ΧΑΡΟΚΟΠΕΙΟ ΠΑΝΕΠΙΣΤΗΜΙΟ
HAROKOPIO UNIVERSITY

HAROKOPIO UNIVERSITY OF ATHENS, GREECE

School of Environment, Geography and Applied Economics
Department of Geography

Master on Applied Geography and Spatial Planning,
Sector: Geoinformatics

Study of spatial phenomena for Europe via night-time images

Master Thesis

Chrysovalantis-Antonios Tsiakos
A.M. 215317

Athens, 2018

ΧΑΡΟΚΟΠΕΙΟ ΠΑΝΕΠΙΣΤΗΜΙΟ
HAROKOPIO UNIVERSITY

HAROKOPIO UNIVERSITY OF ATHENS, GREECE

School of Environment, Geography and Applied Economics
Department of Geography

Master on Applied Geography and Spatial Planning,
Sector: Geoinformatics

Committee

Prof. Christos Chalkias (Supervisor)
Professor, Department of Geography, Harokopio University

Prof Isaak Parcharidis
Associate Professor, Department of Geography, Harokopio University

Prof. Dimitrios Stathakis
Associate Professor, Department of Planning and Regional Development
University of Thessaly, School of Engineering, Volos

Ο Τσιάκος Χρυσοβαλάντης-Αντώνιος, δηλώνω υπεύθυνα ότι:

1) Είμαι ο κάτοχος των πνευματικών δικαιωμάτων της πρωτότυπης αυτής εργασίας και από όσο γνωρίζω η εργασία μου δε συκοφαντεί πρόσωπα, ούτε προσβάλει τα πνευματικά δικαιώματα τρίτων.

2) Αποδέχομαι ότι η ΒΚΠ μπορεί, χωρίς να αλλάξει το περιεχόμενο της εργασίας μου, να τη διαθέσει σε ηλεκτρονική μορφή μέσα από τη ψηφιακή Βιβλιοθήκη της, να την αντιγράψει σε οποιοδήποτε μέσο ή/και σε οποιοδήποτε μορφότυπο καθώς και να κρατά περισσότερα από ένα αντίγραφα για λόγους συντήρησης και ασφάλειας.

''Success is not final,
failure is not fatal:
it is the courage to continue that counts''

Winston churchill

ACKNOWLEDGEMENTS

Firstly, I would like to express my sincere gratitude to my supervisor Prof. Christos Chalkias for the continuous support of my MSc study and related research, for his guidance, motivation, and immense knowledge. Besides my supervisor, I would like to express my appreciation to:

- *Ms. Elena Krikigianni, for her ingenuine, selfless and incentivizing support and love during the whole time of my MSc studies, while standing continuously by me side in times of crisis;*
- *Prof. Dimitrios Stathakis, a friend, colleague and professor, for guiding me and supporting my involvement towards this exciting research field, since my undergraduate studies;*
- *My family, for their love and support throughout writing this thesis and my life in general.*

SUMMARY

Nowadays, the role of Spatial Data Infrastructures and of the relevant directives and policies (i.e. INSPIRE) is becoming more crucial when it comes to earth observation and scientific data. The process of managing the high volume of data provided by the various satellite systems (i.e. Copernicus, NOAA, etc) is complex and demands a standardized approach to cope with storage, processing, access and sharing of these datasets. Focusing on the datasets provided by the Suomi National Polar-Orbiting Partnership (S-NPP), the first tier of this thesis aims to implement a processing chain for these datasets so as to facilitate their utilization by end-to-end applications. Though this process, different aspects hindering the use of these data (such as the existence of background noise, the lack of data due to the removal of stray light in the original composites, etc) have been tackled allowing their utilization for further analysis. Following this, a prototype Spatial Data Infrastructure has been designed and implemented capable to manage these datasets efficiently, while being in compliance with existing directives and policies.

Key words**:** *VIIRS, night-lights, SDI, INSRIRE, OGC*

ΠΕΡΙΛΗΨΗ

Στις μέρες μας , ο ρόλος των υποδομών χωρικών δεδομένων και των σχετικών οδηγιών και πολιτικών (INSPIRE) αποκτά ολοένα και μεγαλύτερη σημασία όσον αφορά δεδομένα παρατήρησης της γης και επιστημονικά δεδομένα. Η διαδικασία διαχείρισης του μεγάλου όγκου δεδομένων που παρέχεται από τα διάφορα δορυφορικά συστήματα (δηλ. Copernicus, NOAA κ.λπ.) είναι πολύπλοκη και απαιτεί μια τυποποιημένη προσέγγιση για την αντιμετώπιση της αποθήκευσης, της επεξεργασίας, της πρόσβασης και της κοινής χρήσης αυτών των δεδομένων. Εστιάζοντας στα σύνολα δεδομένων που παρέχονται από το Suomi National Polar-Orbiting Partnership (S-NPP), η πρώτη βαθμίδα αυτής της εργασίας αποσκοπεί στην υλοποίηση μιας αλυσίδας επεξεργασίας για αυτά τα σύνολα δεδομένων, ώστε να διευκολυνθεί η αξιοποίησή τους από τελικές εφαρμογές. Μέσω αυτής της διαδικασίας, διάφορες πτυχές που παρεμποδίζουν τη χρήση αυτών των δεδομένων (όπως η ύπαρξη θορύβου, η έλλειψη δεδομένων λόγω της απομάκρυνσης του αδέσποτου/διάχυτου φωτισμού στα αρχικά σύνθετα υλικά κ.λπ.), αντιμετωπίστηκαν, επιτρέποντας τη χρησιμοποίησή τους για περαιτέρω ανάλυση. Στη συνέχεια, σχεδιάστηκε και υλοποιήθηκε ένα πρωτότυπο υποδομής χωρικών δεδομένων ικανό να διαχειρίζεται αποτελεσματικά αυτά τα σύνολα δεδομένων, το οποίο είναι παράλληλα και σύμφωνο με τις ισχύουσες οδηγίες και πολιτικές.

Λέξεις Κλειδιά**:** *VIIRS, νυχτερινά φώτα, SDI, INSRIRE, OGC*

Table of Contents

Table of Tables

Table of Figures

1. Introduction

Nowadays, there is a rapid growth in the production of digital spatial data which are collected, processed and used in various and diverse domains. In addition, the diffusion of the Internet and the widespread computer literacy have opened a genuinely new paradigm in spatial data handling, promoting data sharing across different communities and applications. However, challenges regarding the lack of availability, quality, organization, accessibility and sharing of spatial information remain common to many policies and activities and are experienced across various levels of public governance in Europe[1].

Spatial data Infrastructures (SDIs) have emerged as a logical consequence of these technological and societal challenges, aiming to facilitate search, acquisition, re-use and sharing of geographic information. Despite their possibilities towards this goal, SDIs' role is hampered by incomplete documentation, lack of compatibility among spatial datasets, inconsistencies in data collection, and cultural, linguistic, financial and organizational barriers. INPIRE Directive come into force in 2007 to address these challenges by establishing an Infrastructure for Spatial Information in the European Community (INSPIRE).

The role of Spatial Data Infrastructures and of the corresponding directives and policies is becoming more crucial when it comes to earth observation and scientific data. The process of managing the high volume of data provided by the various satellite systems (i.e. Copernicus, NOAA, etc.) is complex and demands a standardized approach to cope with storage, processing, access and sharing of these datasets. For instance, datasets provided by the Defense Meteorological Satellite Program (DMSP) and the Suomi National Polar-Orbiting Partnership (S-NPP) have been managed and processed in various ways and used for a wide range of applications. Literature provides many cases were different techniques were applied for the production of common base products whilst the latter are not available to be used by new applications and demand form the users the repetition of known processes. Based on this, the design and implementation of a system is required that will promote the use and to enhance the distribution of satellite data among the scientific community and the public institutions.

In particular the aim of this thesis is twofold:

1. To implement a processing chain for VIIRS DNB datasets so as to facilitate their utilization by end-to-end applications;
2. To design a Spatial Data Infrastructure capable to manage these datasets efficiently as well as to implement its key components in compliance with existing directives and policies;

---

[1]https://inspire.ec.europa.eu/documents/Data_Specifications/IES_Spatial_Data_Infrastructures_(online).pdf

2. Nighttime Lights Imagery

Night-time light remote sensing provides an accurate, cost-effective and direct way to locate settlements or human activities and to visualize their spatial distribution and density. The DMSP Night Lights provide the largest continuous time series (since 1992) of global urban remote sensing products and have proven valuable in a wide range of scientific applications. However, low spatial resolution[2] coupled with the large pixel overlap in the data acquisition process[3] and the accumulation of geometric errors in the process of composing images[4], results in the exported night lights occupying a greater extent than the actual dimensions of terrestrial light sources. This phenomenon is known in the literature as 'glowing' or 'blooming' and occurs in high brightness lighting areas such as natural gas firing facilities or large urban and industrial centres. In addition, it is enhanced by the dispersion of electromagnetic radiation (scattering) into the atmosphere from suspended particles or large molecules of atmospheric gases (Elvidge et al., 2007; Small et al., 2005). Finally, the DMSP lights cannot be used to distinguish between different types of illumination within an urban environment or between different because of their low radiometric resolution (Elvidge et al., 2007). The problem of saturation, which occurs mainly in urban cores, is a major challenge in using night lights to identify differences between urban areas. Correcting or decreasing the saturation of night light values would improve their utility in intra-urban and inter-urban applications (Zhang et al., 2014).

Elvidge et al. (2011) identify 11 significant deficiencies, characterizing OLS sensor data:

1. Coarse spatial resolution;
2. Lack of on-board calibration;
3. Lack of systematic recording of changes in gain levels;
4. Limited dynamic range;
5. Limited radiometric capability (6-bit);
6. Operation at high gains contributes to the saturation of the signal in areas of intense brightness such as urban centres. There is a lack of low light imaging spectral bands suitable for discriminating lighting types;
7. Lack of ideal thermal infrared channel for detecting erratic fire. The available spectral channels are not suitable for discrimination of thermal sources of light;
8. Limited data capture and storage;

[2] The OLS has a spatial resolution of 2.7km, but in the final product (stable lights) the pixel values are redefined to 1km (resampling).
[3] This means that the light at one location can be recorded in more than one pixel.
[4] The data are recorded by the sensor in rows and are spatially determined by the navigation data. The values of the recorded data are displayed in a 1km analysis grid, which is essentially an approximation of the Earth's surface. Each transformation process involves errors.

9. Overestimation of illuminated areas (blooming effect) due to the spatial analysis of the data and the reflection of light from the adjacent areas. There is a lack of a well-defined Point Spread Function (SPF);
10. Lack of a well-defined instantaneous field of view (IFOV);
11. Lack of many narrow detection zones for light detection (small spectral resolution).

2.1 Suomi National Polar-Orbiting Partnership (S-NPP)

Suomi National Polar-Orbiting Partnership (S-NPP) launched on 28 October 2011 and is considered a bridge mission between NASA's Earth Observing System (EOS) and the Joint Polar-Orbiting Satellite System (JPSS) (Hillger et al., 2013). S-NPP data are mainly used in NOAA's operational weather forecasts, whilst provide continuity to NASA's research in climate change, Earth's energy budget, and the global cycling of water and carbon (Straka et al., 2015).

S-NPP orbits the earth at roughly 834km altitude and performs a single orbit in approximately 101minutes, achieving a daily global coverage at 14 orbits. It carries five earth-observing sensors: The Visible Infrared Imaging Radiometer Suite (VIIRS), the Cross-track Infrared Sounder (CrIS), the Advanced Technology Microwave Sounder (ATMS), the Ozone Mapping and Profiler Suite (OMPS), and the Clouds and the Earth's Radiant Energy System (CERES).

The Visible Infrared Imaging Radiometer Suite (VIIRS), the primary imaging instrument on S-NPP, includes an expanded set of visible and infrared spectral bands (0.4 – 1.2μm) with improved characteristics in respect to its operational predecessor (Advanced Very High-Resolution Radiometer). In addition, the day/night band of VIIRS has the ability to collect visible/near-infrared (500–900 nm spectral response) imagery during both day and night.

Figure 1 below, provides an overview of the bands of the VIIRS instrument. Specifically, it consists of moderate resolution bands (M) and imagery resolution bands (I). M bands highlighted in yellow are available by default as environmental data records (EDRs) in addition to sensor data records (SDRs). Components of true-color visible imagery are highlighted in red, green, and blue and M-band and I-band components of natural-color imagery are noted with R, G, and B (Hillger et al., 2013).

Calibrated and validated VIIRS imagery are provided in Hierarchical Data Format version 5 (HDF5) format via the National Oceanic and Atmospheric Administration (NOAA)'s Comprehensive Large-Array Data Stewardship System (CLASS; www.class.ncdc.noaa.gov).

VIIRS band	Central wavelength (μm)	Wavelength range (μm)	Band explanation	Spatial resolution (m) at nadir
M1	0.412	0.402–0.422	Visible	750 m
M2	0.445	0.436–0.454		
M3	0.488	0.478–0.488		
M4	0.555	0.545–0.565		
M5 (B)	0.672	0.662–0.682		
M6	0.746	0.739–0.754	Near IR	
M7 (G)	0.865	0.846–0.885		
M8	1.240	1.23–1.25	Shortwave IR	
M9	1.378	1.371–1.386		
M10 (R)	1.61	1.58–1.64		
M11	2.25	2.23–2.28		
M12	3.7	3.61–3.79	Medium-wave IR	
M13	4.05	3.97–4.13		
M14	8.55	8.4–8.7	Longwave IR	
M15	10.763	10.26–11.26		
M16	12.013	11.54–12.49		
DNB	0.7	0.5–0.9	Visible	750 m across full scan
I1 (B)	0.64	0.6–0.68	Visible	375 m
I2 (G)	0.865	0.85–0.88	Near IR	
I3 (R)	1.61	1.58–1.64	Shortwave IR	
I4	3.74	3.55–3.93	Medium-wave IR	
I5	11.45	10.5–12.4	Longwave IR	

Figure 1 VIIRS bands overview (Source: Hillger et al., 2013)

2.2 VIIRS Day/Night Band (DNB)

The day/night band (DNB) of VIIRS captures signals from the Earth and atmosphere—daytime solar and night-time lunar reflectance as well as both natural and artificial night-time light emissions. Its low-light imaging capabilities at night is particularly useful when lunar illumination is present to illuminate surface and atmospheric properties in the same way as sunlight during the day. The DNB offers nearly constant 750-m spatial resolution across the full 3000km wide swath and operates three independent stages of gain (low/middle/high) to provide calibrated radiances at 14-bit radiometric resolution (compared to the 6-bit digitization of DMSP/OLS) during the day, across the day/night terminator, and into the night side of Earth. The low-gain stage is used for the daytime portions of the orbit, the medium-gain near the day/night terminator (twilight), and the high-gain on the night side of the orbit (Hillger et al., 2013). The images are taken approximately at 01:30 in local solar time but they can vary by over an hour depending on latitude and the corresponding cycle. As an outcome of its overpass time (near midnight) high latitude sites are imaged during night for a greater portion of the year (Kyba et al., 2015).

The primary mission of DNB band lies in the production of images of clouds constantly for the day, night and near terminator (twilight) scenes observed by VIIRS. However, like OLS, it provides a wide range of applications due to its enhanced capabilities, including the use of on-board solar diffuser calibration (Mills et al., 2010). Thus, DNB provides opportunities that were not possible with OLS. In particular, one of the key advances of DNB over the OLS is its spatial resolution. OLS's "smooth" operation mode captures data at 0.56km horizontal sampling interval (pixel-to-pixel spacing). The significantly smaller horizontal spatial resolution of DNB improves its spatial detail in comparison with OLS (Miller et al., 2013).

The DNB capitalizes on the collocation with the multispectral measurements on VIIRS and the other S-NPP sensors in order to provide accurate geolocation. In addition, DNB's high spatial resolution is the outcome of a unique scan angle dependent sub-detector aggregation strategy, establishing an improved pointing accuracy over OLS. The backside-illuminated charge couple device that forms the pixel elements is aggregated into 32 zones through each half of the instrument swath on either side of the nadir. The aggregation zones near the end of scan have fewer pixels than the ones near nadir, due to the fact that the footprint of a single charge couple device detector element on the ground is much larger at the end of scan (Shao et al., 2013).

DNB is considered as a de facto radiometer, as it uses an on-board calibration system, in comparison with DMSP-OLS, which operates as an imager, without a calibration system on-board (Jing et al., 2016). VIIRS collects data in four separate view windows, Earth view (EV), blackbody (BB), solar diffuser (SD), and space view (SV), successively at each scan (figure 2). The adopted radiometric calibration framework utilizes data constantly collected from VIIRS on-board calibrators' files (BB, SD, SV) to determine the gain and offset parameters over time (Lee et al., 2016).

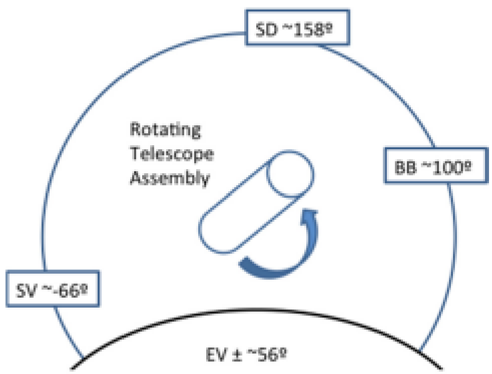


Figure 2 Schematic of VIIRS data collection windows and their approximated angles

The DNB sensor consists of four sectors of detectors: the low gain stage (LGS), the intermediate gain stage (MGS) and two redundant arrays, high-gain A (HGA) and high-gain B (HGB), are the high gain stage (HGS). During the daytime operation, the DNB takes advantage of the on-board calibration achieved via the Solar Diffuser. Observations recorded while the Solar Diffuser is not fully illuminated are used to compute cross-stage gain ratios (MGS/HGA, MGS/HGB and LGS/MGS) using signals within the dynamic ranges of both the higher and the lower gain stages (Lee et al., 2014). The calibrated nature of DNB night-time visible data, provided in units of in-band radiance, represents a significant advance over the uncalibrated OLS—enabling the first quantitative applications from low-light visible observations.

Moreover, DNB utilizes time delay integration (TDI) and a panchromatic band allowing the detection of dim signals within its spectral pass. As of its nominal minimum detection in-band radiance (3*10-5*W*m-2 s*r-1) DNB has very high sensitivity to small amounts of lights present in its band pass and can detect the light emitted from a single isolated point (Miller et al., 2013) (figure 3).

Attribute	DMSP/OLS	VIIRS/DNB
Orbital Details	Sun-synchronous, ~850 km	Sun-synchronous, 824 km
Nighttime Nodal Overpass Time	~1930 UTC	~0130 UTC
Swath Width	3000 km	3000 km
Spectral Passband Bandwidth	Panchromatic 500–900 nm	Panchromatic 500–900 nm
Spectral Passband Center	~600 nm	~700 nm
Horizontal Sampling Interval	2.8 km (Nighttime "Smooth" Data)	0.740 ± 0.043 km (Scan) 0.755 ± 0.022 km (Track)
Horizontal Spatial Resolution	5 km (Nadir)/~7 km (Edge)	<0.770 km (Scan) <0.750 km (Track)
Geolocation Uncertainty	~450 m–5.4 km	266 m (Nadir), 1151 m (Edge)
Minimum Detectable Signal	4×10^{-5} $W{\cdot}m^{-2}{\cdot}sr^{-1}$	3×10^{-5} $W{\cdot}m^{-2}{\cdot}sr^{-1}$
Noise Floor	~5×10^{-6} $W{\cdot}m^{-2}{\cdot}sr^{-1}$	~5×10^{-7} $W{\cdot}m^{-2}{\cdot}sr^{-1}$ (Nadir)
Radiometric Quantization	6 bit	13–14 bit
Accompanying Spectral Bands	1 (Thermal IR)	11 (Night)/21 (Day)
Radiometric Calibration	None	On-Board Solar Diffuser
Saturation, Stray Light Artifcats	Urban Cores, Substantial, Uncorrectable	None, Near-Terminator, Corrected

Figure 3 Comparison between the Suomi National Polar-orbiting Partnership (S-NPP) Visible Infrared Imaging Radiometer Suite (VIIRS)/DNB and heritage DMSP/Operational Linescan System (OLS) (nighttime smooth-mode data) (Miller et al., 2013)

The main challenge of DNB's on-board calibration lies on the fact that the most useful observations have radiance values that are below the sensor's designed minimum observable radiance. Severe degradation of the quality of DNB data is occurred from stray light[5] contamination. The most intense stray light pollution is observed in the medium to high latitude regions, where the spacecraft crosses the northern and southern day/night terminators (Mills et al., 2013).

The regions affected by stray light can be determined by the Earth-Sun-spacecraft geometry and orbital inclinations, which change over time. In these regions, the spacecraft is under direct solar illumination due to its elevated orbital track at approximately 830 km over the night-time Earth surface. The large difference between the Sun and night-time Earth view radiances, leads to tiny fraction of sunlight entering the optical system and causing significant contamination of the images. This results to the domination of stray light in large swaths of the night-time scenes. The current VIIRS

[5] The stray light is unintended light entering the optical path and results in an overall increase in the recorded radiance values.

calibration framework includes stray light correction to remove the mid-to-high latitude stray light contamination (Lee et al., 2016).

Baugh et al. (2013) produced night time lights composites that combine high quality cloud-free data components over a period of time, aiming to improve sensitivity and coverage. The VIIRS DNB and M15 bands were pre-processed in order to create flag images. Flag images are 32-bit and are used to classify DNB pixels into classes. Each pixel can belong to more than one class by turning specific bits on. The flag categories include: daytime, twilight (terminator zone), stray light, lunar illuminance, noisy edge of scan data, clouds and no data (Baugh et al., 2013).

The daytime and twilight flags are calculated based on solar zenith angle[6], which is provided as a layer in the VIIRS DNB geolocation file. The daytime flag bit is set for solar zenith angles less than 96, while the terminator flag bit is defined for solar zenith angles between 96 and 101. This region covers the terminator or the transition zone from night-time to daytime, and is of reduced quality compared to the darker night time data. In order to define the pixels subject to stray light contamination (stray light flag), the ground based solar zenith angle is used as a proxy for space craft's solar zenith angle. In this case an entire scan light is flagged if the solar zenith angle at nadir is between 90 and 118.5. In addition, the DNB composites don't include the last 4 aggregation zones at the edges of each scan, due to the fact that show a visible increase in noise ('no data' flag). The same goes for fill scans that contain no real data. These regions are flagged as 'no data'. The lunar illuminance at the earth's surface is calculated as a function of lunar phase, azimuth and elevation. These parameters are computed based on the latitude and longitude of each DNB pixel and the time of the nadir pixel of each scan. Pixels with returned lunar illuminance less than 0.0005lux are placed in the zero-lunar illuminance flag (Baugh et al., 2013).

Last but not least, a critical aspect of the final composite is to include cloud-free data. The clouds affect both the intensity and location of lights. Thick clouds can 'hide' a light completely, while thinner clouds diffuse the signal, resulting in lights to appear larger but dimmer than they would be on clear conditions. Baugh et al (2013) investigated the use of IICMO, a JPSS retained intermediate product cloud mask. However, the IICMO cloud mask tends to recognise gas flares as clouds and thus it was not selected for the production of the DNB composite (Baugh et al., 2013). The latter, modify the algorithm used for DMSP OLS. Based on this procedure, observations with the presence of clouds are removed by comparing the data on the OLS thermal recording channel with the surface temperature of the earth. Data corresponding to areas with colder temperature in comparison with Earth's surface temperature represent clouds and are removed as the presence of clouds affects the intensity and

[6] It refers to angle defined between the zenith and the centre of the Sun's disc.

location of the lights on the visible recording channel (Baugh et al., 2010). VIIRS consists of 3 longwave infrared channels (M14, M15, and M16) as depicted in Table 1. M15 and M16 channels cover the first and second half of the OLS thermal recording channel. The M15 brightness temperature channel was utilised in the cloud detection algorithm. The reference surface temperature grids used for this procedure were interpolated both spatially and temporally, in order to meet the resolutions of VIILRS M15 band. However, the coarse spatial resolution of the reference surface temperature grid results in diffusion of the land/sea boundaries and subsequently in mis-identification of clouds along the land/sea surface. To this end, a look-up table was created covering each pixel of the land/sea boundary, which identifies the location of all-land and all-sea grid cells. Then for each VIIRS M15 pixel the position was compared against a 30 arc second land/sea mask in order to determine the use of either the land or the sea look up table. Subtracting the VIIRS M15 brightness temperature from the reference surface temperature, a thermal difference image is produced. The clouds are identified via various thresholds (Baugh et al., 2013).

Before the development the final composite, the flag images from DNB and M15 are re-projected into 15 arc seconds grids. The DNB and M15 detector arrays are slightly offset from each other and use different pixel aggregation schemes. Thus, they are combined into a final flag image.

3. Data Processing Method

In the context of this study, the VIIRS DNB monthly composites have been utilized, that have been filtered to exclude data impacted by stray light, lunar illumination, and cloud-cover. Cloud-cover is determined and removed via the VIIRS Cloud Mask product (VCM) and the data near the edges of the swath are not included in the composites (aggregation zones 29-32). The DNB monthly composites cover the globe from 75N latitude to 65S and are produced in 15 arc-second geographic grids. They are made available as a set of 6 tiles, where each tile is actually a set of images containing average radiance values and numbers of available observations. However, the version 1 series of monthly composites available has not been filtered to screen out lights from aurora, fires, boats, and other temporal lights. (https://ngdc.noaa.gov/). Also, the background noise has not been subtracted. For example, weak light reflected by snowcapped mountains and dry lake beds is recorded in the data too (Shi et al., 2014a).

In addition, there are several areas of the globe with data coverage of low quality for some months. This can be due to cloud-cover, especially in the tropical regions, or due to solar illumination, as happens toward the poles in their respective summer months. Thus, for cloud-free observations with a value of zero in the average radiance image, should not always be considered that no lights are observed (https://ngdc.noaa.gov/).

DNB monthly composites are available in two different configurations: The first excludes any data impacted by stray light (denoted as "vcm"), whilst the second includes these data if the radiance values have undergone the stray-light correction procedure (denoted as "vcmsl"). Due to the reduced quality of the "vcmsl" version, the first configuration was used in order to achieve better results in our seasonal monitoring application.

Initially, the 2nd tile was used where each image was cut off at the boundaries of the study area (European regions). Subsequently, the geodetic reference system was transformed into the ETRS'89-LAEA system. The European Terrestrial Reference System ETRS'89 is an adaptation of the global ITRS (International Terrestrial Reference System) reference system in the European space. The ellipsoid that uses ETRS'89 is GRS'80, while the cartographic projection used is Lambert's azimuthal equal-area.

It is noticed that there are some pixels with negative DN values in original NPP-VIIRS data. It is assumed that these negative values are caused by background noise and outliers from data processing (Shi et al., 2014a).

As a first step in the data processing method, an algorithm is used to remove the abnormal lights (outliers) that may be associated with fires and gas flares. Jing et al. (2016) used a mean algorithm to remove the abnormal lights (outliers) that may be associated with fires and gas flares. To this end, the highest DN value from the biggest cities in the study area is selected. Theoretically, the pixel values of the other areas should not exceed the latter and thus this value is used as a threshold to correct the outliers. Each pixel with DN value greater than this threshold value, is assigned as a new value. Shi et al. (2014a) assigned to these pixels (as a new value) the maximum DN value within pixel's immediate eight algorithms.

In our case, the maximum DN value of the 5 biggest European cities (London, Berlin, Rome, Paris, Madrid)[7] in calculated and assigned as the threshold value, based on which the outliers will be removed. Then a mean algorithm is used, calculating the new value based on the surrounding pixel values in 5x5 block area. However, the outliers of the image are not removed via this single application of the algorithm whilst they correspond to a quite small fraction of the whole image. While an option could be to set these pixels as NaN (not a number), it is preferred to iterate the algorithm multiple times (until the complete removal of the abnormal values) due to the fact that these pixels are located in the center of large cities and it would be "un-natural" to have blanc spots. Last but not least, the negative values of the images have been set up to 0.

Following this preliminary correction, two approaches are basically applied in the literature, in order to remove the background noise in DNB monthly composite data. The first one is referred as the noise masking method (NMM) and involves the removal of the background noise via a mask generated by DMSP/OLS stable lights that is applied to DNB composite data (Shi et al., 2014; Li et al., 2013; Jing et al., 2016). The DNB data are resampled to the same resolution as DMSP (30 arc seconds). Then, the pixels of DMSP stable lights with a positive value (DN>0) are extracted in order to generate a mask. The pixels of the DNB data that fall outside the mask are set as NaN (not a number), whilst the pixel value is kept the same for pixels inside the mask (Jing et al., 2016). Following a similar approach, Li et al (2013) multiplied the mask of DMSP/OLS data with the VIIRS image in order to derive noise-free VIIRS data.

The basic drawback of this method is the fact that relies on the DMSP stable lights to generate the mask. The lack of on-board calibration in combination with the saturation in areas of intense brightness and the blooming effect may lead to inaccurate results. In addition, the mask will exclude some low-light emission sources like small towns and road features that the DNB product is sensitive enough to pick up.

Thus, Jing et al. (2016) introduced the optimal threshold method to remove background noise. This method relies on the use of an object function to determine the optimal threshold. Given the fact that DMSP/OLS studies have already proved the close relationship between the light intensity and the type

[7] These cities are selected based on their population

of land use /cover, the latter attempt to use the correlation between light area and built-up area as the object function.

In the context of this analysis, the background noise will be removed via the well-established correlation of night light emissions with the population. The radiance values ranging from 0 to 1 10^{-9}Watts/cm2-sr and with pace of 0.1 are used as potential threshold values. Then, the sum of lights for each region is calculated for each of the selected threshold values. The correlation coefficient R^2 is calculated between the population of January 2016 and the EU countries. The threshold value with the highest R2 is used to remove background noise. In our case, pixels with value smaller than 1 nano-Watts/cm2-sr are removed from the monthly composites.

VIIRS monthly composites for May, June, July and August contain numerous pixels in high-latitude regions of the northern hemisphere with no data, as solar illumination seriously contaminates these regions in the summer. The situation is clearly reflected in figures 4 & 5 below, which presents a timeseries of monthly composites of 2013, focusing on United Kingdom, that have undergone the previous processing steps. In addition, a more quantitative analysis is undertaken, calculating the sum of lights in UK for the monthly composites of 2013 (figure 6). It is assumed that a country or region will not have major changes in the brightness of night lights throughout a year. Thus, in order to undertake this issue, computational techniques must be applied so as to forecast the no-data values in the images of the abovementioned months.

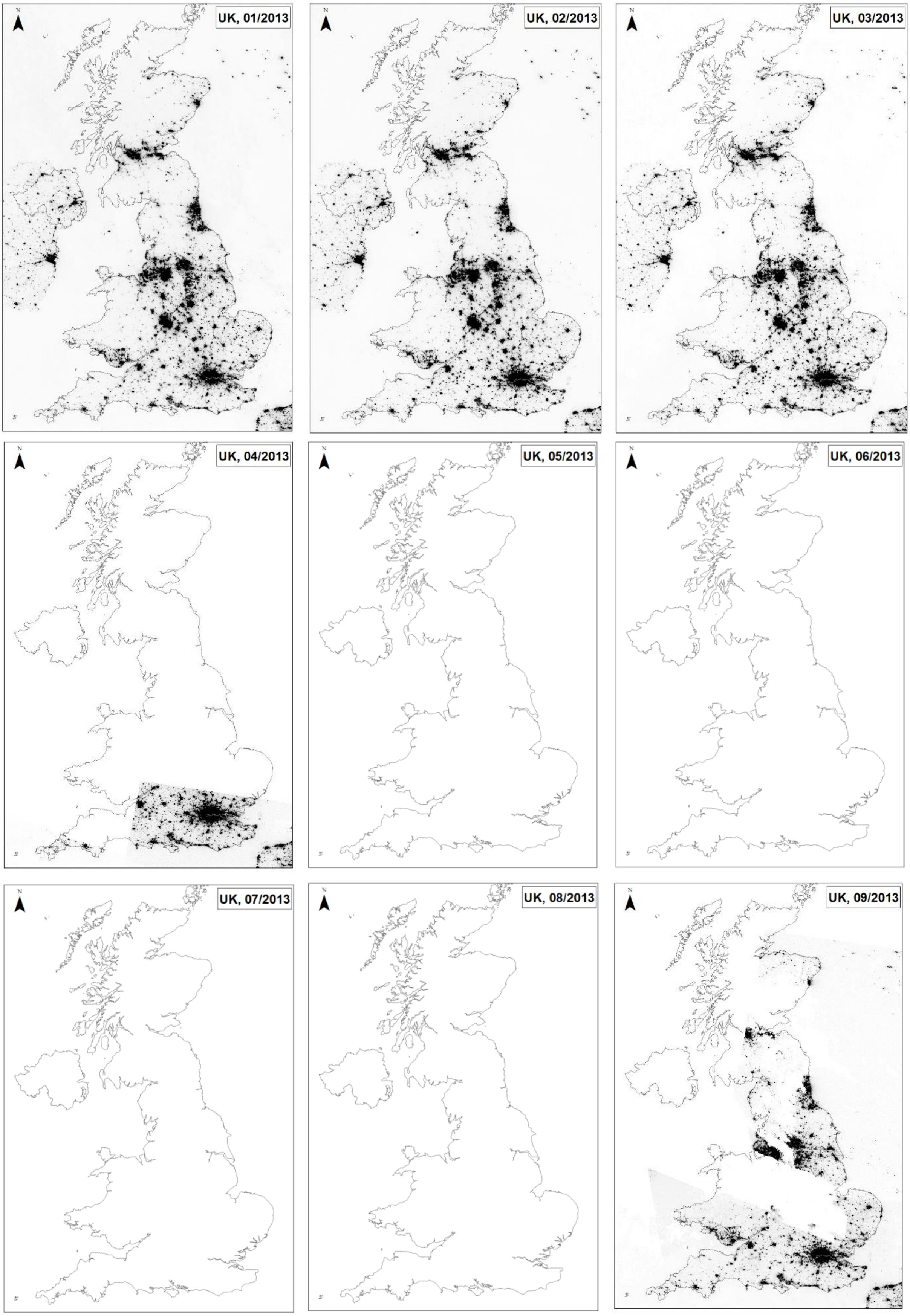


Figure 4 VIIRS Monthly Pre-processed Composites of UK for 2013 (01/2013-09/2013)

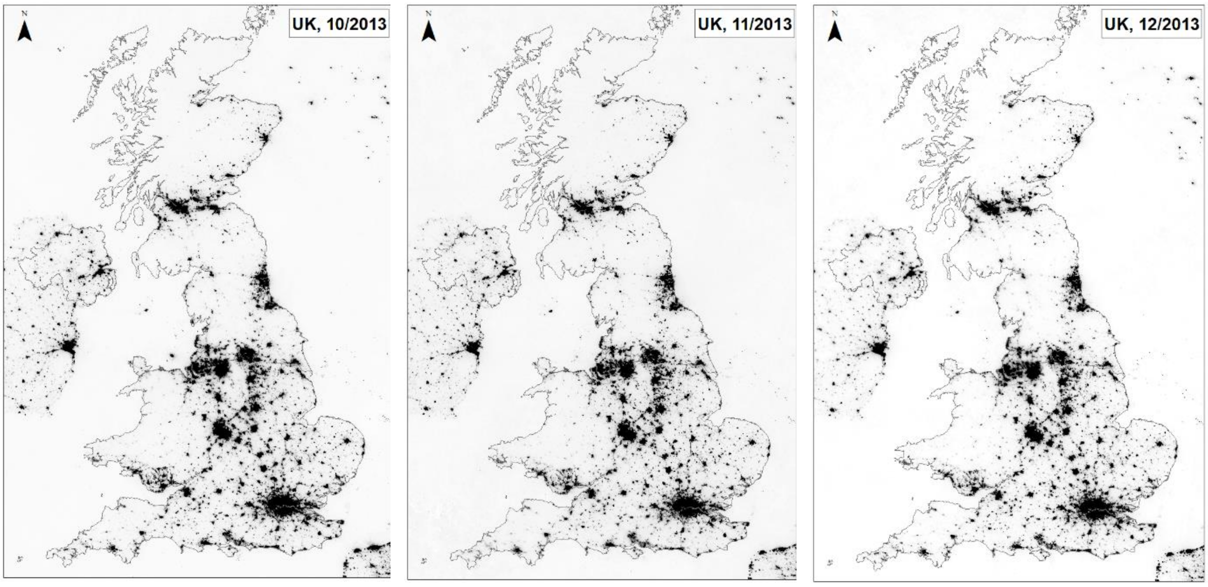


Figure 5: VIIRS Monthly Pre-processed Composites of UK for 2013 (10/2013-12/2013)

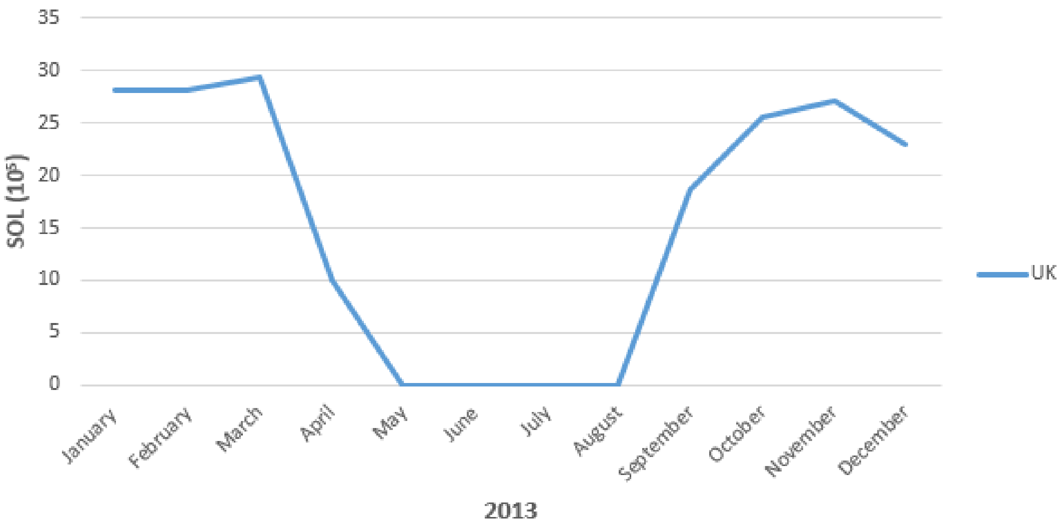


Figure 6: Monthly variation of Sum of Lights in UK for 2013

In the context of this analysis linear regression and exponential smoothing have been used and evaluated towards their applicability in calculating the missing values in the VIIRS imagery. The linear regression method is used to investigate the potential relationship between a variable of interest (often called the response variable but there are many other names in use) and a set of one of more variables (known as the independent variables or some other term).

Forecasts are obtained via the use of a simple linear model (1) as follows:

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x \qquad (1)$$

where x is the value of the predictor for which we require a forecast (i.e. pixels with known values) and y is the value of the corresponding forecast (pixels with missing values).

When this calculation is done using an observed value of x from the data, we call the resulting value of y a "fitted value". This is not a genuine forecast as the actual value of y for that predictor value was used in estimating the model, and so the value of y is affected by the true value of y. When the value of x is a new value (i.e., not part of the data that were used to estimate the model), the resulting value of y is a genuine forecast (Hyndman & Athanasopoulos, 2013).

The following image 7 presents an example of the output of forecasting for April 2013 via the use of a linear regression model. Due to the existence of no data pixels in the northern hemisphere, the municipality of Athens is selected as case study for the evaluation of the method.

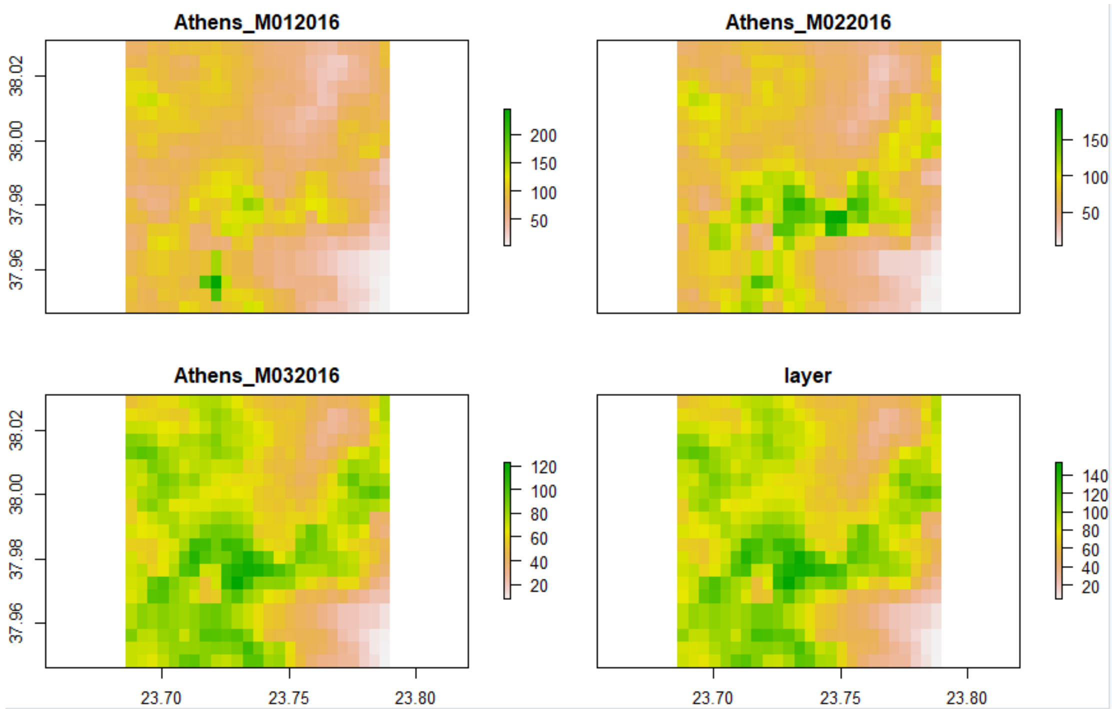


Figure 7: Forecasting missing values with linear regression

The second method implies the use exponential smoothing. Forecasts produced using exponential smoothing methods are weighted averages of past observations, with the weights decaying exponentially as the observations get older. In other words, the more recent the observation the higher the associated weight. This framework generates reliable forecasts quickly and for a wide spectrum of time series which is a great advantage and of major importance to applications in industry. The simplest of the exponentially smoothing methods is naturally called "simple exponential smoothing" (SES). (In some books, it is called "single exponential smoothing"). This method is suitable for forecasting data with no trend or seasonal pattern (Hyndman et al. 2014). In the context of this study, the Holt Winters exponential smoothing function was utilised.

Original monthly VIIRS DNB composites of January, February, and March, were stacked sequentially and then, each pixel had three-time series DN values applied. The following exponential smoothing equation (2) described by Holt (2004) is used in the context of this analysis:

$$\begin{aligned} S_t &= ax_t + (1-a)S_{t-1} \\ &= a\left[x_t + (1-a)x_{t-1} + (1-a)^2 x_{t-2} + \cdots + (1-a)^{t-1}x_1\right] + (1-a)^t x_0 \quad (0<a<1). \end{aligned} \tag{2}$$

where x denotes an observation value (i.e. the radiance of a pixel in an original monthly VIIRS DNB image), S represents a smoothed (or predicted) value, t denotes a period (i.e. a month in this study), and α is the smoothing factor. The abovementioned equation demonstrates that in exponential smoothing, a new smoothed (or predicted) value St is affected by a current observation value xt and the last period smoothed value St−1. Old observation values have an exponentially declined effect on the predicted value depending on period t. The value of α determines the smoothing effect. Large values of a (i.e. close to 1) contribute to a small smoothing effect, whilst smaller values of a (i.e. close to 0) indicate that the prediction is influenced considerably by both current and the previous observations.

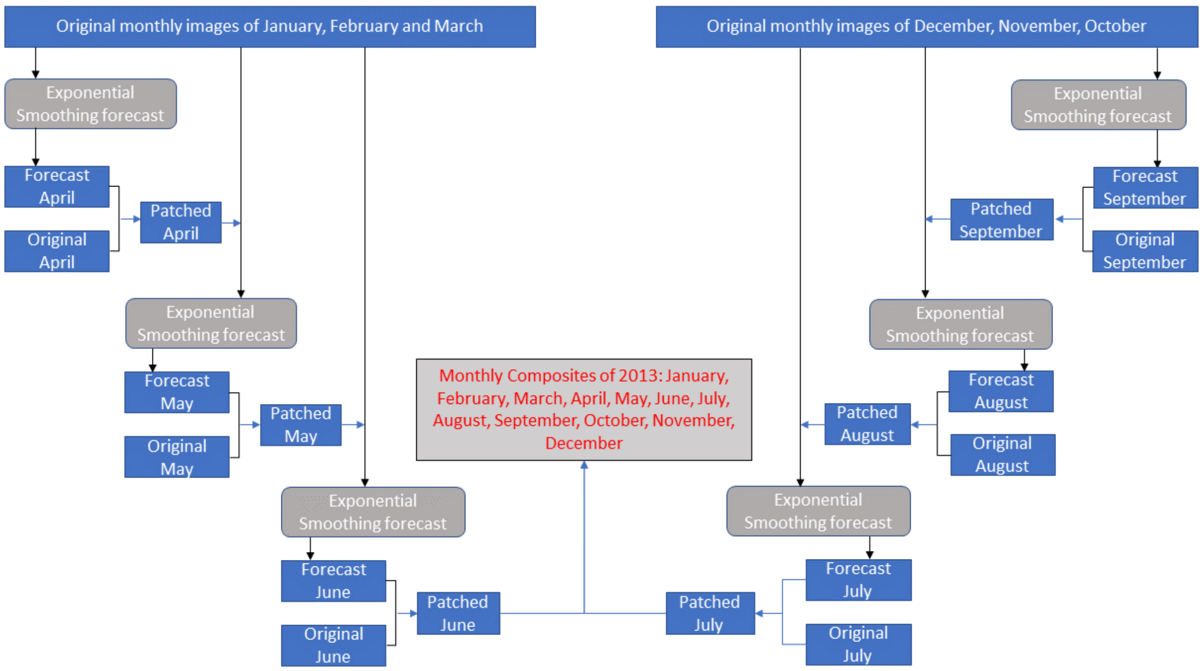


Figure 8: Flow chart for producing a "complete" time series of VIIRS DNB monthly composites

Usually, a region's economic level and population in 1 month are closer to its adjacent months than in later/earlier months. For instance, night time lights in April should be closer to the brightness levels of March, than the ones in February or January.

Since in the original April image, only a portion of pixels suffer from no data values, a patched image for April was produced by combining the original and the estimated April image. The patched April image was added into the stacked images of January, February, and March resulting in each pixel having four-time series radiance values. The same process was repeated for the images of May and June. Following the same approach, the original images of December, November, and October, were sequentially stacked to produce a patched image of September. The patched image of September was added into the stacked images of December, November, and October so as to produce a patched image of August. Similarly, the image of July was calculated. The overall flow chart of producing a "complete" time series of monthly VIIRS DNB composites is exhibited in figure 8 above.

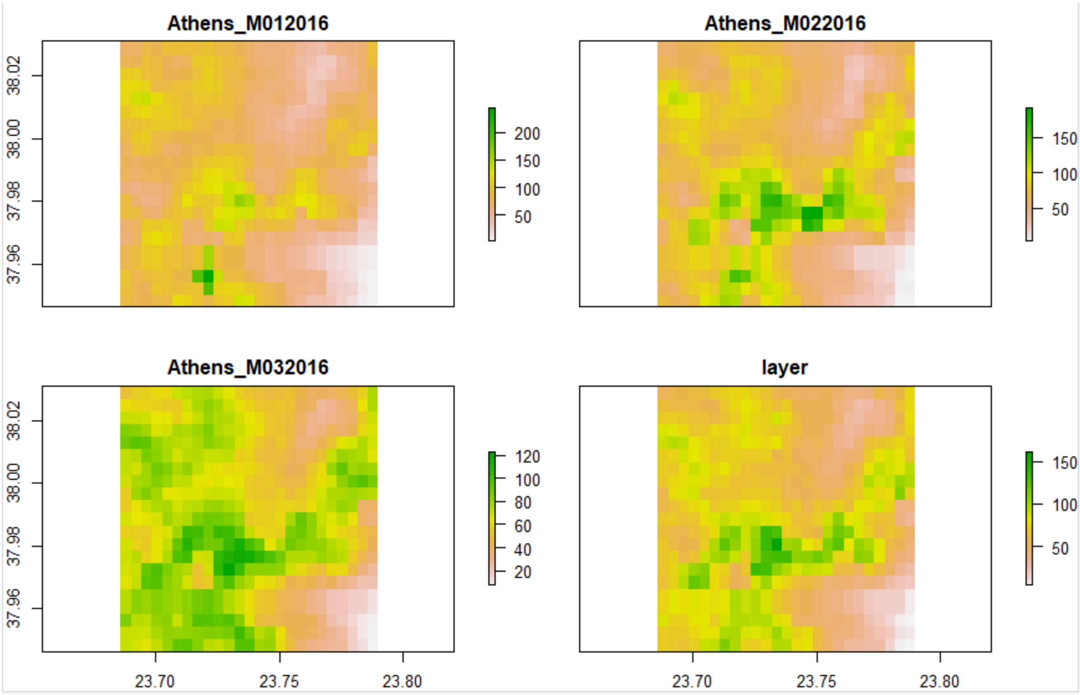


Figure 9: Forecasting missing pixel values with Holt Winters exponential smoothing

To support the evaluation of the Holt Winters forecasting method, the image of April 2013 is calculated for the Municipality of Athens, as also performed in the linear regression (figure 9). A visual comparison between the forecasted images from exponential smoothing and linear regression and the original image can be achieved in figure 10.

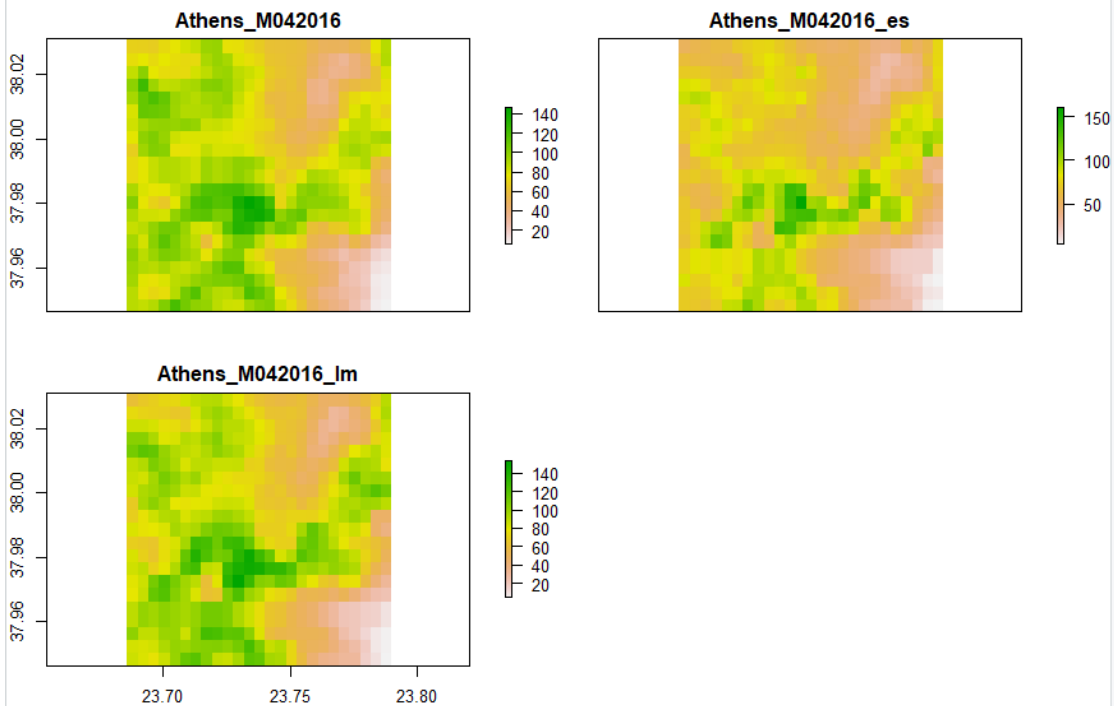


Figure 10: Visual comparison of the forecasted images from exponential smoothing and linear regression with the original image

It seems that the linear regression method provides results closer to the original image of April than the one of exponential smoothing. However, in order to obtain more accurate results, the Pearson correlation is calculated for each of the forecasted images with the original image (Table 1). In particular, a set of 100 random points was created in which values of cells from the 3 images were stored, for the defined locations. The linear regression appears higher values of R^2 (0.9142) in comparison with the Holt Winter's exponential smoothing. Thus, the linear regression method is used for the calculation of no data values in the monthly composites.

Pearson Correlation			
Variables	**M042016_Original**	**M042016_ Exponential Smoothing**	**M042016_ Linear Regression**
M042016_Original	1	0.8863	0.9142
M042016_Exponential Smoothing	0.8863	1	0.9604
M042016_Linear Regression	0.9142	0.9604	1

Table 1 Evaluation of the exponential smoothing and linear regression models - Calculation of Pearson correlation

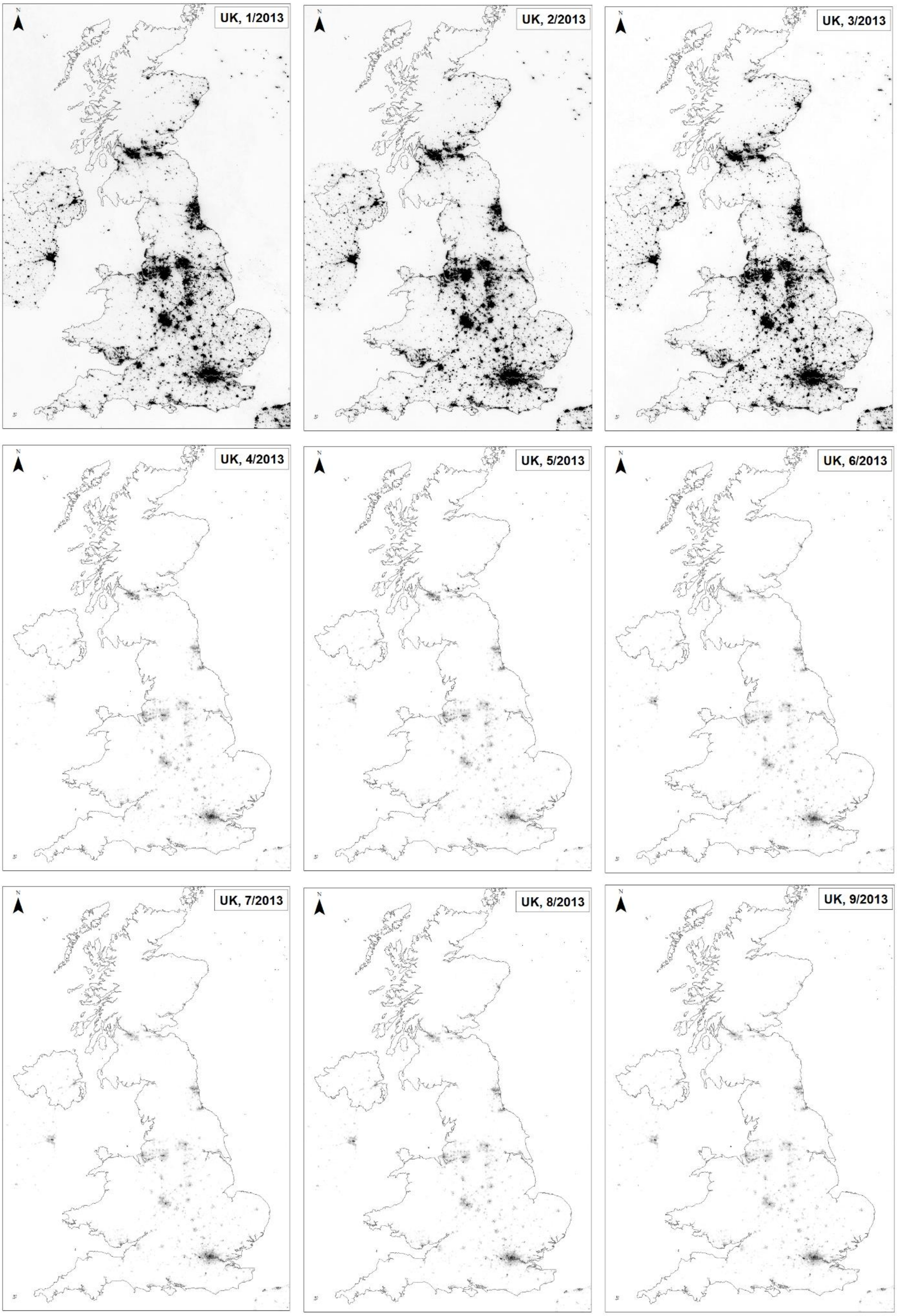


Figure 11 VIIRS monthly pre-processed Composites of UK for 2013, including forecasted/patched images produced via the linear regression method (01/2013-09/2013)

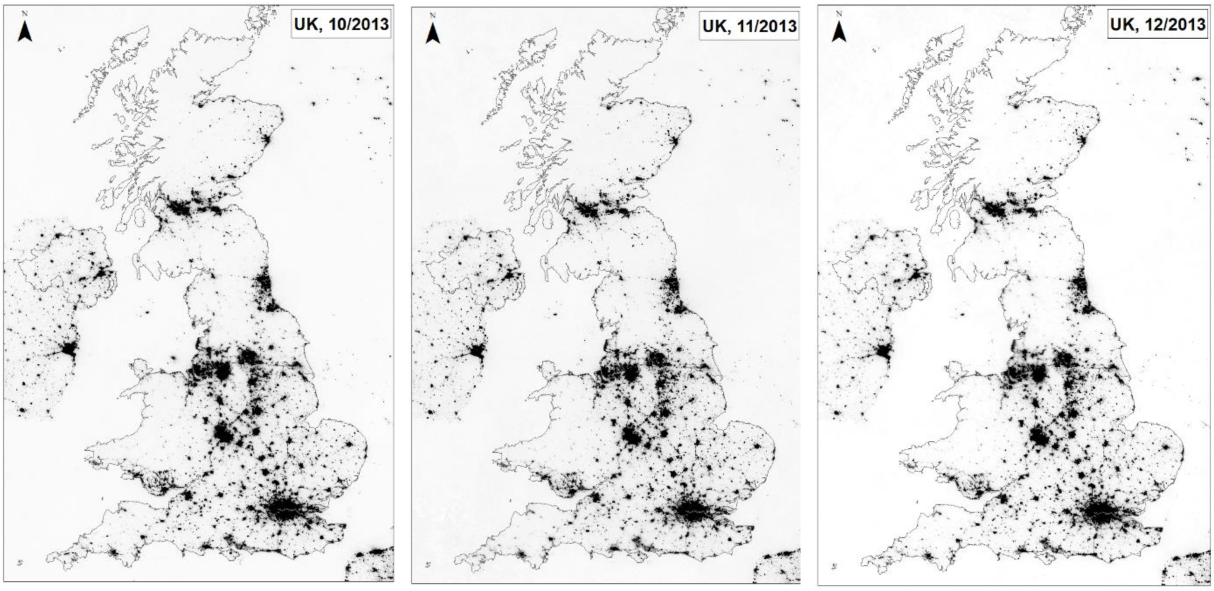


Figure 12: VIIRS monthly pre-processed Composites of UK for 2013, including forecasted/patched images produced via the linear regression method (10/2013-12/2013)

For the calculation of the final images the same approach as depicted in figure 8 was implemented for the linear regression method. Thus, forecasted and patched images for April, May, June, July, August and September were produced for the pre-processed monthly composites (figures 11 & 12). The following figure 13, presents the sum of lights calculated for the final monthly composites of UK for 2013. It is obvious that the new images for April, May, June, July and August lead to a smoother variation of the variance values across different months of a year.

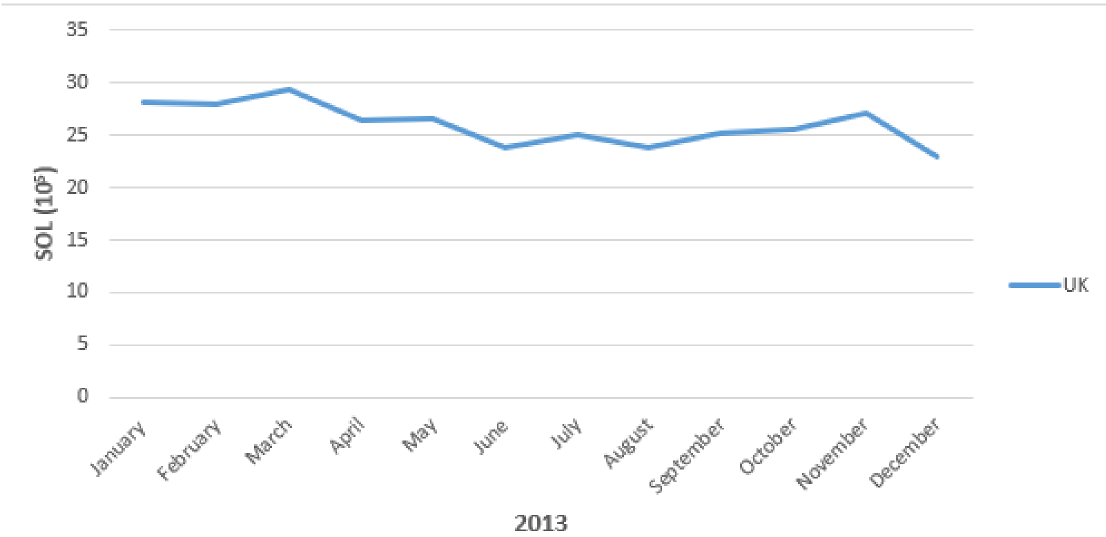


Figure 13 Monthly variation of Sum of Lights of the final processed images of UK for 2013

4. Spatial Data Infrastructures

Initially, Geographic Information systems were developed as closed applications, with their own software packages, data structures and programming languages and were focused on individual use within separate organizations. A GIS not only enables the user to manage spatial datasets as separate map layers in a digital environment but allows the combination and interpretation of information on a certain location from different layers, including textual, statistical and map data. However, despite their various advantages, the need for bridge between the isolated GIS islands soon became apparent (Dessers, 2013).

Searching for spatial data available at other organizations is a time-consuming process, whilst similar datasets may be developed by different organizations. In addition, access to spatial data is not always regulated and institutional barriers often inhibit the potential use of GIS, by restricting spatial data access and availability. Nowadays, there is an increasing shift towards the shared use and exchange of spatial data across different organizations. During the last years, various initiatives have been taken to promote and optimize spatial data access, use and sharing. These initiatives are referred to as a whole with the term ***Spatial Data Infrastructure*** (SDI).

An SDI is typically defined as a set of interacting resources for facilitating and coordinating spatial data access, use, and sharing. However, literature provides a wide range of definitions of SDIs (Dessers, 2013; Rajabifard et al., 2002; Brand, 1998; Groot et al., 2000; Vandenbroucke et al., 2009; Chan and Williamson, 1999).

These definitions are differentiated by 2 basic elements. The first class of definitions focuses on the description of the SDI components, whilst the second class focuses on listing of SDI objectives. The first-class ranges from broad identification of an SDI (description of a framework) to listing of specific components and technologies. The second class consists of definitions that include data-related objectives (i.e. data sharing) and user-related objectives (i.e. supporting processes) (Dessers, 2013).

In the context of this study, a Spatial Data Infrastructure is defined as a set of technological and non-technological set-ups within and between organizations to facilitate access, sharing and use of spatial data, thereby contributing to the enhanced performance of the business, policy making and service provision processes.

It is important to establish that the actual goal of the SDI is not to serve the data handling functions per se, but to serve the needs of the user community, with regard to issues of globalization, sustainable development, economic reform, political unrest and war, urbanization, environmental awareness and human rights (figure 14) (Rajabifard et al., 2002).

Objectives / Components	B0: no objectives	B1: only data-related objectives	B2: also user-related or broader objectives
A0: no components	-	(3) E.g. "Spatial Data Infrastructure (SDI) is about the facilitation and coordination of the exchange and sharing of spatial data between stakeholders in the spatial data community" (Crompvoets et al., 2004)	(6) E.g. "An infrastructure for accessing and sharing spatial data to reduce the duplication of spatial data collection by both users and producers, and enable better utilization of spatial data and associated services" (Grus, 2010)
A1: general typification of components	(1) E.g. "Spatial Data Infrastructures are foremost social networks of people and organisations, in which technology and data play a supportive role. The technology is cheap, data is expensive, but social relations are invaluable" (Craglia and Campagna, 2009a)	(4) Of this class, no examples were found.	(7) E.g. "An SDI is a set of technological and non-technological set-ups [components] within and between organisations [network] to facilitate access, exchange and use of spatial data [narrow objectives], thereby contributing to the performance of the business processes [broader objectives]" (Vandenbroucke et al., 2009)
A2: list of components	(2) E.g. "The components of a spatial data infrastructure should include sources of spatial data, databases and metadata, data networks, technology (dealing with data collection, management and representation), institutional arrangements, policies and standards and end-users"(McLaughlin and Nichols, 1992)	(5) E.g. "National Spatial Data Infrastructure (NSDI) means the technology, policies, standards and human resources necessary to acquire, process, store, distribute, and improve utilization of geospatial data" (Executive Office of the President, 1994)	(8) E.g. "Infrastructure for spatial information means: metadata, spatial data sets and spatial data services; network services and technologies; agreements on sharing, access and use; and coordination and monitoring mechanisms, processes and procedures, established, operated or made available in accordance with this Directive. (. . .) INSPIRE should assist policy-making in relation to policies and activities that may have a direct or indirect impact on the environment" (European Commission, 2007)

Figure 14 Sorting scheme for SDI definition (Source: Dessers, 2013)

4.1 INSPIRE Directive

Geographic Information is needed in many policy domains. Despite the fact that a vast number of spatial datasets exist, many challenges have emerged in the EU community (figure 15):

- Lack of real reference and authentic data;
- Difficulty to obtain spatial data (difficult to find, cumbersome process to obtain, often expensive);
- Gaps in availability of spatial data and a lot of duplication at the same time;
- Lack of harmonized spatial data between countries and within countries (different coordinate reference systems, different scales, lack of consistency etc);
- Difficulties in the interpretation and understanding of spatial data as they are often not documented;
- A lot of existing datasets are “copyrighted” datasets[8];
- Dissemination of end products, can sometimes be restricted (sometimes end-products have to “degraded” in order to be used in other applications);
- Quality “label” is not commonly adopted: uncertainty about the products.

The absence of a real European Spatial Data infrastructure results in higher investment costs, delayed implementation of projects, uncertain quality and dissemination constrains.

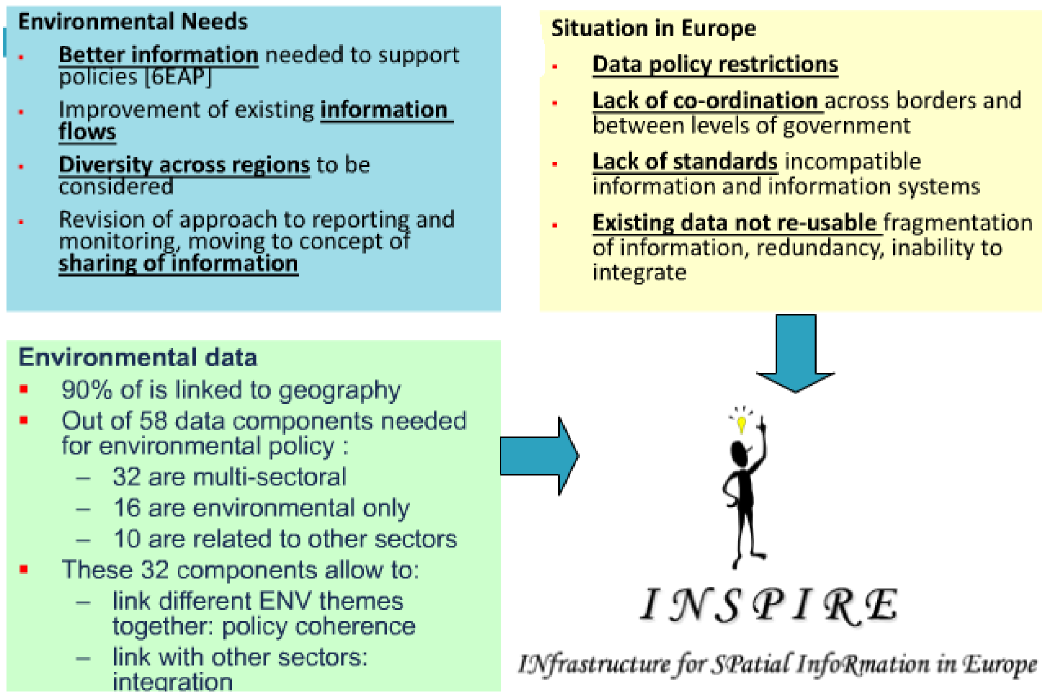


Figure 15 The need for INSPIRE (Source: EEA)

[8] You do not buy information itself, but the right to use it (“license”)

All the above-mentioned considerations led to the set-up of a legal initiative called **INSPIRE**. In May 2007, the Directive 2007/2/EC establishing an Infrastructure for Spatial Information in the European Community (INSPIRE)[9] came into force. INSPIRE lays down general rules aimed at the establishment of the Infrastructure for Spatial Information in the European Community for the purposes of EU environmental policies and policies or activities which may have an impact on the environment[10]. The actions under INSPIRE aim at removing obstacles hampering the exchange and use ('sharing') of spatial data and services between public authorities at all levels of government and across jurisdictional and administrative borders in the EU when such is necessary for performing their public tasks that may have a direct or indirect impact on the environment.

INSPIRE adheres to the following 5 principles which formulate the vision of the Directive (https://inspire.ec.europa.eu/inspire-principles/9):

1. Data should be collected once and maintained at the most effective way possible;
2. It should be possible to combine seamlessly spatial data from different sources across the EU and share it between many users and applications;
3. It should be possible for spatial data collected at one level of government to be shared between all levels of government;
4. Spatial data needed for good governance should be available on conditions that are not restricting its extensive use;
5. It should be easy to discover which spatial data is available, to evaluate its fitness for purpose and to know which conditions apply for its use.

INSPIRE doesn't set any obligations on the private sector. In addition, the Directive does not require from the Member States to collect new datasets neither to give up on their national data models and standards, as long they provide them in an INSPIRE compliant version. Last but not least, the Directive doesn't affect ownership rights. INSPIRE view services and metadata must be freely available via discovery services through the INSPIRE geoportal. On the other hand, data download can be subject to different conditions depending on the data provider's policy.

The spatial information considered under INSPIRE is extensive and includes 34 themes, divided into three Annexes as presented in the following figure 16. INSPIRE aims at thematic domains that concern environmental policies. The classification in three Annexes is resulted from the different priorities, timelines and deadlines to implement data components. Themes under Annex I, are considered to be

[9] Directive 2007/2/EC of the European Parliament and of the Council of 14 March 2007 establishing an Infrastructure for Spatial Information in the European Community (INSPIRE)
[10] https://inspire.ec.europa.eu/inspire-directive/2

reference data that are needed by any geospatial application[11], whilst Annex II and Annex III contain more specific themes. The individual themes as defined in the Annexes I, II and III of the Directive, are refined in document D2.3 'Definition of Annex Themes and Scope'.

Figure 16 INSPIRE Themes (Source: https://inspire.ec.europa.eu)

4.1.1 Linkage with other directives

INSPIRE framework and implemented actions is linked and should be coherent with the provisions laid down in Directive 2003/4/EC on public access to environmental information (PAEI) and Directive 2003/98/EC revised by Directive 2013/37/EU on the re-use of public sector information (PSI), where they regard access by third parties to environmental information and the re-use of public sector data falling within the scope of both INSPIRE and PSI. It is worth highlighting that PSI Directive identifies geospatial data, environmental data and Earth observation data as strategic assets for the economy and society at large.

4.1.2 Inspire Implementing Rules (IRs)

INSPIRE Implementing Rules aim to create the necessary conditions (rules) for the establishment of each INSPIRE component (Metadata, Data Specifications, Network Services, Data and Service Sharing and Monitoring and Reporting) within the context of the Directive. These IRs are adopted as Commission Decisions and are binding in their entirety. The Commission is assisted in the process of adopting such rules by a regulatory committee composed by representatives of the Member States and chaired by a representative of the Commission (this is known as the Comitology procedure).

The Implementing Rules will address among other things the implementation of a common framework for the unique identification of spatial objects. Each spatial object will be identified through a unique and persistent identifier (this identifier will not change over time).

[11] Themes for the other Annexes are depending on them

4.1.3 Metadata

Based on Article 5 of the Directive 2007/2/EC, Member States have to create[12] and keep up to date metadata for the spatial data and services. Metadata shall include the following information about: (i) Conformity with IR on interoperability / harmonization, (ii) Conditions for access and use of spatial data and services, (iii) Quality and validity of spatial data, (iv) The public authorities responsible for the establishment, management, maintenance and distribution of spatial data sets and services and, (iv) Limitations on public access and corresponding reasoning (Directive 2007/2/EC).

4.1.3 Data Specifications

Data specifications are not binding (figure 17) and aim to establish the framework for interoperability and actual data harmonization to common standards. They are used to specify common data models, code lists, map layers and additional metadata on the interoperability to be used when exchanging spatial datasets[13].

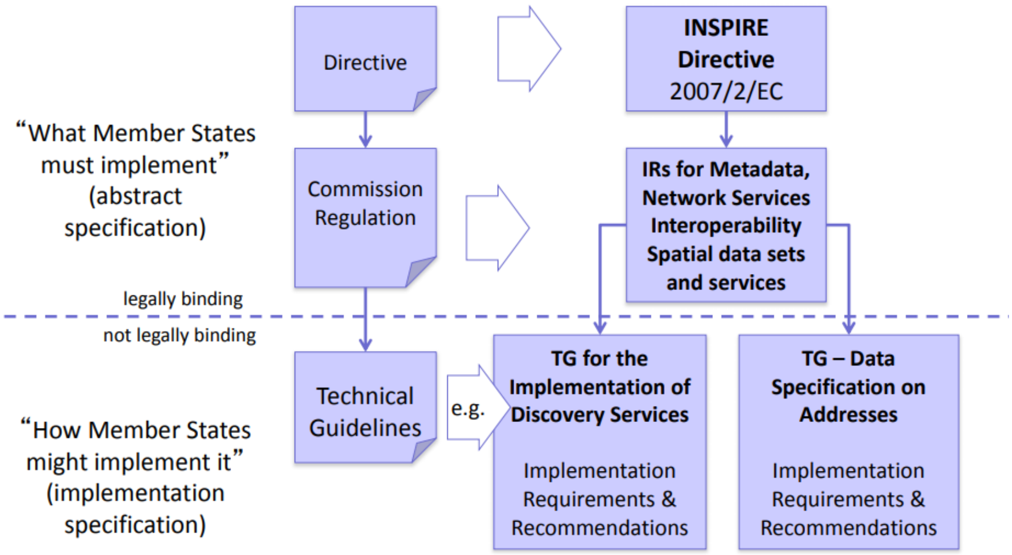


Figure 17 Implementing Rules vs. Technical Guidelines, (Source: JRC, 2013[14])

In particular, the common Data Specifications goal is to make available data within the scope of INSPIRE utilizing:

1. the same spatial object types (and definitions);
2. the same attributes (and definitions, types, code lists) and relationships to other types;
3. a common encoding (GML application schemas);
4. common portrayal rules

[12] *Metadata have to be produced within 2 and 5 years from the adoption of the IRs for Annex I & II and Annex III respectively.*

[13] https://inspire.ec.europa.eu/data-specifications/2892

[14] https://www.unece.org/fileadmin/DAM/energy/se/pp/unfc_egrc/egrc4_april2013/25_april/5_Thomas_INSPIRE.pdf

The following figure 18 presents the layered modelling framework adopted for the compilation of the INSPIRE data specifications. The basis lies in the use of international standards (i.e. ISO 1919, observation and measurements schemas). On top of this there is the Generic Conceptual Model acting as a middle layer which provides some useful common application schemas that define essential INSPIRE base types (i.e. INSPIRE identifier) as well as generic spatial objects that can be used by multiple teams. It was created by the Data Specifications Drafting Team. The third layer, established by the Thematic Working Group, contains the developed application schemas of the data models.

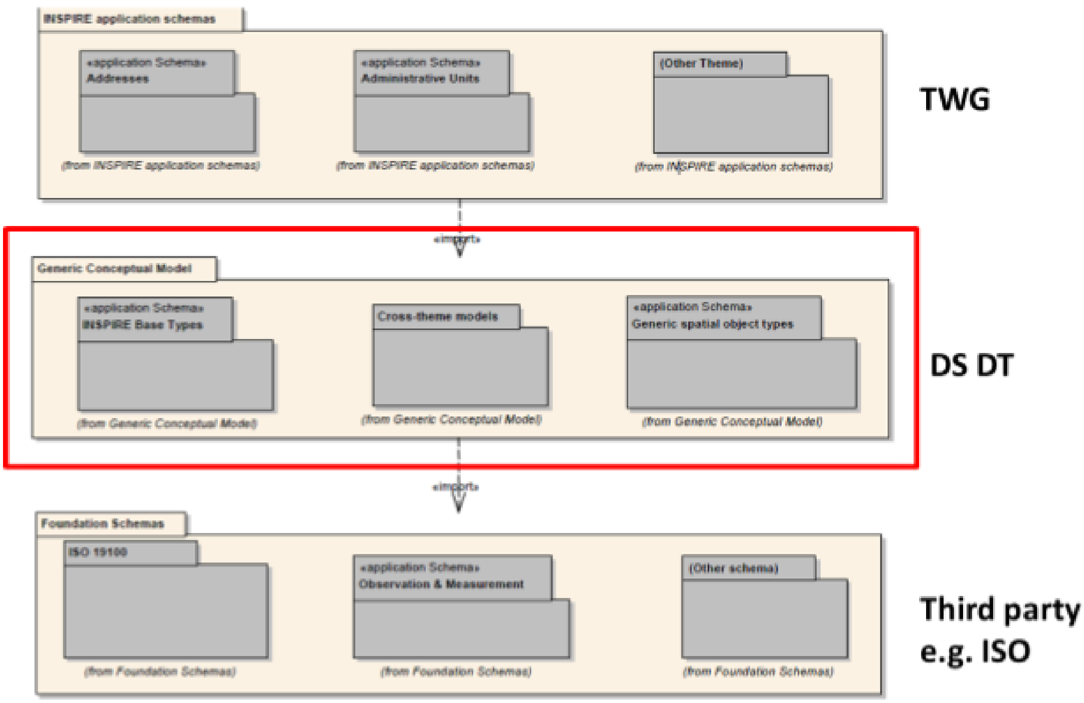


Figure 18: Consolidated UML modelling framework

The INSPIRE modelling framework is based on 4 cornerstones that ensure data interoperability. The most important one is the pillar of the *Conceptual Data Models*. This pillar defines for each theme, common spatial objects, their properties and their relationships to other object types and are managed in a common UML repository. The second pillar is *Encoding* which applies a common way to encode the conceptual data models. The standard approach in INSPIRE is to encode all data sources as GML, which is the most common file for exchanging geographical data[15]. The use of *Harmonized vocabularies* is the third pillar, which is important to overcome different semantic meanings. It concerns common classifications and code lists that list the definitions and values translated in different official languages that should be used by all data providers. These vocabularies will be managed in an online *Register* (4th pillar) which will also provide unique identifiers for reference to resources.

[15] For coverage data, alternative encoding can be used (i.e. GeoTIFF)

These 4 key pillars are described in INSPIRE Conceptual Framework documents[16]. The INSPIRE conceptual framework aims to provide a repeatable data specification development methodology and general provisions for the data specification process, which is valid for all spatial data themes. The Generic Conceptual Model is considered one of the most important document tackling different aspects towards data interoperability and harmonization (figure 19). It is not a "ready-to-use" common data model for INSPIRE, but involves:

- **a.** common terminology and basic concepts;
- **b.** requirements and recommendations for the creation of the Theme data specifications;
- **c.** a set of base types to use in the Theme data models (called application schemas in INSPIRE);
- **d.** Specification of cross-theme concepts including a INSPIRE identifier, Generic Network model and Gazetteers;

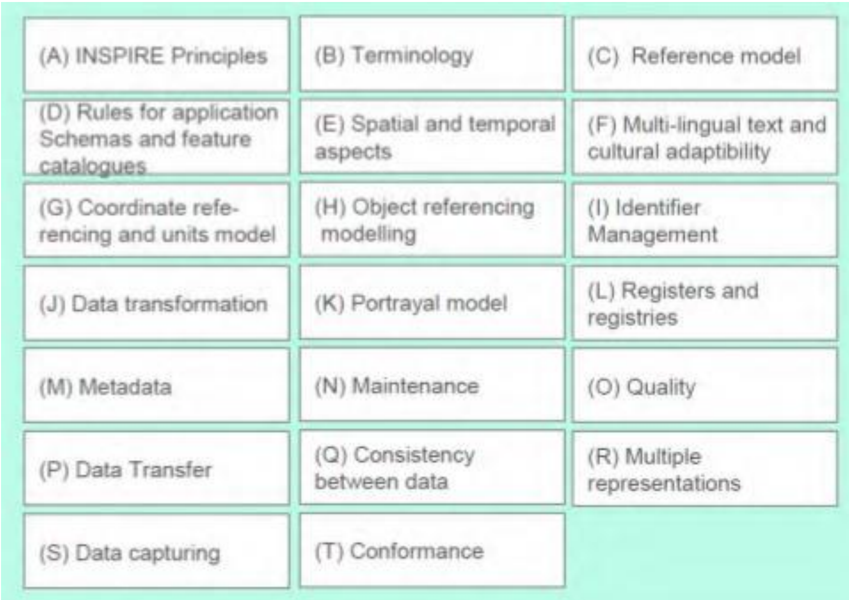


Figure 19: INSPIRE data harmonization aspects in Generic Conceptual Model, (Source: INSPIRE DS-D2.5, Generic Conceptual Model, v3.0)

[16] Generic Conceptual Model (D2.5), Methodology for specification development (D2.6), Guidelines for the encoding of Spatial Data (D2.7), O&M guidelines (D2.9), Common data models (D2.10.x)

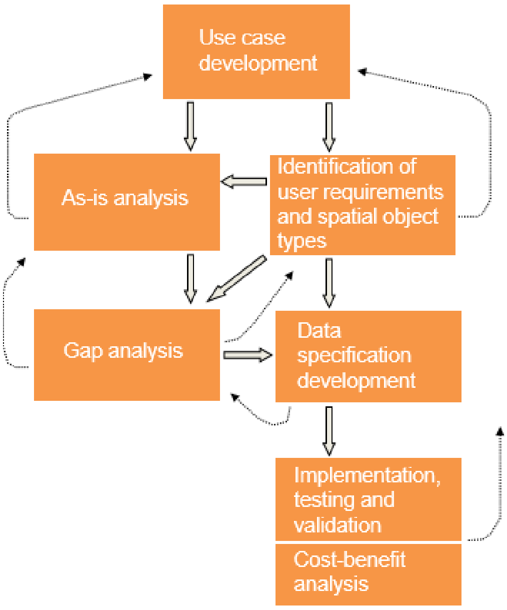


Figure 20: Step-wise methodology for data specification development (Source: INSPIRE DS-D2.6, Methodology for the development of data specifications)

The creation of data specifications followed a step-wise approach (figure 20) involving the interaction with relevant stakeholders. It starts with the development of a use case based on which user requirements (data content, metadata, etc) and spatial data object types are identified. Many stakeholders had the chance to propose use cases to be taken into consideration in the next steps. Simultaneously, an as-is analysis was made of the existing datasets and procedures. Based on the requirements and the description of the real situation, gaps were identified, and the results of the abovementioned steps were used to draft the first version of the data specifications. Then, an implementation testing, and validation phase was followed after which a new round of feedback was collected and was used to fine tune the data specifications[17].

- **Foreword**
- **General Executive Summary**
- **Theme-specific Executive Summary**

1. **Scope**
2. **Overview (incl. 2.2 informal description)**
3. **Specification scopes**
4. **Identification information**
5. **Data content and structure**
   - 5.1. **Application schemas - overview**
   - 5.2. **Basic notations**
   - 5.3. **Application schemas (incl. UML diagrams and feature catalogues**
6. **Reference Systems**
7. **Data quality**
   - 7.1. **DQ elements**
   - 7.2. **Minimum DQ Requirements**
8. **Metadata**
   - 8.1. **Common MD elements**
   - 8.2. **MD elements for DQ**
   - 8.3. **Theme-specific MD elements**
   - 8.4. **Guidelines for common elements**
9. **Delivery (incl. encodings)**
10. **Data capture**
11. **Portrayal (incl. layers, styles)**

- **Annex A: Abstract Test Suite**
- **Annex B: Use Cases**
- **Annex C: Code lists**
- **Other annexes (e.g. examples)**

All data specifications documents are based on a common template which means that they have the same structure as depicted in figure 21.

Chapter 5 is actually the core part of each Data Specification. In its beginning it states that the theme specific applications schemas defined by the TWG are included in the IR which means that the "Types for the Exchange and Classification of Spatial Objects" concerning that theme are the ones described in the specific application schemes dealing with that theme.

Figure 21 INSPIRE Data Specification Template (Source: INSPIRE DS-D2.8.II/III.x, Data Specification on <Theme Name> – Technical Guidelines, v3.0rc3)

[17] Drafting Team "Data Specifications" – deliverable D2.6: Methodology for the development of data specifications

In the second subchapter some general basic notions are explained which are valid for each theme, it concerns relevant parts taken from the Generic Conceptual Model (i.e. UML Notation of the Application schemas as a common conceptual schema language; List of Stereotypes that can be used in the application schemas; Explanation of the concept of “voidable properties”; Difference between enumerations and code lists; Concept of Identifier management; Geometry representation (allowed geometry types); Temporality representation- the concept of life-cycle of spatial objects; Information on how to deal with “coverages”). It provides essential knowledge needed for the understanding of the following paragraphs of the Data Specification. In Subchapter 5.3 a detailed description of each application schema is provided. It contains a narrative description of the model next to the UML diagrams that demonstrate the different types/objects and their relationships. The models themselves are then accompanied by the full description of all objects, attributes, the relationships and the used classifications also called the feature catalogue. It provides the information needed to perform a mapping from existing resources to the related Data Specification. Package diagrams are used to show dependencies between application schemas whilst class diagrams depict the different classes with their attributes and their relationships within one application schema. The complete description of the objects, including their properties is presented in the feature catalogue (INSPIRE DS-D2.8.II/III.x, Data Specification on <Theme Name> – Technical Guidelines).

Chapter 7 delineates some important aspects for evaluating the fit for use of a dataset. The objective here is to give end-users some assurance about the reliability of the data. This chapter defines some standard indicators and data quality measures to be used to report the quality of a theme specific dataset. There are general data quality elements that apply to all themes but in some cases theme specific Data Quality elements were defined by the Thematic Working Group. For all the defined data quality elements there must be a corresponding metadata element for documenting the quality, in accordance to the IR on Metadata. But in some cases additional metadata elements should be taken from the ISO 19158 standard (which is replacing the former standards 19113 and 19138) and which is specifically designed for documenting data quality (INSPIRE DS-D2.8.II/III.x, Data Specification on <Theme Name> – Technical Guidelines).

Chapter 8 specifies the dataset-level metadata elements, which should be used for documenting metadata for the complete dataset. In general, we can distinguish three levels in metadata:

1. Discovery metadata which are considered typical elements used to find and search for data sets;
2. Evaluation metadata;
3. Metadata concerning the use of data.

In INSPIRE we speak of Discovery metadata, metadata for interoperability (corresponds to the metadata for evaluation) and theme-specific metadata elements (corresponds to the metadata for use of data).

Last but not least, for all themes there is Annex A, the Abstract Test Suite (ATS) with a normative part that describes some tests for testing the conformity of a transformed dataset with the IR on interoperability, and an informative part describing tests for testing the conformity with the TG requirements. Annex B is informative and lists the all the Use Cases that were taken into account while developing the Data Specification. Annex C includes the code list values defined in the Application schemas of each theme. Other annexes can be defined depending on the need of that specific theme, like specific encodings rules, examples of implementation, etc (INSPIRE DS-D2.8.II/III.x, Data Specification on <Theme Name> – Technical Guidelines).

4.2 INSPIRE Architecture

In this section an overall analysis of the INSPIRE architecture is undertaken (figure 22). At the bottom there are the data components of the Infrastructure containing the spatial data itself as well as the metadata that describe the data and support the users in searching, finding and understanding the offered datasets. The data components side also holds the various registers to store information commonly used by different datasets, such as code lists, thesauri etc. On top of this data layer, lies the service layer that contains services that can be called by users and applications to retrieve information on the data. All these services are grouped and collected by the Service Bus. Between the service layer and the service bus, there is a layer (Geospatial Rights Management layer) managing the access to the services. On the top of the architecture there are the applications and the geoportals that support various functionalities on the data (search, consume, view, download etc).

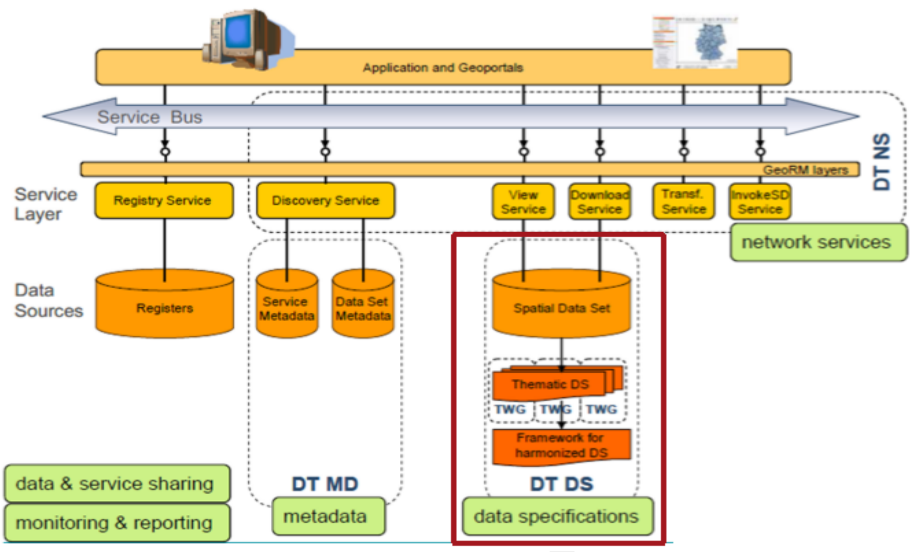


Figure 22: INSPIRE Architecture

4.3 Web mapping services

INSPIRE Network Services specify common interfaces for web services that can be consumed by generic clients to support search, download and visualization of datasets.

A web service can be defined as a software system designed to support (interoperable) machine-to-machine interaction over a network (https://www.w3.org/TR/ws-gloss/#defs). The web services work in the background, as they do not have a Graphical User Interface (GUI) and they are used to support the development of client/server applications. The main advantages of the web services are described as follows (https://www.w3.org/TR/ws-gloss/#defs):

1. They provide Interoperability between various software components. This means that they can effectively communicate with other s/w components using the same standards
2. They are accessible via a network
3. They are self-describing, which means that you can always ask a web-service what it can do (i.e. the list of maps that can provide).

There are 2 ways to request a service from the web service. The 1st is based on the use of remote procedure calls. In this case, in the classical client-server paradigm, the client sends a request (i.e. for map data and/or services) to the server based on standard HTML methods through a URL. This request contains information about the server address, the type of the service (i.e. WMS, WFS, etc) and other keywords describing the function and separated by the characters (?) and (&). Then the request is processed by the server, which generates a response including the requested map and/or data. The second way is based on the adoption of a service-oriented architecture, in which XML messages are used as means of communication between the client and the server.

To facilitate spatial data sharing and interoperability, Open Geospatial Consortium (OGC) has established an arsenal of standards based on which geographic information interoperates.

Web Map Service (WMS)

A Web Map Service (WMS) is a web interface that allows publishing and deploying of geographical maps on the internet. The WMS in an official specification from OGC and ISO, that defines a set of interface specification that provide uniform access by web clients to maps rendered by map server on the internet. Upon a map request, the WMS server first loads the data for which it takes the request for. The data itself it may not be inside the WMS server, but the server has to be connected in the network. Then the WMS server applies specific styling rules and renders the image. The data leaves the server only as an image (figure 23).

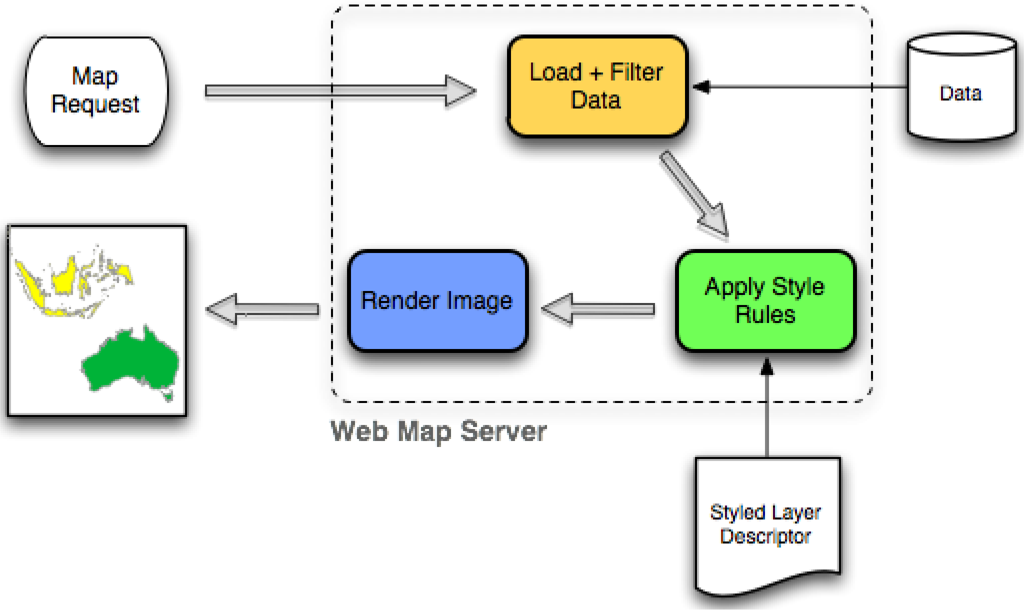


Figure 23 Workflow diagram of Web Map Service (Source: https://boundlessgeo.com/)

WMS Server has the following functionalities:

- *Through the getMap request, t*he client requests from the server a snapshot of the data (map) satisfying the parameters defined in the query. The server responds with an image file (.tif, png., etc);
- Answer basic queries about the datasets on the server satisfying certain parameters via *GetFeatureInfo request,* which returns to the client an XML with the attributes description;
- Tell other programs what maps it can produce and which of those can be queried further via *GetCapabilities request* which returns an XML file with metadata. The GetCapabilities request provides the following information:
    - Image formats it can serve;
    - List of the map layers;
    - List of layers supporting the optional GetFeatureInfo interface;
    - List of available spatial reference systems;
    - List of exception formats for return of exceptions;
    - SLD styles (optional);
    - Vendor specific capabilities (optional).

Web Feature Service (WFS)

The Web Feature Service (figure 24) provides an interface for describing data manipulation operations (create, delete, update and get feature) on geographical feature stored in a database, using HTTP as the distributed computing platform. The WFS interface exposes the data in these repositories as Geography Mark-up Language (GML). A standard WFS request constitutes a description of query, or data transformation operations which are applied to one or several features. The client generates the request, which is then posted to a web feature server through HTTP. The web feature server then reads and executes the request. Due to the fact that the data leave the server as data (GML), the client chooses the style and presentation details. The Transactional Web Feature Service (WFS-T) enables the client to create, update and delete data on the server. The WFS Server has the following basic operations:

- *GetCapabilities request:* Discover what services and what data types the WFS supports (capabilities of the service);
- *DescribeFeatureType request:* Retrieves a description of the structure of any feature type (schema description) the server supports in a XML schema;
- *GetFeature request:* Retrieves the actual data (GML presentation of the data) that the client has requested (query defines filters to limit the number of features returned).

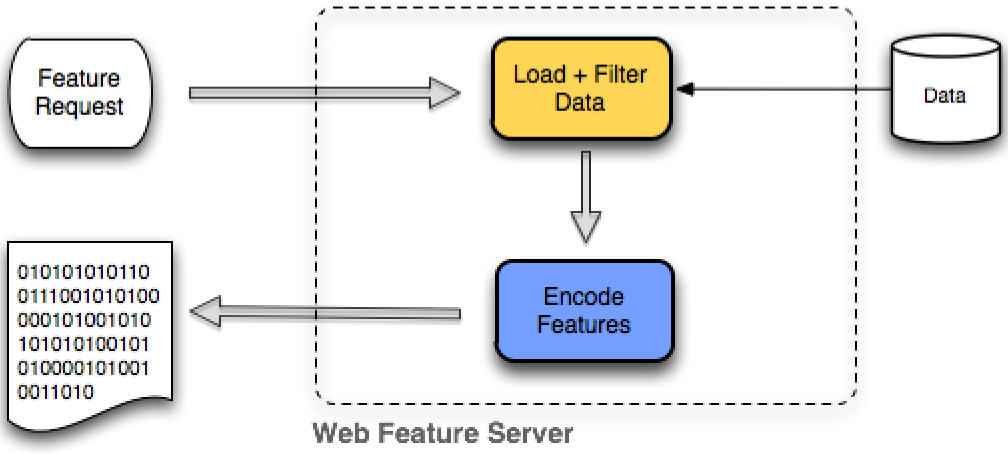


Figure 24: Workflow diagram of Web Feature Service (Source:** https://boundlessgeo.com/**)

Catalogue service for the web (CSW)

A catalogue server publishes collections of descriptive information (metadata) about geospatial data. It supports several operations relevant to metadata management (including transaction services) such as search, discovery, insert, update and delete. In addition, CSW enables metadata harvesting by querying the server to "pull" them from a specific target. Some of the main operations of the CSW are the following:

- Provides metadata about the service itself through the *GetCapabilities request*

- Executes a search through the *GetRecords request, which returns a list with all the metadata satisfying the search criteria;*
- Returns the metadata of a specific record using their identifier through *GetRecordById request*

The interaction between a client and a server is accomplished using a standard request-response model of the HTTP protocol. A client sends a request to the server using HTTP, and expects to receive a response to the request or an exception message. An XML HTTP POST operation is performed (Catalogue Services Specification Version 2.0.2 OGC 07-006r1).

4.4 INSPIRE Network Services

INSPIRE aims to provide access to spatial data via network services and according to harmonized data specifications to achieve interoperability of data.

- INSPIRE Discovery services

INSPIRE Discovery Services are used to discover spatial datasets and services (like Catalogue Services). They provide operations on metadata (i.e. search, register and update of metadata) and they are based on OGC™ Catalogue Services Specification 2.0.2 - ISO Metadata Application Profile for CSW 2.0.

The INSPIRE guidelines specify the operations that the Discovery Services must provide. In particular, each discovery service must support:

- ✓ Provision of information about the service itself via the *Get Discovery Metadata Service* operation
- ✓ Requesting INSPIRE metadata elements based on a query statement via the *Discover Metadata* operation
- ✓ Allow to insert, update and delete metadata elements via the *publish Metadata* operation
- ✓ Allow discovery of resources though the Member State while maintaining the resource metadata at the owner location via the *Link Discovery Service* operation.

The following table 2 provides a preliminary comparison between OGC and INSPIRE discovery services operations. Discovery Services include also additional parameters such as ‘Language’ for GetCapabilities operation (i.e. REQUEST=GetCapabilities&SERVICE=CSW& LANGUAGE=ENG)

INSPIRE Operation	OGC CSW Operation
Get Discovery Service Metadata	OGC_Service.GetCapabilities
Discover Metadata	CSW Discovery GetRecords
Publish Metadata	CSWT Manager Transaction or CSWT Manager Harvest

Link Discovery Metadata	Combination of OGC_Service.GetCapabilities and CSW Discovery GetRecords OR Using Publish Metadata operation: CSWT Manager Transaction or CSWT Manager Harvest

Table 2: Mapping between the INSPIRE metadata elements and the OGC Capabilities metadata elements

Some of the softwares that can be used to produce an INSPIRE compliant CSW are GeoNetwork, MapServer, ArcGIS for INSPIRE, Geomajas, deegre, pyCSW etc.

- INSPIRE View Service

INSPIRE View services aim to display, navigate, zoom in/out, pan and overlay viewable and legend information. WMS 1.3.0, WMS 1.1.1 (INSPIRE profile) as well as WMTS 1.0.0 for tile services are considered INSPIRE compliant View services. Each INSPIRE View service must provide one of the following European Coordinate Reference Systems:

- ETRS89 Lambert Azimuthal Equal Area;
- ETRS89 Lambert Conformal Conic;
- ETRS89 Transverse Mercator

The operations of a View Service are presented as follows:

- *Get Map* operation: Returns a map for a specified area;
- *Get View Metadata Service*: Get metadata about a specific view service;
- *Link View Service*: Allows the linking of view services together
- INSPIRE Download Services

They are used to provide direct access (download) to the datasets or to parts of them. They are based on the OGC WFS 2.0 and must provide conformity with ISO 19142, 19143 for direct access.

The Network Services Regulation describes the following four download operations that must be implemented by Download Services:

- Get Download Service Metadata
- Get Spatial Dataset
- Describe Spatial Dataset
- Link Download Service

The Network Services Regulation also states that where practicable, the following two operations shall be implemented by Download Services:

- Get Spatial Object

- Describe Spatial Object Type

Furthermore, if the Get Spatial Object and Describe Spatial Object Type operations are implemented, then particular search capabilities shall also be implemented. These capabilities include the ability to search by:

- URI of Spatial Dataset
- Key attributes of spatial objects, including URI and date/time of update
- Bounding Box
- Spatial data theme
- Combinations of the above

➢ <u>Transformation services</u>

They are used to transform existing data on the fly to INSPIRE conformant data. The transformation process is usually undertaken offline with ETL tools (i.e. HALE, FME, etc). Three implementation strategies have been proposed for the development of INSPIRE compliant datasets and services (figure 25):

1. On the fly transformation of spatial data though a download service that automatically transforms the source data to the schema of the INSPIRE data specification
2. Offline transformation of spatial data: Each data provider transforms its own data based on the relevant INSPIRE data specification and offers the result to the user by means of INSPIRE download service
3. External Transformation of spatial data by separate network service: The user first download the source dataset and then makes use of INSPIRE transformation service to harmonise the data based on the INSPIRE data specifications

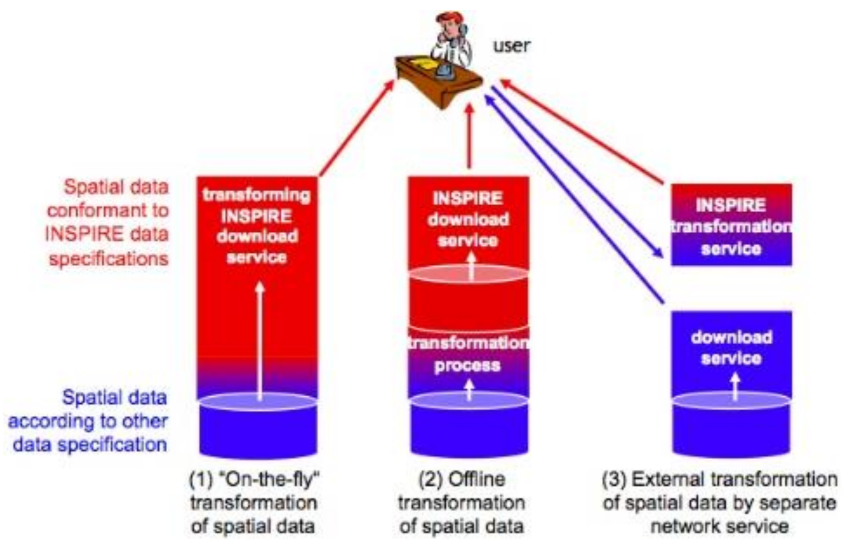


Figure 25: INSPIRE Transformation alternatives

➢ Invoke spatial data services

They are used to call (invoke) individual spatial services as well as combinations of them. Invoking of spatial services can be achieved through the use middleware. Web Processing Service can be used as basic building block for the implementation of invoke spatial data services

4.5 Data Interoperability

In the context of implementing a geospatial application, a user had to search different data by contacting different data producers, all having their own rules for access and use of their data products. In many cases these datasets are retrieved in different formats and the end user has to deal with interpreting heterogeneous data in different formats, identify, extract and post process these data in order to end up with a harmonised dataset that could be used for further analysis. The whole process of searching retrieving and combining different data sources should be made much easier and preferably in an automatic way.

Syntactic heterogeneity: Data may be implemented in a different syntax which can be translated as data stored in different formats (i.e. shapefile, ascii, relational or object- oriented models). Syntactic heterogeneity is also related to the geometric representation of geographic objects (i.e. raster and vector representations).

Structural or schematic heterogeneity: It refers on how you model the information in a certain structure inside the database. Even if you use the same syntax for database format and you deal with object concepts with a semantic meaning, then still you can model them in a different way and different schemas. Objects in one database can be considered as properties in another, or object classes can have different aggregation or generalisation hierarchies, although they might describe the same Real World concepts.

Semantic heterogeneity: Different meaning can be given to the same real world applications. It is not strange and it depends on the point of view/ discipline based on which someone interprets the application. For example, a road network for pavement management has different semantic descriptions from transportation data maintained in a GIS database designed for traffic management. In the first case the municipality needs to be aware of the quality of the pavement whilst in the second case the relevant authorities need information about the type the road, the incoming flows, its characteristics etc. Semantic interoperability should guarantee that data content is understandable by all in the same way via the use of harmonized definitions and classifications for objects.

Based on Article 3(7) of the INSPIRE Directive, interoperability is defined as "the possibility for spatial data sets to be combined, and for services to interact, without repetitive manual intervention, in such a way that the result is coherent, and the added value of the data sets and services is enhanced."

INSPIRE aims to facilitate data use and interoperability by adopting common cross-domain models for data exchange.

4.6 INSPIRE Implementation

INSPIRE comes with a strict implementation roadmap for the different data themes. Today, two of the most important and difficult milestones are upcoming (figure 26):

- All Annex I spatial datasets must be available in accordance with Implementing Rules for Annex I (23/11/2017)
- All Annex II & III spatial datasets must be available in accordance with Implementing Rules for Annex II & III (21/10/2020)

In total 34 data themes need to be offered in a harmonised way, meaning according to common data specifications. The data harmonization milestone includes providing interoperable datasets, as well as the related interoperable network services, metadata for these harmonised datasets and services and a catalogue to discover.

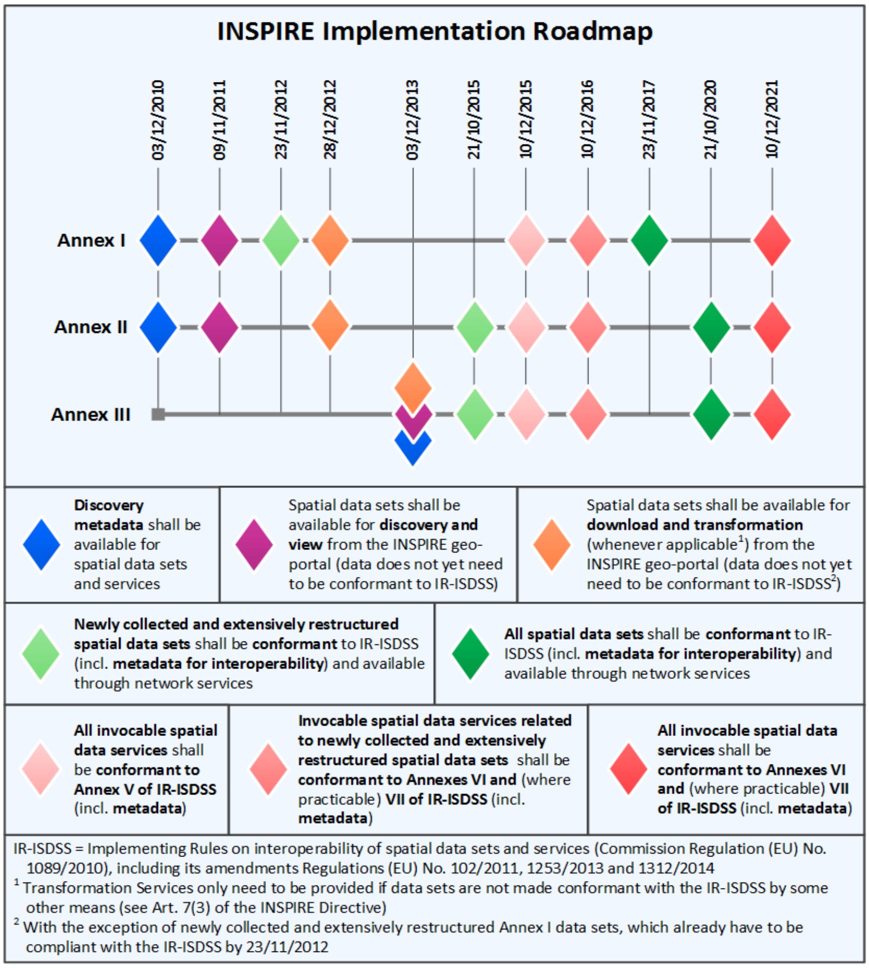


Figure 26 INSPIRE Implementation Roadmap (Source: https://inspire.ec.europa.eu/road-map-graphic/32443)

5. SDI Design and Implementation

Geographic Information Systems are integrated collections of computer software and data used to view and manage geographic information in order to analyse spatial relationships and to model spatial processes (Wade, T. and Sommer, S., 2006). Nowadays, the developments in the field of information and communications technologies lead to the replacement or traditional desktop GIS applications by web-based architectures and paradigms, facilitating spatial data handling, promoting data sharing across different actors and various applications.

Depending on the SDI participants (data providers, service providers, end-users) different software components may be used when building an SDI. A Spatial Data Infrastructure should support the discovery and delivery of spatial data from different repositories via the use of web services. Furthermore, data providers should be able to create and update spatial data stored in a repository (even remotely). Hence, the following software components are considered essential for a basic SDI architecture (figure 27):

1. a client software able to display, query, and analyse spatial data;
2. a catalogue service for the discovery, browsing, and querying of metadata or spatial services, spatial datasets and other resources,
3. a web service enabling the delivery of spatial data via the Internet, and/or processing services such as datum and projection transformations,
4. a spatial data repository, and
5. a GIS software (client or desktop) that permits the creation and maintenance of data.

The catalogue service functions in a similar manner to analogue directories such as the Yellow Pages. The extent of processing services is somewhat unlimited, but transformation services are generally fundamental to a distributed system (Steiniger S. and Hunter A.J.S., 2012). The thesis will focus on the use of open source market leading technologies for the design and development of a prototype Spatial Data Infrastructure.

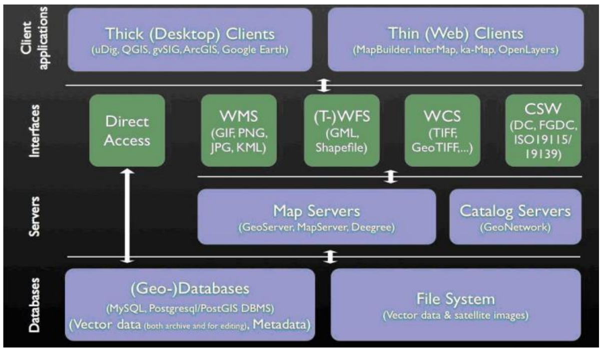


Figure 27: GeoFOSS based SDI Software Architecture

5.1 Database layer

Spatial databases are usually used to store geographic information such as point locations, boundaries and grids. They provide a series of special functions and indexes for querying and manipulating that data, which can be called from a query language such as Structured Query Language (SQL). A spatial database is developed and maintained from a group of application programs aiming at supporting its functionalities, which constitute the database management system. The Database Management Systems (DBMS) is a set of programs that facilitate the definition (i.e. specification of data types, structures and data constrains), construction (i.e. store the data), and manipulation (i.e. data retrieval, data update and production of reports) of a database (Stefanakis, 2014).

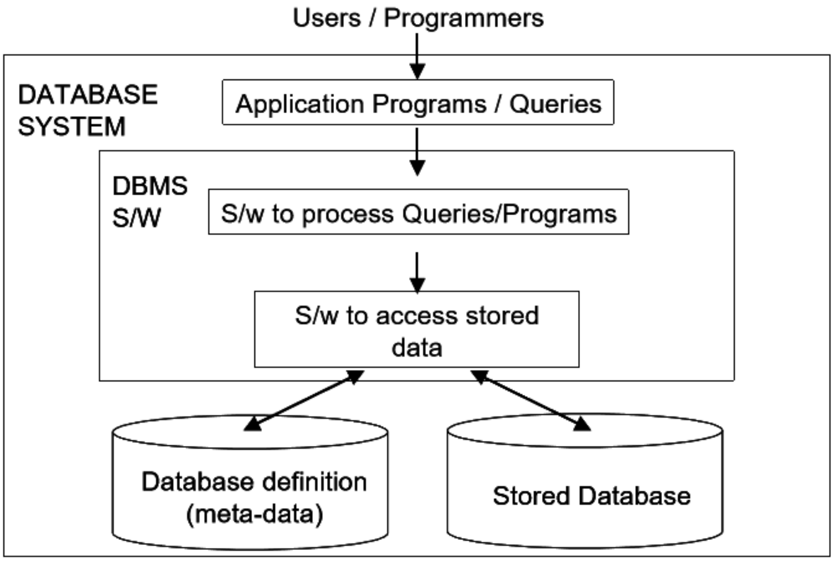


Figure 28: The architecture of a Database System (Stefanakis, 2014)

A database in combination with the database management system consist the Database System. Figure 28 presents a simplified overview of the architecture of the Database System, comprising of 4 basic modules: i) the users and programs interface, ii) the database management system, which is divided into software tools for processing queries coming from the interface and s/w tools to access the data and the metadata of the system, iii) the database and iv) the metadata catalogue. Prior to accessing the database, the DBMS searches the database catalogue to find the physical location of the data and the index structures on top of this data (Stefanakis, 2014).

A variety of database technologies are available in the market (either open source or proprietary) such as PostGIS, Oracle Spatial, Microsoft SQL Server, SpatiaLite, MySQL, ArcGIS Spatial Data Engine (SDE).

PostGIS

PostGIS is a free and open source library that spatially extends the object-relational database management system PostgreSQL, allowing the storage of geographic objects. A geographic object is a special type of data that enables the storage of a geographic position or a set of them as part of a line or a polygon. Essentially, PostGIS is a powerful tool that supports complex handling of geographical data including data visualization when utilised along with other graphical tools, such as QGIS. The key features of PostGIS are presented as follows:

- It is capitalizing on the Generic Index Structure (GIST) functionality of PostgreSQL, which allows the development of indexes and custom functions for almost any type of dataset.
- It is enhanced and supported by other well-known projects such as Proj4, GEOS, GDAL, etc.
- It provides an arsenal of spatial functions to search, analyse, convert and manage spatial data;
- It supports both vector and raster datasets
- It's based on open standards as defined by the Open Geospatial Consortium.
- It is compatible with almost all open source and/or proprietary GIS software tools

Raster data types in PostGIS

Raster datasets in PostGIS are stored in a table with a column of type raster. Pixels are organized in rows and columns to form tiles. Data is usually evenly tiled so that one row holds the same rectangular size of pixels as other rows. To achieve faster processing, large raster datasets should be broken into tiles for storage in multiple rows rather than keeping them in a single row. For instance, it is recommended that each row should keep between 50 and 500 pixels for both width and height.

The ***raster2pgsql*** loader packaged with PostGIS is capable of taking larger raster datasets and chunking them into smaller tiles for database storage.

Each raster tile (a row in the raster column) has a width and height which is measured in pixels. Rasters can only store numeric values in their pixels. The number of values that each pixel can store, depends on the number of bands of the raster. For example, an RGB raster can store 3 values in each pixel. Pixel types describe the type of numbers that a given band in a pixel can accommodate. Each band has a single pixel type defined for all pixels. Some of the possible choices are:

- 1-bit Boolean, abbreviated as 1BB
- Unsigned integer of 2, 8, 16, or 32 bits, abbreviated as 2BUI, 8BUI, 16BUI, 32BUI
- Signed integers of 8, 16, or 32 bits, abbreviated as 8BSI, 16BSI, 32BSI
- Two float types of 32 bits and 64 bits, abbreviated as 32BF and 64BF

VIIRS spatial database

In the context of this thesis, a dedicated spatial database is designed and created for storing and efficiently managing the processed monthly composites of VIIRS. As the envisioned application is simple and intends to establish a prototype SDI, conceptually, the database consists of one table pertaining the information about the images. The following diagram (figure 29) provides an overview of the technologies used and steps followed towards the creation and manipulation of VIIRS spatial database.

Last but not least, it has to be mentioned that there is potentially a performance penalty with storing rasters inside a database, and in particular if the database is not tuned and optimized.

Due to the potential performance penalty, it is worth considering some of the reasons why it was decided to store rasters inside a database:

- **Portability**: When migrating servers, instead of having to move hundreds, or even thousands, of files around the network, we can use database backup, restore, and replication functions.
- **Shared Database**: Organizations might choose to store rasters in databases for cataloguing purposes. This is particularly true if they have a large library of raster files, for example, terabytes of aerial survey files. It is often easier to store and catalogue this volume of raster files in a database.
- **Faster Searching**: One of the biggest features of a database is its ability to rapidly search indexes. If the coverage contains a large number of files, then the index searching performance of a database will be of benefit.

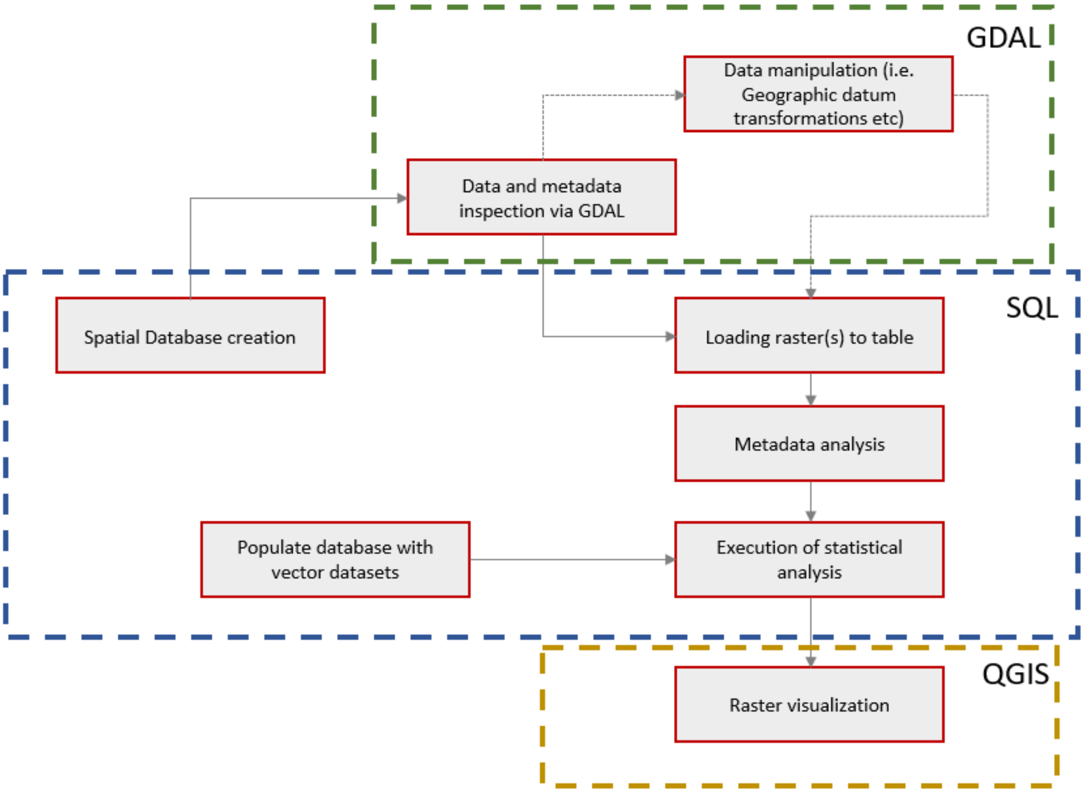


Figure 29: Technological and functional overview of VIIRS spatial database creation and manipulation

The spatial database can be created either via the *pgAdmin* interface or via command line. In this case, the database is build using the command line by following a series of steps. Initially, VIIRS database is created [***command***: created -U postgres VIIRS], indicating the name of the owner of the database. Then, the databased is selected [***command***: psql –d VIIRS –U postgres] and is spatially extended [***command***: CREATE EXTENSION postgis;].

Having created a spatially enabled database, the raster datasets can be inserted into the database via the following command.

```
raster2pgsql -s 3035 -c -b 1 -I -M \tsiak\Desktop\final_images\images_2013\03_2013.tif -
F -t 256x256 public.VIIRS_032013 | psql -U postgres -d viirs -h localhost -p 5432
```

The -s switch defines the spatial reference system of the dataset, which in our case is the Lambert Azimuthial Equal Area (EPSG:3035). As the database is currently empty, -c option creates a new table which is populated by the desired dataset. In addition, *raster2pgsql* loader is provide with information about the number of bands to be extracted by the raster (-b 1), the creation of a GiST index on the raster column (-I) and the vacuum analysis of the table (-M). The relative path of the raster is also required, which in our case points to the VIIRS monthly composite of March 2013. The -F option means that we want to create a field that holds the file name of the raster file, whilst the raster is divided to tiles to be inserted one per table row (-t). The size of the tile is expressed as WIDTHxHEIGHT (256X256).

GIST stands for Generalized Search Tree, and is used in PostGIS to index irregular structures, instead of basic data types (such as integers, characters, strings, and so on). The syntax used to build a GIST index for a geometry column is:

```
CREATE INDEX [indexname] ON [tablename] USING GIST ( [geometryfield] );
```

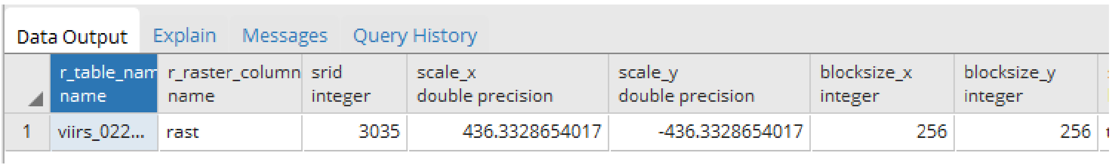
In our case the following SQL query is used to index the raster data. This query is automatically created when the raster data are loaded to the database.

```
CREATE INDEX viirs_032013_st_convexhull_idx
    ON public.viirs_032013 USING gist
    (st_convexhull(rast))
    TABLESPACE pg_default;
```

A high-level summary of the raster columns in the database is provided by the ***raster_columns*** view. Running the following SQL query will provide the available information in the VIIRS_022013 table.

```
SELECT
        r_table_name,
        r_raster_column,
        srid,
        scale_x,
        scale_y,
        blocksize_x,
        blocksize_y,
        same_alignment,
        regular_blocking,
        num_bands,
        pixel_types,
        nodata_values,
        out_db,
        ST_AsText(extent) AS extent
FROM raster_columns WHERE r_table_name = 'viirs_022013';
```

The SQL query returns a record similar the following:



Data Output | Explain | Messages | Query History

	r_table_name name	r_raster_column name	srid integer	scale_x double precision	scale_y double precision	blocksize_x integer	blocksize_y integer
1	viirs_022...	rast	3035	436.3328654017	-436.3328654017	256	256

In addition, geo-statistical analysis can be achieved to the database content via the use of other datasets. For instance, a shapefile of the administrative units of Europe (Nuts 1) is loaded to the database and is utilized for the calculation of statistics per country polygon (zonal operations) via the following query:

```
SELECT europe.fips_cntry, (
        ST_SummaryStats(
                ST_Union(
                        ST_Clip (rast, 1, geom, TRUE)))
      ).*
      FROM public.viirs_022013, public.europe
      WHERE ST_Intersects(rast,geom)
      GROUP BY europe.fips_cntry;
```

The output of the previous query is presented in the following figure 30, indicating important indexes (i.e. sum of light) of each country for February 2013.

Data Output

	fips_cntry character varying (2)	count bigint	sum double precision	mean double precision	stddev double precision	min double precision	max double precision
1	AL	181115	94014.94976025	0.519089803496397	2.47956598365912	0	149.539733886719
2	AU	616548	749317.80241112	1.21534382142367	3.50829442421416	0	164.199401855469
3	BE	167457	815764.955067918	4.87148912895799	9.26788000875203	0.235181465744972	321.850006103516
4	BK	309724	257949.137578378	0.832835484426065	3.46985728330771	0	180.989379882813
5	BU	820673	377604.758952356	0.460115976707356	2.00771163745692	0	245.309127807617
6	DA	553912	554462.194437167	1.00099328853169	4.29422770652497	0	321.850006103516
7	EI	389494	281156.659909144	0.721851068075873	2.64652640794504	0	173.812957763672
8	EN	396095	284470.341471694	0.718187155787613	3.45765609182387	0	321.850006103516
9	EZ	424940	751450.343560025	1.76836810740346	4.38303164676994	0	321.850006103516
10	FR	4333186	6025079.41254567	1.39045021666406	5.11119327914415	0	321.850006103516
11	GM	2093637	3323807.29744858	1.58757573421208	3.97695705417551	0	321.850006103516
12	GR	2167248	963017.038964248	0.444350180027504	2.97046463182821	0	253.437484741211
13	HR	637937	691139.326623325	1.08339746185489	4.08274890280499	0	180.989379882813
14	HU	614917	645288.382878137	1.04939102818451	3.34967118963151	0	209.954772949219
15	IC	660290	754311.368658304	1.14239405209575	3.12587660242006	0.177053540945053	321.850006103516
16	IT	2952274	5588577.71216432	1.89297392862733	6.56532807136318	0	321.850006103516
17	LG	477293	242385.393584793	0.507833539533981	1.82275310016964	0	117.113006591797
18	LH	357812	222819.759980924	0.62272858367222	2.03493267436498	0	117.113006591797
19	LO	273999	360188.618032487	1.31456179778936	3.78548645007906	0	209.954772949219

Figure 30 Data output indicating important indexes of each country for February 2013

Last but not least, the database content can be easily visualized (figure 31) though the use of desktop GIS software like QGIS.

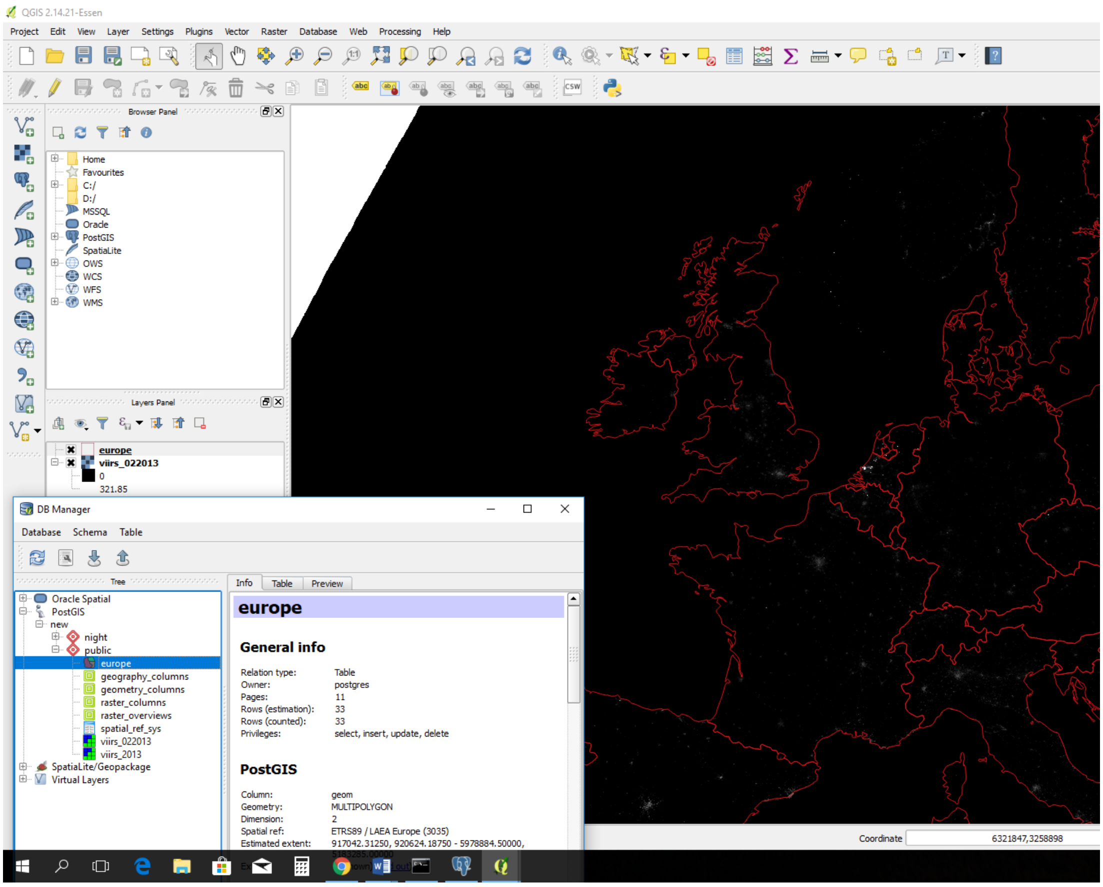


Figure 31 Visualization of the database content

5.2 Application layer (Middleware)

The application layer includes all the features that will allow for interoperability between the platform and external systems. In other words, it consists of a set of technologies that enables search, access and processing of the datasets contained in the database layer, by the presentation layer, the software sub-system that permits users to deal with geographical data (figure 32). In the context of this application, middleware consists of the following technologies:

- The **Map Server** and the implementation of the data and service interoperability protocols, which are provided as services conforming to the Open Geospatial Consortium (OGC) and INSPIRE standards and specifications. The map server is the module of the system that undertakes publishing of geospatial information. This component communicates with the spatial database so as to retrieve and/or store data. The data is provided through services that are conformant to the INSPIRE Guidelines and adhere to the relevant OGC (WMS, WFS, WCS, etc.) standards to internal and external client applications and to end-user interfaces that can visualize them. In addition to retrieving data from the SDI repositories, map server enables publishing of geospatial information by consuming corresponding OGC services from remote external data sources and systems, operating essentially as a Cascading WMS Service.
- The **Web Server** that supports hosting of the applications and web services that utilize the content of the database layer. The web server (also called application server) acts as a "mediator", linking the information originating from the database server to the client. Its necessity and importance in spatial data infrastructures and web applications development lies on its ability (i) to reduce the size and complexity of the client software components, (ii) to control and manage the data flow with better performance, and (iii) to meet the requirement for both data and user security.
- The **Metadata Catalogue Server**, the functional module of the application layer that implements the OGC CSW (Catalog Service for Web) protocol, as defined by OGC specifications and in accordance with the INSPIRE guidelines. This server will expose a catalog of metadata records via interfaces to other SDI components and to external applications through standard protocols (i.e. CSW, OpenSearch). The Catalogue Server will define specific common interfaces that allow the discovery, browsing, and searching of metadata related to the data and services of the SDI.

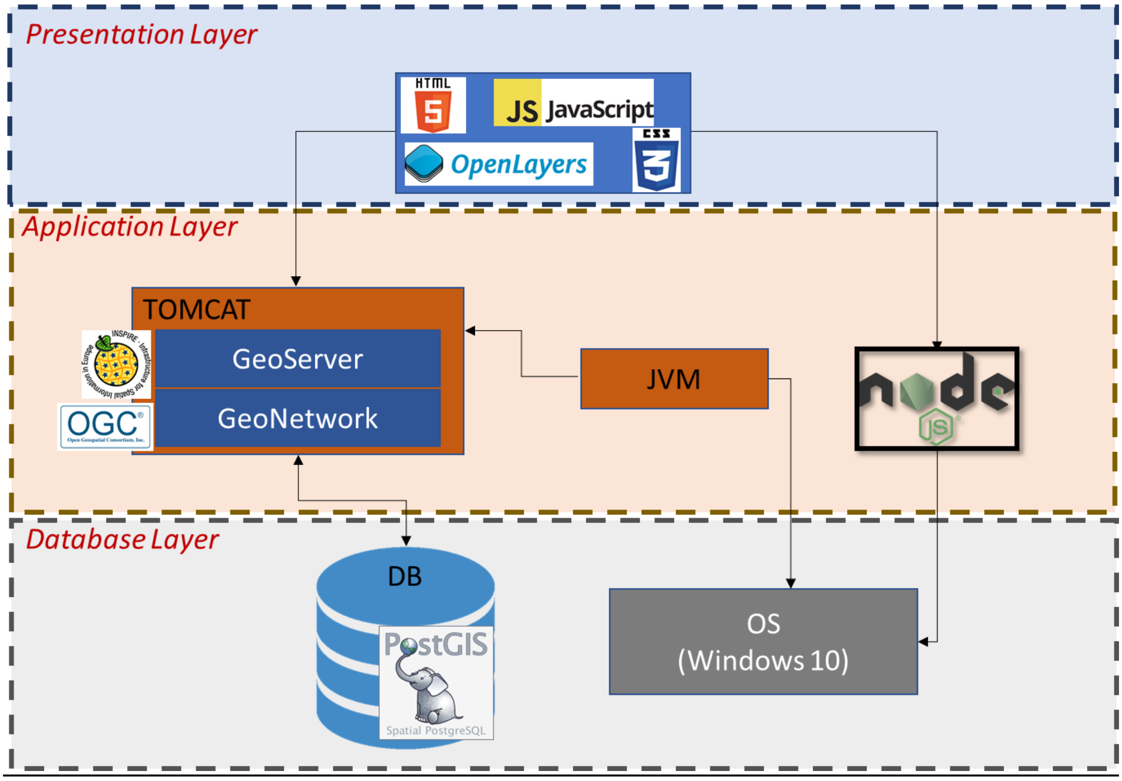


Figure 32: Architectural overview of Spatial Data Infrastructure for night-time imagery – Software components communication diagram

<u>*Apache Tomcat*</u>

Apache Tomcat is an application server and web server developed by the Apache Software Foundation. Tomcat consists of an Implementation of Java Servlet, JSP technologies – typically run in conjunction with Apache or other HTTP servers (http://tomcat.apache.org/).

This tool provides the capabilities of:

- Setting up the Java security manager;
- Using Java Database Connectivity (JDBC) data sources;
- Secure Sockets Layer (SSL) support for encrypted connections;
- Using proxies;
- Virtual hosting;
- Running Common Gateway Interface (CGI) scenarios.

In the context of the proposed Spatial Data Infrastructure, Tomcat operates both as Application Server and Web Server. As an application server, Tomcat provides the required environment for creating and running web applications (GeoServer and GeoNetwork).

GeoServer

GeoServer is an open-source, Java-based server that is interoperable and compatible with OGC Standards (WMS, WFS, WCS, WMTS, WPS, CSW, etc.) and the specifications of INSPIRE Directive, either inherently or through extensions that add this functionality to it. It is worth mentioning that GeoServer consists the reference implementation for WFS. In addition, it supports the SLD standard for the standardization of the symbols used in the maps. GeoServer allows both users and client applications (web browsers, GIS software) to share and process geospatial data. It has the ability to publish data from any source of spatial data using open standards. It supports all widely used reference systems, including the Greek Reference System, and allows for automatic transformation between them. It can read a variety of different data formats but also retrieve data from many relational databases, such as: PostGIS, Oracle Spatial, ArcSDE, DB2, MySQL, Shapefiles, GeoTIFF, GTOPO30, ECW, JPEG2000. Through standardized processes, it can produce data in various formats such as: KML, GML, Shapefile, GeoRSS, PDF, GeoJSON, JPEG, GIF, SVG, PNG and others (http://geoserver.org/).

GeoServer 2.12.1 was downloaded as a Web application Archive (WAR) file type and subsequently deployed on Apache Tomcat servlet container. To meet the needs of the current SDI, the base GeoServer installation was extended with different external modules (add-ons) such as the INSPIRE and Image Mosaic JDBC extensions.

The INSPIRE extension enables the creation of INSPIRE compliant view and download services. This is achieved through the inclusion of additional elements into the capabilities requests of WMS, WFS and WCS. These are the following (http://geoserver.org/):

- **Language**: Enables the definition of the response language (mandatory field);
- **Service Metadata URL**: Contains a URL that points to the metadata associated with the service (mandatory field);
- **Service Metadata Type**: Allows the selection of the appropriate metadata type, which could be either a CSW or a standalone metadata file (optional field).

The Image Mosaic JDBC extension supports the efficient serving of the VIIRS monthly composites that have been previously stored in a geo-database, though GeoServer. The JDBC Image Mosaic extension has been designed to support a range of different databases, irrespective of whether they have a spatial extension. The extension reads the raster data from a column in the database where the data has been stored in its binary form, usually as a BLOB (http://en.wikipedia.org/wiki/Binary_large_object). To make use of the extension, the following steps have to be conducted:

1. Prepare and load the raster data into the database;
2. Create a metadata table for the extension;
3. Create extension configuration files;
4. Create a data store in GeoServer linked to the configuration files;
5. Publish the contents of the data store as a layer.

The first step has already been undertaken since the development of the database layer of the Spatial Data Infrastructure. The use of PostGIS, facilitates the integration process, as the JDBC Image Mosaic plugin has support for the built-in raster data type of the database. Moving to the second step, the JDBC mosaic plugin requires a metadata table to be present in the database. To this end, a table called MOSAIC is developed inside the VIIRS database though the following SQL statement:

```
CREATE TABLE mosaic (
        name varchar(254) not null,
        tiletable varchar(254) not null,
        minx FLOAT8,
        miny FLOAT8,
        maxx FLOAT8,
        maxy FLOAT8,
        resx FLOAT8,
         resy FLOAT8,
        primary key (name, tiletable)
);
```

Following the table’s creation, it is populated with some values. The values need to describe the tables where the raster data can be found.

```
INSERT INTO public.mosaic(
        name, tiletable, maxx, maxy, minx, miny, resx, resy)
        VALUES (‘viirs_012013', 'viirs_012013', 7082533.209106532, 5843057.225498975,
999180.3996763239, 992344.7608285099, 436.3328654017, -436.3328654017);
```

For the plugin to function correctly, it needs to have an XML file that describes the mapping necessary to allow it to get the raster files from the database. The configuration file has three sections to it describing the database connection details, table mappings, and coverage properties. Following the best practice, the configuration file is divided into three separate files, and an XML inclusion is used to bring them all together into a mosaic configuration. Thus, inside the GeoServer data directory a new text file called *connect.pgraster.xml.inc* is created and populated with the following XML fragment.

```
<connect>
  <!-- value DBCP or JNDI -->
  <dstype value="DBCP"/>
  <!--   <jndiReferenceName value=""/>  -->
  <username value="postgres" />
  <password value="tsiakos" />
  <jdbcUrl value="jdbc:postgresql://localhost:5432/viirs" />
  <driverClassName value="org.postgresql.Driver"/>
  <maxActive value="10"/>
  <maxIdle value="0"/>
</connect>
```

Then the table *mapping pgraster.xml.inc* is created that includes the table mappings fragment as follows. This fragment provides all the necessary information about the metadata table to the JDBC Image Mosaic plugin. It is an important element of the configuration since it allows us to specify our own metadata table.

```
<spatialExtension name="pgraster"/>
<mapping>
    <masterTable name="MOSAIC" >
        <coverageNameAttribute name="name"/>
        <maxXAttribute name="maxX"/>
        <maxYAttribute name="maxY"/>
        <minXAttribute name="minX"/>
        <minYAttribute name="minY"/>
        <resXAttribute name="resX"/>
        <resYAttribute name="resY"/>
        <tileTableNameAtribute name="tiletable" />
    </masterTable>
    <tileTable>
        <blobAttributeName name="rast" />
    </tileTable>
</mapping>
```

Last but not least, an XML file is created, importing the separate components that are required for a complete configuration. This file will be the one that we point GeoServer when creating the data store. The name of the file can be random, but it is suggested to use the following naming convention: [coverage_name]_[spatial_type].xml

```
<?xml version="1.0" encoding="UTF-8" standalone="no"?>
<!DOCTYPE ImageMosaicJDBCConfig [
    <!ENTITY mapping PUBLIC "mapping"  "mapping.pgraster.xml.inc">
    <!ENTITY connect PUBLIC "connect"  "connect.pgraster.xml.inc"> ]>
<config version="1.0">
    <coverageName name="01_2013"/>
    <coordsys name="EPSG:3035"/>
    <scaleop  interpolation="1"/>
    <axisOrder ignore="false"/>
    &mapping;
    &connect;
</config>
```

An example of WMS request for the VIIRS monthly composite of January 2013 to be displayed as a PNG map image in SRS EPGS:3035 and using default styling is presented below.

```
http://127.0.0.1:8080/geoserver/viirs/wms?
service=WMS
&version=1.1.0
&request=GetMap
&layers=viirs:viirs_012013
&styles=
&bbox=999180.3996763239,992344.7608285099,7082533.209106532,5843057.2254975&width
=768
&height=612
&srs=EPSG:3035
&format=image/png
```

Java Virtual Machine (VM)

Java is designed to support network applications. A network, however, consists of a variety of different systems, with different CPUs and operating systems. In order for Java applications to run everywhere on the network, the Java program must go through two processes to make it executable. First, a compiler converts the source code of the program into an intermediate language called Java bytecodes. Then an interpreter is used to convert each bytecode command into a suitable binary format to run on the computer. Compilation occurs only once for each Java program, whilst interpretation is performed every time the program is running.

Each Java interpreter (i.e. a web browser that can run java applets) is a Java Virtual Machine (JVM) software application. JVM undertakes to convert the bytecodes into appropriate executable form, depending on the underlying software and hardware.

The technique described above is called "write once, run anywhere". The Java program is compiled into Java bytecodes with the Java compiler only once. Then, the bytecodes can run on any machine that has an installed JVM (https://en.wikipedia.org/wiki/Java_virtual_machine).

In the context of this thesis, Java Development Kit (JDK) 8u161 was used as provided from Oracle Corporation or by the OpenJDK open source project. JDK provides several tools for Java application programming such JVM, javac compiler, etc.

Node.js

Node.js is an open source, cross-platform runtime environment for server-side application programming and is built in a JavaScript environment. Node.js provides an event-based architecture and a non-blocking I/O API designed to optimize the performance and scalability of an application. Unlike most modern network deployment environments, a node process relies on an asynchronous input / output communication model and contains an embedded library to allow applications to act as an autonomous web server (https://nodejs.org/en/).

Node.js is used to to build and run the web portal of the Spatial Data Infrastructure. The main reason for using Node.js lies on the fact that Openlayers web mapping library, which is utilised for the development of the presentation layer, is based on Node.js.

GeoNetwork

GeoNetwork is a free and open catalogue server application for geo-spatial resources that implements the OGC CSW standard. It provides a standardized and decentralized spatial information management environment designed to allow access to spatial databases, geospatial information and related metadata from a variety of sources whilst improving the exchange and sharing of spatial information between information systems and their users. GeoNetwork utilizes an easy-to-use web interface and offers advanced custom search capabilities with all the fields included in metadata files. The catalogue within this application may include metadata referring either to geospatial data sets or to web-based geospatial services(https://geonetwork-opensource.org/).

GeoNetwork is downloaded as a Web application Archive (WAR) file and then is deployed to the servlet container, which is Tomcat in our case. GeoNetwork supports the creation of different metadata profiles and in particular:

- **Dublin Core** metadata standard which was developed by the Federated Geographic Data Committee and is used to describe different data (not only spatial data);
- **ISO 19139/19115** metadata standard which is dedicated to the spatial data documentation (ISO 19139:2007 and ISO 19115:2005). The following figure 33 presents the relationship between ISO 19139 and the other ISO metadata standards. The ISO 19139 Schema (ISO 19139:2007) defines the technical encoding specifications of the metadata. ISO 19115 “provides a comprehensive vocabulary and structure of metadata used to describe geographic data, whereas ISO 19119 provides the vocabulary and metadata elements used to describe the published (geo) web

services" (GSDI Association, 2009). ISO 19115 standard can be extended to accommodate the characteristics of different spatial data structures such as gridded data and imagery or it can be extended to reflect the data documentation requirements specific to different countries: e.g. North American profile of 19115.

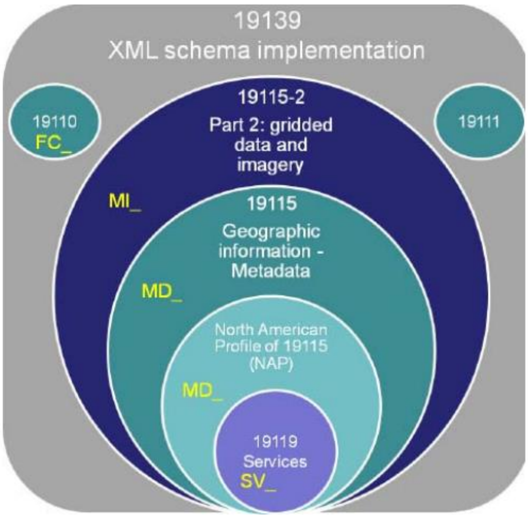


Figure 33 ISO 19139 other International standards (Longley et al., 2005)

- **ISO 19139/19119** metadata standard describes the structure of a metadata model for spatial web service instances to help searching, discovery and using available services (ISO 19119:2003)

Through the available templates metadata have been created for both the datasets and services of the Spatial Data Infrastructure. Then the metadata have been validated though the INSPIRE validator (http://inspire-geoportal.ec.europa.eu/validator2/) so as to check the conformance with the Directive and relevant metadata regulations (figure 34).

To claim INSPIRE metadata conformance, two steps are required:

1. "Discovery Metadata" validation: Conformance to Commission Regulation (EC) N°1205/2008 also referred to as "Implementing Rules for Metadata". To this this, each metadata record should include the following elements:
    - Resource title (mandatory)
    - Resource abstract (mandatory)
    - Resource type (mandatory)
    - Resource locator (conditional)
    - Unique resource identifier (mandatory)
    - Resource language (conditional)
    - Topic category (mandatory)
    - Keyword (mandatory)
    - Geographic bounding box (mandatory)

- Temporal reference (mandatory)
- Lineage (mandatory)
- Spatial Resolution (conditional)
- Conformity (mandatory)
- Conditions for access and use (mandatory)
- Limitation on public access (mandatory)
- Responsible organization (mandatory)
- Metadata point of contact (mandatory)
- Metadata date (mandatory)
- Metadata language (mandatory)

2. “Metadata for interoperability” validation: Conformance to Commission Regulation (EU) N°1089/2010 also referred to as “Implementing Rules for Interoperability of Spatial Datasets and Services “or ISDSS Regulation. In order to be conformant to Commission Regulation (EU) N°1089/2010, the metadata describing a spatial data set shall include the following metadata elements:
    - Coordinate reference system (mandatory)
    - Temporal reference system (conditional)
    - Encoding (mandatory)
    - Character encoding (conditional)
    - Spatial representation type (mandatory)
    - Data Quality – Logical consistency – Topological consistency (conditional)

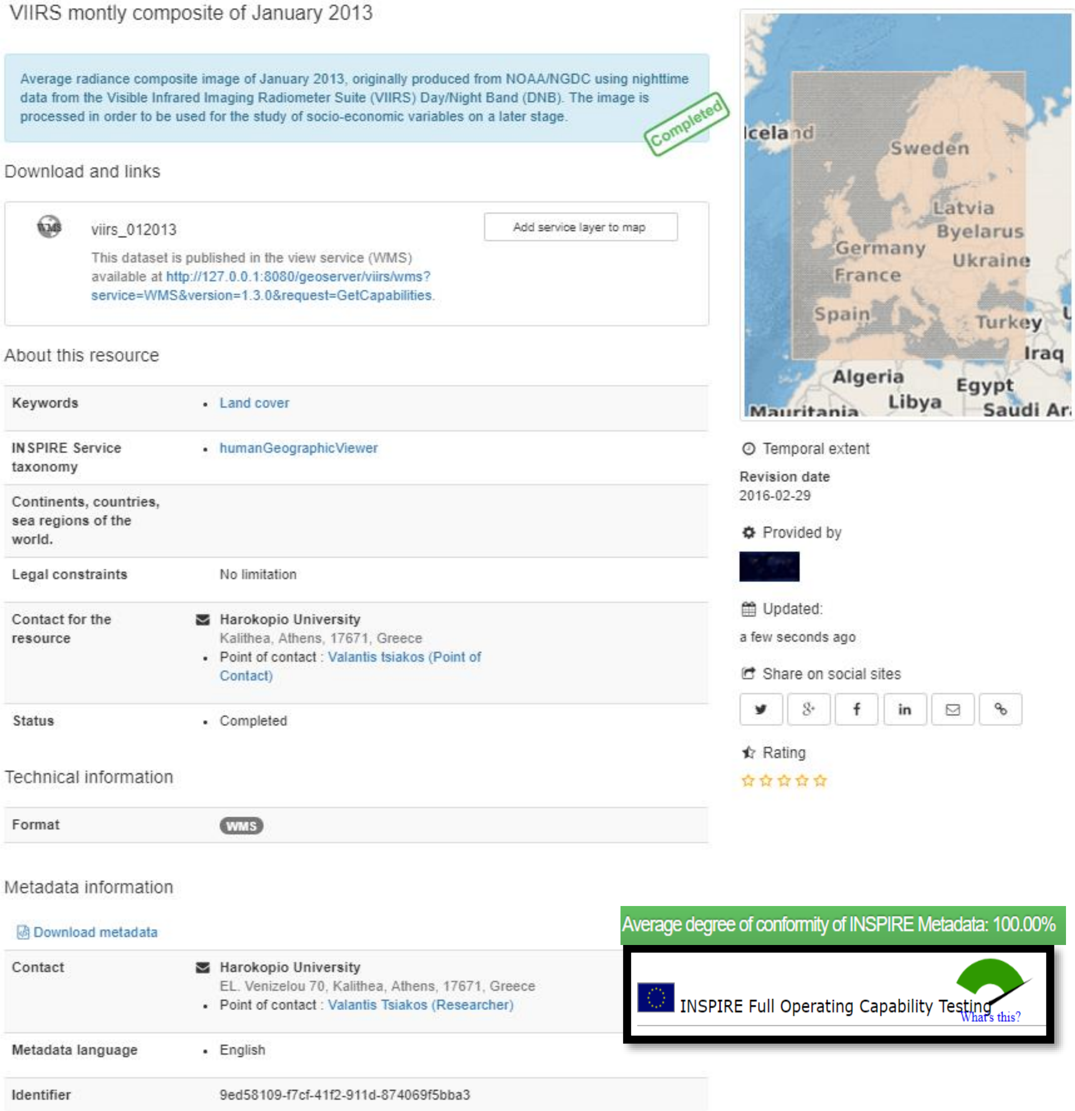


Figure 34: Example of INSPIRE compliant metadata record in GeoNetwork catalogue

5.3 Presentation layer

The presentation layer consists of a combination of software tools and technologies aiming at providing access to the services and datasets developed in the context of this thesis. An overview of the utilized technological stack is presented below.

HTML5

HTML (Hyper Text Markup Language) is a markup language for the World Wide Web, that is, the language with which we build web pages. The websites we visit on the Internet are nothing more than files that contain code written in the HTML language. Browsers read these files and display the result of the HTML code on our screen. HTML5 is the new standard for HTML, XHTML and HTML DOM, developed from the collaboration of the World Wide Web Consortium (W3C) and the Web Hypertext

Application Technology Working Group (WHATWG). WHATWG worked on web forms and web applications, while W3C, which created and manages HTML and XHTML templates, dealt with the development of the new XHTML 2.0 standard.

JavaScript

JavaScript (JS) is an interpreted programming language for computers. Its logic is based on the concept of prototype-based (object oriented). Its syntax is influenced by C language. JavaScript copies many names and naming conventions from Java, but generally these two languages are not related and have very different semantics. JavaScript is a language based on different programmatic examples (multi-paradigm), supporting object-oriented imperative and functional programming styles.

Typically, it is used for client-side programming, which is the user's browser, and in these cases is characterized as client programming language. This means that the processing of JavaScript code and the production of the final HTML content is not done on the server but in the guest browser, and can be incorporated into static HTML pages.

Unlike most programming languages, JavaScript does not come with I / O capabilities. It is designed as a scripting language in an environment that hosts it, and the existence of mechanisms of communication with the outside world is the responsibility of this environment. The most common hosting environment is the browser, but JavaScript interpreters can be found in various programs (eg Adobe Acrobat, Adobe Photoshop, etc.).

CSS

CSS (Cascading Style Sheets) is a formatting language used to control the appearance of a document written with a markup language. With CSS, a homogeneous formatting (style) is added in a uniform and easily verifiable way to all pages of a site or portal. Formatting features can be related to fonts, colors, text alignment points, etc. For a stylish and well-designed website the use of CSS is deemed necessary.

Using CSS, we can define colors and styles, and then apply them to the elements of our site. This way, each time we change the color of a style, it changes the color of all elements that have a reference to that style. Further to the ease of managing a site, another important advantage of using CSS on pages is the "cleaner" code, without many labels that make it unreadable. In addition, it makes navigating faster as the file in which styles are defined is "read" by the browser only once and then stored in the cache memory, thus reducing the amount of information downloaded by browsers.

OpenLayers

OpenLayers is an open source, client-side JavaScript library for making interactive web maps, viewable in nearly any web browser. OpenLayers is a map engine that provides an API (Application Program

Interface) that can be used to develop various web mapping applications. Instead of building a mapping application from scratch, OpenLayers can be used for the mapping part, which is maintained and developed by a dedicated community. Since it is a client-side library, it requires no special server-side software or settings.

Night light imagery Geoportal

Some of the functionalities offered by the web geoportal are the following:

- **Map content**: Interface to display the connected data starting from layers, supporting flagging / deflagging of the relative check box in order to add /remove the related data on the map (figure 35);
- **Geocoding**: Search field that converts addresses into geographic coordinates, which can be placed on a map;
- **Coordinate reference systems**: The portal is able to manage different coordinate reference system;
- **Map navigation**: The following basic map navigation functionalities are supported:
    - Zoom-in
    - Zoom-out
    - Display attribute feature
- **User authorization:** The portal supports registration and login functionalities (figure 37);
- **Overlays**: Different layers can be overlaid and displayed in different order through the portal (figure 36);
- **Layer switching**: The user can select between different base-maps in the context of its application.

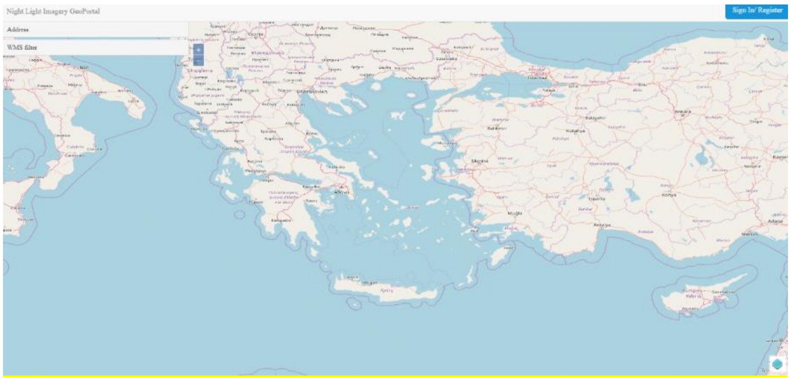


Figure 35: General arrangement of Night Lights Imagery Geoportal

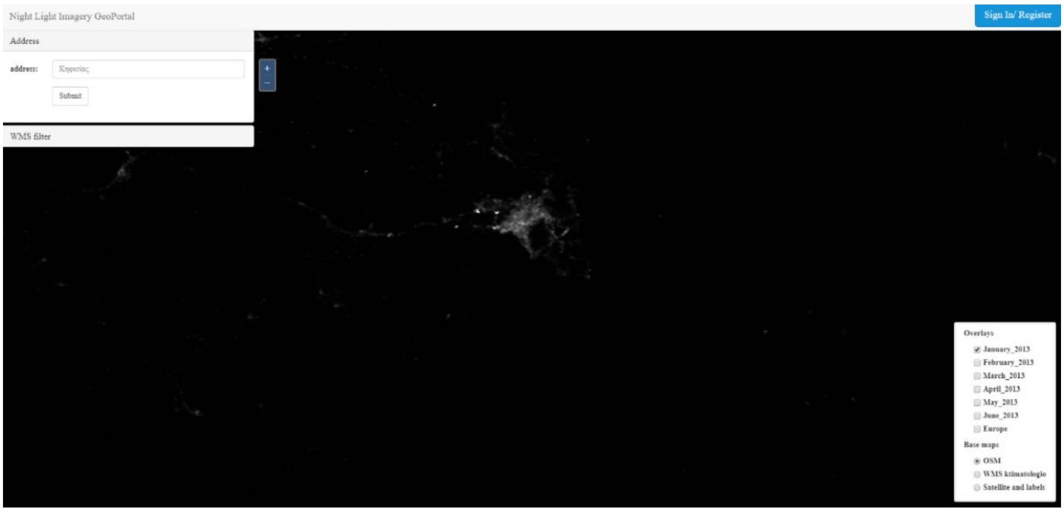


Figure 36: Display of VIIRS monthly composite for January 2013

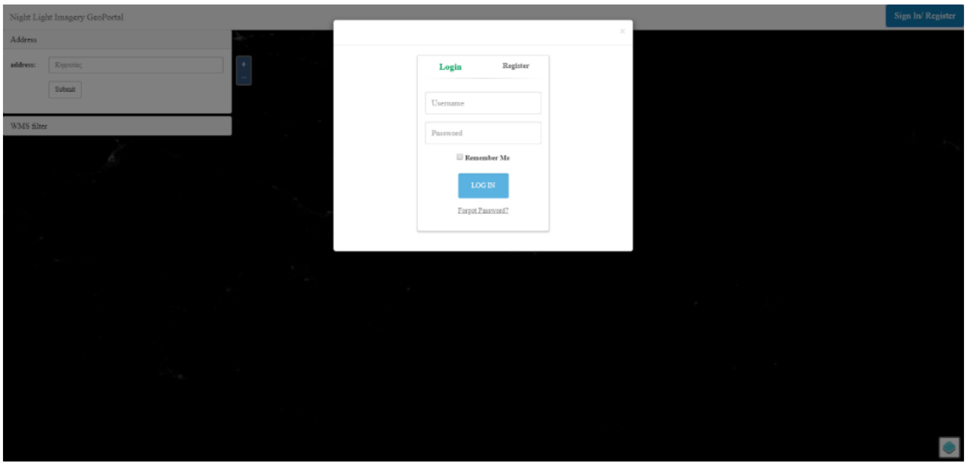


Figure 37: User registration/login interface

6. Conclusions

The analysis and results coming from this thesis clarify the landscape around the development of the necessary tools and methods for the utilization of VIIRS DNB monthly composites for scope-varying applications. Firstly, the implementation of a processing chain for VIIRS datasets, provides positive results for their utilization in studying over time various socioeconomic, and other phenomena. On the other hand, these processed datasets need to be made available to the end users, whilst adhering to well-known standards and policies. To this, a prototype Spatial Data Infrastructure has been designed and implemented according to INSPIRE and OGC standards. Future research should focus on more complex systems that will support the synergistic use of earth observation products with sensor-based measurements and/or crowdsourced information.

7. References

Elvidge C.D., Cinzano P., Pettit D.R., Arvesen J., Sutton P., Small C., Nemani R., Longcore T., Rich C., Safran J., Weeks J. and Ebener S., (2007), *The Nightsat mission concept*, International Journal of Remote Sensing, 28 (12), pp. 2645-2670.

Baugh K., Elvidge C.D., Ghosh T. and Ziskin D., (2010), *Development of a 2009 stable lights product using DMSPOLS data*, Proceedings of the Asia-Pacific Advanced Network, 30, pp. 114-130.

Baugh K, Hsu F. C., Elvidge C. and Zhizhin M., (2013), *Nighttime Lights Compositing Using the VIIRS Day-Night Band: Preliminary Results*, Proceedings of the Asia-Pacific Advanced Network, 35, pp. 70-86.

Brand M., (1998), Theme Paper: Global Spatial Data Infrastructure: Policy & Organisational Issues, available at http://www.eurogi.org/gsdi/canberra/theme.html [last accessed 10 January 2018].

Chan T. O. and Williamson I. P., (1999), *Spatial data infrastructure management: lessons from corporate GIS development*, AURISA 99 - The 27th Annual Conference of AURISA. Blue Mountains, New South Wales, 22-26 November 1999.

Directive 2003/4/EC of the European Parliament and of the Council of 28 January 2003 on public access to environmental information and repealing Council Directive 90/313/EEC, https://inspire.ec.europa.eu/ [last accessed 10 January 2018].

Directive 2003/98/EC of the European Parliament and of the Council of 17 November 2003 on the re-use of public sector information, https://inspire.ec.europa.eu/ [last accessed 10 January 2018].

Directive 2007/2/EC of the European Parliament and of the Council of 14 March 2007 establishing an Infrastructure for Spatial Information in the European Community (INSPIRE), https://inspire.ec.europa.eu/ [last accessed 10 January 2018].

Directive 2013/37/EU of the European Parliament and of the Council of 26 June 2013 amending Directive 2003/98/EC on the re-use of public sector information, https://inspire.ec.europa.eu/ [last accessed 10 January 2018].

EEA, European Environmental Agency, https://www.eea.europa.eu/ [last accessed 10 January 2018].

Elvidge C.D., Baugh K.E., Sutton P.C., Bhaduri B., Tuttle B.T., Ghosh T., Ziskin D. and Erwin E.H., (2011), *Who's in the dark: Satellite based estimates of electrification rates*, In Yang, X. (ed.) Urban remote sensing: Monitoring, synthesis and modeling in the urban environment. Chichester: John Wiley & Sons, pp. 211-224.

Ezra Dessers, (2013), *Spatial Data Infrastructure at Work: Analysing the Spatial Enablement of Public Sector Processes*, Leuven University Press, ISBN 978 90 5867 937 6.

Groot R. & J. McLaughlin, (2000), *Chapter 1. Geospatial Data Infrastructure, Concepts, Cases and Good Practise*, Groot & Mclaughlin, Oxford University Press.

GSDI Association, (2009), *SDI Cookbook* [online]. Available at http://www.gsdi.org/gsdicookbookindex [last accessed 27 November 2017].

Hillger D., Kopp T., Lee T, Lindsey D., Seaman C., Miller S., Solbrig J., Kidder S., Bachmeier S., Jasmin T., (2013), *First-light imagery from Suomi NPP VIIRS*, Bull. Am. Meteorol. Soc. 2013, 94, pp.1019–1029. ISSN 2072-4292.

Holt C. C., (2004), *Forecasting Seasonals and Trends by Exponentially Weighted Moving Averages*, International Journal of Forecasting 20 (1): 5–10. doi:10.1016/j.ijforecast.2003.09.015.

Hyndman J. R. and Athanasopoulos G., (2013), *Forecasting: principles and practice*, OTexts: Melbourne, Australia, ISBN 978-0-9875071-0-5.

Hyndman R. J., G. Athanasopoulos, S. Razbash, D. Schmidt, Z. Zhou, Y. Khan, and E. Wang, (2014), R Package: Forecast, http://cran.r-project.org/web/packages/forecast/index.html [last accessed 10 January 2018].

INSPIRE DS-D2.3, Definition of Annex Themes and Scope, v3.0, http://inspire.jrc.ec.europa.eu/reports/ImplementingRules/DataSpecifications/D2.3_Definition_of_Annex_Themes_and_scope_v3.0.pdf [last accessed 10 January 2018].

INSPIRE DS-D2.5, Generic Conceptual Model, v3.3, http://inspire.jrc.ec.europa.eu/documents/Data_Specifications/D2.5_v3_3.pdf [last accessed 10 January 2018].

INSPIRE DS-D2.6, Methodology for the development of data specifications, v3.0, http://inspire.jrc.ec.europa.eu/reports/ImplementingRules/DataSpecifications/D2.6_v3.0.pdf [last accessed 10 January 2018].

INSPIRE DS-D2.7, Guidelines for the encoding of spatial data, v3.2, http://inspire.jrc.ec.europa.eu/documents/Data_Specifications/D2.7_v3.2.pdf [last accessed 10 January 2018].

INSPIRE DS-D2.8.II/III.x, Data Specification on <Theme Name> – Technical Guidelines, v3.0rc3, http://inspire.ec.europa.eu/documents/Data_Specifications/INSPIRE_DataSpecification_Template_v3.0rc3.pdf [last accessed 10 January 2018].

Jing X., Shao X., Cao C., Fu X. and Yan L., (2016), *Comparison between the Suomi-NPP Day-Night Band and DMSP-OLS for Correlating Socio-Economic Variables at the Provincial Level in China*, Remote Sens. 2016, 8, 17; doi:10.3390/rs8010017.

Kyba C., Garz S., Kuechly H., Sánchez de Miguel A., Zamorano J., Fischer, J., Höelker F., (2015), *High-resolution imagery of earth at night: new sources, opportunities and challenges,* Remote sensing, vol. 7 (1). pp. 1-23. ISSN 2072-4292.

Lee C., Cao C., (2016), *Soumi NPP VIIRS Day/Night Band Stray Light Characterization and Correction Using Calibration View Data,* Remote Sens. 2016, 8, 138; doi:10.3390/rs8020138.

Lee S., Chiang K., Xiong X., Sun C. and Anderson S., (2014), *The S-NPP VIIRS Day-Night Band On-Orbit Calibration/Characterization and Current State of SDR Products*, Remote Sens. 2014, 6, 12427-12446; doi:10.3390/rs61212427.

Li X., Xu H., Chen X., Li C., (2013), *Potential of NPP-VIIRS nighttime light imagery for modeling the regional economy of China*, Remote Sens. 2013, 5, 3057–3081.

Longley P., M. Goodchild, D. Rhind and D. Maguire, (2005), *Geographical Information Systems and Science*, 2nd edition, Chichester, UK, J. Wiley & Sons Ltd.

Miller S. D., S. P. Mills, C. D. Elvidge, D. T. Lindsey, T. F. Lee, and J. D. Hawkins, (2012), *Suomi satellite brings to light a unique frontier of nighttime environmental sensing capabilities*, Proc. Natl. Acad. Sci. U. S. A., 109(39), 15,706–15,711.

Miller S.D., Straka W., Mills S.P., Elvidge C.D., Lee T.F., Solbrig J., Walther A., Heidinger A.K., Weiss S.C., (2013), *Illuminating the Capabilities of the Suomi National Polar-Orbiting Partnership (NPP) Visible Infrared Imaging Radiometer Suite (VIIRS) day/night band,* Remote Sens. 2013, 5, 6717–6766.

Mills S., Jacobson E., Jaron J., McCarthy J., Ohnuki T., Plonski M., Searcy D., Weiss S., (2010), *Calibration of the VIIRS Day/Night Band (DNB)*, In Proceedings of American Meteorological Society 6th Annual Symposium on Future National Operational Environmental Satellite Systems-NPOESS and GOES-R, Atlanta, GA, USA, 16–21 January 2010.

Mills S., S. Weiss, and C. Liang, (2013), *VIIRS day/nightband (DNB) stray light characterization and correction*, Proc. SPIE, 8866, 886663.

NOAA https://ngdc.noaa.gov/eog/VIIRS/download_dnb_composites.html [last accessed 10 January 2018].

OGC 06-042, OpenGIS® Web Map Service (WMS) Implementation Standard, version 1.3

OGC 07-006r1, OpenGIS® Catalogue Services Specification, version 2.0.2

Rajabifard A., Feeney M.E. F. and Williamson I. P., (2002), *Future directions for SDI development*, International Journal of Applied Earth Observation and Geoinformation, 4 (1): 11-22.

Shao X., Cao C., Uprety S., (2013), *Vicarious calibration of S-NPP/VIIRS day-night band*, Proc. SPIE 2013, 8866.

Shi K., Huang C., Yu B., Yin B, Huang Y, Wu J., (2014b), *Evaluation of NPP VIIRS night-time light composite data for extracting built-up urban areas*, Remote Sensing Letters, 5 (4), pp. 358-366.

Shi K., Yu B., Huang Y., Hu Y., Yin B., Chen Z., Chen L., Wu J., (2014a), *Evaluating the Ability of NPP-VIIRS Nighttime Light Data to Estimate the Gross Domestic Product and the Electric Power Consumption of China at Multiple Scales: A Comparison with DMSP-OLS Data*, Remote Sensing, 2014a,6, pp. 1705-1724.

Small C., Pozzi F. and Elvidge C.D., (2005), *Spatial analysis of global urban extent from DMSP-OLS night lights*, Remote Sensing of Environment, 96 (3-4), pp. 277-291.

Stathakis D., Perakis K. and Savin I., (2012), *Efficient segmentation of urban areas by the VIBI*, International Journal of Remote Sensing, 33 (20), pp. 6361-6377.

Stefanakis E., (2014), *Geographic Databases and Information Systems*, Edition 1.0, pp. 386, Fredericton, NB, Canada

Steiniger S., Hunter A.J.S., (2012), *Free and Open Source GIS Software for Building a Spatial Data Infrastructure*, In: Bocher E., Neteler M. (eds) Geospatial Free and Open Source Software in the 21st Century. Lecture Notes in Geoinformation and Cartography. Springer, Berlin, Heidelberg.

Straka W., Seaman C., Baugh K., Cole K., Stevens E. and Miller S., (2015), *Utilization of the Suomi National Polar-Orbiting Partnership (NPP) Visible Infrared Imaging Radiometer Suite (VIIRS) Day/Night Band for Arctic Ship Tracking and Fisheries Management.* Remote Sensing, vol. 7. pp. 971-989. ISSN 2072-4292.

Vandenbroucke D. and Janssen K., (2008), *Spatial Data Infrastructures in Europe: State of Play 2007 Leuven*, Belgium: K.U.Leuven.

Wade T. and Sommer S., (2006), *A to Z GIS. An illustrative dictionary of geographic information science,* (2nd edition pp. 1-265), Redlands: ESRI Press

Zhang Q., Schaaf C. and Seto K.C., (2014), *The vegetation adjusted NTL urban index: A new approach to reduce saturation and increase variation in nighttime luminosity*, Remote Sensing of Environment, 129, pp. 32-41.