prerequisite for high-quality outputs (Zhu and Woodcock, 2012; Ju and Roy, 2008).

4. Elevation and Slope Derivation

Elevation and slope data were derived from the Shuttle Radar Topography Mission (SRTM) and were used to represent terrain conditions that influence both natural and human land use. These features are critical in differentiating upland, mountainous zones from lowland agricultural or urban areas. Slope and elevation also supported volcano masking, helping to exclude high-altitude terrain that may cause misclassification. The application of topographic parameters in LULC studies has long been standard practice (Farr et al., 2007).

5. Volcano Terrain Masking

High-altitude volcanic regions present classification challenges due to their unique reflectance patterns and rugged topography. To reduce misclassification, a terrain-based mask was applied to exclude pixels above 2500 meters in elevation or with slope exceeding 30 degrees. This terrain masking ensures that classification and regression models focus on stable and interpretable surfaces. Similar strategies have been applied in volcanic landscape analyses to prevent erroneous LULC mapping (Wessels and Kargel, 2004).

6. Clipping to Study Area

Restrict imagery to Klungkung Regency boundaries for focused analysis. In the preprocessing phase, the LULC datasets were acquired and clipped to match the Area Of Interest (AOI). (Khachoo, Cutugno, Robustelli, & Pugliano, 2024)

7. Training Sample Collection

A comprehensive set of training samples for the four primary LULC classes was manually digitized using the highest quality basemaps in Google Earth Engine. To support visual interpretation during sample selection, official

NASA-recommended RGB band combinations were applied, allowing enhanced visualization of different land cover types. Additionally, spectral indices NDVI and NDBI, derived from Landsat 8's official bands, were used as further guidance to differentiate vegetated, built-up, and bare surfaces based on known spectral behaviors (Huete et al., 2002; Zha et al., 2003).

II.6 Landsat 8 Satellite

The Landsat program, jointly operated by NASA and the USGS, has been a cornerstone of Earth observation since its inception in 1972.

Figure II-4. Landsat 8 satellite.

Landsat satellites provide multispectral imagery that enables researchers to monitor land use and land cover (LULC) changes over time, making it a valuable tool for environmental management and land resource studies (Williams, 2001; Wulder et al., 2008).

Landsat 8, launched in 2013, is equipped with the Operational Land Imager (OLI) and the Thermal Infrared Sensor (TIRS), which deliver improved spectral and radiometric performance compared to its predecessors (Figure II-3 Landsat 8 Satellite) . The satellite features 11 spectral bands, including two thermal bands and a coastal aerosol band. Most bands offer a spatial resolution of 30 meters, with the panchromatic band achieving 15 meter resolution, making Landsat 8 ideal for detailed land cover analysis (Roy et al., 2014).

Landsat 8 data have been pivotal in tracking deforestation, urban expansion, and agricultural changes globally. Its enhanced capabilities have been effectively used to monitor land use changes across Bali, revealing the significant impacts of tourism-driven urban growth on natural ecosystems (Rimba et al., 2020).

II.7 Landsat 7 Satellite

The Landsat program, operated jointly by the National Aeronautics and Space Administration (NASA) and the United States Geological Survey (USGS), has long served as a critical source of Earth observation data, particularly for the analysis of land use and land cover (LULC) changes. Among its key missions, Landsat 7 launched in 1999 has made a substantial contribution to the global effort to monitor environmental change over time. The satellite is equipped with the Enhanced Thematic Mapper Plus (ETM+), which includes eight spectral bands covering visible, near-infrared, short-wave infrared, and thermal portions of the electromagnetic spectrum (Figure II-4). It offers a 30-meter spatial resolution for multispectral bands, a 60-meter resolution for thermal data, and a high-resolution 15-meter panchromatic band, which enhances its capacity to detect fine-scale surface features (Chander et al., 2009)

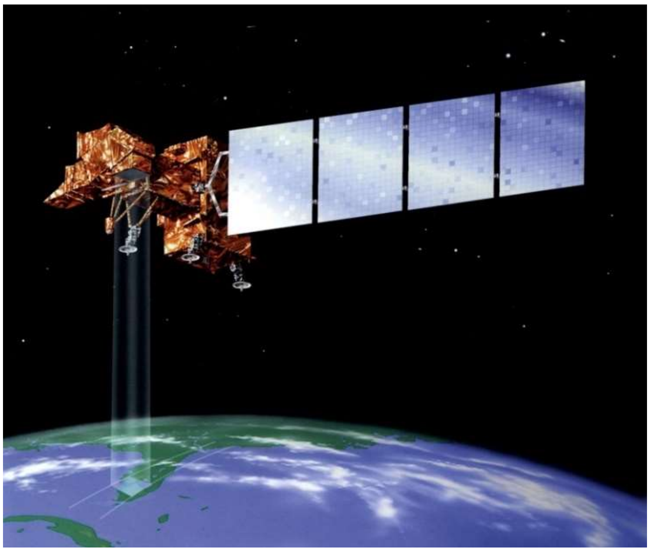

Figure II-5. Landsat 7 satellite.

Landsat 7's multispectral imaging capabilities have enabled consistent and systematic monitoring of various land surfaces, making it an invaluable tool for assessing changes in vegetation, urban expansion, agricultural activities, and water bodies. Its imagery has supported diverse applications ranging from biodiversity conservation to urban planning and agricultural resource management. One of the satellite's key advantages lies in its long-term data continuity, which supports time-series analyses crucial for understanding trends and transformations in land use. The satellite's global coverage and free access through USGS data portals have facilitated wide adoption across scientific and policy domains (Goward et al., 2001; Wulder et al., 2008, 2022).

In summary, Landsat 7 continues to offer vital contributions to the study of LULC changes, particularly in regions undergoing rapid transformation. Its integration with cloud-based platforms such as Google Earth Engine enhances the accessibility and efficiency of processing large geospatial datasets, making it a valuable asset in the global effort to monitor and manage environmental changes.

II.8 Sentinel 5-P

The monitoring of atmospheric composition is a cornerstone of modern environmental science, providing essential data for air quality management, climate studies, and public health policy (Sekiya et al., 2022). Within the global framework for Earth observation, the Copernicus Sentinel-5P (S5P) mission stands as a critical asset. Launched on October 13, 2017, by the European Space Agency (ESA), S5P is a dedicated atmospheric monitoring satellite designed to provide timely and high-resolution data on a global scale, bridging the observational gap between past and future atmospheric chemistry missions (Veefkind et al., 2022). Its primary objective is to perform detailed measurements of trace gases and aerosols that influence air quality and climate, thereby supporting operational services and scientific research worldwide (European Space Agency, 2015). An llustration of the Sentinel-5P satellite and its main instrument, TROPOMI, is shown in Figure II-5.

Figure II-6. Sentinel-5P satellite.

The sole instrument aboard the Sentinel-5P satellite is the Tropospheric Monitoring Instrument (TROPOMI), a state of the art passive imaging spectrometer. TROPOMI measures solar radiation reflected and scattered by the Earth's atmosphere and surface across the ultraviolet, visible, near infrared, and shortwave infrared spectral ranges (van Geffen et al., 2022). This broad spectral coverage allows for the retrieval of numerous key atmospheric constituents, including nitrogen dioxide (NO_2), ozone (O_3), sulfur dioxide (SO_2), and carbon monoxide (CO). A key advancement of TROPOMI is its combination of a wide 2,600 km swath, which enables daily global coverage, with a high native spatial resolution (as fine as 3.5 km x 5.5 km), representing a significant improvement over previous-generation instruments like the Ozone Monitoring Instrument (OMI) (Sekiya et al., 2022; Wang et al., 2020).

For air quality applications, the primary data product derived from TROPOMI is the tropospheric vertical column density (VCD) of NO_2. This product quantifies the total number of NO_2 molecules per unit area within the troposphere, the atmospheric layer where most anthropogenic emissions occur (Eskes et al., 2022). The retrieval process is based on the Differential Optical Absorption Spectroscopy (DOAS) method, a three-step algorithm that first determines the total slant column of NO_2, then separates it into its stratospheric and tropospheric components, and finally converts the tropospheric slant column into a vertical column using an Air Mass

Factor (AMF) (van Geffen et al., 2022). The resulting tropospheric NO_2 VCD is a direct indicator of near-surface pollution, making it an invaluable tool for environmental monitoring.

The scientific validity and reliability of the TROPOMI NO_2 data product are well-established through extensive validation campaigns and its widespread use in peer-reviewed research. Numerous studies have compared TROPOMI retrievals with ground-based reference measurements from global networks like MAX-DOAS and Pandora, confirming a high degree of consistency, with reported correlation coefficients often exceeding 0.90 (Verhoelst et al., 2021; Wang et al., 2025) While these studies note a systematic negative bias (underestimation) of approximately -20% to -30% in polluted regions, the data's precision and stability make it highly suitable for tracking pollution patterns and trends (Verhoelst et al., 2021). The instrument's superior spatial resolution and data quality have proven to be a major advantage for monitoring pollution from cities, traffic, and industrial activity with unprecedented detail (Ngo et al., 2023).

The proven capabilities of TROPOMI make it an indispensable tool for air quality research in Southeast Asia, a region characterized by rapid development and complex emission sources. The instrument has been successfully used to assess NO_2 emissions from major cities across the region and to develop high resolution pollution maps (Ngo et al., 2023). Of particular relevance to the Indonesian archipelago, TROPOMI has demonstrated its ability to detect plumes from volcanic eruptions, such as those from Anak Krakatau, by measuring gases like sulfur dioxide (Theys et al., 2019). This capability to monitor both anthropogenic and natural emission sources validates Sentinel-5P TROPOMI as a robust and reliable data source for mapping NO_2 concentrations and understanding air quality dynamics over Bali.

II.9 Machine Learning

Machine Learning (ML) is a subfield of artificial intelligence that has become a cornerstone of modern data analysis. It is broadly concerned with the design and development of algorithms that allow computers to learn complex patterns and make predictions from empirical data (Bishop, 2006). Unlike traditional algorithms