within a 10km radius around all of these points was visually inspected - using a Sentinel-2 cloudless mosaic from 2019 as a background layer and data from Google Earth and Microsoft Bing as auxiliary sources in case of ambiguity - and all land cover features mentioned above were delineated [31]. While the 10km buffer is likely to also enable the correct delineation of large mines which can spread out over many kilometres, the dataset is unlikely to be a perfectly random sample of global mining sites as mines in less transparent jurisdictions might systematically be underrepresented. Consequently, this limitation is also likely present in the dataset proposed in this thesis.

The accuracy of the delineations were independently validated by a second group of people categorising randomly selected points within the area of interest as either being within the mining area or not [32]. This classification was then compared with the delineations from the dataset. From all evaluated points, 88.3% were classified congruently to the area delineated by the polygons in the dataset. However, most errors resulted from missing mines - approximately 20% of the points classified as mines were not within the polygons of the dataset - whereas the likelihood that a mapped mine is, indeed, part of a mining area was with 97.3% very high [32]. The later metric suggests that there is a low risk of having light emission from non-mining infrastructure included in our dataset. There is, however, the risk that nighttime light emission from mining activities of locations which are not included in the area of the polygons might be missed.

2.2 Nighttime Light Emissions

The second input used for the dataset proposed in this thesis is data on nighttime light emissions. In this section, we will take a look at the origin and features of this data.

As discussed in the introduction, measurements of anthropogenic nighttime light emissions constitute a valuable source of information which allows to explore human activity on a spatially explicit way. However, the potential of this data was discovered only by chance in the mid 1970's during an attempt by the U.S. Air Force to improve the accuracy of meteorological predictions within the Defense Meteorological Satellite Program (DMSP) [36]. The initial goal was to measure moonlight reflection from clouds during the night which would allow for better cloud coverage detection but it soon became apparent that these instruments were also capturing light emissions from human activity [36]. Over time, a community of researchers emerged, which were primarily interested in this 'byproduct'.

The nighttime light emissions used in this thesis are measured by the Suomi-National Polar-orbiting Partnership (S-NPP) and Joint Polar Satellite System (JPSS) satellite platform using the Day/Night Band (DNB) sensors contained in the Visible Infrared Imaging Radiometer Suite (VIIRS) [37]. This new generation of NTL measurement instrument provides major improvements with regards to resolution, sensitivity and also avoids known problems with the DMSP data such as sensor saturation on bright lights, for example in city centers [38]. S-NPP and JPSS is an ongoing satellite program by the National Aeronautics and Space Administration (NASA) and the National Oceanic and Atmospheric Administration (NOAA) which started in 2011 with the launch of S-NPP [38]. The JPSS system consists of 4 satellites two of them have been launched in 2017 and 2021 respectively while the other two are scheduled for 2027 and 2032 [39].

All of the satellites in S-NPP and JPSS do already or will have the VIIRS onboard [40] which conducts daily measurements in 22 spectral bands ranging from visible to the long-wave parts of the electromagnetic spectrum using a scanning radiometer with a rotating telescope assembly.

For this thesis of particular interest is the panchromatic Day/Night Band (DNB) designed to enable low light detection. The benefit of using a panchromatic band for this task is an increased resolution which is in the case of the VIIRS/DNB products 742 on 742 meters on the ground per pixel [38]. This level of resolutions allows for a detailed representation of human activity as can be seen in Figure 2. However, this comes at the cost of losing information on colour. In the figures of this thesis, a colour map has been applied to the single band values in order to improve the readability of the visualisations.

The orbit of the satellites are polar, one circumnavigation of the earth takes 102 minutes and images are taken at around 1:30 a.m. local time [38]. In our case, this means that the images obtained from this instrument can only capture mining activities taking place at that specific point of the night.

As Román et al. [37] point out, 'intrinsic surface optical properties' such as NTL emissions cannot be inferred from 'at-sensor top-of-atmosphere (TOA) radiances' directly, since there are various light sources and chemical phenomena present between the light source on the ground and the sensor which can potentially distort the data collection. As an example, during nights where the moon is visible, its light will inevitably add to the radiance measured by the DNB an, thus, cause biased measurements of human activity. Another factor which needs to be taken into con-

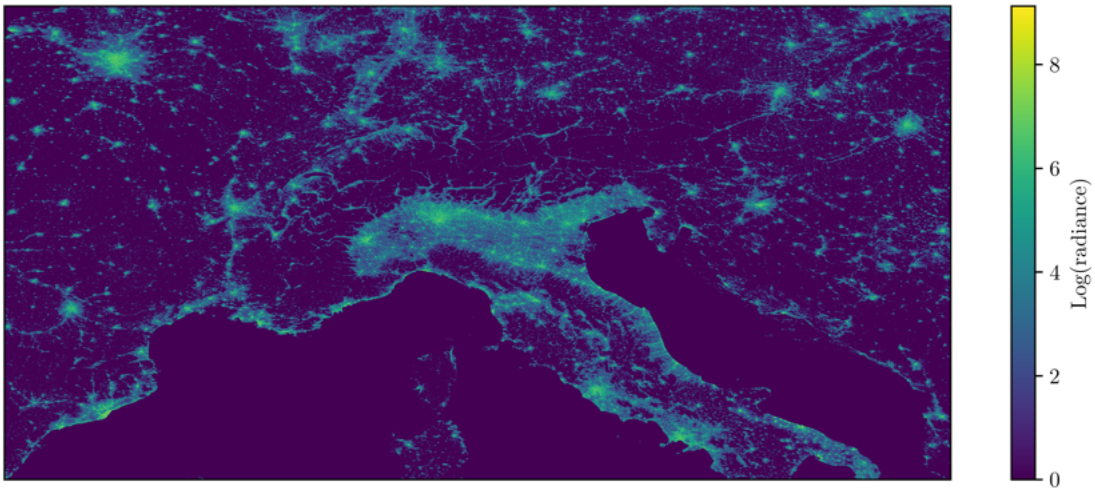

Figure 2: Yearly composite from 2020 of nighttime light emissions captured by VIIRS/DNB. Note the stark contrast of the Swiss Alps and the Po Valley. Cities such as Paris and Rome are easily discernible. Data Source: NASA [41].

sideration are chemical properties of the atmosphere which can cause phenomena such as air glow [37]. All of these factors need to be controlled and corrected for which makes it necessary to conduct substantial post-collection processing on the data gathered by the sensor.

For this thesis, the nighttime light emissions data from NASA's Black Marble product suite were used. This suite has a level 3 processing which, in particular, includes atmospheric correction, bidirectional reflectance distribution function (BRDF) correction and seasonal vegetation correction[37] [42]. From the Black Marble suite, the products called VNP46A3 and VNP46A4 were selected. VNP46A3 stands for 'VIIRS/NPP Lunar BRDF-Adjusted Nighttime Lights Monthly L3 Global 15 arc-second Linear Lat Lon Grid' [43]. The long name for VNP46A4 is analogue with the sole difference of 'yearly' being replaced by 'monthly' which also reflects the main difference between the products.

VNP46A3 and VNP36A4 products are composites which take the mean of the values from all available daily measurements within their respective interval from the VNP46A2 product [44], to give an overview of the NTL emissions in their indicated frequency. Both products also include multiple datasets with different options of compilation such as off-nadir, near-nadir and alternatives with or without snow coverage a given location and date [44]. As the aim of this thesis is to obtain an impression of mining activities over the year and considering the low risk of bias due to high rise building covering up lights emitted horizontally, we opted for the

the snow-free composite consisting out of data from all capture-angles.

The measuring unit of the NTL is radiance and given by $nWatts * cm^{-2} * sr^{-1}$, were sr stands for steradian and is a measure of solid angle [44]. The scale of the values ranges from zero, i.e. no light emissions were detected, to 65'534 which is the highest radiance which the sensor can captured in a differentiated manner; the filler value is 65'535 [44].

Both the monthly and yearly products from the Black Marble Suite are accessible online over NASA's Level 1 and Atmosphere Archive and Distribution System Distributed Active Archive Center (LAADS-DAAC) [43] [41].