



UNIVERSIDAD  
**COMPLUTENSE**  
MADRID

FACULTAD DE CIENCIAS ECONÓMICAS  
Y EMPRESARIALES

**A Methodology for Early Prediction and Proxies of Global and Local  
Evolution Patterns of Economic Activity and Crises for European Union  
Countries through Spaceborne Earth Observation Physical Data:  
The Case of Cyprus**

Giovanni De Angelis  
Valentín Edo Hernández

**Working Papers / Documentos de Trabajo. ISSN: 2255-5471**  
DT CCEE-2401 diciembre 2024

**Abstract:**

This paper is based on research carried out in the framework of the European project "Economy bY spacE (EYE)", although the part analyzed here has its specific development. The economic data assessment methodology is original, based on economic theory and on the concepts of National Accounting applied in European Union countries and by EUROSTAT. In particular an economic analysis model was built up and used to estimate the degree of correlation between economic variables and variables observable from space. This approach has been already developed in the framework of the MICROS project, in which Cyprus national-level and local-level economic data have been tested for correlation vs. physical data observed with satellites from space for Cyprus locations, city and districts, with significant correlations observed in spite of the physical data noise patterns. In the present work the same datasets have been tested for correlations vs. the national-level economic data collected from EUROSTAT. The model aims to infer in advance, with respect to official statistical publications, the evolution of economic variables, and is applied to the case of Cyprus, by using the methodology developed for MICROS. The main contribution of this work is that, in the approach applied, by selecting the appropriate aggregate variables, there is no need for detailed economic local-level data to obtain significant correlations.

**Resumen:**

Este trabajo se basa en investigaciones realizadas en el marco del proyecto europeo "Economy bY spacE (EYE)", aunque la parte analizada aquí tiene un desarrollo específico. La metodología de evaluación de datos económicos es original, basada en la teoría económica y en los conceptos de Contabilidad Nacional aplicados en los países de la Unión Europea y en EUROSTAT. En particular, se elabora un modelo de análisis económico y se estima el grado de correlación entre variables económicas y variables observables desde el espacio. Este enfoque ya ha sido desarrollado en el marco del proyecto MICROS, en el cual se probaron datos económicos a nivel nacional y local de Chipre para correlacionarlos con datos físicos, recogidos por satélites, en diferentes localizaciones, ciudades y distritos de Chipre, obteniendo correlaciones significativas a pesar de los patrones de ruido en los datos físicos. En **este** trabajo, se han utilizado los mismos conjuntos de datos para correlacionarlos con los datos económicos a nivel nacional recopilados de EUROSTAT. El modelo desarrollado busca inferir por adelantado, con respecto a las publicaciones estadísticas oficiales, la evolución de variables económicas y se aplica al caso de Chipre utilizando la metodología desarrollada para MICROS. La principal aportación de este trabajo es que, con el enfoque aplicado, seleccionando las variables económicas agregadas adecuadas, no es necesario disponer de datos económicos detallados a nivel local para obtener correlaciones significativas.

**JEL classification:** E1, E3, E6.

**Keywords:** Macroeconomic magnitudes, Aggregate supply, Economic cycles, Economic crises, Economic forecasts, Economic outlook.

**Dr. Giovanni De Angelis**  
Senior Scientist  
[gianni.deangelis@spacesystems.solutions](mailto:gianni.deangelis@spacesystems.solutions)  
<https://orcid.org/0000-0001-6720-4516>  
Space Systems Solutions S3 Ltd.  
1016 Nicosia  
Cyprus

**Valentín Edo Hernández**  
[vedohern@ucm.es](mailto:vedohern@ucm.es)  
<https://orcid.org/0000-0002-0967-8251>  
Department of Applied and Public Economics,  
and Political Economy  
Faculty of Economics and Business  
Complutense University. Somosaguas  
Campus. Madrid. 28223

Este trabajo ha sido editado por la Biblioteca de la Facultad de CC Económicas y Empresariales de la UCM, de acuerdo con los requisitos de edición que figuran en la Web institucional. Las opiniones expresadas en este documento son de exclusiva responsabilidad de los autores.

## **SUMMARY**

1. INTRODUCCIÓN.
2. THE MACROECONOMIC MODEL
3. METHODOLOGY AND SELECTION OF ECONOMIC VARIABLES.
4. DATA COLLECTION AND SELECTION OF PHYSICAL VARIABLES  
OBSERVABLE FROM SPACE
5. THE CASE OF CYPRUS.
6. OUR ANALYSIS FOR CYPRUS: DATA AND METHODOLOGY.
7. OUR ANALYSIS FOR CYPRUS: RESULTS.
8. OUR ANALYSIS FOR CYPRUS: DISCUSSION.
9. CONCLUSIONS.
10. REFERENCES

## **1.INTRODUCTION**<sup>1</sup>

In recent years, most economies around the world have been affected by a new crisis, resulting from the COVID-19 pandemic. This has forced all countries to adopt measures to curb its negative effects on health and consequently on their economy and their people's welfare. Some of the measures adopted have drastically affected people's mobility, generating major issues around economic activity. The effects of an economic crisis can be observed in movements in the relevant economic variables but can also be seen in the changes occurring in physical variables observable from space.

This research is a contribution to the EYE Project and develops a model to estimate the evolution of certain economic variables. Its main objective is to select economic and physical variables and estimate their degree of statistical correlation. Consequently, if they present a high degree of correlation, predictive models can be built with observable variables from space anticipating the possible evolutionary patterns of the economic variables. The EYE project is the ideal follow-up of MICROS, a project aiming at performing the same analysis, economic variables vs. physical variable observable from space with instrumentation onboard Earth-orbiting instrumentation, for the island of Cyprus. Economic data in the latter case was collected at the national and the local level through sources from Cyprus government ministries and other institutions, whereas physical data were observed with satellites from space for Cyprus locations, city and districts, with significant correlations observed despite the physical data noise patterns. In the present work the same physical datasets from MICROS project have been tested for correlations vs. national-level-only economic data collected from EUROSTAT.

This paper first presents a theoretical macroeconomic model. Second, it describes the statistical sources and characteristics of the selected economic variables. Third, it demonstrates the method for the selection of the physical variables and their selection philosophy is described, as from the MICROS project. And finally, it shows the results of the estimates made in the case of Cyprus between the years 2018 and 2021, the time

---

<sup>1</sup> This document is part of a research project called "Economy bY spacE (EYE)" **European Commission (2021). G.A.101007638. Type of action MSCA-RISE-2020.**

interval for which data for both economic variables and physical variables observed from space are available.

The authors are personally indebted to Geomatics Cyprus Ltd. (GMCY) for the use of the physical data of the MICROS project. GMCY was the PI organization of the MICROS project.

## **2.THE MACROECONOMIC MODEL**

The economic model presented in this chapter has been developed as the framework for the selection of suitable economic variables to establish correlations with physical variables observable from space.

The model is based on economic theory but is also designed to facilitate the selection<sup>2</sup> of variables with different levels of disaggregation. First, those with the highest degree of aggregation will be presented. Second, new variables can be defined by breaking down the previous ones. It should be borne in mind that the variable selection process will not only be based on their intrinsic value, but also on the availability of statistical information.

An essential element of the model will be its internal cohesion as it is based on an exhaustive review of all the possible relevant variables at each level of disaggregation. One outstanding characteristic of the model is that the sum of the values of the variables' components in any level is equal to the value of their respective immediate aggregate variable of the level above.

This process endeavours not only to select the appropriate variables, but also to guarantee two outcomes: no redundancy in the selection of variables; and no variables left without being considered. In other words, the methodology establishes that the aggregate value of the selected variables exactly matches the value of the corresponding aggregate variable at the level immediately above, ensuring that they complete the model correctly.

The effects of economic crises are frequently measured by analysing the evolution of the main economic variables, such as Gross Domestic Product (GDP, hereafter), Unemployment, Prices, Disposable Income, etc. All these can be defined here as First-level variables, as they are the result of the highest aggregation of certain components. Although these First-level variables give an overview of the economy's evolution, they reflect the changes in its components. The economic variables that best fit the variables observable from space should be chosen, since they can be predicted to give higher correlation results.

---

<sup>2</sup> These variables must satisfy not only criteria based on economic theory, but also others related to statistical information, as described in the next chapter. Moreover, to select the most appropriate variables, data for a sufficiently long period and for all European Union countries are needed.

The First-level variables, such as GDP, are those with the highest level of aggregation, such as those listed above, and which represent general aspects of a country's economy. Second-level variables are those comprising first-level variables. In addition, it is possible to obtain variables at a third level and continue the process of breaking them down as necessary.

From the point of view of measuring economic activity, GDP is generally recognised as the most representative variable; in relation to this research, it will be considered as the single and most important First-level<sup>3</sup> variable.

The model defines GDP as the First-level variable, and its components will be termed the Second-level variables. It is important to point out that the GDP value can be obtained in different ways, which gives us the opportunity to look for variables among different kinds of aggregates.<sup>4</sup>

When selecting the Second-level variables, it is helpful to make an approximation from the economic theory, to better understand the role they play in the model. Consequently, economic theory allows us to identify variables. In relation to this task the variables to be identified can be found, in principle, on the demand side or on the production side.

Following the economic theory, on the demand side, we can differentiate between those related to private consumption expenditure [C], gross capital formation (Investments) [K], public consumption expenditure [G], and net exports [X-M] ("exports" minus "imports"). A symbolic representation of these variables indicating their relationship with one another is as follows:

$$\text{GDP} = \text{C} + \text{G} + \text{K} + (\text{X}-\text{M}) \quad [ 1 ]$$

---

<sup>3</sup> Since the most important variables in this research are those related to GDP, in the initial phase we are going to select variables that are components of GDP, calling them First-level variables. Therefore, the procedure for selecting the appropriate GDP-related variables consists of selecting macroeconomic variables regarding their role in economic activity and to their position relative to GDP, according to whether they are demand-side or supply-side

<sup>4</sup> There is another common procedure for estimating GDP. That is adding economic incomes such as "Compensation of employees" and "Operating surplus and mixed income, gross", complemented by "Taxes on production and imports less subsidies". However, there are no variables that are useful for our approach in this research.

On the supply or Production side, we need to distinguish the variables related to the main production activity; in an initial approximation, these can be grouped into three classes. In particular, the aggregates of these Second-level variables will be “Agricultural production and others related to the primary sector” (AA, hereafter), “Industrial production and others related to the Energy and Construction sector” (II, hereafter), and the “Production of Services, especially Trade, transport and accommodation, catering and public administrations” (SS, hereafter). In addition, a closing variable is necessary to calculate GDP and is defined as “Taxes less subsidies on products” (Ti, hereafter). This variable broadly corresponds to what are usually termed indirect taxes (for more details and introductory view see Edo, 2018) They can be represented symbolically in an equation as follows.

$$\text{GDP} = \text{AA} + \text{II} + \text{SS} + \text{Ti} \quad [ 2 ]$$

The aggregate value of the two sets of Second-level disaggregated variables is in both cases equal in construction to GDP, and in both equations their respective addition gives the same GDP value. They are accounting identities and GDP, in each case, is equal to their respective aggregate values.

Considering the set of equations presented here, GDP change can be said to depend on the evolution of its Second-level components. Therefore, it is possible to estimate correlations between components of GDP with variables observable from space. Furthermore, the Second-level variables can be broken down into their components to obtaining what we shall call Third-level variables, allowing us to estimate new, likely more precise, correlations between the two types of series.

Therefore, since the variables observable from space may adjust better to one of its components than to GDP, it is advisable to disaggregate the variables in all possible components until we find the series that presents a higher correlation with the physically observable ones from space.

In principle, the search for the disaggregated economic variables can be carried out on the demand side or on the supply side. Nevertheless, by studying the data availability, as argued in the next chapter, we will see that the appropriate area for exploration is the supply side.

Moreover, in searching for new variables it is important to consider that, because of the way in which the model has been defined, we can establish a set of mutual relationships among the variables, that can be of proximity, complementarity and/or hierarchy. By doing this we can develop a broad set of variables and can choose the most appropriate among them. The process of looking for new variables depends on the economic theory, on the accounting concepts and on the available information. These relationships will make it easier to understand the results and will guarantee that our review of the variables is comprehensive, enabling us to select the most appropriate ones.

However, this theoretical description requires additional information about certain methodological aspects and statistical support. Therefore, the following chapter presents the methodology for developing the model, the accounting concepts and the statistical sources that allow us to carry on with this research.

### **3.METHODOLOGY AND SELECTION OF ECONOMIC VARIABLES.**

A proper analysis of a country's economy requires detailed information about all its economic activities. Information on the economic agents is needed in order to identify their operations and the products that are manufactured and/or consumed. Consequently, any study in a country or among countries about the evolution of their economies entails a detailed review of the information and a selection of the most relevant variables to conduct this study.

Throughout the twentieth century and until the present time, accounting information systems have developed significantly. In addition, accounting systems have been standardized and applied in most countries, so a lot of comparisons can now be made over time and among countries. Thus, very similar and homogeneous accounting criteria are now applied in Europe and in most countries around the world.

The identification of economic agents and the measurement of economic activity is carried out within the framework of national accounting systems. Nowadays, accounting criteria are based on international agreements, which standardize the information, facilitating comparisons between different economies. The general purpose of a national accounting system is to make a systematic and detailed description of an economy (country, region, locality) and of the variables that are important in order to understand its evolution and for comparison with other economies.

The basic statistical official information about a country's economy, apart from the sectorial or local statistics that can be carried out by different kinds of agencies, is usually collected by official agencies. Although some accounting criteria and methods began to be developed in earlier times, the development of economic theory and particularly the development of macroeconomics by the mid-twentieth century, gave a great boost to statistics. Economic authorities' need to take economic policy decisions greatly promoted the production of increasingly detailed and homogeneous statistics.

In Europe and more specifically in the European Union the standardized criteria of the United Nations are followed, although a particular version is adapted and applied by Eurostat (the statistical office of the European Union). In particular, the criterion set out

is the so-called European System of Accounts (ESA-2010, hereafter)<sup>5</sup>, which is not only an accounting criterion, but also legally a regulation of the European Union, so it has the force of the law (Official Journal of the European Union (the OJEU, 2013). Therefore, its application is mandatory in every country of the European Union (EU hereafter).

The regulatory criteria defined in the ESA-2010 provide abundant and precise information about countries' economic activity and requires EUROSTAT to present homogeneous data for all EU countries so that comparative analysis can be made. In particular, the institution's official sector accounts provide a systematic description of the different phases of the economic operations (consumption, etc.), and describe the industry in detail, structured by branches of activity, the production process and the flows of goods and services.

Although there are units of economic activity of a "legal, administrative or accounting nature open to statistical treatment", "the ESA-2010 system uses two types of units and two corresponding ways of subdividing the economy". The first group consists of the institutional units, grouped into "institutional sectors". The second group are those units (ESA-2010, p.12) that produce goods and services, and are defined as units of local economic activity, whose grouping gives rise to "branches of activity".

An Institutional Unit allows us to measure the economic activity of agents. It is defined as an "economic entity characterized by decision-making autonomy in the exercise of its principal function (ESA-2010, p.27). The aggregation of institutional units by category according to their function allows the establishment of behavioural relationships and the interaction between them, giving rise to a set of operations. The definition of institutional units makes it possible to delimit the institutional sectors, whose analysis will enable the most appropriate economic variables to be selected for the purpose needed, to measure the impact of any economic change from the perspective of institutional agents.

---

<sup>5</sup> In the EU the statistical service of each country collects statistical data, applying the same ESA-2010 criteria. This has allowed the creation of various integrated databases at the European Union level, among those of EUROSTAT stand out. It provides European statistics in partnership with National Statistical Institutes and other national authorities in the EU Member States. This partnership is known as the European Statistical System (ESS), and Eurostat coordinates statistical activities at the Union level and more particularly inside the European Commission.

For the purposes of the ESA-2010, the institutional units are grouped together into five mutually exclusive domestic institutional sectors: (a) non-financial corporations; (b) financial corporations; (c) general government; (d) households; and (e) non-profit institutions serving households. The five sectors together make up the total domestic economy.

Thus, the dimensions of each of the most important accounting concepts can be measured for the entire economy and for each of the five institutional sectors and identifying them enables us to select the most significant variables, such as consumption, production, and any others that may be relevant.

From the productive point of view, we can identify the Units of Local Economic Activity. Considering these units allows us to analyse the economy from the perspective of the production process. The information system must be capable of indicating the value of production and certain other outstanding characteristics for each local kind-of-activity. (ESA-2010, p.49). Therefore, groups of similar productive activities can be grouped if they are added and are called branches of activity. These branches can allow us to choose the most appropriate economic variables, to measure the impact of any change in the economy of a country from the perspective of the productive branches.

In this paper, we will select GDP and other variables with the aim of making correlations with variables observable from space. Although some variables can be measured from the demand or supply side, the variables that will be studied in this paper, mainly for the case of tourism, are best analysed by selecting those located on the production or supply side, as indicated in Chapter 2.

In national accounting, demand data about economic activity can be linked with institutional agents, but this activity can be performed by different kinds of agents. To measure and aggregate demand variables and to relate them with observable variables from space is a difficult task. Economic activity can be atomised around different places and take place at different moments in time, while the observable data from space can usually be seen at the exact time and place when the satellite collects information. They are therefore not the most helpful variables for estimating correlation with variables observable from space.

Statistics of production economic variables are easier to associate with observable data from space, as they can more easily be correlated with some proxy variables, that are also observable from space. For instance, in the case of consumption, an economic agent such as households can consume services, such as tourism, at different moments in time, or in different places, even in different countries. Therefore, it is quite difficult to establish any relationship with variables observable from space (for example, residents of a given country can demand services such as tourism, but the consumption data can be compiled in different places, even in different countries).

When consumers and producers are from different countries, it is rather complicated, since tourism consumption is treated in Eurostat as an export. This and some other circumstances, depending on the products, can add more difficulties to finding economic variables and series that can be correlated with observations from space.<sup>6</sup> Therefore, it is more appropriate to select economic variables from the production or supply side to correlate with variables observable from space.

In this research, the supply side selection process of the economic variables to be correlated with variables observable from space has three stages. The process is based on the degree of aggregation of the variables. First, we have defined the First-level variables; in this study this is GDP, as defined in Chapter 2. In the next stage, we can obtain the Second-level variables by breaking down GDP into its components. Finally, by breaking down the Second-level variables it is possible to obtain the Third-level variables.

Before selecting the appropriate variables for this work, the general restrictions that will be required of all the variables should be applied.

Among them, the availability of statistical data for a sufficiently long period to enable correlation analysis between the series of economic variables and the series of variables observable from space should be highlighted.

---

<sup>6</sup> Example. The provision of tourism services by a hotel in a Mediterranean country clearly does not consist of moving to a Nordic country to provide the tourist service. This economic activity is in the same accounting category as exporting goods. When the resident of a Nordic country must travel, for example, to a Mediterranean country to receive the tourism services, this is classified as an export in the Mediterranean country's accounting. And, conversely, residents in Mediterranean countries may receive tourism, medical attention, or of any other kind of services from residents in foreign countries; that activity is also considered exports in those countries' accounting systems

Here, following ESA-2010 and NACE-2008 methodology, it is possible to identify various relevant variables for which information is available for a sufficiently long period to carry out the necessary studies. For the purposes of this research and with the information available, we have chosen the 2015-2024 period as the most appropriate for the economic variables.

In the case of tourism there is an additional very important consideration, because it is seasonal, giving rise to annual cycles that can facilitate the estimation of certain correlation analyses. Thus, it is advisable to have quarterly data for any year selected.

The economic classification of branches is largely based on the original contributions of the economist Colin Clark (1957). He was one of the first to analyse economic growth and compare economic development between countries, studying the evolution of different types of activity. Here, following him in his classification of economic activities, we begin by selecting the three variables as presented in chapter 2, defining Second-level variables and represented in this essay by the symbols [AA], [II], and [SS]. Although we ordinarily refer to these three Second-level variables by their symbols, for practical purposes and for ease of writing in some cases we will also refer to them respectively as agriculture, industry and services.

The First-level and Second-level variables have been presented in Chapter 2 and in this chapter we have provided further explanations. We now propose the selection of some new variables, that are to be identified as the Third-level variables.

The Third-level variables may be among the most useful for estimating correlations, since they enable us to establish a narrower statistical relationship with observable variables from space.

The methodology applied in this study requires that the sum of the values corresponding to the variables of a more disaggregated level be equal to the corresponding value of the variable of the level immediately above. In this case, as indicated for the Second-level variables and GDP, the sum of the values of the Third-level variables must be equal to the value of the aggregate Second-level variable.

As we have seen in Chapter 2 and according to the Statistical Classification of Economic Activities in the European Community, Rev.2. (hereafter NACE 2008)<sup>7</sup>, to obtain the value of GDP on the supply side we must add the Gross Value Added of the three aggregated Second-level variables plus the value of the indirect taxes, as shown in the equation [2]. This can be expressed thus:

$$\text{GDP} = \text{GVA-}[\text{TOTAL}] + \text{Ti} \quad [3]$$

If we do not take into account the variable [Ti], the GVA-[TOTAL] is equal to the Gross Value Added of its components, that are: GVA-[AA], GVA-[II] and GVA-[SS].

Nevertheless, if we do not consider indirect taxes [Ti], which is not a relevant variable for the analysis being made in this study, a new aggregate variable can be obtained, that we can call “Gross Value Added-Total” (GVA-[TOTAL], hereafter).

This can be represented symbolically with the following equation:

$$\text{GVA-}[\text{TOTAL}] = \text{GVA-}[\text{AA}] + \text{GVA-}[\text{II}] + \text{GVA-}[\text{SS}] \quad [4]$$

Once we have arrived at these aggregates of the Second-level variables, we can begin to look for new variables in order to find which are appropriate ones we should study to see how they correlate with variables observable from space. To do this, we must start by breaking down the Second-level variables into their components.<sup>8</sup>

In NACE-2008 (and in NACE-2022), there are some explanations of the concepts and contents of different branches that contain the GVA-[TOTAL] components. This

---

<sup>7</sup>NACE-2008 and NACE-2025 constitute the basic instrument that ensures the coordination of statistical information and allows for the categorization of productive units and institutions - according to the economic activities they perform - into 21 branches. However, this study does not require a high degree of detail. An updated NACE Classification, adopted by the European Commission in October 2022, will be used for European statistics from 2025 onwards; as it does not impact on this research.

NACE rev. 2.1 introduces new concepts at all levels of the classification and restructures several of the existing headings. At the same time, efforts have been made to maintain the structure of the classification in all areas that do not explicitly require change based on new concepts. One outstanding difference is in the number of concepts at level 1, which has increased by one more concept, increasing the number of divisions to 22.

<sup>8</sup>The Second-level variables [AA] can be related to the Third-level variables. In this case the NACE Third-level variable, defined as [A], coincides with the Second-level variable [AA]. The Second-level variable [II] is defined here as the result of the aggregation of the NACE concepts [B, C, D, E, F]. Finally, the Second-level variable [SS] is the result of the aggregation of the NACE concepts [G, H, I, J, L, M, N, O, P, Q, R, S, T, U].

aggregate is broken down into 21 variables.<sup>9</sup> Therefore, the sum of the GVA of the 21 NACE variables is equal to the value of the GVA-[TOTAL].

By this stage it is possible to find economic variables that are sufficiently specific to be able to relate them to variables observable from space.

On the one hand, the most immediate economic variables are those defined in NACE-2008. Regarding the productive branches of this classification, we will proceed to select variables that could give rise to high correlations with variables observable from space. Once the different NACE-2008 variables have been reviewed, at least one variable can be selected that can offer us the possibility of being related to variables observable from space. This is the variable that, in NACE terminology, is represented by [I] and called "Accommodation and food service activities"; we will shorten it to "tourism".

On the other hand, different variables that may be related to tourism, such as light intensity, can also be observed from space. Although these variables will be described more thoroughly in a later chapter, in this chapter we use them merely as a reference to guide the selection of economic variables that could be used to relate them to those observables from space.

---

<sup>9</sup> THIRD-LEVEL VARIABLES ON THE PRODUCTION SIDE (NACE-2008)

		Total - all NACE activities
1	A	Agriculture, forestry and fishing
2	B	Mining and quarrying
3	C	Manufacturing
4	D	Electricity, gas, steam and air conditioning supply
5	E	Water supply; sewerage, waste management and remediation activities
6	F	Construction
7	G	Wholesale and retail trade; repair of motor vehicles and motorcycles
8	H	Transportation and storage
9	I	Accommodation and food service activities
10	J	Information and communication
11	K	Financial and insurance activities
12	L	Real estate activities
13	M	Professional, scientific and technical activities
14	N	Administrative and support service activities
15	O	Public administration and defense; compulsory social security
16	P	Education
17	Q	Human health and social work activities
18	R	Arts, entertainment and recreation
19	S	Other service activities
20	T	Activities of households as employers; undifferentiated goods- and services-producing activities of households for own use
21	U	Activities of extraterritorial organizations and bodies

Of all the Third-level variables, the tourism variable is one of the most appropriate for studying its correlation with variables observable from space. It is represented by [I] and categorised “Accommodation and food service activities”; this characteristic would make it very appropriate for making estimates that would produce high correlations.

However, it does not fully satisfy all the other requirements demanded in the methodology described. Although there is data for all the years in the 2015-2024 period, quarterly data are not available. Therefore, we must either replace the tourism variable with other proxy variables or make some approximation in order to make reliable estimates of quarterly data.

We must search Eurostat to identify the proxy variable or variables for tourism [I], for which quarterly information is available for the years 2015-2024. One possible solution is to use the variable, identified here, as “G-I”, that could be suitable. On studying the characteristics of the aggregate [G-I] we deduce that it is a good candidate to make estimations substituting the [I] variable.

Although the [G-I] aggregate has the inconvenience of being an aggregate of [I] with additional Third-level variables, if it behaves like [I], it could be an appropriate variable to stand for [I]. Therefore, it is a good candidate for testing and correlation with variables observable from space. This will therefore be one of the Second-level series selected to correlate with variables observed from space.

Another way to overcome the fact of not having quarterly data for [I], is to estimate them based on the behaviour of other variables. Among the most suitable variables to estimate the quarterly data for the concept [I] would be the use of the quarterly data of the variables GVA-[TOTAL] and the variable GVA-[G-I].

The procedure to follow consists, firstly, in each case, of identifying the quarterly values of each variable (GVA-[TOTAL] or GVA-[G-I]). Secondly, the annual values of GVA-[I] is broken down into quarters, following the same quarterly distribution as in the previous variables.

In this way, two new quarterly proxy series of the tourism variable [I] become available to estimate correlations with variables observable from space. The new series are represented in the tables as  $(GVA-[I]^{GVA-[TOTAL]})$  and  $(GVA-[I]^{GVA-[G-I]})$ .

Finally, the five variables selected to estimate correlations with variables observable from space are the following: 1° GDP; 2) GVA-[TOTAL]; 3) GVA-[G-I]; 4)  $(GVA-[I]^{GVA-[TOTAL]})$ ; and 5)  $(GVA-[I]^{GVA-[G-I]})$ .

Once the information has been extracted from EUROSTAT and the estimates made to approximate the quarterly values corresponding to the indicated variables, they have all been presented in Table 1 (at current prices) and in Table 2 (in chain linked prices).

Consequently, in this chapter, we have obtained five series of economic variables values that can be related to tourism. Their analysis and the study of correlations with the observable variables from space, in this case or in similar cases, enable us to predict the effects on the economic variables from the changes in the observable variables from space.

Table 1 a. Economic Variables to Correlate with Variables Observable from Space

Current prices, million euros

TIME	GDP QUARTERLY EU OFFICIAL DATA	GVA-[TOTAL] QUARTERLY EU OFFICIAL DATA	GVA -[G-I] QUARTERLY EU OFFICIAL DATA	GVA -[ I ] <sup>GVA-[TOTAL]</sup> ESTIMATED as % GVA-[TOTAL]	GVA -[ I ] <sup>GVA-[G-I]</sup> ESTIMATED as % GVA-[G-I]
2015-Q1	4.169,4	3.651,9	777,1	237,1	203,23
2015-Q2	4.543,5	3.997,4	1.000,1	259,5	261,55
2015-Q3	4.663,9	4.062,2	1.172,1	263,7	306,53
2015-Q4	4.567,2	4.003,3	951,6	259,9	248,86
2016-Q1	4.355,8	3.853,6	817,4	271,9	231,00
2016-Q2	4.789,4	4.206,2	1.059,1	296,8	299,31
2016-Q3	4.951,6	4.317,9	1.265,2	304,7	357,55
2016-Q4	4.917,0	4.268,7	1.014,3	301,2	286,64
2017-Q1	4.716,0	4.146,7	874,3	297,0	246,34
2017-Q2	5.133,9	4.511,4	1.164,7	323,1	328,16
2017-Q3	5.284,9	4.582,4	1.376,2	328,2	387,76
2017-Q4	5.177,5	4.499,9	1.094,7	322,3	308,44
2018-Q1	5.041,1	4.382,3	975,8	322,7	276,02
2018-Q2	5.557,7	4.768,7	1.276,3	351,2	361,02
2018-Q3	5.585,9	4.898,1	1.478,4	360,7	418,18
2018-Q4	5.490,5	4.842,9	1.187,8	356,6	335,98
2019-Q1	5.350,0	4.702,0	1.006,2	332,1	283,65
2019-Q2	5.880,9	5.122,7	1.319,1	361,8	371,85
2019-Q3	6.080,9	5.293,6	1.531,3	373,9	431,67
2019-Q4	5.866,0	5.209,4	1.236,6	368,0	348,60
2020-Q1	5.568,0	4.928,7	970,6	111,2	116,08
2020-Q2	5.093,3	4.558,3	745,6	102,9	89,17
2020-Q3	5.730,7	5.019,4	1.015,0	113,3	121,39
2020-Q4	5.694,7	5.164,4	981,2	116,6	117,35
2021-Q1	5.742,2	5.071,4	894,9	222,0	180,98
2021-Q2	6.078,3	5.393,6	1.133,5	236,1	229,24
2021-Q3	6.523,8	5.761,4	1.439,9	252,2	291,20
2021-Q4	6.583,3	5.829,5	1.304,9	255,1	263,90
2022-Q1	6.462,1	5.679,8	1.164,0	325,1	281,89
2022-Q2	6.781,4	6.067,2	1.441,5	347,3	349,09
2022-Q3	7.398,3	6.399,9	1.719,6	366,3	416,44
2022-Q4	7.135,5	6.298,0	1.452,2	360,5	351,68
2023-Q1	7.018,6	6.183,4	1.289,5	:	:
2023-Q2	7.288,6	6.491,3	1.573,8	:	:
2023-Q3	7.845,7	6.814,7	1.834,6	:	:
2023-Q4	7.654,4	6.682,7	1.532,5	:	:
2024-Q1	7.469,8	6.592,0	1.347,9	:	:
2024-Q2	7.824,8	6.905,5	1.651,2	:	:

EUROSTAT: Data extracted on 04/09/2024

Table 1 b. Economic Variables to Correlate with Variables Observable from Space

Chain linked volumes (2015), million euros

TIME	GDP QUARTERLY EU OFFICIAL DATA	GVA-[TOTAL] QUARTERLY EU OFFICIAL DATA	GVA-[G-I] QUARTERLY EU OFFICIAL DATA	GVA-[I] <sup>GVA-[TOTAL]</sup> ESTIMATED as % GVA-[TOTAL]	GVA-[I] <sup>GVA-[G-I]</sup> ESTIMATED as % GVA-[G-I]
2015-Q1	4.190,8	3.668,0	793,4	238,1	207,49
2015-Q2	4.553,2	3.993,9	999,6	259,3	261,42
2015-Q3	4.614,4	4.047,3	1.153,6	262,7	301,69
2015-Q4	4.585,6	4.005,6	954,3	260,0	249,57
2016-Q1	4.471,2	3.911,1	837,8	268,3	231,25
2016-Q2	4.844,2	4.248,9	1.064,5	291,5	293,82
2016-Q3	4.899,9	4.298,7	1.237,8	294,9	341,65
2016-Q4	4.908,5	4.289,2	1.022,5	294,3	282,23
2017-Q1	4.751,4	4.157,3	884,5	283,3	237,52
2017-Q2	5.132,8	4.502,8	1.161,3	306,8	311,85
2017-Q3	5.181,6	4.546,4	1.344,4	309,8	361,02
2017-Q4	5.154,5	4.504,2	1.104,1	306,9	296,49
2018-Q1	5.031,3	4.402,0	989,8	307,4	264,28
2018-Q2	5.400,1	4.737,4	1.269,4	330,8	338,94
2018-Q3	5.472,8	4.803,0	1.442,4	335,3	385,13
2018-Q4	5.458,1	4.772,6	1.192,1	333,2	318,30
2019-Q1	5.276,8	4.622,2	1.020,1	313,1	269,97
2019-Q2	5.672,4	4.983,1	1.305,1	337,6	345,39
2019-Q3	5.825,6	5.120,1	1.498,5	346,8	396,57
2019-Q4	5.770,3	5.052,4	1.238,6	342,3	327,79
2020-Q1	5.444,3	4.774,1	977,3	105,8	110,49
2020-Q2	5.049,8	4.439,3	750,9	98,4	84,89
2020-Q3	5.566,1	4.894,3	1.025,4	108,5	115,93
2020-Q4	5.708,7	4.999,6	992,3	110,8	112,18
2021-Q1	5.594,3	4.906,0	916,8	214,8	176,80
2021-Q2	5.870,4	5.160,5	1.143,1	226,0	220,45
2021-Q3	6.210,7	5.460,5	1.435,5	239,1	276,84
2021-Q4	6.251,0	5.474,1	1.273,4	239,7	245,57
2022-Q1	6.009,5	5.269,6	1.113,6	296,8	261,00
2022-Q2	6.263,0	5.505,5	1.324,5	310,0	310,43
2022-Q3	6.468,8	5.687,4	1.550,4	320,3	363,37
2022-Q4	6.395,6	5.600,7	1.312,7	315,4	307,66
2023-Q1	6.187,1	5.425,3	1.181,5	:	:
2023-Q2	6.401,5	5.627,1	1.416,9	:	:
2023-Q3	6.635,1	5.833,4	1.621,3	:	:
2023-Q4	6.533,1	5.720,9	1.379,5	:	:
2024-Q1	6.404,1	5.615,6	1.219,9	:	:
2024-Q2	6.631,2	5.829,0	1.464,8	:	:

EUROSTAT: Data extracted on 04/09/2024

#### **4. DATA COLLECTION AND SELECTION OF THE PHYSICAL VARIABLES OBSERVABLE FROM SPACE**

Earth Observation (hereafter EO) data from space has recently become a new source of information that allows addressing various research questions. In economics it can be used as a tool for the observation of the economic situation in selected areas: e.g., trade, employment, migration, production, and tourism all affect the surrounding environment, with a modulation on light intensity and on the concentration of pollutants with a direct effect on the air quality, all quantities and consequently variables observable from space. Those data, collected by satellites, can help as proxies for economic activity and show structural changes and/or breaks in the economic processes.

Among those data provided by EO are night-time light radiance, air and water quality, transport movements and naval traffic. When considering nightlights and air quality data, we can observe the dynamics of economic activity and its seasonal patterns: regions with higher economic activity produce higher nightlights radiance and, since mass tourist flows usually occur during mid-spring – summer months, values for nightlight data in resort areas will be much higher in those months than in winter. Nightlights might be used as a proxy for economic activity and to consider economic activity at specific locations such as regions, municipalities, city compounds, ports, and coastal areas. Even if the correlation between nightlights vs. economic activity may suffer from top-censoring and bottom-censoring, in the literature a strong positive correlation between any measure of it vs. luminosity density is shown. Air quality parameters from EO could be proxies for economic activity. The overwhelming majority of factories and vehicles produce particulate grains and pollutant gases, so usually industrial development and air pollution are positively correlated. We can consider some types of gases and particulate size, corresponding to different pollution sources and human activities.

Data is collected by EO satellites and processed for calibration and distortion, then the images are selected and data extracted, polished and made available for analysis. In the following chapters the data used, the processing infrastructure and methodologies, the software dependencies and the data products results will not be discussed. Data statistical issues will be addressed later.

## **5. THE CASE OF CYPRUS.**

The spaceborne EO data sets used to test our hypotheses were selected in the framework of the MICROS (“MonItoring CypRus ecOnomy by Space”) project, based on multiband data collected by several international EO-dedicated satellites. The project was conceived on the basis that, as seen in previous studies, there are a significant number of indirect parameters observable from space (EO parameters) that can be correlated to macro and micro economy indicators providing proxies for economy trends, to be used as indicators during normal time as well as during crisis periods to assess the impact of policy guidelines and laws as well as major crisis events to the economy at different scale and geographic areas. The idea was that classic environmental parameters and man induced impact on the environment (e.g., air pollution, land use change) can be combined with population information and economic parameters of human activities impacted by policies and events including industry, tourism and trades. These EO parameters may include monitoring of the levels of air pollution changes, suspended pollutants, percentage of transport container occupancy in the yard of industrial harbors etc., which could be utilized for specific business sectors (e.g. tourism) and city growth – economy growth. At the next level, all these “observed parameters” need to be correlated to indicators related to economy at different geographical scales. At medium- and long-term scale, the methodology may enable a near real-time monitoring of macro and microeconomic indicators.

MICROS addressed the use of remote sensing assets and EO data, combined with image processing techniques and economy modeling to assess a series of economic indicators suitable for monitoring near real-time economy trends at different geographical scales and so analyzing the impact of policies, laws and crisis events to real economy. MICROS had the potential to provide these near real-time indicators of the impact on the economy in all countries and geographical areas in case of emergencies where direct information and data are not fully reliable or available and/or provided with large time delay or otherwise. In fact, current econometric models are based on statistical data provided with several months of delay in the best case and thus MICROS will investigate filling this temporal gap. MICROS aimed at providing a well-documented analysis of the relationship between economic statistical data and data from EO processing methodologies for the region of Cyprus, chosen both for its small size and for the good data availability for both statistics and EO observations. In particular, the project aimed

at testing whether the economic effects of COVID-19 pandemic were observable from space and whether the economic patterns could be reconstructed from space data.

If the economy of a region undergoes a shock, EO satellite data patterns should change with those parameters assumed as proxies for socio-economic factors. In the MICROS analysis they considered the effects of COVID-19 pandemic from its onset in March 2020. This crisis greatly affected Cyprus economy: tourist flow dropped severely, major employment sectors, linked to hospitality and food services, suffered from the demand decrease, trade and production rates felt due to quarantine regime for the employees. The economy of Cyprus is closely linked to tourism, so the cancelation of all tourist flights was a real blow to it. In the current analysis we aim at estimating the effect of COVID-19 pandemic on the socio-economic situation in Cyprus via econometric modeling and correlation analysis of EO data vs. socio-economic data.

The MICROS project built observed dataset dealing with five different general applications: Ship Detection, Truck (Freight Transport) Detection, Water Quality Assessment, General Nighttime Activity (Lights), Atmosphere Quality Assessment. The data used are large scale data, selected for different sectors but mainly focusing on tourism and trade, given the clear orientation of the Cyprus economy, to highlight the economic impact of lockdowns, specifically for the island of Cyprus which relies heavily on its maritime trade and activities. The MICROS project succeeded in estimating the effect of COVID-19 pandemic on the socio-economic situation in Cyprus via econometric modeling and correlation analysis of EO data vs. socio-economic data. In the current analysis we limit our attention to these data, both statistical and EO-based, related to tourism activities and our analysis limits our analysis to correlation patterns, with no use of time-series and/or multivariate econometric models.

## **6. OUR ANALYSIS FOR CYPRUS: DATA AND METHODOLOGY.**

The socio-economic statistical data for Cyprus used in this work are not the same as in the MICROS project. Whereas in the latter project the data came directly from Cyprus government institutions, i.e., ministries and statistical survey institutes, in this work data have been extracted from the EUROSTAT database. The EUROSTAT data for Cyprus are only at the national level, whereas in the MICROS project the EO data and several of the econometric data are at different geographical scales, i.e., selected areas of Interests (AOI), municipalities, at the district level, and at the global level.

As from the previous chapters, in this analysis we used the data for the Total GDP, the GVA-[TOTAL], and the GVA-[G-I], i.e., data for wholesale and retail trade, repair of motor vehicles and motorcycles, accommodation and food service activities, used as a proxy for tourism-related activities. We tried to match these national-level data with the EO data from the MICROS project, available for Cyprus for various geographical sets, i.e. selected areas, municipalities and at the district level.

The EO datasets for our analysis were obtained from satellite raw data as final products from the data processing infrastructure for the case study. To bridge the gap between the different scaling and usage of the economic and satellite data, we had to cope for every instrument of each satellite with its spatial and temporal resolution. Final data sets are obtained as monthly aggregated time-series during the lockdown period for the years 2018-2021, ready to be correlated with similar economic statistical data. A longer time range (2015-2024) is available for socio-economic statistical data, not for EO data.

The EO data used in our analysis, as any other experimental data, are affected by experimental error, defined as the difference between a measured or estimated value for a quantity and its true value, inherent to all measurements. Two sources of errors in measurements should be considered. The systematic error in a measurement is a consistent and repeatable bias or offset from the true value, typically the result of miscalibration of the equipment, or problems with the experimental procedure, or just the effect of the experimental set-up and the involved instrumentation, without any major mistake. On the other hand, variations between successive measurements made under apparently identical experimental conditions are called random errors and can occur in either the physical quantity being measured, or the measurement process, or both. There is no way to know where they originate in the measure cycle and must be dealt with by statistical techniques. The experimental error is the sum of the two error sources, and sometimes its size is comparable with that of the quantity being measured. Reported data from experimental results should always include a realistic estimate of their error, either explicitly, as plus/minus an error value, or implicitly, using the appropriate number of significant figures, ideally with an explanation of the reasoning and calculations that went into the error estimate. This means that each individual point of our EO data sets comes with an uncertainty over its actual numerical value, evaluated as the error bar, in the form of the standard deviation (i.e., the square root of the variance) of the measurement result.

Data from sophisticated instruments such as EO satellites always comes with their individual standard deviation.

Even the easiest correlation analysis is not an easy task, since we need a thorough analysis of variance and an effect size evaluation, both necessary when so noisy data such as EO are to be correlated with socio-economic data that show no experimental standard deviation. Whereas statistical significance shows that an effect exists in a study, practical significance shows that the effect is enough to be meaningful in the real world. Statistical significance is denoted by p-values, whereas practical significance is represented by effect sizes. Statistical significance alone can be misleading because it is influenced by the sample size. Increasing the sample size always makes it more likely to find a statistically significant effect, no matter how small the effect truly is in the real world. In distinction, effect sizes are independent of the sample size, and that is why it is necessary to report effect sizes to indicate the practical significance of a finding. by reporting effect sizes and confidence intervals wherever possible. Such large standard deviation values can allow for and explain any correlation even when this shows statistical significance due to a low value for the p coefficient. All correlations were built on average-only data, with no use and consideration of standard deviation for individual data points of physical observable variables. This means that at first all correlations were built by using the EO data average data values, so the results should be taken with caution.

The net effect of the experimental error is that scientists get data with a large amount of additional meaningless information in it, the “noise”, defined as any data that a user system cannot understand and interpret correctly. We need to evaluate the dependence of the results of our modeling effort on experimental noise and develop a procedure to take the noise in the data provided through EO into proper account. The width of the noise distribution at a given experimental condition provides a measurement of the reliability of the individual data points, and we used it to attain fits to the data (for both models and correlations) as statistically unbiased as possible.

A way to evaluate the effect size for EO-provided data makes use of the variance of each variable, as obtainable, e.g., in a standard ANOVA analysis. White noise in ANOVA simply means a sequence of samples uncorrelated with zero mean and finite variance, with no restriction on the distribution from which the samples are drawn. For our data instead we can obtain a pooled variance for those datasets involved in any fit analysis, where the mean and the variance of the samples of each dataset may be different, and

each individual sample may come with its own standard deviation. The effect size is often measured in terms of the proportion of variance that can be explained by variables, and we assessed the way to measure effect size from the ANOVA process for modeling and correlational studies. This way we can describe the proportion of variance explained in the dependent variable by a predictor while controlling for other predictors and assembling these individual proportion into functions able to compute the proportion of fit in the correlations than can be explained by the variables' behavior. For our analysis we used an approach in which countermeasures for most of the possible sources of bias in the effect size analysis are proposed and duly discussed. We do not describe here the mathematical procedure for the effect size analysis and evaluation, nor do we report all intermediate parameters needed to evaluate effect size. For each correlation we obtain two parameters:  $P_{\%}$ , i.e., the proportion (percent) of correlation that can be explained as a statistical fluke and therefore can just be generated by the experimental error in the form of the standard deviation for each individual point, and  $R_{\text{eff}}$ , i.e., the R-squared parameter rescaled for the effects due to the experimental standard deviations. The technique is then applied to our correlation results, to any individual fit. We show the results for the effect size in the next chapter of this paper.

## **7.-OUR ANALYSIS FOR CYPRUS: RESULTS.**

As said above, in this paper we limit our analysis to economic activities related to overall and tourism. Economic data for Cyprus as specified before at the national level from EUROSTAT have been tested vs. the EO data at different geographical scales. The results are evaluated both through higher correlation coefficients and through lower noise level generated by data pattern variance, as also shown by the effect size evaluation. This means that first all correlations were built by using, for the EO data, average data values, then a correction factor is applied to the results to deal with effect size.

For this paper we report results for Total GDP and GVA-[G-I] variables since the GVA-[TOTAL] variable gives similar results to those from Total GDP. Only one of the two chain-linked values will be used: the first (obvious) result is that the two chain-linked volumes units of measure modes, since they differ only by a constant factor, give the same correlation results for any location, with differences arising due to numerical roundup (e.g.,  $R = 0.775$  vs.  $R = 0.777$  in the two cases).

Meaningful correlation results for Total GDP data have been found only vs. nighttime mean radiance, i.e., nightlights and their time variations, analyzed with standard error (noise). For all other EO data available to us we found no meaningful correlation patterns. The EUROSTAT GDP data for Cyprus at the national level were tested against MICROS EO data at different geographical scales, i.e., selected areas of interest (AOI), municipalities and at the district level. As Areas of Interest (AOIs), we adopted the same areas as in the MICROS project for our analysis (Larnaca international airport, Limassol port, Limassol coastal area and Zygi Industrial Area), selected as potential locations for different kinds of economic activity: tourism activity (Larnaca international airport), international trade (Limassol port), residential activity (Limassol coastal area) and industrial activity (Zygi Industrial Area). The most interesting result is that for inhabited regions such as e.g., districts, the current values not corrected for inflation give significantly better correlation results. This pattern shows that people live in current times with current prices and scenarios, so theoretical corrections seem not to be relevant to their life and worsen correlation patterns. The opposite happens when special economic zones such as ports, airports, industrial zones are considered, where data corrected for inflation gave significantly better correlation results. The underlying logic is that these locations are more susceptible to the economic patterns and to the actual (i.e., adjusted) values rather than to current prices conditions. Higher increases are evident for pre-COVID epochs but are also present in post-COVID scenarios. The highest results are: **HIGH POSITIVE CORRELATION** of Nightlight Mean Radiance vs. Total GDP for the Larnaca Airport AOI before COVID-19 onset; **HIGH CORRELATION** of Nightlight Mean Radiance vs. Total GDP for the Limassol Coastline after COVID-19 onset; **HIGH POSITIVE CORRELATION** of Nightlight Mean Radiance vs. Total GDP for the Nicosia Municipality after COVID-19 onset; **HIGH NEGATIVE CORRELATION** of Nightlight Mean Radiance vs. Total GDP for the Limassol Municipality after COVID-19 onset; no significant correlation results have been obtained for the district of Ammochostos. The full result set is shown in Tables 2a, 2b and 2c for special zones, municipalities and districts respectively. In red is reported the highest correlation value for each analysis. [(AOIs) = Areas Of Interest]

**Table 2a. Correlation of National-level Quarterly Total GDP data vs. Nightlight Mean Radiance in Cyprus Special Zones (AOI).**

Current prices				Chain-linked volumes			
Total GDP vs. Nightlight Mean Radiance in AOIs				Total GDP vs. Nightlight Mean Radiance in AOIs			
AOI	Overall	Before Covid	After Covid	AOI	Overall	Before Covid	After Covid
Larnaca International Airport	0,25	0,71	0,22	Larnaca International Airport	0,35	0,82	0,28
Limassol Port	-0,48	0,47	-0,64	Limassol Port	-0,49	0,52	-0,80
Limassol Coastline	0,31	0,47	0,75	Limassol Coastline	0,40	0,60	0,75
Zygi Industrial Area	0,34	-0,06	0,07	Zygi Industrial Area	0,40	-0,06	0,15

**Table 2b. Correlation of National-level Quarterly Total GDP data vs. Nightlight Mean Radiance in Cyprus Main Municipalities.**

Current prices				Chain-linked volumes			
Total GDP vs. Nightlight Mean Radiance in Municipalities				Total GDP vs. Nightlight Mean Radiance in Municipalities			
Municipality	Overall	Before Covid	After Covid	Municipality	Overall	Before Covid	After Covid
Larnaca	-0,31	0,20	0,60	Larnaca	-0,22	0,11	0,50
Limassol	-0,33	0,55	-0,74	Limassol	-0,24	0,52	-0,60
Nicosia	0,61	0,18	0,74	Nicosia	0,57	0,10	0,73

**Table 2c. Correlation of National-level Quarterly Total GDP data vs. Nightlight Mean Radiance data in Cyprus Districts.**

Current prices				Chain-linked volumes			
Total GDP vs. Nightlight Mean Radiance in Districts				Total GDP vs. Nightlight Mean Radiance in Districts			
District	Overall	Before Covid	After Covid	District	Overall	Before Covid	After Covid
Larnaca	-0,38	-0,12	0,15	Larnaca	-0,29	-0,02	0,12
Limassol	-0,30	0,19	-0,26	Limassol	-0,21	0,10	-0,25
Nicosia	-0,28	-0,05	-0,09	Nicosia	-0,19	0,04	-0,07
Paphos	0,45	0,37	0,57	Paphos	0,41	0,31	0,55

As for the GVA-[G-I] data, meaningful results for correlation have been found not only for nighttime mean radiance, i.e., nightlights and their time patterns, but also for particulate (both PM2.5 and PM10, even if such data were available for the Nicosia district only), all variables analyzed with their standard error (noise). For all other EO data available we found no meaningful correlation patterns. The EUROSTAT GVA-[G-I] data for Cyprus at the national level were tested against MICROS EO data at different geographical scales, i.e., selected areas (AOI), municipalities and at the district level. For nightlights, again for inhabited regions current values not corrected for inflation give better correlation results, with theoretical corrections not relevant and worsening correlation patterns. The opposite happens when special economic zones such as ports, airports, industrial zones are considered, where data corrected for inflation gave significantly better correlation results with respect to the uncorrected ones. The effects are the same as for Total GDP but to a lower extent: the improvement in the results is not so evident as in the previous cases. For particulate we found slightly better results for the current price scenario rather than for chain linked values, with smaller enhancement than that found for the nightlights for GDP data. The full result set is shown in Tables 3a, 3b, 3c and 3d for special zones, districts, municipalities and particulate respectively.

	Current prices	Chain-linked volumes
<b>Larnaca Intern. Airport</b>	0.71	0.80
<b>Limassol Port</b>	0.61	0.65
<b>Limassol Coastal Area</b>	0.62	0.63
<b>Zygi Industrial Area</b>	0.22	0.29

**Table 3a. GVA-[G-I] vs. Nightlight  
Mean Radiance in AOIs**

	Current prices	Chain-linked volumes
Larnaca Intern. Airport	0.71	0.80
Limassol Port	0.61	0.65
Limassol Coastal Area	<b>0.62</b>	<b>0.61</b>
Zygi Industrial Area	0.22	0.29

**Table 3b. GVA-[G-I] vs. Nightlight  
Mean Radiance in Districts**

	Current prices	Chain-linked volumes
Limassol District	0.66	0.64
Larnaca District	0.66	0.63
Ammochostos District	0.78	0.75
Paphos District	0.70	0.68
Nicosia District	0.52	0.50

**Table 3c. GVA-[G-I] vs. Nightlight Mean Radiance in Municipalities**

	Current prices	Chain-linked volumes
Nicosia Municipality	0.28	0.24
Limassol Municipality	0.36	0.33
Larnaca Municipality	0.27	0.25

**Table 3d. GVA-[G-I] vs. Particulate (PM10 and PM2.5) for the Nicosia District**

	Current prices			Chain-linked volumes			
	Overall	Before Covid	After Covid	Overall	Before Covid	After Covid	
PM10	0,20	0,28	-0,89	PM10	0,19	0,27	-0,89
PM2.5	0,42	- 0,52	-0,04	PM2.5	0,41	- 0,51	-0,03

A theoretical explanation of the need for effect size evaluation in our statistical analysis has already been given in the previous chapter. Now we report the effect size evaluation for the correlation analysis we carried out. For each correlation we computed two parameters:  $P\%$ , i.e., the proportion (percent) of correlation that can be explained as a statistical fluke and therefore can just be generated by the experimental error in the form of the standard deviation for each individual point, and  $R_{\text{eff}}$ , i.e., the R correlation coefficient rescaled for the effects due to the experimental standard deviations. Results are given below. In the tables a red background shows highly reliable results even after an effect size correction. We report effect size results for correlations with data for which experimental error bars are available for each individual point, i.e., nightlights mean radiance. In these tables we only report values for the  $P\%$  whereas in order to get values for the  $R_{\text{eff}}$  the above obtained R parameter need to be aptly rescaled. The results set for Total GDP is shown in Tables 4a, 4b and 4c for special zones, municipalities and districts respectively, whereas that for GVA-[G-I] is shown in Tables 5a, 5b and 5c for special zones, municipalities and districts respectively.

**Table 4. Results for Total GDP of Effect Size (P%) for nightlights mean radiance.**

	Current prices	Chain-linked volumes
Larnaca Intern. Airport	0.44	0.38
Limassol Port	0.78	0.73
Limassol Coastal Area	0.46	0.44
Zygi Industrial Area	0.89	0.81

**Table 4a. Results of Effect Size (P%) for Total GDP in Cyprus special zones (AOIs).**

	Current prices	Chain-linked volumes
Larnaca Intern. Airport	0.44	0.38
Limassol Port	0.78	0.73
Limassol Coastal Area	0.46	0.44
Zygi Industrial Area	0.89	0.81

**Table 4b. Results of Effect Size (P%) for Total GDP in Cyprus Municipalities.**

	Current prices	Chain-linked volumes
Nicosia Municipality	0.46	0.53
Limassol Municipality	0.45	0.51
Larnaca Municipality	0.65	0.69

**Table 4 c. Results of Effect Size (P%) for Total GDP in Cyprus Districts.**

	Current prices	Chain-linked volumes
Limassol District	0.91	0.94
Larnaca District	0.85	0.91
Paphos District	0.67	0.75
Nicosia District	0.92	0.96

**Table 5. Results for GVA-[G-I] of Effect Size (P%) for nightlights mean radiance.**

	Current prices	Chain-linked volumes
Larnaca Intern. Airport	0.38	0.35
Limassol Port	0.60	0.58
Limassol Coastal Area	<b>0.62</b>	<b>0.63</b>
Zygi Industrial Area	0.92	0.90

**Table 5a. Results of Effect Size (P%) for GVA-[G-I] in Cyprus special zones (AOIs).**

	Current prices	Chain-linked vol
Larnaca Intern. Airport	0.38	0.35
Limassol Port	0.60	0.58
Limassol Coastal Area	<b>0.62</b>	<b>0.63</b>
Zygi Industrial Area	0.92	0.90

**Table 5b. Results of Effect Size (P%) for GVA-[G-I] in Cyprus Municipalities.**

	Current prices	Chain-linked volumes
Nicosia Municipality	0.36	0.37
Limassol Municipality	0.51	0.52
Larnaca Municipality	0.46	0.47

**Table 5c. Results of Effect Size (P%) for GVA-[G-I] in Cyprus Districts.**

	Current prices	Chain-linked volumes
Limassol District	0.56	0.56
Larnaca District	0.56	0.57
Ammochostos District	0.37	0.39
Paphos District	0.50	0.51
Nicosia District	0.70	0.72

As said above, effect sizes are independent of the sample size, and that is why it is necessary to report effect sizes to indicate the practical significance of a finding by reporting effect sizes and confidence intervals wherever possible. Such large standard deviation values can allow for and explain any correlation even when this shows statistical significance due to a low value for the p coefficient. Just a few values are still significant after an evaluation of the effect size. These cases are discussed below.

## **8.-OUR ANALYSIS FOR CYPRUS: DISCUSSION.**

Results for the correlation analysis were reported in the previous chapter. The most affected sector of Cyprus economy during the pandemic was tourism: Cyprus did not accept foreign tourists from March to May 2020. We analyzed the Total GDP and GVA-[G-I] as related to nightlights and air quality data and find significant correlations for these factors vs. EO satellite data. As a main indicator of economic activity, GDP reflects structural breaks. As a consequence of the pandemic, earnings of the population decreased during the COVID-19 pandemic, especially for people employed in sectors such as those linked to accommodation services and all other activities connected to tourism and travel.

The choice of specific pairs of socio-economic parameters and EO observables is based on their relationships from theoretical economic and empirical analyses. We focused our analysis mostly on datasets and observables, which might illustrate significant changes during COVID-19 pandemic. As a starting point we use nightlights mean radiance and air quality data from EO. Economic factors analyzed and mostly affected by the pandemic are GDP, tourism activities and consequently its associated employment. The same as for the MICROS project, but in the present work only economic data from EUROSTAT and therefore at the national level have been used. These data are more easily obtained and collected than data from local authorities and ministries, and anyhow good significant correlation results and patterns have been obtained vs. satellite data at different geographic scales. This is a sound advantage with respect to the MICROS approach.

From our pilot exercise we pointed out the detrimental effect of experimental errors on correlation results, especially in the case of large error bars, through the effect size analysis. Our correlations still show indications that some EO-provided parameters may act as proxies for economic patterns and be used in correlations. In some cases, the fit was good so that even after the effect size evaluation those correlations could be retained as reliable representations. All correlations fits were built by using averaged values for

EO-provided data, therefore in such analysis the choice of the right size, timeframe, and location of the sample windows on EO imaging is crucial, to be able to use the full statistical power from the sampled data. We have shown that, with all its intrinsic limitations we pointed out, this is a viable approach to obtain proxies for economic activities and build correlations vs. EO data. This is an encouragement to apply this data analysis approach to other cases and realities.

## **9.-CONCLUSIONS.**

The procedure followed has consisted of establishing a series of requirements to select the appropriate variables. On the one hand, economic variables have been required, which can capture the impact of the Covid-19 crisis and also have abundant statistical information on them, for European countries for a long period of years.

To meet the first requirement, the macroeconomic variables have been selected, since they reflect very well the effects of economic crises. To do this, we have started by taking GDP and then all the others that are directly related to it. In this way, all the variables that have an immediate relationship with the GDP have been selected, either on the Demand or Supply side.

In this process, five variables of different levels have been chosen, both in current prices and in chain-linked volumes (adjusted for inflation). GDP is the first level variable, but when it is broken down into its components, other variables appear. These new variables, which we have called second and third level, are of outstanding importance, because they have a closer relationship with certain economic activities of interest in this work and could give better correlations.

Finally, two economic variables have been selected to make correlation estimates with variables observable from space. These are GVA-[G-I] and GVA-[I].

The GVA-[G-I] has the advantage that there are statistical data for all the quarters of the period analyzed, but it has the disadvantage that it includes effects of other variables that are not tourism itself, although the effect of the variable that reflects tourism activity may be dragging the variations in the set GVA-[G-I].

On the other hand, the variable GVA-[I], refers exclusively to tourism activity, but has the disadvantage of not having quarterly data. In this case, an estimate has been made in

two ways, which should reflect a very realistic evolution of the variable GVA-[I] and therefore be suitable for correlations with other variables observable from space.

We tested for correlation data from EUROSTAT for Cyprus (Total GDP, GVA-[G-I]) vs. MICROS EO data. Meaningful results for all economic data were found only for correlations vs. nighttime mean radiance i.e., nightlights and their time patterns, whereas only for GVA-[G-I] also for particulate (both PM2.5 and PM10, even if for the Nicosia district only). All EO variables were analyzed with their standard error (noise). For all other EO data available to us we found no meaningful correlation patterns. The EUROSTAT data for Cyprus at the national level were tested against MICROS EO data at different geographical scales, i.e., selected areas (AOI), municipalities and at the district level.

For nightlights, for inhabited regions current values not corrected for inflation give better correlation results, with theoretical corrections not relevant and worsening correlation patterns. The opposite happens when special economic zones such as ports, airports, industrial zones are considered, where data corrected for inflation gave significantly better correlation results. The effects are evident for Total GDP, lower for GVA-[G-I] data. For particulate we found slightly better results for the current price scenario rather than for chain-linked values with a smaller enhancement than that found for the nightlights for GDP data. Such increases are also evident for both the pre-COVID and the post-COVID scenarios as well. The underlying logic seems to be that inhabited zones are more susceptible to current prices conditions, since people live in current times with current prices and scenarios, so theoretical corrections seem not to be relevant to their life and worsen correlation patterns. On the opposite the special economic zones are more susceptible to economic patterns and actual (i.e. adjusted) values rather than to current prices conditions.

The most important result of our work was showing that with wisely collected EO data at different space scale and resolution (e.g., special zones, cities, regions) good correlation patterns could be established vs. national-level economic data from EUROSTAT. This is a remarkable asset of the present methodology with respect to models from previous work. Another asset of this work is that no use has been made of AI tools: with usual techniques good results could be anyhow achieved. We showed how far, with simple and cheap techniques and tools, such a study can go when its rationale is solid. The proved direct association between EO-provided parameters and national-level economic

parameters makes our conclusions about socioeconomic parameters stronger and further demonstrates the correctness and resiliency of our approach.

The main limitation of this example is that data and results are only for Cyprus and its peculiar situation (EO and economic data not collected at the same scale). It would be great to test this overall approach in the framework of EYE for some different and larger countries with different data patterns and specificities.

**REFERENCES:**

C. CLARK, 1940. The Conditions of Economic Progress. Macmillan, London, United Kingdom.

G. DE ANGELIS, E.V. SEMERIKOVA, A.O. BLOKHINA, 2022. “Analysis of Econometric Models”. In CH. KONTOPOULOS, V. CHARALAMPOPOULOU, G. DE ANGELIS, S. ZERVOS, 2022. MICROS – Monitoring CypRus ecOnomy by Space, ESA Contract No. 4000136060/21/NL/SC, Final Report, pp. 13-28, European Space Agency, ESA/ESTEC, Noordwijk, The Netherlands.

G. DE ANGELIS, E.V. SEMERIKOVA, A.O. BLOKHINA, 2022. “Finalized Correlation Analysis”. In CH. KONTOPOULOS, V. CHARALAMPOPOULOU, G. DE ANGELIS, S. ZERVOS, 2022. MICROS – Monitoring CypRus ecOnomy by Space, ESA Contract No. 4000136060/21/NL/SC, Final Report, pp. 29-52, European Space Agency, ESA/ESTEC, Noordwijk, The Netherlands.

V. EDO HERNANDEZ 2018. Introducción a la Economía: Del Dinero a los Recursos Naturales (2º Edition). Editorial Delta Publications, Madrid, Spain.

ESA-2010. European System of Accounts, Official Journal of the European Union (OJEU) 26-06-2013. Product code: KS-GQ-23-002. ISBN 978-92-76-98618-8. European system of national and regional accounts in the European Union

CH. KONTOPOULOS, V. CHARALAMPOPOULOU, G. DE ANGELIS, S. ZERVOS, 2022. MICROS – Monitoring CypRus ecOnomy by Space, ESA Contract No. 4000136060/21/NL/SC, Final Report, pp. 29-52, European Space Agency, ESA/ESTEC, Noordwijk, The Netherlands.

CH. KONTOPOULOS, TH. PAPAKOSMAS, V. CHARALAMPOPOULOU, G. DE ANGELIS, S. ZERVOS, 2022. “Addressing the use of EO and remote sensing technologies to assess economic indicators and identify economy trends - MonItoring CypRus ecOnomy by Space (MICROS)”, paper presented at the 41st European Association of Remote Sensing Laboratories (EARSeL) ‘Earth Observation for Environmental Monitoring’ Symposium (Paphos, Cyprus, 13-16 September 2022).

NACE-2008/NACE-2022. Statistical classification of economic activities in the European Community NACE. Rev. 2. Luxembourg: Office for Official Publications of the European Communities, ISBN 978-92-79-04741-1. European Communities.