters for Barakholka bazaar. Frequency is the number of mapped features per indicator..

	Year/Month	2012/05	2013/03	2014/05	2014/11	2015/05	2015/06	2016/09	2017/03	2017/10	2018/04	2019/08	2020/04
BazaarA	SUM_Area	1272676	1273965	1634253	1635944	1680348	1702315	1757227	1814759	1810403	1813664	1833373	1837842
BuildingA	SUM_Area	177616	182287	225102	230586	219324	221239	262789	273309	312973	340592	373785	393431
ContainerBuildingA	SUM_Area	724890	744529	775056	794801	738505	746007	747572	740739	733316	715093	698331	684601
VehicleLotA	SUM_Area	116516	111827	142796	210162	292879	294178	313188	320742	303392	328326	344675	335610
ConstructionSiteA	SUM_Area	0	0	7793	2886	25288	34347	37843	56678	48728	19724	17864	21075
TotalRoofedA	SUM_Area	902506	926815	1000158	1025387	957829	967246	1010361	1014047	1046289	1055685	1072115	1078032

	Year/Month	2012/05	2013/03	2014/05	2014/11	2015/05	2015/06	2016/09	2017/03	2017/10	2018/04	2019/08	2020/04
BazaarA	FREQUENCY	12	13	13	13	13	13	16	16	16	16	16	16
BuildingA	FREQUENCY	84	86	108	113	127	123	143	144	153	163	169	184
ContainerBuildingA	FREQUENCY	129	118	141	124	145	120	130	124	124	119	109	111
VehicleLotA	FREQUENCY	24	20	22	30	29	27	35	35	31	33	34	34
ConstructionSiteA	FREQUENCY	0	0	2	2	4	7	7	9	7	2	5	4

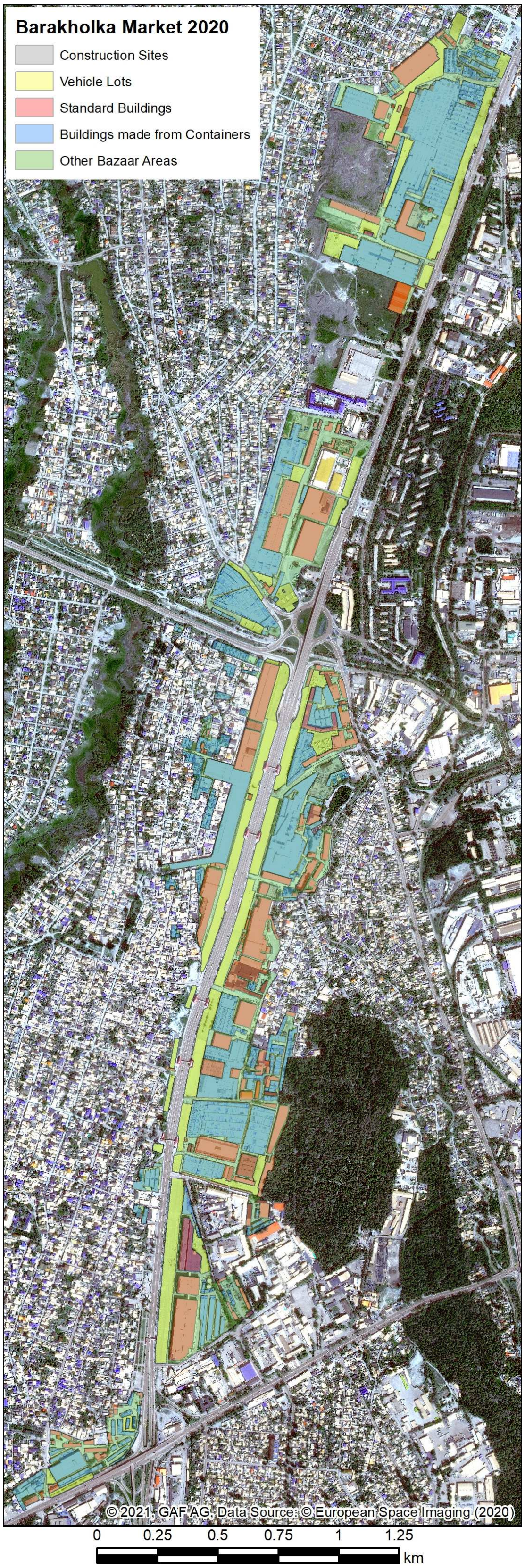

Figure 2: Barakholka bazaar in 2020

Barakholka bazaar in its recent appearance:

The bazaar is located along a motorway-like road. As it consists in fact of many separate bazaars for different goods and different clients it comprises several separated areas.

Formal parking lots with markings have been established along the road in the southern and northernmost part. They have evolved from more informal multi-purpose open spaces and container buildings.

Although a transformation towards fixed buildings occurred througout the previous years, the surface area occupied by container buildings is still high.

The center section, located north and northwest to the roundabout has developed from manufacturing plants into bazaar area.

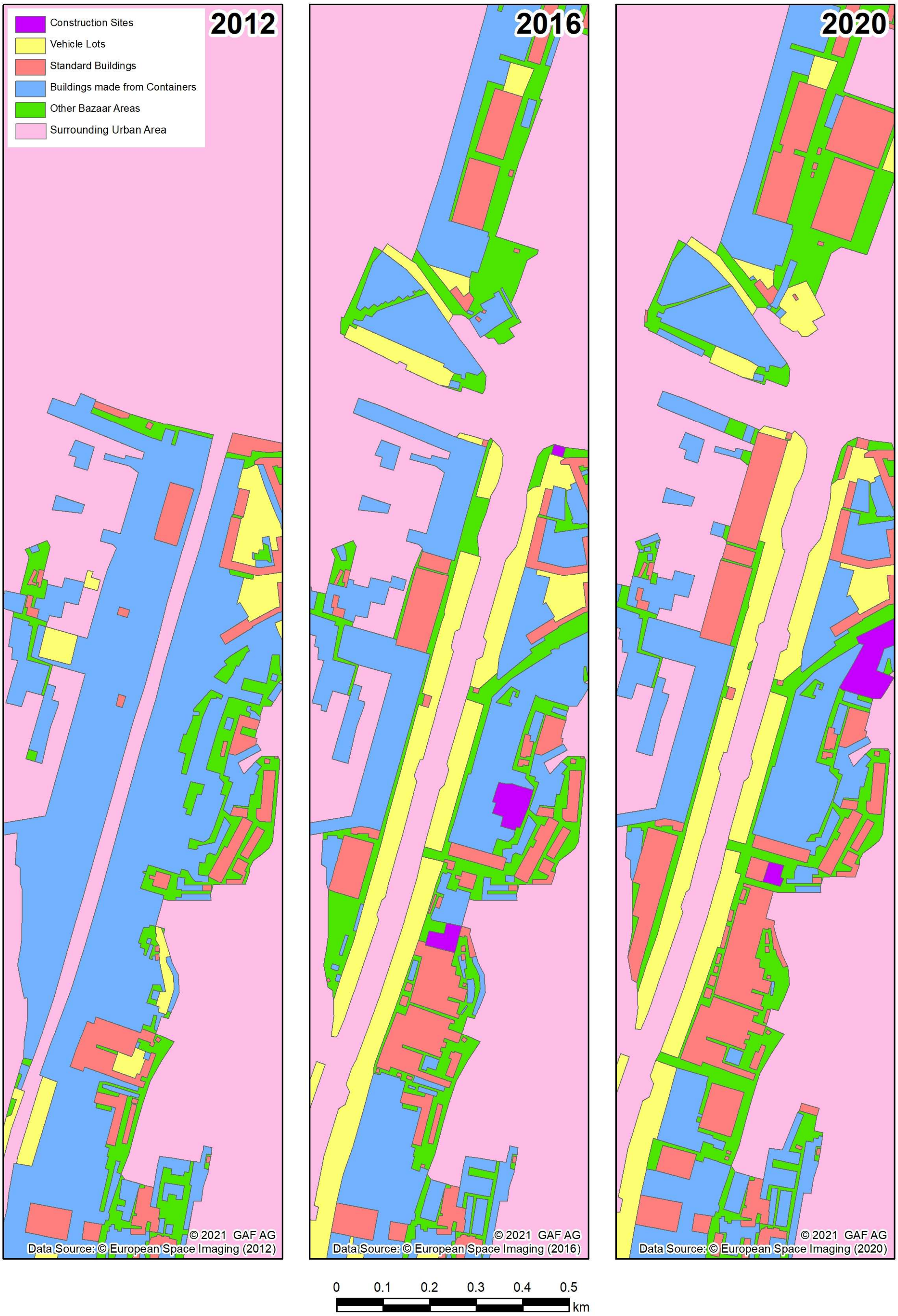

Figure 3: Center of Barakholka bazaar in 2012, 2016 and 2020

Dordoi Bazaar Results

For Dordoi one satellite scene was available for all years 2012 - 2020. As changes are small in surface area, this rate of observations is fully sufficient. The growth of the total bazaar area is relatively small. However, the growth is on top of a very high level already extant in 2012. Acquisition of new land for the bazaar to the north is visible in 2019. The nearly-constant growth of total roofed over area until ca. 2016 is coming from larger Container buildings. After 2017 growth is coming from fixed buildings, also showing a qualitative change. Slowly growing total area with medium growing building area results in a loss in Vehicle Parking Lot area. No larger construction activity could be observed.

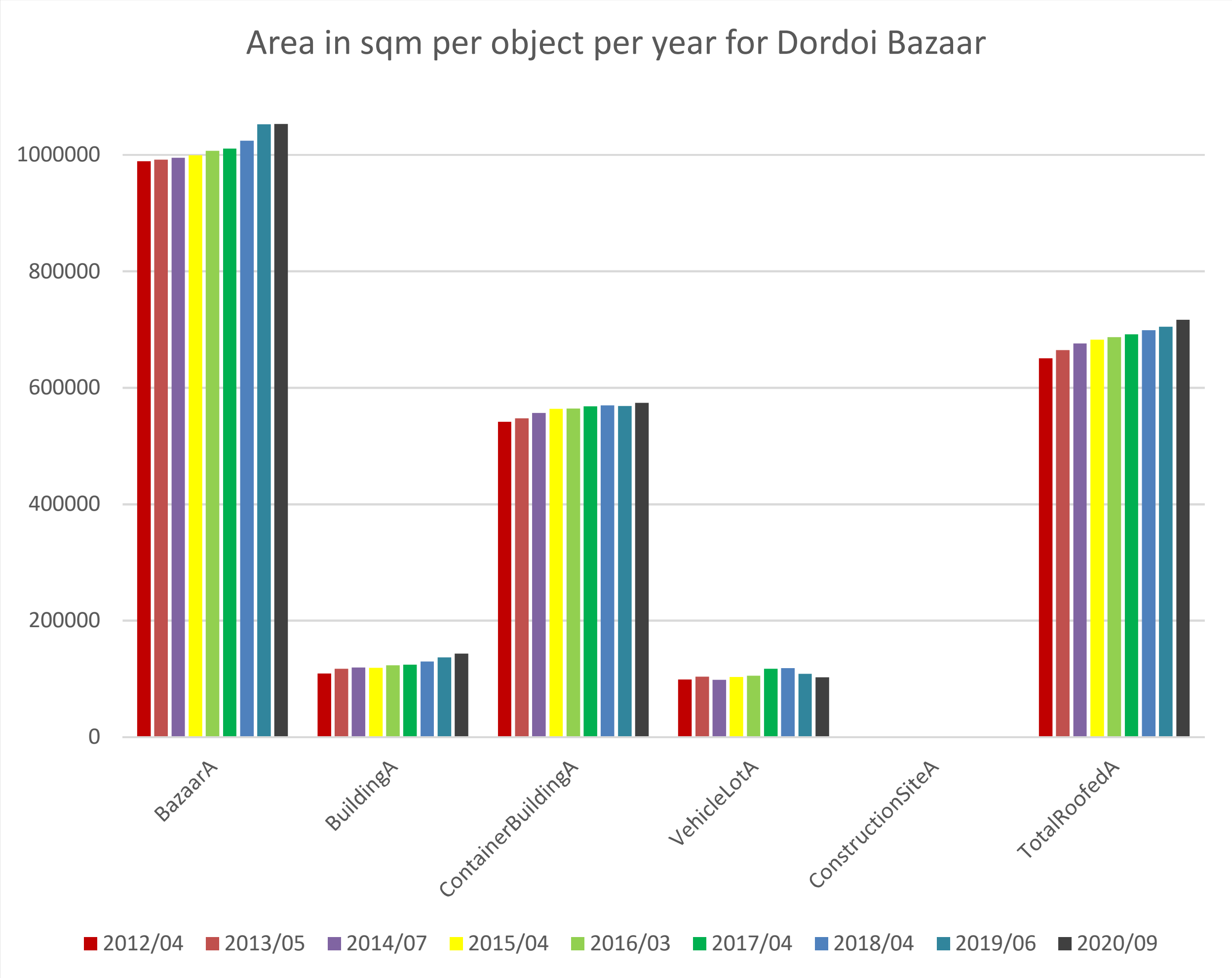

Figure 4: Area per indicator between 2012 and 2020 for Dordoi bazaar

Table 4: Surface areas indicated in square meters for Dordoi Bazaar. Frequency is the number of mapped features per indicator.

m²	**Year_Month**	**2012/04**	**2013/05**	**2014/07**	**2015/04**	**2016/03**	**2017/04**	**2018/04**	**2019/06**	**2020/09**
BazaarA	SUM_Area	988841	991496	995099	999242	1007140	1010656	1024218	1052452	1053029
BuildingA	SUM_Area	108782	117285	119453	118962	122958	124151	129650	136366	143088
ContainerBuildingA	SUM_Area	541530	547499	556374	563450	563984	567718	569333	568248	573700
VehicleLotA	SUM_Area	98802	103716	97988	103181	105180	116903	117951	108146	102520
ConstructionSiteA	SUM_Area	0	0	0	0	0	0	0	0	0
TotalRoofedA	SUM_Area	650313	664785	675827	682412	686942	691869	698982	704614	716788

Number of features	**Year_Month**	**2012/04**	**2013/05**	**2014/07**	**2015/04**	**2016/03**	**2017/04**	**2018/04**	**2019/06**	**2020/09**
BazaarA	FREQUENCY	1	1	1	1	1	1	4	4	4
BuildingA	FREQUENCY	116	116	121	121	127	129	140	150	157
ContainerBuildingA	FREQUENCY	87	80	87	88	87	87	93	119	129
ConstructionSiteA	FREQUENCY	0	0	0	0	0	0	0	0	0
VehicleLotA	FREQUENCY	28	29	29	30	31	31	32	30	29

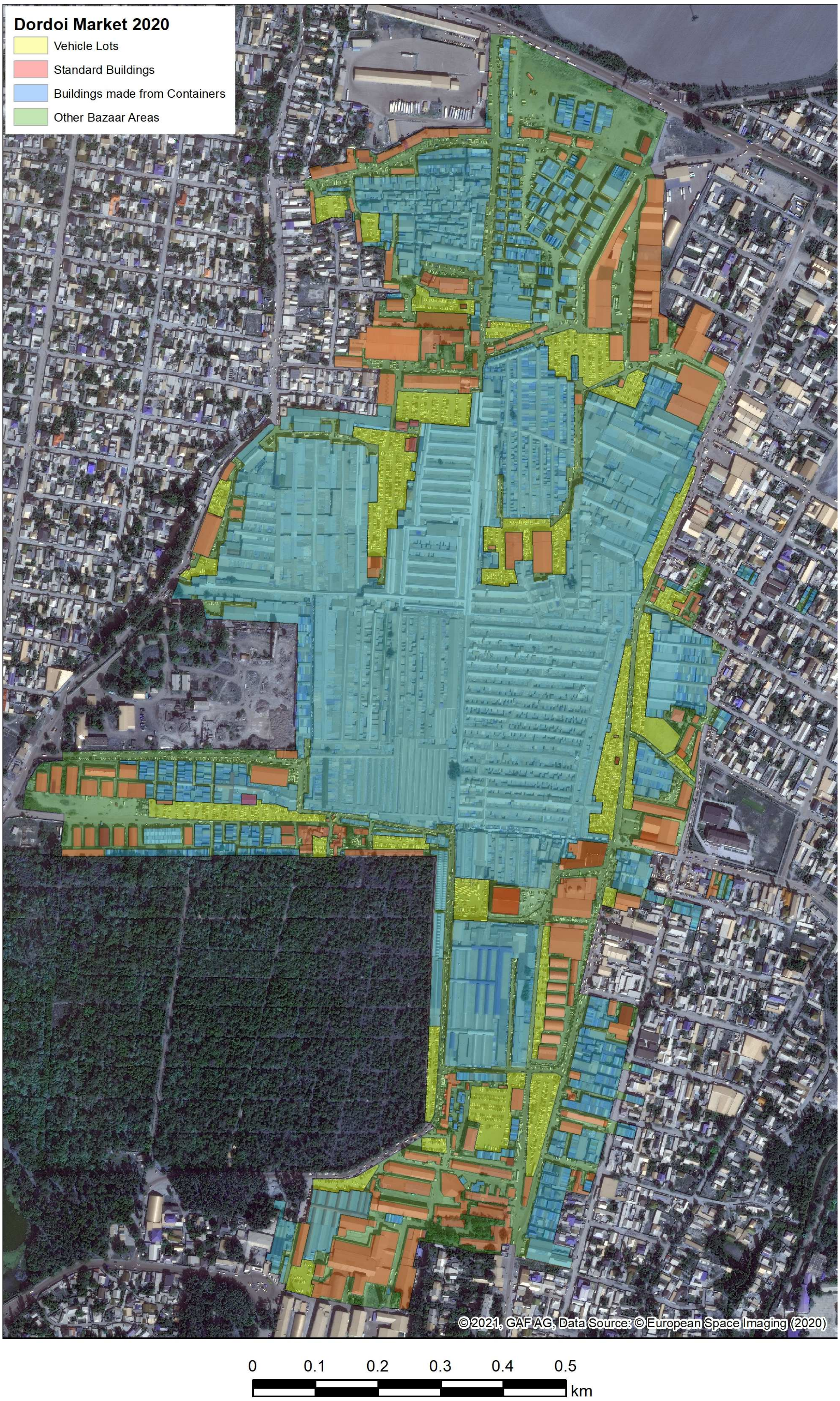

Figure 5: Dordoi bazaar in 2020

The Dordoi bazaar, founded on the site of a former manufacturing plant, is squeezed into its neigbourhoods, consisting of a large cemetery, residential and mixed areas. With hardly any space to grow it appears as a near-monolithic landscape of malls made from stacked containers. The as yet not fully occupied open space to the north is the newest extension area, aquired in ca. 2017. Dordoi has outliers in the backyards of its eastern neighbourhood, used as additional storage space and also made from shipping containers.

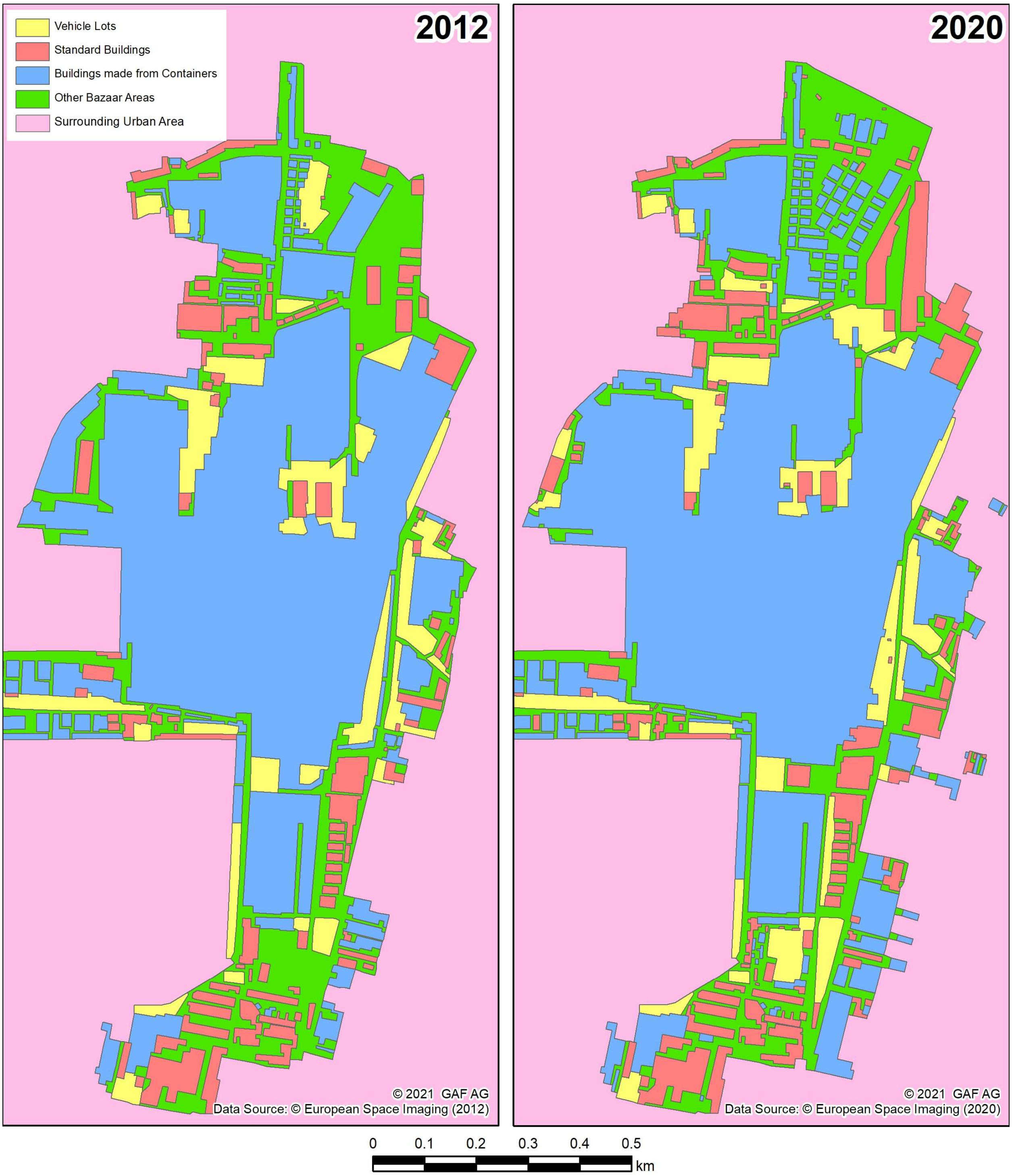

Figure 6: Dordoi bazaar in 2012 and 2020

3.1.1.7 Usage, Limitations and Constraints

Description of indicators, sqm- and feature count numbers of indicators and geometric delineations of indicators (vector data) can be freely used on behalf of ESA. However, no commercial use is possible. The satellite image data is licensed to ESA and WBG. The EULAs apply.

The detailed geospatial analysis of two bazaars in Central Asia indicated a remarkably constant change and development over the years from 2012 until 2020. The change in appearance is clearly indicated by a trend towards a more formal landuse, represented by a transition from informal container buildings towards fixed buildings and additionally in the case of Barakholka in the creation of formal parking space along its main axis, road A-350.

The type of growth observed in Barakholka and Dordoi suggests that a lower number of satellite images should be sufficient to estimate the development of a certain bazaar.

3.1.2 Service 2: Predictive Analysis

3.1.2.1 Specifications

The predictive analysis is performed using machine learning techniques on input data composed by EO and not-EO data. The main problem encountered is the low availability of not-EO data and their low quality, as described in the following section. This is an issue because, in general, good quality input data are required to have a good predictive capability from any AI approach. Moreover, some data are quite old, and they do not match the date of the information derived from EO. In order to overcome this difficulty, it has been decided to assume, the data, at least their relative order of magnitude, are stationary in time. For example, Table 6 and Table 7 report some characteristics and revenues of the main bazaars in Kazakhstan and the Kyrgyz Republic, including the Dordoi and Barakholka ones. In the stationary hypothesis, the relative importance of the latter two bazaars in their countries related to the total of the bazaars in their respective countries can be considered constant in time. Therefore, it is possible to use ratios derived from Table 6 and Table 7 in conjunction with more recent countrywide data.

Two scenarios have been set up to test two different approaches:

<u>Scenario 1</u>: Dordoi Bazaar (Bishkek, Kyrgyz Republic). For this bazaar are used all EO-derived data and the most important economic data.

<u>Scenario 2</u>: Barakholka Bazaar (Almaty, Kazakhstan). For the analysis of this bazaar, besides the data used in Scenario 1, other economic and transportation data are added. This scenario aims to verify the usefulness of additional general economic data and information about mobility at the national level in the predictive analysis.

In the following section, the non-EO input data found are listed, while in section 3.1.2.1.2 are described the methodologies used in the two scenarios.

3.1.2.1.1 Input data analysis

The EO-derived data input for Service 2 are described in section 3.1.1. The present section reports the data survey results of not-EO data performed to feed the AI algorithms properly. It has been found that, in general, the availability of data is low and the ones available are not updated.

In the following tables are reported the data found that are used in the predictive analysis.

Table 5: Unique features of "bazaar" imports into Central Asia in 2006 (Own calculations based on world reporting to the UN COMTRADE database and World Bank's World Development Indicators database.)

	Kazakhstan	Kyrgyz Republic
Share of bazaar goods 'in total imports (in %)	20.1	54.7
share of a country in CA 'bazaar' imports (in %)	65.3	26.7
Imports of 'bazaar' goods per capita (in US dollars)	318.0	385.8
Imports of 'bazaar' goods in % of GNI	6.3	65.4
Share of a country in total CA imports (in %)	72.4	10.9
Share of a country in total non-bazaar 'CA imports	74.4	6.4
Memorandum: Imports in % of GNI Atlas method	31.2	119.6

Table 6: Some characteristics of the principal bazaars of Kazakhstan and the Kyrgist Republic (source: "Bazaars and Trade Integration in CAREC Countries", Report prepared by the World Bank May 13, 2009)

bazaar	country	Total number of sales outlets	Total breakeven sales(monthly) (in millions USD)	Total estimated sales(monthly) (in millions USD)	Employment (in '000)	Share in local employment (%)
Barakholka (Almaty)	KAZ	15450	176	211	39.9	5.1
Dordoi (Bishkek)	KGZ	40300	301	311	54.6	7.8
Karasuu (Osh)	KGZ	10200	58	94	16.3	3.0
Altyn Orda (Almaty)	KAZ	3181	7.7	14.3	8.2	1.3
Madina (Bishkek)	KGZ	1030	2.9	24.2	1.8	0.4
Shanghai (Astana)	KAZ	3127	9.4	22.9	4.8	1.7
Artem (Astana)	KAZ	1750	12.3	15.1	2.6	0.9
Karkara (Almaty)	KAZ	624	3.1	4.4	0.9	0.2
Sary-Arka (Almaty)	KAZ	282	1.2	2.2	0.4	0.1

Table 7: estimates of annualized revenues and expenditures of bazaars " and sales outlets" owners and fees paid in surveyed bazaars in 2008 in millions of US dollars (source: "Bazaars and Trade Integration in CAREC Countries", Report prepared by the World Bank May 13, 2009)

bazaar	country	Total fixed cost	labor related outlays	lease or lease equivalent	official bazaar fees	informal bazaar fees
Barakholka (Almaty)	KAZ	314.7	57.0	66.5	49.8	141.4
Dordoi (Bishkek)	KGZ	855.7	252.6	540.4	62.7	0.0
Karasuu (Osh)	KGZ	123.9	44.4	35.2	31.7	44.3
Altyn Orda (Almaty)	KAZ	29.5	19.3	5.8	4.4	0.0
Madina (Bishkek)	KGZ	7.4	3.4	2.0	1.3	0.6
Shanghai (Astana)	KAZ	31.5	23.7	7.6	0.2	0.0
Artem (Astana)	KAZ	54.2	16.8	21.7	15.8	0.0
Karkara (Almaty)	KAZ	10.1	4.7	5.4	0.0	0.0
Sary-Arka_A (Almaty)	KAZ	2.6	1.4	0.6	0.6	0.0

Total	Total fixed cost	Labor related outlays	lease or lease equivalent	official bazaar fees	informal bazaar fees
Kazakhstan	442	123	107	71	141
Kyrgyzstan	987	300	578	96	45

WBG provided more recent information about the gap between exports of consumer goods reported by China and the rest of the world, and imports reported by Kazakhstan and the Kyrgyz Republic economies. This gap can be related to the informal trades and can give an estimation of the bazaar's economy. The data refers to the years 2015 till 2019 included, and they are classified as "confidential", therefore not reported here.

Another type of data explored is related to the mobility of the citizens in the two countries. The most suitable data for service 2 are probably the numbers of cars that pass through the roads around the bazaars. Unfortunately, such data are not available; however, for Kazakhstan, it has been found some data related to citizen mobility, listed in the following tables and used in scenario 2:

Table 8: The number of millions of Vehicle-kilometres per year in Kazakhstan (source: https://knoema.com)

Date	Total road motor vehicle traffic of Kazakhstan in millions of vehicle-kilometers
2017	1676.00
2016	1161.00

2015	1137.00
2014	665.00
2013	648.53

Table 9: The number of Commercial Vehicles per years in Kazakhstan (source: https://knoema.com)

Date	Kazakhstan commercial vehicles per years
2020	1,068.00
2019	868.00
2018	5417.00
2017	4925.00
2016	7351.00
2015	15969.00
2014	12137.00
2013	12046.00
2012	11272.00
2011	5000.00
2010	5000.00
2009	5000.00

Table 10: Length of paved roads of regional importance in Kazakhstan, km (source: Transport in Kazakhstan 2013-2017. Statistical collection. Astana, 2018. http://stat.gov.kz)

	2013	*2014*	*2015*	*2016*	*2017*
The Republic of Kazakhstan	29133	32694	32620	32242	31995
Akmola region	2649	2649	2665	2661	2661
Aktobe region	1099	1099	1048	1061	1048
Almata region	3321	6670	6674	6675	6138
Atyrau region	691	691	690	691	691
West-Kazakhstan region	1580	1581	1493	1493	1931
Jambyl region	2241	2210	2210	1943	1944
Karaganda region	3458	3458	3458	3458	3458
Kostanay region	2077	2077	2077	2077	2009
Kyzylorda region	268	470	471	466	466
Mangystau region	883	883	883	831	831
Turkestan region	4291	4330	4315	4291	4222
Pavlodar region	1173	1173	1173	1133	1133
North-Kazakhstan region	2356	2356	2416	2416	2416
East Kazakhstan region	3046	3047	3047	3046	3047

Table 11: Length of paved roads of local importance in Kazakhstan, km (source: Transport in Kazakhstan 2013-2017. Statistical collection. Astana, 2018. http://stat.gov.kz)

	2013	*2014*	*2015*	*2016*	*2017*

The Republic of Kazakhstan	42851	38780	38946	38936	38567
Akmola region	2965	2965	2965	2965	2956
Aktobe region	3602	3602	3602	3995	3602
Almata region	3589	-	-	-	-
Atyrau region	1089	1089	1089	1089	1089
West-Kazakhstan region	3383	3383	3383	3380	3380
Jambyl region	1918	1904	1904	2048	1913
Karaganda region	2522	2522	2522	2522	2522
Kostanay region	5898	5671	5671	5671	5671
Kyzylorda region	1976	1769	1770	1797	1867
Mangystau region	541	541	541	717	717
Turkestan region	2065	2031	2081	1632	1643
Pavlodar region	2964	2964	2965	2791	2803
North-Kazakhstan region	5103	5103	5043	5043	5043
East Kazakhstan region	5236	5236	5411	5286	5361

Table 12: The principal macroeconomic indicators of Kazakhstan (source: LOGISTICS AND TRANSPORT COMPETITIVENESS IN KAZAKHSTAN, UN 2019, ISBN: 978-92-1-117205-8)

Indicator	2010	2011	2012	2013	2014	2015	2016	2017
GDP, mln. US$	148052.4	192627.6	208002.1	236633.3	221417.7	184387.0	137278.3	158180.3
As a percentage of the previous year	107.3	107.4	104.8	106.0	104.2	101.2	101.1	104.0
GDP per capita, US$	9071.0	11634.5	12387.4	13890.8	12806.7	10509.9	7714.8	8769.5
Population, mln. people	16203.0	16440.1	16673.1	16909.8	17160.8	17417.7	17670.6	18157.1
The number of permanent residents, as a percentage of the previous year	101.4	101.5	101.4	101.4	101.5	101.5	101.5	101.3
Average nominal monetary incomes, US$	264.8	313.2	347.8	371.1	347.5	303.6	223.8	246.1
Volume of industrial output, bln. tenge	12105.5	15929.0	16851.8	17834.0	18529.2	14903.1	919026.8	122659.0
The agricultural products, bln, tenge	1 822.1	2720.4	2393.6	2949.5	3143.7	3307.0	3684.4	4097.4
Capital Investment, mln. US$	31581.5	34171.5	36953.3	33293.2	36784.9	31681.4	22686.2	26838.4
Food price index	110.1	109.1	105.3	103.3	108.0	110.9	109.7	106.5

Price index for non-food products	105.5	105.3	103.5	103.3	107.8	122.6	109.5	108.9
Foreign trade turnover, mln. US$, including	91397.5	121241.7	132807.2	133506.0	120755.3	76523.5	62113.6	78102.9
export, mln. US$	60270.8	84335.9	86448.8	84700.4	79459.8	45955.8	36736.9	48503.3
import, mln. US$	31126.7	36905.8	46358.4	48805.6	41295.5	30567.7	25376.7	29599.6

3.1.2.1.2 Methodologies

The predictive analysis contains a vast set of methods and algorithms and falls within different contexts. In fact, it includes statistical approaches, data mining algorithms, and also machine learning tools. In particular, machine learning methodologies are able to combine different heterogeneous data to perform tasks like classification or regression to forecast future events. Among all these methodologies, artificial neural networks, decision trees, support vector machines are the tools most adopted in predictive analytics tasks. In this context, they made extensive use of probability theory and regression analysis. The basic idea of the regression methods is to use input data to build a model to perform forecasts with a detailed level of granularity. Furthermore, using artificial intelligence methodologies is possible to build a non-linear connection between data. In other words, these methods can find complex links (non-linear functions) using input data. In the literature, it is possible to distinguish between two main approaches, linear-methods (i.e., linear regression) and non-linear-methods (i.e., Random Forest Regression).

The main idea for the analysis of the two bazaars in Khazakstan and the Kyrgyz Republic is to use indicators coming from EO-data and other ancillary information to estimate informal trade. Due to the multiple information to be used to predict future events, the regression method used is called Multiple Regression (MR) and, when it is linear (MLR), it can be expressed by the following equation:

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_n x_{in} + \varepsilon$$

Where *i=1,…,p* (*p* number of observation/Bazar), y_i are the dependent variable (informal trade), x_i are the explanatory variables (indicators coming from EO-data and ancillary information), β_0 is the *y-intercept* (costant term, to be found to estimate future y_i), β_p is the slope coefficients for each explanatory variable while ε is the model's error term.

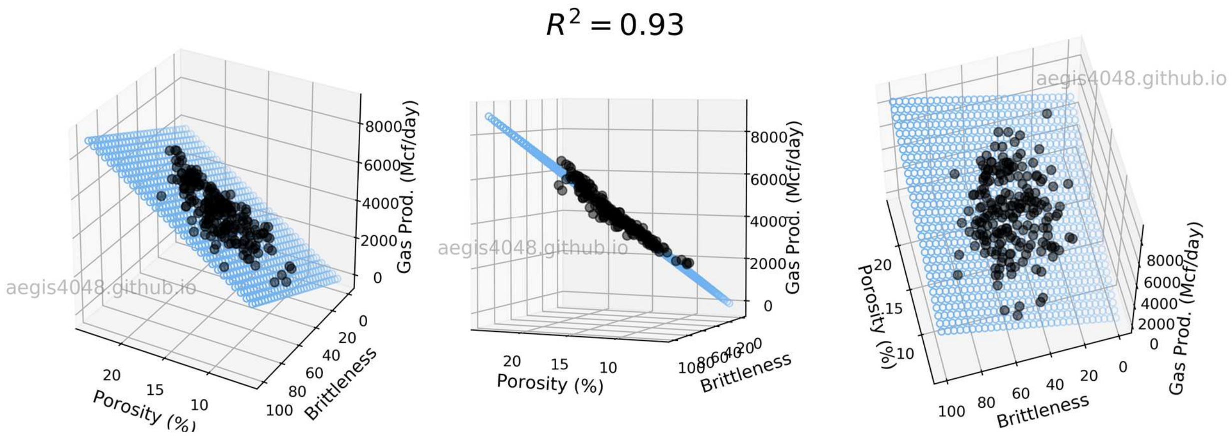

Figure 7: Visual example of linear regression (in the middle) and multi-linear regression (left and right side).[1]

[1] https://aegis4048.github.io/mutiple_linear_regression_and_visualization_in_python

The easiest mode to transform the linear regression model in a non-linear regression model is to change the dependencies between y and β modifying the x_i (i.e., x_i^2, $x_i x_j$, etc.).

In the last period, machine learning approaches have allowed us to build a more complex relationship between data in a simpler way. For example, using different kernels (linear, polynomial or, RBF), is possible to find the function that better fits your data, and consequently, to conduct the prediction in the best possible way (Figure 8).

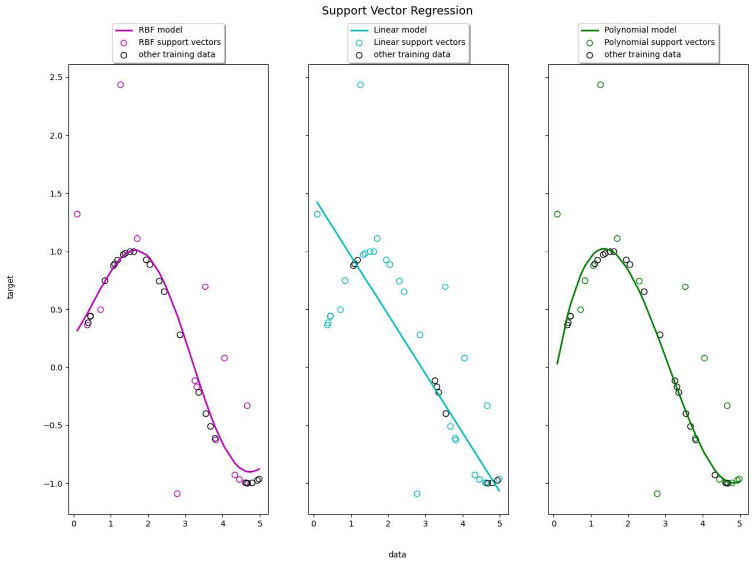

Figure 8: Example of fitting function using different Kernel in SVM approach[1]

This model is a supervised learning algorithm and, as shown in the figure above, is possible to reach different accuracy, changing the Kernel in the SVM approach. In the literature, there are also other approaches like Random Forest Regression (Figure 9), a very useful approach, able to overcome the overfitting problem and also Multi Layer Perceptron Regressor model, which is a supervised deep learning approach, able to create a complex non-linear function using backpropagation with no activation function in the output layer.

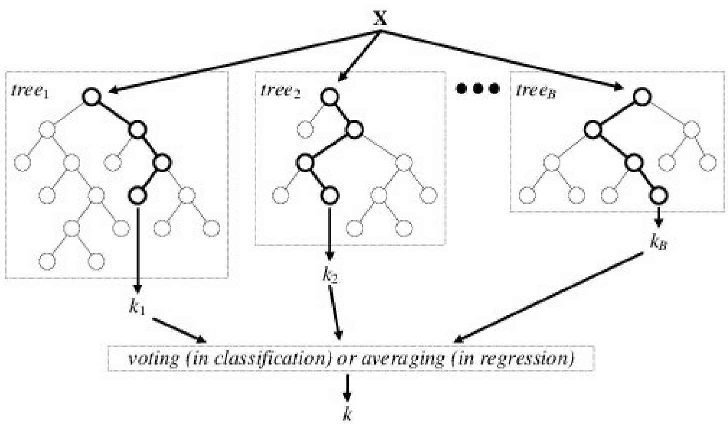

Figure 9: Random Forest Regression. Image credits: Electromyographic Patterns during Golf Swing: Activation Sequence Profiling and Prediction of Shot Effectiveness

[1] https://scikit-learn.org/stable/auto_examples/svm/plot_svm_regression.html

In this project, the four methods mentioned before will be used:

1. Linear Regression (LR);
2. SVM Regressor, with linear and non-linear kernel (SVR);
3. Random Forest Regressor (RFR);
4. Multi Layer Perceptron Regressor (MLPR).

All the above methods are characterized by an accuracy level with respect to the test and the validation set. The best method in the validation set will be used to perform prediction.

Before starting with the analysis, Figure 10 shows a timetable of all variables useful for the regression. Note that due to the nature of the data (single value), not all tables described before can be used in the analysis (T1, T2, T3 and T4) furthermore, in this context, not the entire period is covered from data, and, for this reason, an extrapolation method will be used. The latter could introduce an error in the regression algorithms.

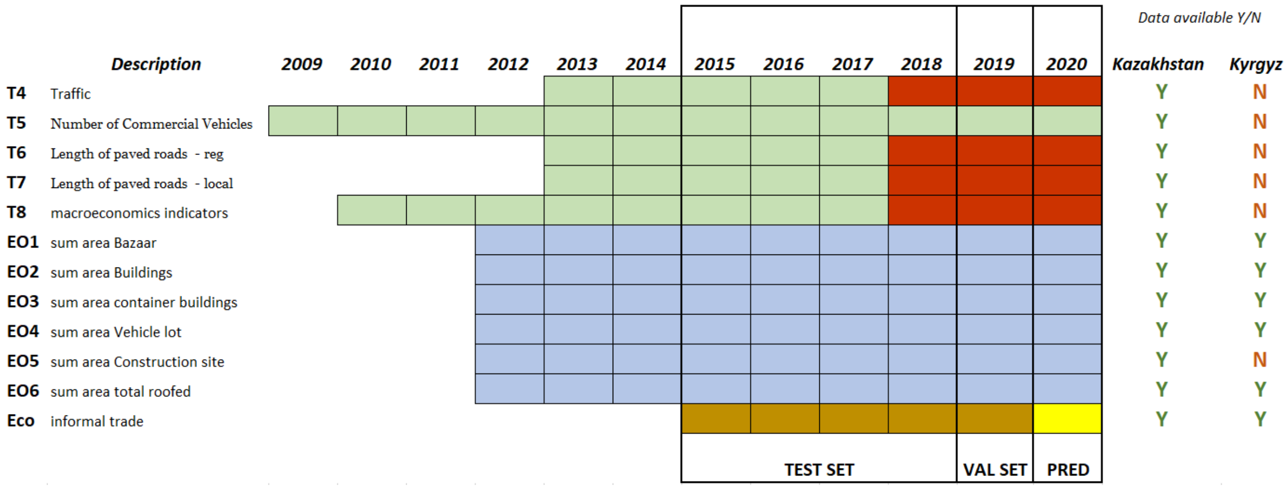

Figure 10: Data recap: Green boxes - no-EO data available; Blu boxes - EO data available; Red boxes - data to be extrapolated in order to use regression steps; Gold boxes - Test and validation set; Yellow box - period to be predicted using results in test and validation set.

Furthermore, some input data analysis inspection has to be performed before the use of regressors because some data can correlate in time. In this case, not all of them are needed during the training phase, and a dimensionality reduction method has to be used to remove the unnecessary information.

To summarize, the performed steps are:

1. *Normalization* of time series data (EO and not-EO);
2. *To fill* and *to reconstruct* time-series data to cover the same interval time;
3. Use of *dimensionality reduction* method (i.e., PCA) to reduce the dimension of the created space, keeping the same information. This step is also able to remove data that are linear correlated;
4. Application of Linear and Non-linear approaches to *train regression* algorithm;
5. Application of single regression methods on the last reference year to extract the accuracy of each method in test and validation set;

These steps will be applied in the two different scenarios defined before (Scenario 1: Dordoi Bazaar (Bishkek, Kyrgyz Republic) and Scenario 2: Barakholka Bazaar (Almaty, Kazakhstan)) performing 4 different tests related to the available data (EO and no- EO):

TEST1. Prediction in Dordoi Bazaar (Bishkek, Kyrgyz Republic) using only EO data (the only available and useful data) - Scenario 1;

TEST2. Prediction in Barakholka Bazaar (Almaty, Kazakhstan) using only EO data - Scenario 2;

TEST3. Prediction in Barakholka Bazaar (Almaty, Kazakhstan) with EO and no-EO data (to understand if auxiliary data can be useful to grow up the accuracy);

TEST4. Prediction in Barakholka Bazaar (Almaty, Kazakhstan) with dimensionality reduction step to remove correlated data.

In each test will be described the used set-up (data used, extrapolation step, dimensionality reduction step, etc.) and, the results of all methodologies will be analyzed. Furthermore, each AI regressor will be trained using the test set (from 2015 to 2018) and assessed using data from 2019. The best method in terms of distance (in million USD) from the informal trade of 2019 will be selected and then, the AI regressor will be re-trained with all available dates (from 2015 to 2019) to predict the next year (2020).

Normalization Phase

Figure 11 and Figure 12 show the effect of the Normalization phase applied on some time series presented in Table 12. As is possible to see, the normalization transforms the data and, the similar behavior between GDP (black line) and GDP per capita (the yellow one) is similar as aspected. This is important because we can use only GDP or GDP per capita in the AI Regressors.

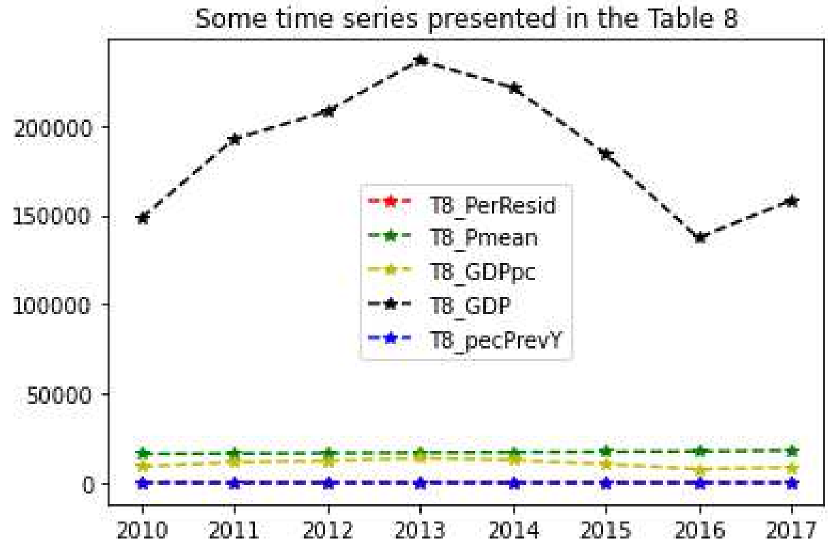

Figure 11: Time series presented in Table 12 without normalization

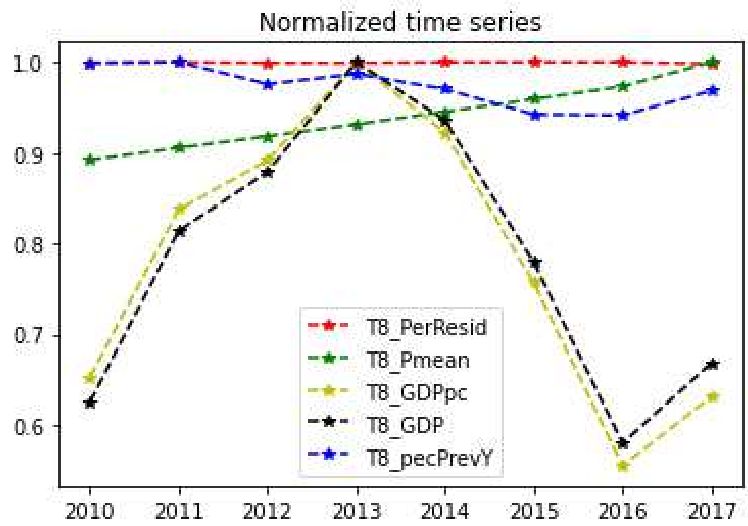

Figure 12: Normalized time series. The behavior between T8_GDPpc and T8_GDP is very similar

Data Extrapolation

Figure 13 shows the effect of the data extrapolation. This step is fundamental to use the AI regressor (due to the different lengths of the time series) but, as shown, this could introduce errors. The same step has been applied to all-time series as will be described better in the following TEST3.

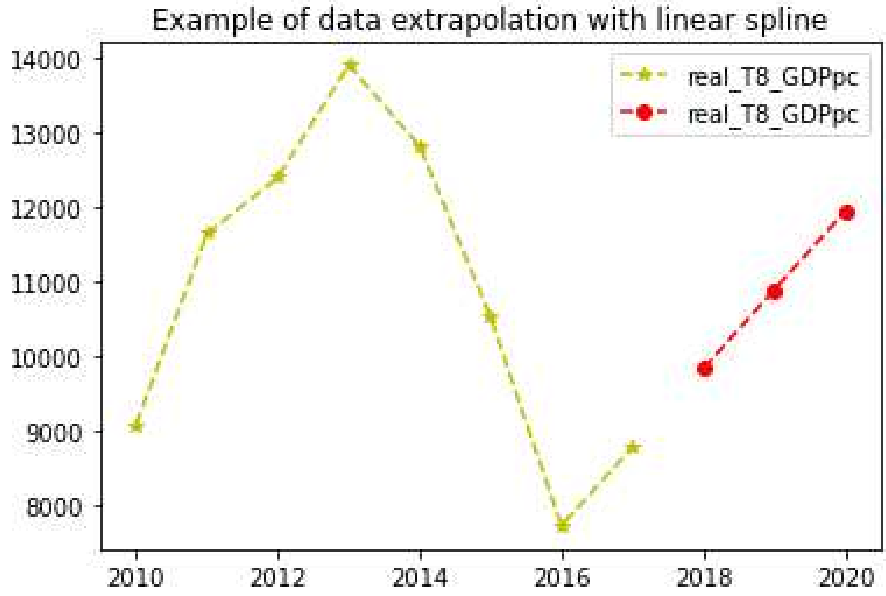

Figure 13: Example of data extrapolation

TEST1. Prediction in Dordoi Bazaar (Bishkek, Kyrgyz Republic) using only EO data

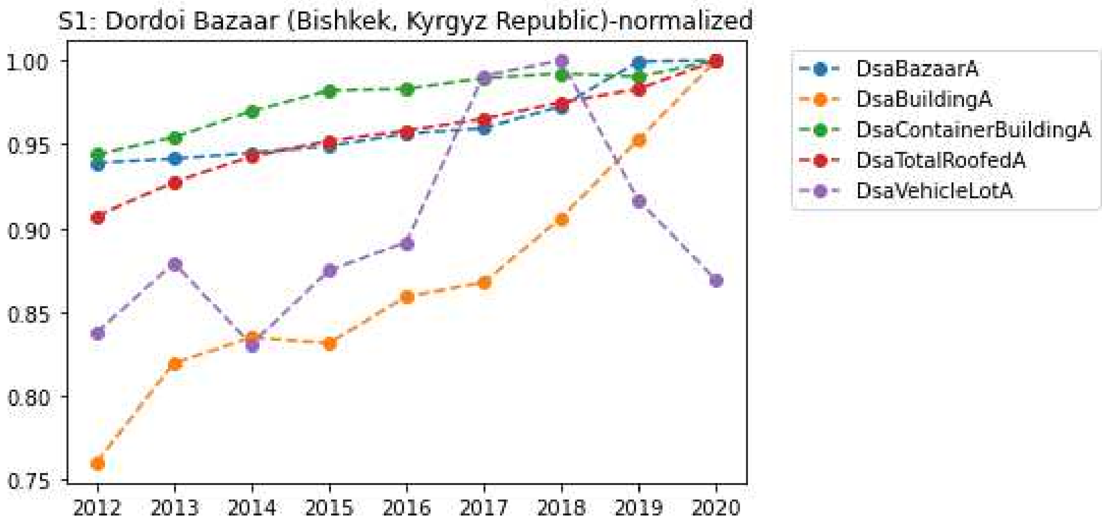

Figure 14: Normalized EO data used for TEST1

Set-up:

1. ***Data used***: EO data [EO1, EO2, EO3, EO4, EO6];
2. ***Extrapolation step***: unneeded;
3. ***Dimensionality reduction***: Not performed;
4. ***Test set***: 2015-2016-2017-2018;

5. ***Validation set***: 2019;
6. ***Data dimension***: 5 (date) x 5 (EO) -Test 4 (date) x 5 (EO) - Validation 1 (date) x 5 (EO)

Results:

1. ***Accuracy in Test set***: LR: 1.0; SVR: 0.99; RFR: 0.79; MLPR: 0.77;
2. ***Abs error (million USD) in Validation set***: LR:2740.39; SVR: 329.93; RFR: 1080.2; MLPR: 319.43;

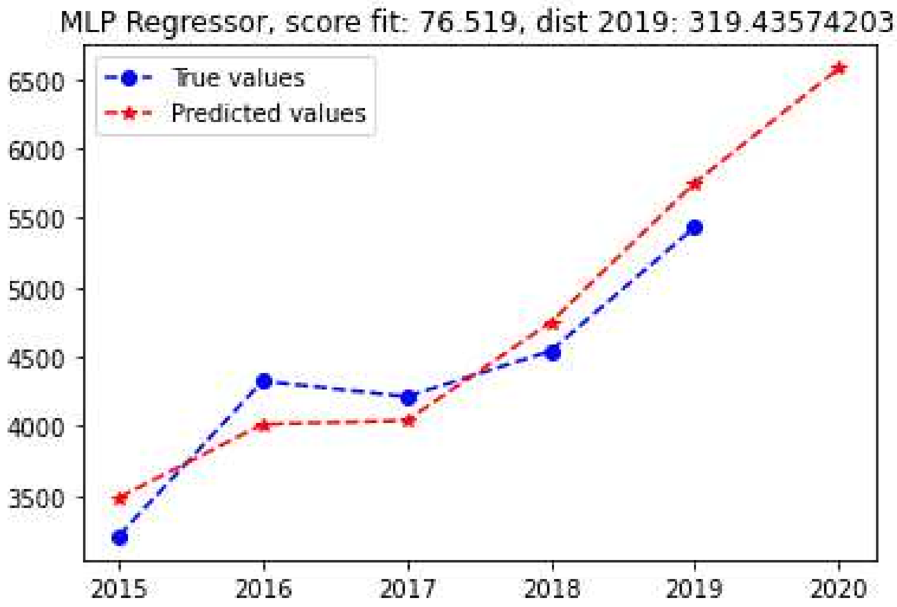

Figure 15: Results of MLP Regressor

The results show that the models LR and SVR are in "overfitting" in fact, despite the accuracy in the Test-set is very high (close to 1.0), they have a not negligible error in the Validation-set. The major problem is that the time series has not enough length to ensure high accuracy in the Validation set (only 4 data in the time are used to fit the model). On the other hand, MLPR seems to have enough accuracy to be selected as the best model. Figure 15 shows the distance between the predicted informal trade output with respect to the real one from 2015 to 2019 while, Figure 16 shows the prediction in 2020 using data from 2015 to 2019. The predicted value is 6000.2 so that, in 2020 an increase of ~10% is expected to the informal trade flows.

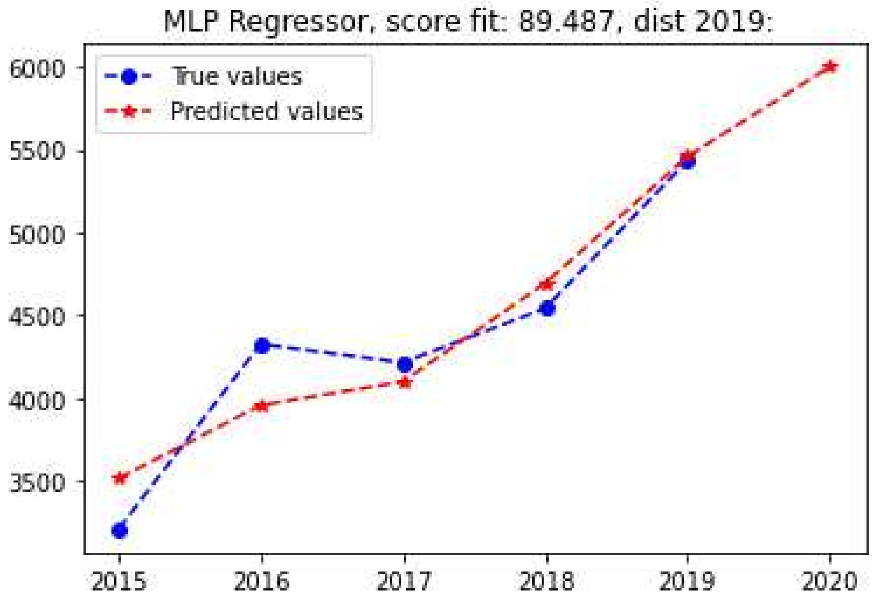

Figure 16: Prediction in 2020 using data from 2015 to 2019. The score fit reached is 89.49%

TEST2. Prediction in Barakholka Bazaar (Almaty, Kazakhstan) using only EO data

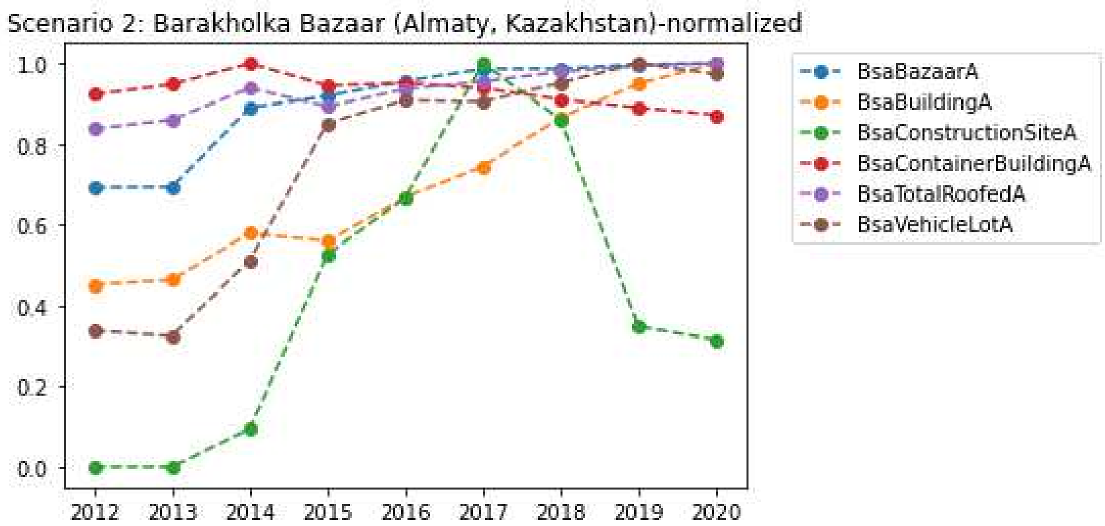

Figure 17: Normalized EO data used for TEST2

Set-up:

1. ***Data used***: EO data [EO1, EO2, EO3, EO4, EO5, EO6];
2. ***Extrapolation step***: unneeded;
3. ***Dimensionality reduction***: Not performed;
4. ***Test set***: 2015-2016-2017-2018;
5. ***Validation set***: 2019;
6. ***Data dimension***: 5 (date) x 6 (EO) -Test 4 (date) x 6 (EO) - Validation 1 (date) x 6 (EO)

Results:

1. ***Accuracy in Test set***: LR: 1.0; SVR: 0.99; RFR: 0.79; MLPR: 0.99;
2. ***Abs error (million USD) in Validation set***: LR:166.56; SVR: 680.75; RFR: 222.93; MLPR: 124.96;

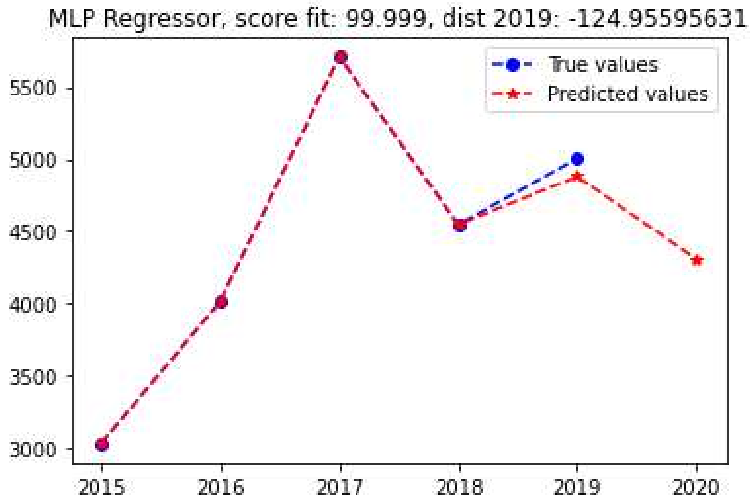

Figure 18: Results of MLP Regressor

Also in this test case, MLPR present the best performances. The results suggest that there are linear dependencies from data in fact, LR has a perfect fit in Test-set and "low" error in Validation-set while, due to non-linear behaviors, SVMR and RFR have some problems both in Validation-set (SVMR) and in Test-set (RFR).

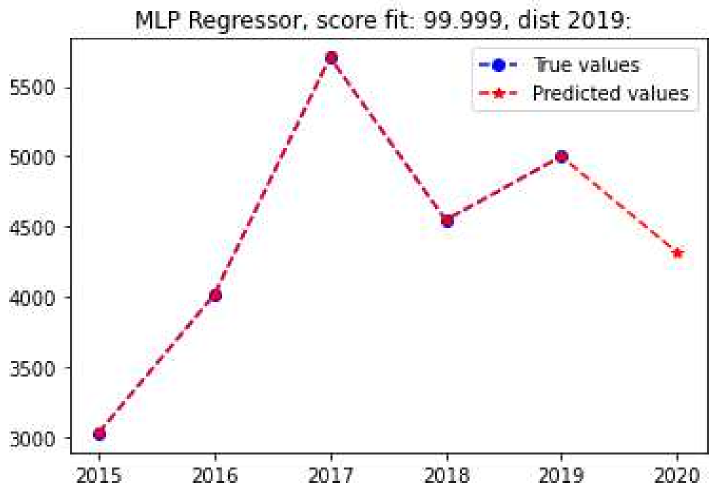

Figure 19: Prediction in 2020 using data from 2015 to 2019. The score fit reached is >99%

Figure 19 shows that the informal trade flow predicts in 2020 is lower than the previous year. This fact is confirmed to all methods used for the test and, in particular, the MLPR predicts a value ~4316.03, -13.7% in comparison with the year 2019.

TEST3. Prediction in Barakholka Bazaar (Almaty, Kazakhstan) with EO and no-EO data

Set-up:

1. ***Data used***: EO data [EO1, EO2, EO3, EO4, EO5, EO6] + no-EO data [T4,T5,T6,T7,T8];
2. ***Extrapolation step***: applied on T4, T6, T7 and T8 from 2018 to 2020;
3. ***Dimensionality reduction***: Not performed;
4. ***Test set***: 2015-2016-2017-2018;
5. ***Validation set***: 2019;
6. ***Data dimension***: 5 (date) x 23 (EO) -Test 4 (date) x 23 (EO) - Validation 1 (date) x 23 (EO)

Results:

1. ***Accuracy in Test set***: LR: 1.0; SVR: 0.99; RFR: 0.82; MLPR: 0.99;

2. ***Abs error (million USD) in Validation set***: LR:143.10; SVR: 1164.48; RFR: 386.313; MLPR: 99.35;

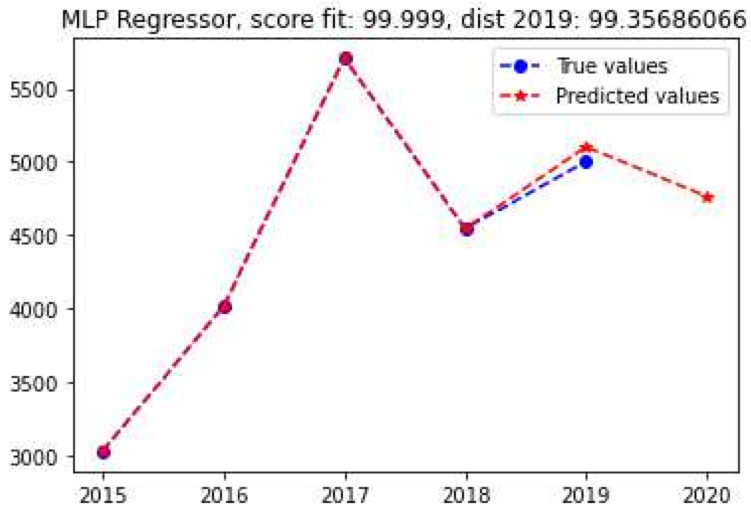

Figure 20: Results of MLP Regressor

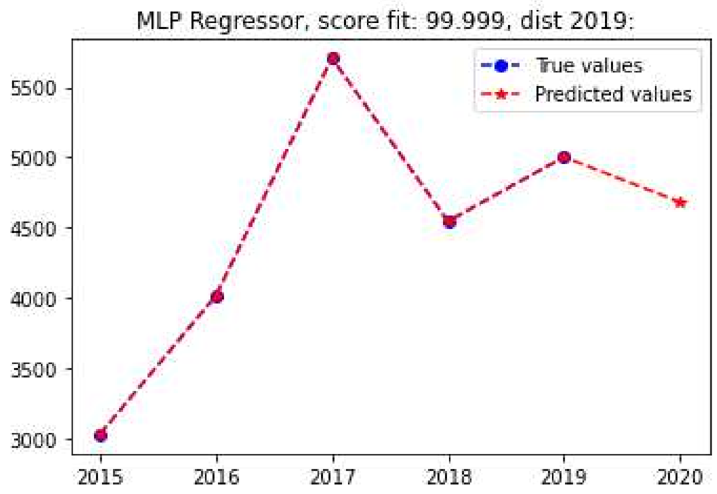

Figure 21: Prediction in 2020 using data from 2015 to 2019. The score fit reached is >99%

This test, with the use of auxiliary no-EO data, confirms that the predicted value for 2020 is lower than the informal trade of 2019. The expected value is 4682,59, which means -6,85%. Furthermore, the use of no-EO data reduces the error in Validation-set for LR and MLPR.

TEST4. Prediction in Barakholka Bazaar (Almaty, Kazakhstan) with dimensionality reduction step

Set-up:

1. ***Data used***: EO data [EO1, EO2, EO3, EO4, EO5, EO6] + no-EO data [T4,T5,T6,T7,T8];
2. ***Extrapolation step***: applied on T4, T6, T7 and T8 from 2018 to 2020;
3. ***Dimensionality reduction***: Performed;
4. ***Test set***: 2015-2016-2017-2018;
5. ***Validation set***: 2019;
6. ***Data dimension***: 5 (date) x 23 (EO and no-EO) -Test 4 (date) x 23 (EO and no-EO) - Validation 1 (date) x 23 (EO and no-EO)

Results:

1. ***Accuracy in Test set***: LR:1.0 ; SVR:0.99; RFR:0.74; MLPR:0.99;
2. ***Abs error (million USD) in Validation set***: LR:98.65; SVR: 999.14; RFR: 621.81; MLPR: 1013.33;

Due to the dimension of the data that contains 23 information for each year, the dimensionality reduction is performed to understand if it is possible to reduce the dimension keeping the same information. In particular, Principal Component Analysis (PCA) is applied to the entire dataset. The explained variances described using principal components, says that is possible to apply the regressor methods only on the first 5 components (explained variance PC1: 93.87% - PC2: 5.70% - PC3: 0.34% - PC4:0.06% -PC5: 0.02%; Total explained variance of the first 5 components: ~100%). In this way, the dataset is transformed from 6 (date) x 23 (EO and no-EO) to 6 (date) x 5 (PCs).

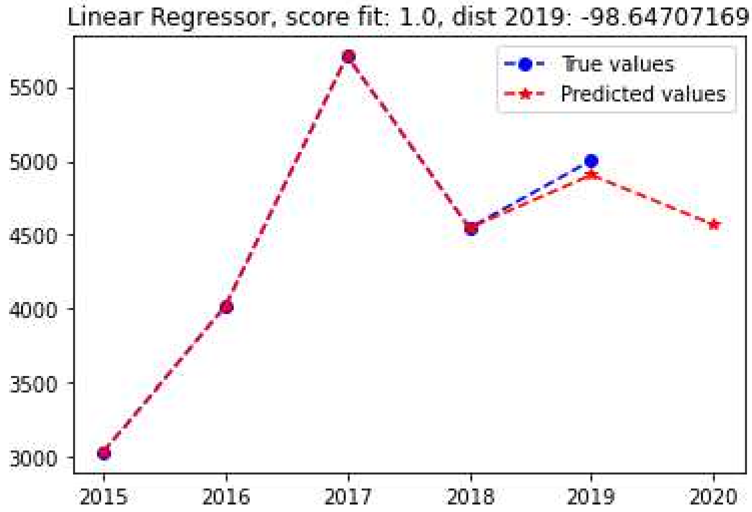

Figure 22: Results of Linear Regressor

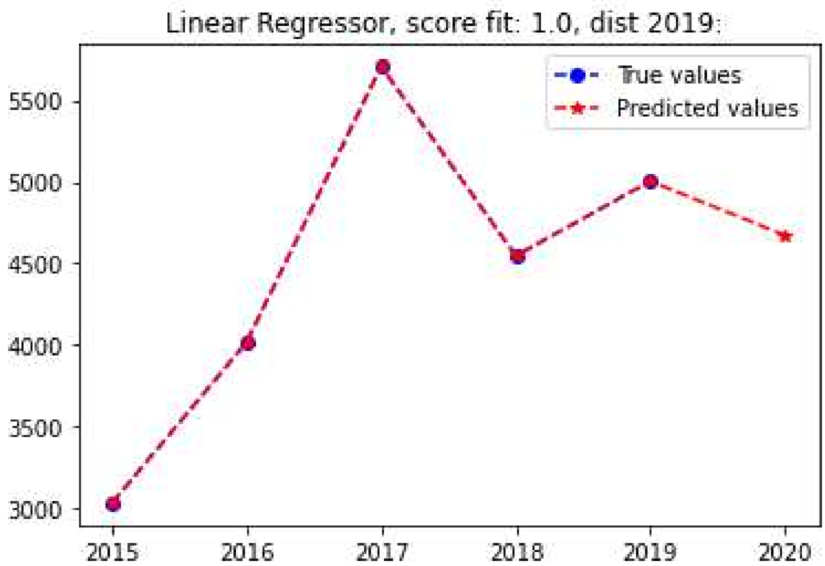

Figure 23: Prediction in 2020 using data from 2015 to 2019. The score fit reached is >99%

Figure 22 shows that LR is the best method with the best accuracy in Validation-set. The results show that, also in this case, all methods suggest a decrease of the informal trade flows in 2020 and, the estimated prediction remain more or less the same respect to the other methods (4671.86, that means ~ -6.64%, Figure 23). This analysis confirms that several data are linear correlated and then is possible to use only 5 information instead of 23 for each year keeping the same accuracy.

3.1.2.2 Quality Control and Validation

To test the quality of the methods, the available dataset is split into 2 sets, Test-set and Validation-set. The Test-set consists of data from 2015 to 2018 while the data from 2019 is used as Validation-set. The Test-set is used to train the AI Regressor and after the pre-trained AI method is applied to the Validation-set. In this way is possible to understand what is the real error in the prediction. Finally, the best method is applied to predict 2020 where, of course, we don't have any data to be used. The idea is that if we use the best method in 2019 with data from 2015 to 2018, the prediction in 2020 with data from 2015 to 2019 will be better because the Test-set will contain an added value (2019).

3.1.2.3 Usage, Limitations and Constraints

As mentioned before the most important problem is the length of the time series. Typically, 5-6 data in the time cannot represent the entire behavior of the time series, this means that the available data used in the test are not enough to reach high accuracy in the Validation-set. Another limitation is in the use of the extrapolation step. This approach introduces several errors in the dataset and, of course, this causes prediction errors. For these reasons, in order to have better results in the Validation-set and overcome the overfitting problem, more data in the time and for the entire period is needed and requests for further analysis.

3.2 Dashboard

In order to visualize the information obtained by Service 1 and Service 2, the input data used and the results obtained from the prediction are made accessible also from a Dashboard. This visualization is done using a web application powered by Hexagon's M.App Enterprise. Since the data have not a geographical extension, but are generally related to the two Bazaars, the main part of the dashboard is the map window showing the locations of the Barakholka Bazaar in Kazakhstan and Dordoi Bazaar in Kyrgyz Republic are highlighted through two points. Clicking on one of them, a tooltip appears. It contains links to the data and results obtained from Service 1 and Service 2. Each link opens a new window where the information are displayed as time series graphs.

The web app was designed not only for PC, but is also consultable from portable devices, such as tablets and smartphones, through a web portal and can be reached using the following URL/credentials::

URL: https://services.rheticus.eu

Username: p21s1683_eo_clinic14

Password: mnbbPsqPGAgv

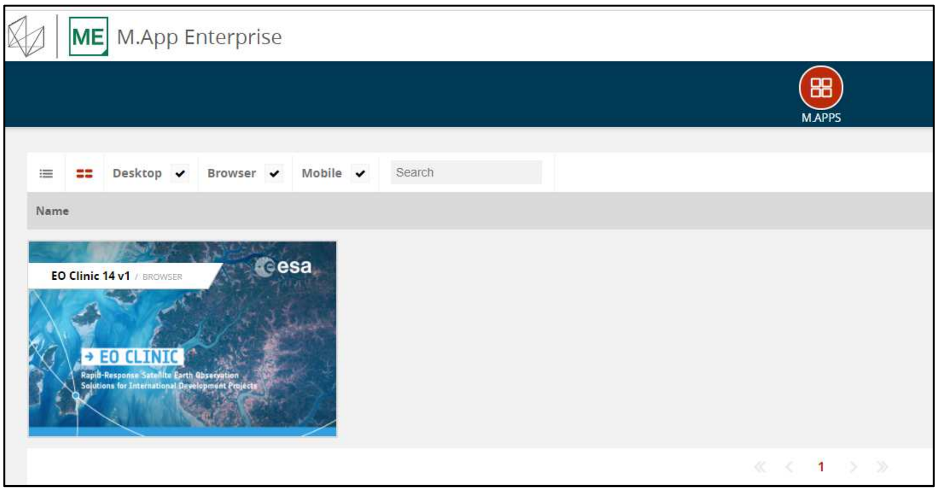

Figure 24: Access portal to the dashboard, accessible after the login

3.3 Conclusions

The results of the analysis described in the previous sections, are cautiously encouraging. Unfortunately, it is not possible to derive solid conclusions, because of the poor availability of input data, which does not allow an accurate understanding of the powerfulness of the predictive techniques in this context. For their nature, they need a big amount of input data.

A first set of analyses has been performed using the data from 2015 to 2018, while, the data related to the year 2019 has been used as a benchmark for the prediction. We have obtained, in this case, an error of about 10% for Dordoi bazaar and around 5% for Barakholka in predicting the informal trade flow of 2019. An interesting result obtained is the improvement of the quality of the prediction for the Kazakh bazaar when the ancillary data are added at the input dataset. In this case, the prediction error for the 2019 data is about 2-3%.

More interesting is the second set of analyses performed. In this case, to increase the amount of data for the prediction, the data related to the year 2019 are also used as input. Of course, in this case, we do not have a benchmark because the predicted year is 2020, about which we do not have any data. However, it is interesting to see how the prediction shows different behavior for the two bazaars: for the Dordoi bazaar it is foreseen an increase of the informal trade flow of about 10% with respect to 2019, while for Barakholka bazaar it is predicted a reduction of the flow. The foreseen flow decrease is about 13% if only the EO data are used, while it is about 7% in case of also the ancillary data are used.

Although it would be interesting to compare the latter results with the data related to the year 2020, and in general there is a coherence among the results obtained in the various tests performed, it should be underlined again that the lack of data creates a general high uncertainty in the results obtained.

APPENDIX A: DELIVERABLES OF SERVICE 1

The deliverables of service 1 are at the same time input into service 2.

Description of indicators, sqm- and feature count numbers of indicators and geometric delineations of indicators (vector data) can be freely used on behalf of ESA. However, no commercial use is possible. The satellite image data is licensed to ESA and WBG. The EULAs apply.

Bazaar	Item	Content	Format	
Barakholka	Vector data	Geospatial Analysis Results, Geometry Features with Attributes	Esri File-Geodatabase. One featureclass per indicator.	1 GDB
Barakholka	Statistics	Calculated Areas and number of all features per observation date and indicator	MS Excel file.	1 xlsx
Barakholka	Satellite scene	Satellite data, licensed to ESA and WBG. Bundle block adjustment with other data of the same bazaar.	Geotiff	1 tif + EULA.pdf
Dordoi	Vector data	Geospatial Analysis Results, Geometry Features with Attributes	Esri File-Geodatabase. One featureclass per indicator.	1 GDB
Dordoi	Statistics	Calculated Areas and number of all features per observation date and indicator	MS Excel file.	1 xlsx
Dordoi	Satellite scene	Satellite data, licensed to ESA and WBG. Bundle block adjustment with other data of the same bazaar.	Geotiff	1 tif + EULA.pdf