Land Subsidence and Groundwater Storage Assessment Using ICOPS, GRACE, and Susceptibility Mapping in Pekalongan, Indonesia

Wahyu Luqmanul Hakim, *Student Member, IEEE*, Muhammad Fulki Fadhillah, *Student Member, IEEE*, Kwang-Jae Lee, Seung-Jae Lee, Sung-Ho Chae, and Chang-Wook Lee, *Member, IEEE*

Abstract— Floods in Pekalongan, Indonesia, often occur due to river water overflowing during heavy monsoon rain. Simultaneously, the northern coastal area of Pekalongan, located adjacent to the Java Sea, has been affected by coastal floods due to sea level rise. The flood conditions in this area were exacerbated by land subsidence, leading to coastal inundation. Monitoring land subsidence in Pekalongan has become essential in predicting other possible land subsidence occurrence areas and mitigating the possible hazards caused by land subsidence. The analysis of land subsidence has been much easier since the introduction of radar satellites. In this study, 124 synthetic aperture radar (SAR) datasets from the Sentinel-1 radar satellite between 2017 and 2022 in descending tracks were used. The data were processed through a time-series interferometry SAR (InSAR) method based on the improved combined scatterers interferometry with optimized point scatterers (ICOPS) algorithm to provide accurate measurements over large areas by improving the selection of measurement points (MPs) from persistent scatterer (PS) and distributed scatterer (DS) points using a deep learning algorithm based on a convolutional neural network (CNN), and the resulting optimized MPs were then spatially clustered using optimized hot spot analysis (OHSA) to estimate significant points statistically and define them as hot spot points. The results of time-series deformation in Pekalongan were compared with the GPS station measurements. From the comparison, a good correlation in terms of deformation patterns between time-series InSAR and GPS measurements was observed. Our study revealed that land subsidence in Pekalongan has occurred mostly in settlement areas under the young alluvium soil, which cannot support many buildings' maximum compression. Another cause of land subsidence in Pekalongan is excessive groundwater extraction in settlement areas. Thus, compaction in the aquifer areas may occur as a result of the reduced effective stress of the pore pressure. Further analysis of this study would involve monitoring groundwater activity using data from the Gravity Recovery and Climate Experiment (GRACE) satellite and comparing them with weather station data. The analysis of the two datasets aims to understand the relationships between groundwater storage data and the monthly precipitation in Pekalongan. Finally, the potential outcomes of land subsidence in Pekalongan will be assessed using the geographic information system (GIS) method based on susceptibility mapping.

Index Terms— Convolutional neural network (CNN), groundwater, improved combined scatterers interferometry with optimized point scatterers (ICOPS), interferometry synthetic aperture radar (InSAR), land subsidence, susceptibility.

I. INTRODUCTION

LAND subsidence may occur due to natural and anthropogenic processes [1], such as Holocene sediment compaction [2], mining activities [3], groundwater and gas extraction [4], and building and construction loads [5]. Groundwater withdrawal may cause a reduction in pore pressure and lead to the compaction of unconsolidated sediment [6]. The sediment compaction process can be worsened by artificial building and construction loads [7], [8]. Land subsidence can expand river flood-prone areas [9] and tidal flood-prone areas due to sea-level rise [10]. Land subsidence can also damage urban infrastructures, such as buildings, roads, and railway tracks [11], lead to coastal inundation in coastal areas, and permanently submerge a city in seawater [4], [12], [13].

Several coastal cities in Indonesia have suffered from land subsidence, including Pekalongan [4]. Excessive groundwater extraction for agricultural use and the natural compaction of quaternary alluvium soil from Holocene sediment are suspected to be the leading causes of land subsidence in Pekalongan [2], [4]. Another factor contributing to land subsidence in this area is related to anthropogenic processes from the construction of high-rise buildings, which induces the compaction of clay, because young alluvium soil cannot bear the loads of such buildings [8]. As a result, the flood conditions in every area in this region are exacerbated by land subsidence, leading to coastal inundation and coastal areas that are below the relative sea level [4], [14].

Monitoring land subsidence has been widely carried out using radar satellite imagery, and the process has been conducted using differential interferometry synthetic aperture radar (DInSAR) processing [15], [16], [17], [18]. The accuracy

of DInSAR processing can be improved to the submillimeter level by using the time-series interferometry synthetic aperture radar (InSAR) method [19], [20]. The time-series InSAR algorithm has been applied to monitor the deformation of the Earth's surface induced by several causes, such as slow-moving landslides [21], [22], volcanic activities [23], [24], crustal deformation [25], human-made deformation such as mining activities [3], [26], excessive groundwater and gas extraction [4], and building and construction loads [4], [27].

A study monitoring land subsidence in Pekalongan was conducted using Advanced Land Observing Satellite (ALOS) Phased Array type-L Synthetic Aperture Radar (PALSAR) satellite data acquired from 2007 to 2009 [4]. A time-series method based on a small baseline subset (SBAS) was implemented to produce a land subsidence map [28]. Land subsidence of up to 10.5 cm/year was found in the northern area of Pekalongan, which is adjacent to the Java Sea. Apart from the coastal area, land subsidence was also found in the western area of Pekalongan, and the land subsidence was approximately 7.7 cm/year [4].

However, in recent studies, the land subsidence in Pekalongan has not been validated against other geodetic measurements, such as GPS surveys. In addition, monitoring land subsidence in Pekalongan needs to be updated due to the importance of monitoring the area affected by land subsidence for preliminary studies for the local government to make development plans. Therefore, this study aims to monitor land subsidence from 2017 to 2022 in Pekalongan using Sentinel-1 data. The data were processed through a time-series InSAR method based on the improved combined scatterers interferometry with optimized point scatterers (ICOPS) algorithm to maximize the density of the measurement points (MPs) by exploiting both persistent scatterer (PS) and distributed scatterer (DS) points. The convolutional neural network (CNN) will be used to analyze and predict the optimum time-series data over time [29]. The significant MPs were identified using optimized hot spot analysis (OHSA) based on spatial clustering by a statistical approach, and the nonsignificant MPs were removed [22]. The utilization of the CNN algorithm as the predicting algorithm in this study was because the CNN algorithm was better in model performance, MP density, and time consumption [29]. However, the ability of the CNN algorithm to optimize the MPs is not yet validated with the ground-truth data due to the GPS measurement only monitoring a stable area in the previous study. Therefore, further study needs to be conducted to validate the integration between ICOPS and CNN algorithm with the ground-truth data. Thus, the novelty of this study would be to identify the performance of the ICOPS algorithm based on the CNN and OHSA method to be compared with the ground-truth data from the GPS station. Several new requirements are presented in this study, especially in the data preparation step for MP optimization to maximize the predictive performance in predicting the optimum data in each MP. Analyzing this study further would contribute to the understanding of the relationship between rainfall and groundwater storage (GWS) anomalies in Pekalongan. Finally, from this study, land subsidence susceptibility maps can be generated to find other possible locations that are prone to land subsidence in the future by analyzing the correlation between the current land subsidence location and its geographic information system (GIS)-related factors using a deep learning algorithm based on the CNN algorithm.

II. STUDY AREA

Pekalongan is a lowland plain on the northern coast of Java Island, between 6°50'42" and 6°55'44" south latitude and 109°37'55" and 109°42'19" east longitude. Pekalongan has a population of approximately 300 000 people, with an annual growth rate of 0.89%. In a total area of 45.25 km^2, Pekalongan has a population density of approximately 6.78 people/km^2 [30]. The study area is shown in Fig. 1(a). Geologically, Pekalongan lies on 80.43% alluvium landforms, 12.42% alluvial fan deposits, and 7.4% damar formations. The distribution of geological formations in Pekalongan city is shown in Fig. 1(b). The alluvium landform that stretches along the coastline of Java Province consists of beach deposits formed by clay and sand from the Quaternary age of the Holocene period [12]. Alluvial fan deposits are triangle-shaped deposits resulting from the deposition of sediment transported by a stream and mainly consist of volcanic debris [31], [32]. The damar formation consists of tuffaceous sandstone, conglomerate, and volcanic breccia [33]. The distribution of land use in Pekalongan [see Fig. 1(c)] is dominated by rice fields (56.04%), followed by settlement areas (27.39%). Pekalongan also consists of 3.55% dryland agriculture and 3.52% mixed dryland and bush agriculture. The other area in Pekalongan is a water body, which consists of 8.70% fishponds and 0.80% rivers.

The climate of Pekalongan was considered to be tropical with a rainy season that occurred between October and March and a dry season from April to September. The average temperature ranges from 27 °C to 29 °C year-round, peaking during the dry season. The city has high humidity throughout the year, and annual rainfall ranges from 2000 to 3000 mm, with the highest precipitation occurring between January and February [34]. Although rainfall in Pekalongan city is relatively high, Pekalongan city still experiences water scarcity due to a lack of access to water sources, especially due to the scarcity of surface water sources in the city. This incident can result in the use of groundwater to be used as the main water source, which, in turn, can cause other problems such as land subsidence and seawater intrusion. Land subsidence in Pekalongan was reported to have a significant impact with an increase in tidal flooding of around 50–100 cm, which caused an increase in inundation area on the coast of Pekalongan and its surroundings in the last decades [35].

III. DATA AND METHODOLOGY

A. InSAR Data

In this study, synthetic aperture radar (SAR) data from Sentinel-1 SAR C-band satellites that emit a wavelength of 5.5 cm were used. The satellite data provided by the European Space Agency (ESA) cover Pekalongan city, and the data coverage is shown in Fig. 1(a). The two different frames (613 and 618) of Sentinel-1A SAR data used in this study
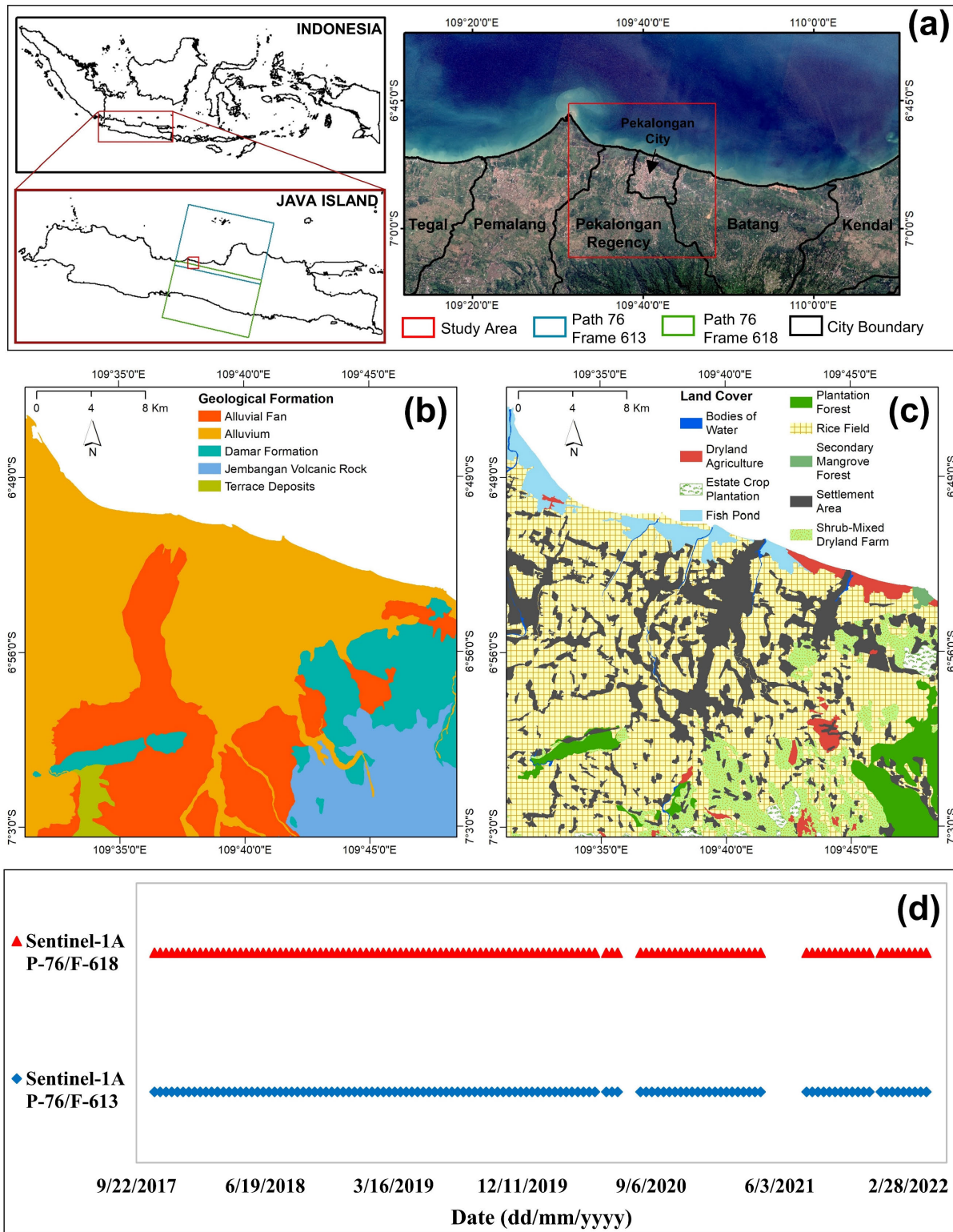
Fig. 1. (a) Pekalongan city location in Indonesia and Sentinel-1 data coverage in descending track path number 76 (blue square frame 613 and green square frame 618). (b) Geological formation of Pekalongan city and its surrounding area (modified from a geological map produced by the Ministry of Energy and Mineral Resources of the Republic of Indonesia). (c) Land use map of Pekalongan city and the surrounding area (modified from a land-use map of Pekalongan produced by the Ministry of Environment and Forestry of the Republic of Indonesia). (d) List of Sentinel-1 datasets used in this study.

consist of a total of 248 images in descending flight direction from October 2017 to April 2022 with path number 76 [see Fig. 1(d)]. A total of 124 satellite images of each frame were combined to cover the whole study area, as shown in Fig. 1(a). The SAR data have a spatial resolution of 5 × 20 m and an incidence angle of 34° in the interferometric wide swath mode,

InSAR Preprocessing

Sentinel-1 SAR Datasets → Orbit Correction → Reference Date Selection

Radiometric Calibration → Thermal Noise Correction → Coregistration Process

Digital Elevation Model Data (Copernicus DEM)

Resampled SLC Data → RSLC Pair Formation → Interferogram Generation

Single-Master Interferograms | Multi-Master Interferograms

PS Processing

Phase Noise Estimation | Final PS Points

PS Candidates Selection | SCLA error Estimation

PS Weeding | PS Correction

DS Processing

Spatial Coherence Estimation | Final DS Points

Temporal Coherence Estimation | DS Candidates

Amplitude Images → SHP Identification using GLR Test

Combined PS+DS Processing

PS+DS Network Triangulation

MCF Unwrapping

Topographic and Atmospheric Correction

Combined PS+DS points and Time-series data

Measurement Points Optimization using CNN and OHSA

Data Preparation

Coefficient Correlation Estimation

Data Labeling

Stratified Random Sampling

Testing dataset (30%) | Training dataset (70%)

ICOPS Result

Training Model using CNN algorithm

Input layer | Convolution layer+ReLU | Pooling layer | Pooling Layer | Fully Connected Layer | ReLU Layer | Fully Connected Layer | Regression Layer | Output

Feature Extraction | **Prediction**

Optimized Hot Spot Analysis

P-Values <0.01

Z-Scores <-3 & >3

Clustering

Global Moran's Index

$$I = \frac{n}{\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}} \frac{\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}(x_i - \bar{x})(x_j - \bar{x})}{\sum_{j=1}^{n}(x_i - \bar{x})^2}$$

Getis-Ord Gi* Statistics

$$G_i^*(d) = \frac{\sum x_j + x_i - n_{ij} \times \bar{x}^*}{\sqrt{S^*\left[\left[(n \times n_{ij}) - n_{ij}^2\right] - (n-1)\right]}}$$
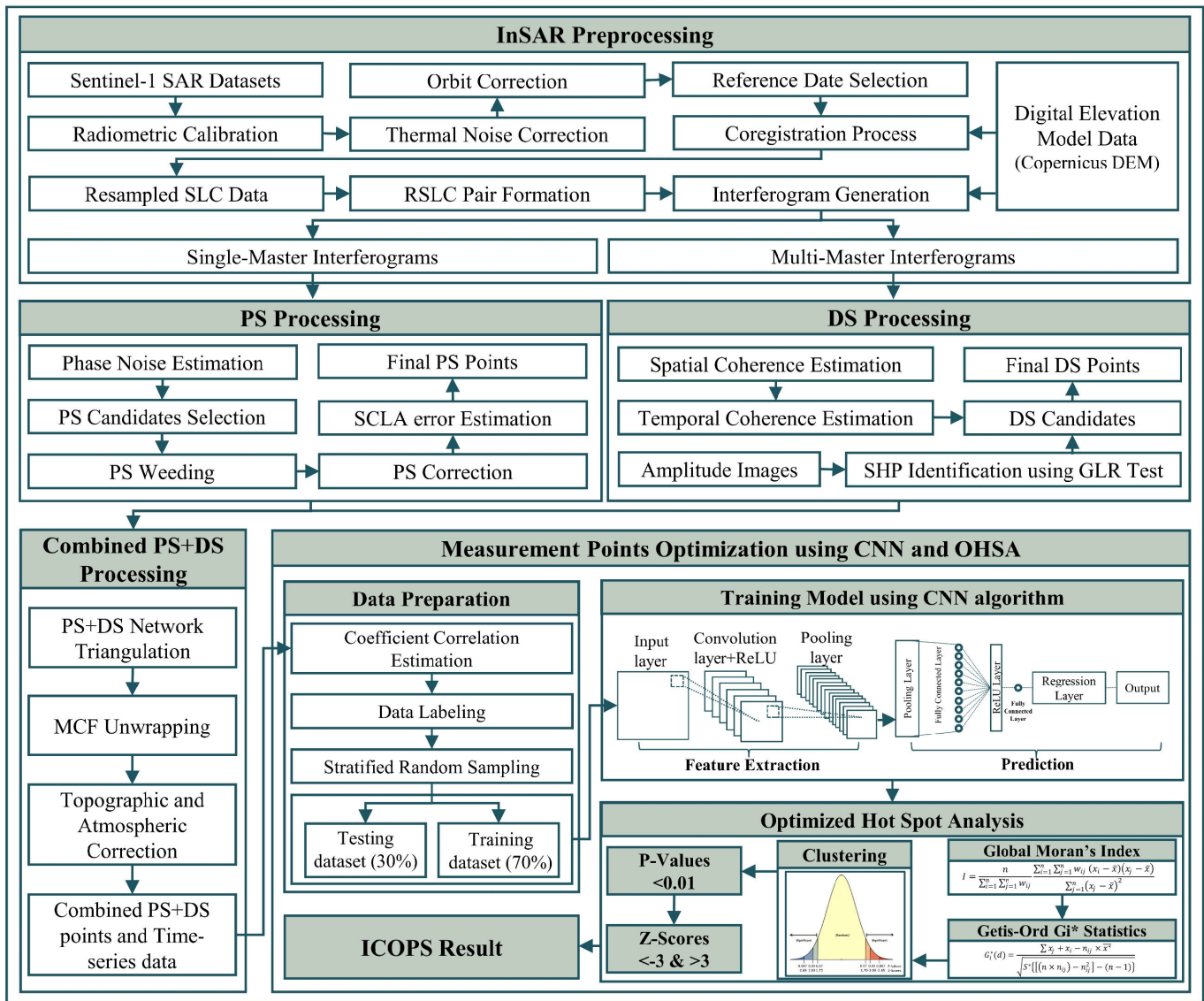
Fig. 2. Flowchart of ICOPS processing methods (this figure was modified from [29] and [36]).

and VV polarization was used in this study. Thus, SAR data were thermal noise corrected, radiometrically calibrated, and orbit corrected before initiating the coregistration process.

B. Time-Series Interferometry SAR Processing

The workflow of the time-series InSAR process in this study is illustrated in Fig. 2. The process of the time-series InSAR processing is initiated by the InSAR preprocessing step, which includes Sentinel-1 data preparation, coregistration data, and interferogram generation for PS and DS processing. The interferogram generated from the preprocessing step will be used as the input of PS and DS processing. After PS and DS processing, the PS and DS MPs were combined and prepared for the MP optimization process. After that, the combined PS and DS MPs were optimized using the CNN algorithm. The result from the optimization from the CNN algorithm then will be further optimized spatially using the OHSA method. Finally, the result from the optimization process from CNN and OHSA method will become the result of ICOPS time-series InSAR processing. Further explanation regarding each subprocess of ICOPS processing will be explained as follows.

1) InSAR Preprocessing: The Sentinel-1 SAR data from 2017 to 2022 were initiated by thermal noise correction and radiometric calibration. After that, all data were orbit corrected by precise orbit ephemerides data, and then, all data were coregistered to a single reference image to increase the image accuracy to subpixel accuracy. The coregistered result was computed to generate the topographic phase interferogram, and the topographic phase was subtracted using the digital elevation model (DEM) from the Copernicus GLO-30 to generate DInSAR. The process of generating DInSAR is divided into two parts, namely, the single-master interferogram and the multimaster interferogram. Thus, the results from single-master and multimaster interferograms are used to be the input for PS and DS processing, respectively. The single-master interferogram processing selects a common master image that is used in the coregistration process.

In this study, a reference date of November 13, 2019, was selected to generate 123 single-master interferograms for PS processing. Meanwhile, the multimaster interferograms select the reference data based on the perpendicular baseline and temporal baseline thresholds of about 200 m and 36 days [21], [37], [38].

2) PS Processing: The results of the topographic phase removal process, which includes a pair of master interferograms, were iteratively analyzed to assess the phase noise value of each pixel candidate before selecting the PS candidate. PS candidates were then selected based on the noise characteristics of each pixel in every interferogram. After that, the selected PS candidates were filtered out to remove the pixels with too much noise. Furthermore, the wrapped phase of the selected PSs that were removed was corrected for DEM error before the phase unwrapping process. After the phase unwrapping process, to estimate the spatially uncorrelated look angle error for the selected PSs from the unwrapped phase, the unwrapped data were high-pass filtered in time and low-pass filtered in space. Then, the PS outputs were generated [39], [40], [41].

3) DS Processing: We used the GLR test to identify SHP from the amplitude values of SAR data with a window size of 15×15 pixels for DS processing. We selected DS points with SHP values >20 for further processing. Then, we computed the adaptive spatial coherence of the resulting interferogram and used DS points with a spatial coherence > 0.2 to increase the density of the DS candidates. The final step in the DS process was to evaluate the temporal coherence and phase stability of the DS candidates. We used a temporal coherence > 0.75 to indicate stability in the DS points. Then, the DS outputs were generated based on the specific parameters of SHP, spatial coherence, and temporal coherence [37].

4) Combination of PS and DS Processing: After the PS and DS MPs were generated from the previous step, we combined PS and DS MPs by adding PS and DS into combined scatterer interferometry (CSI) data. The wrapped CSI data then needs to be unwrapped with a minimum-cost method to retrieve the accurate displacement values. After that, we obtained a deformation time-series map using the singular value decomposition (SVD) approach [37].

5) Preparation Data for MPs' Optimization: The PS and DS point combination process was refined by predicting the optimal MPs using a CNN algorithm, followed by clustering the deformation area based on the confidence level of data from the prediction results using the OHSA method. The first step in optimization is selecting sample MPs for training and testing data based on the coefficient correlation of the time-series deformation of each MP over time. We defined good data candidates by collecting MPs with a coefficient correlation > 0.9. In addition to the coefficient correlation, we also identified the outliers between the time-series data that are more than three scaled median absolute deviations (MADs) and time-series data that are more than three standard deviations (StDs) from the mean value. Thus, the previous sample with a coefficient correlation higher than 0.9 and the data that do not contain any outliers are assigned as "1" data, and the outlier-affected data are assigned as "0" data.

A stratified random sampling method was used to randomly divide the "1" and "0" data into a 70:30 ratio of training and testing data, respectively. Note that optimization in this process aims to find untrusted data that are not labeled as "1" or "0" data during the sample preparation process. Those data will be analyzed by the CNN algorithm to predict the optimality of the pattern from the time-series data. The prediction results with low prediction values will be used for considering the data to be removed from the MPs [29].

6) MPs' Optimization Using CNN Algorithm: The data preparation phase begins the training process for CNN and includes creating training and testing sets. The network architecture consists of 20 convolutional layers with a rectified linear unit (ReLU) activation function to capture the nonlinear characteristics of the input data. The output of the convolutional layers is downsampled using 2-D average pooling to reduce the input feature dimensions. The collected data are then multiplied and weighted to obtain ten fully connected layers. Another ReLU layer is used before the data are connected to a fully connected layer to reduce the computation complexity. The final output of the CNN should be continuous, and a regression layer must be applied before generating the output. Parameters such as the learning rate, the maximum number of epochs, and the batch size were optimized using the gray wolf optimizer (GWO) algorithm.

7) MPs' Optimization Using OHSA Method: Finally, the optimized MPs from the CNN optimization results were spatially optimized to improve the spatial reliability of the MPs using the statistical method based on OHSA. The OHSA method uses Getis-Ord Gi* statistics to measure the spatial clustering of the high and low values of the MP. The optimized term in the hot spot analysis refers to the measurement of spatial autocorrelation using the global Moran's index to provide the overall assessment of the spatial pattern of the MPs to improve spatial reliability. The Z-score and the p-value calculated by the OHSA method will indicate the presence of hotspots between clusters and random patterns in the results. Finally, we selected the data with a p-value < 0.01, which corresponds to the 99% significance level from the OHSA processing Thus, the MPs with p-value > 0.01 were considered to be insignificant MPs and need to be removed after the clustering process [29], [37].

C. GWS Processing

The gridded $1° \times 1°$ monthly Gravity Recovery and Climate Experiment (GRACE) and GRACE Follow-On (GRACE-FO) mission satellite data (RL-06, level-3) were acquired from three institutions: the Center for Space Research (CSR, The University of Texas), the Jet Propulsion Laboratory (JPL, Pasadena, CA, USA), and GeoForschungsZentrum Potsdam (GFZ, Potsdam, Germany). The level-3 data products from the GRACE/GRACE-FO missions were subjected to several preprocessing steps to ensure their quality and usefulness. These steps include the application of destriping filters to remove any instrument artifacts, Gaussian smoothing to remove any high-frequency noise, and glacier isostatic adjustment (GIA) to account for the Earth's response to the melting of glaciers and ice sheets. These preprocessing steps help to enhance the
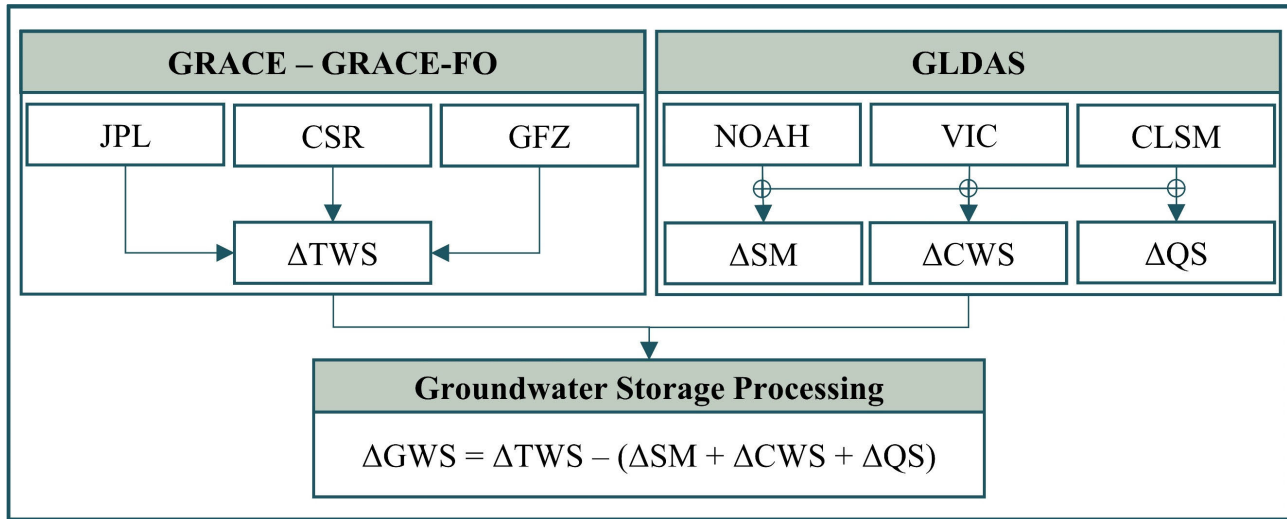
Fig. 3. Method for GWS processing from GRACE/GRACE-FO and GLDAS data.

accuracy and precision of the data products, making them more suitable for use in various scientific and operational applications. Pixel averaging of the terrestrial water storage (TWS) data from the three different institutions could improve the accuracy of the GRACE satellite measurements [1], [42], [43], [44], [45].

Further processing of the TWS data needs to be conducted since the change in TWS is the sum of several components, including GWS, soil moisture (SM), canopy water storage (CWS), quickflow surface runoff (QS), and snow water equivalent (SWE). These parameters can be assessed from the Global Land Data Assimilation System (GLDAS). The monthly average GLDAS data, with a spatial resolution of $1° \times 1°$ (V2.1, level 4), which is equivalent to the GRACE data, were used in this study. Three different land surface models were used in this study: the catchment land surface model (CLSM), the variable infiltration capacity (VIC) model, and the Noah model. Three parameters (SM, CWS, and QS) were utilized to estimate the GWS of the study area. The effect from the SWE parameter was not considered in the analysis, as the study area is located near the equator, where snow is absent; thus, the SWE value was zero [42], [43], [45].

Finally, the GWS can be estimated by the GWS processing workflow, as shown in Fig. 3, and using the water balance (1). Since the TWS data from the GRACE satellite were anomalous data, we needed to convert the SM, CWS, and QS parameters to anomalous data by subtracting the current data at a specific time from the average of the total data. This step must be conducted to minimize any biases during the final processing to generate GWS data [46]

$$\Delta GWS = \Delta TWS - (\Delta SM + \Delta CWS + \Delta QS). \quad (1)$$

D. GIS-Related Factors

The selection of the land subsidence-related factors can be assessed by studying previous work on land subsidence susceptibility mapping [29], [47], [48], [49]. Based on previous studies and data availability, we considered three different categories of land subsidence-related factors to generate a land subsidence susceptibility map. These categories are topo-hydrological, land cover, and geological factors. Thus, 12 land subsidence-related factors were used in this study, and they were categorized into topo-hydrological factors (altitude, slope, aspect, total curvature, topographic wetness index (TWI), normalized difference water index (NDWI), precipitation, distance to river, and drainage density), land cover factors (normalized difference vegetation index (NDVI) and land use), and geological factors (lithology), as shown in Fig. 4(a)–(l). These factors were evaluated and investigated using the information gain method as the attribute evaluator to identify the relative importance of the land subsidence-related factors [50]. The information gain result will be discussed further in Section III-B. Thus, all selected land subsidence-related factors in Fig. 4 are effective for generating land subsidence susceptibility maps.

The altitude, slope, aspect, total curvature, and TWI factors [see Fig. 4(a)–(e)] were generated by extracting the DEM from the Copernicus DEM. Both NDVI and NDWI maps [see Fig. 4(f) and (g)] were generated by calculating the average of Landsat 8 imagery from 2017 to 2022 using Google Earth Engine over the Pekalongan area. The precipitation map [see Fig. 4(h)] was created by calculating the annual average precipitation using monthly data obtained from the Central Bureau of Statistics in Pekalongan. These data represent the monthly accumulation of precipitation across subdistricts in Pekalongan. To interpolate the precipitation data, the central area of the subdistrict in Pekalongan was selected as the reference for storing the precipitation data. The inverse distance weighting (IDW) method was employed as the interpolation algorithm to generate the final precipitation map because this method is known to be suitable for interpolating precipitation data [47], [48]. The land use map [see Fig. 4(i)] and the lithology map [see Fig. 4(j)] were acquired from the Ministry of Environment and Forestry of Indonesia and the Ministry of Energy and Mineral Resources of Indonesia, respectively. Finally, the distance to the river map [see Fig. 4(k)] and the drainage density map [see Fig. 4(l)] were generated using the Euclidean distance tool and the kernel density tool,
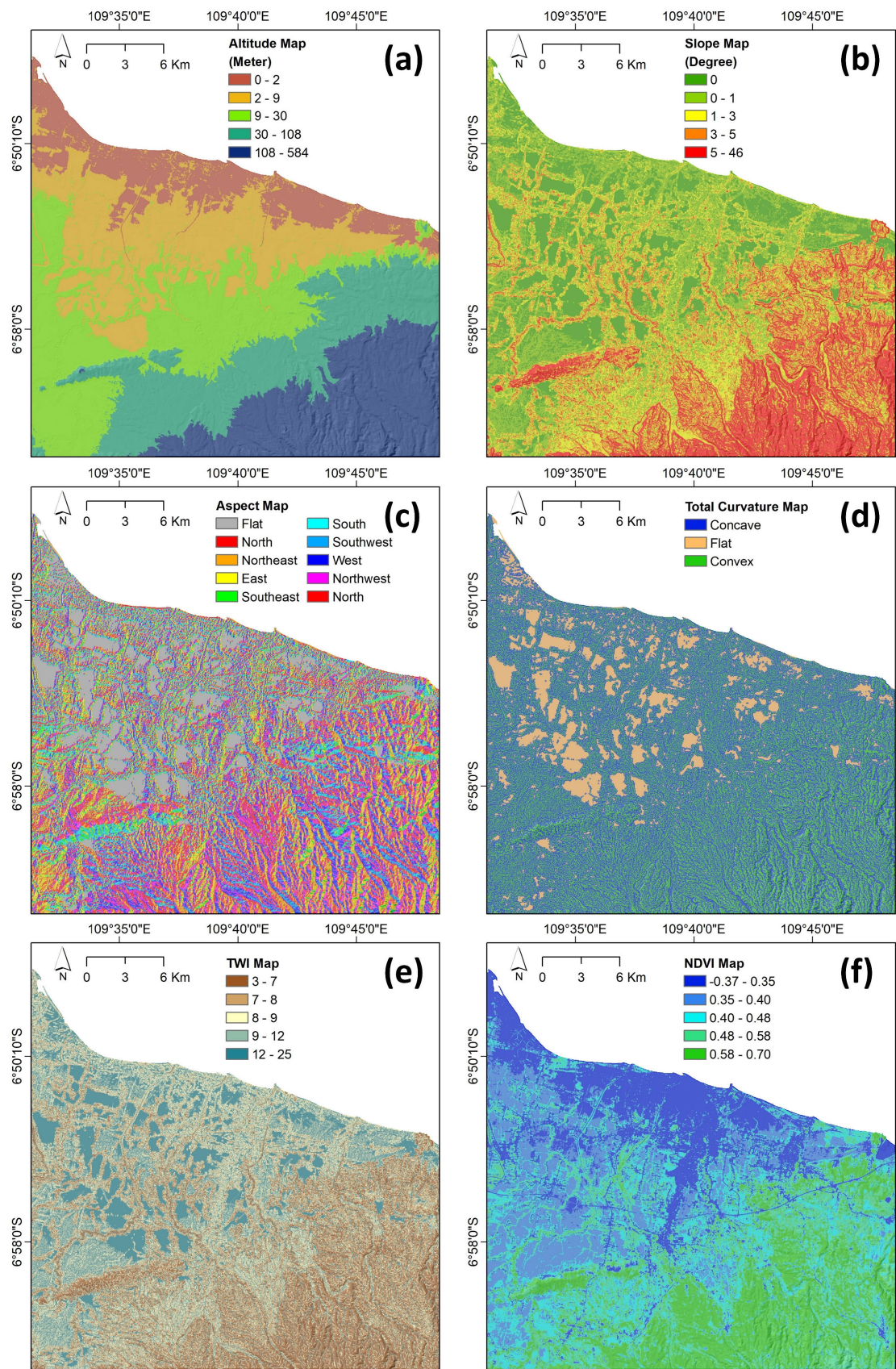
Fig. 4. Land subsidence susceptibility factors: (a) altitude, (b) slope, (c) aspect, (d) total curvature, (e) TWI, (f) NDVI, (g) NDWI, (h) precipitation, (i) land use, (j) lithology, (k) distance to the river, and (l) drainage density.
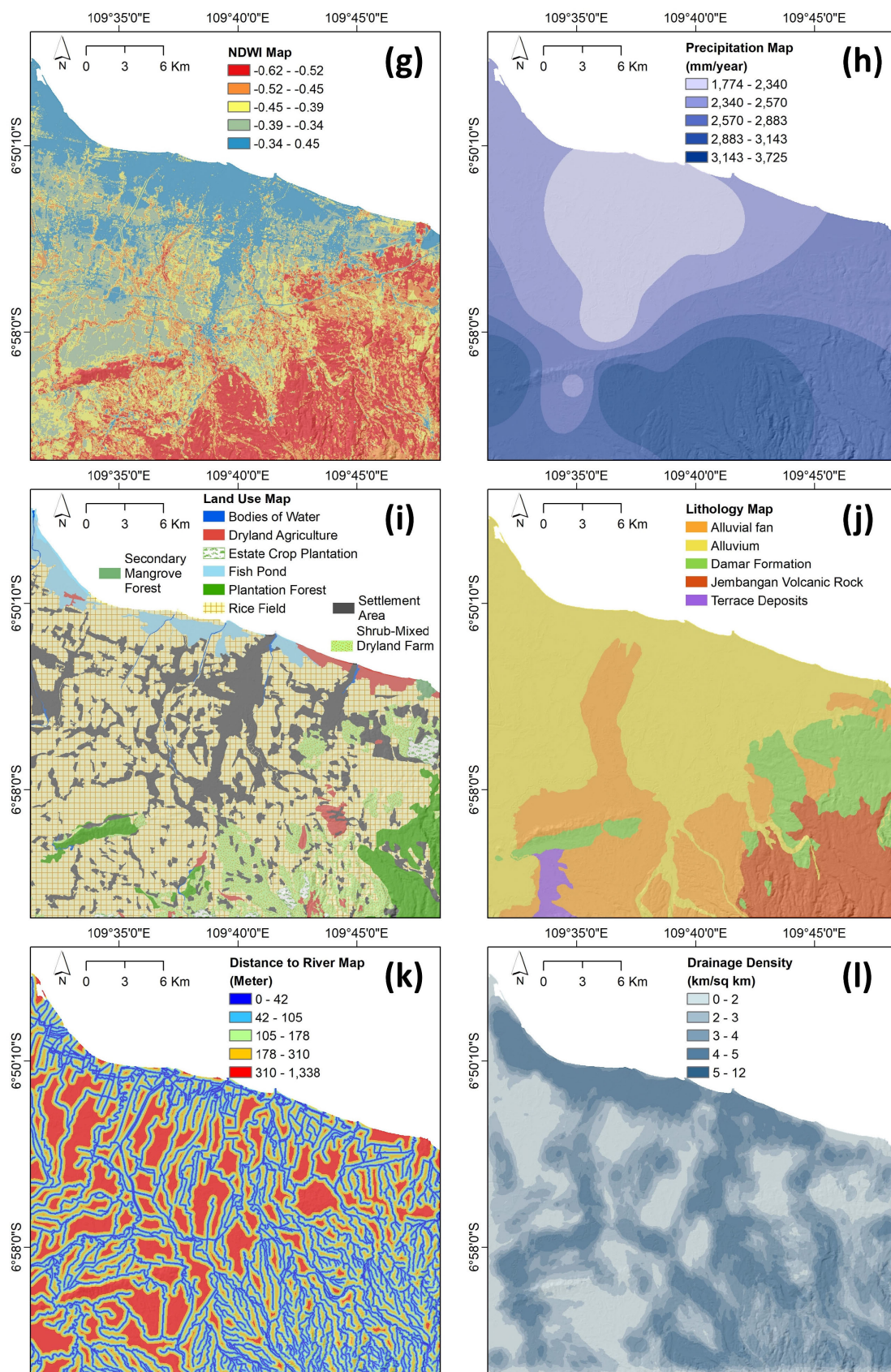
Fig. 4. *(Continued.)* Land subsidence susceptibility factors: (a) altitude, (b) slope, (c) aspect, (d) total curvature, (e) TWI, (f) NDVI, (g) NDWI, (h) precipitation, (i) land use, (j) lithology, (k) distance to the river, and (l) drainage density.

respectively [43]. Information on the river location was acquired from the Geospatial Information Agency of Indonesia.

The importance of GIS-related factors used in this study has a direct and indirect influence on land subsidence. The topographic factors that include altitude, slope, aspect, total curvature, and TWI were indirectly related to land subsidence by the direct influence of topographic factors with the anthropogenic factors that related to urbanization [51]. The area in Pekalongan with a lower elevation mostly has a gentler slope and flatter aspect. This area is mostly occupied by the higher population as shown by the settlement area in the land use map [see Fig. 4(i)]. This area has a higher urban development than the higher altitude and steeper slope, which is located in a mountainous area. The higher urban activity in an area that has a scarcity of surface water could lead to the overexploitation of groundwater that will cause land subsidence occurrence in that area [29], [47], [51]. Total curvature is related to the combination of profile and plan curvature that is indirectly associated with land subsidence by affecting the convergent and divergent flow of water across the surface. Meanwhile, the TWI factor is related to the accumulation of water over an area, and the highest TWI factor was usually indicated by the land subsidence occurrence due to the surface runoff during the dry season with lower soil moisture [47], [52]. The NDVI and NDWI maps were associated with the presence of vegetation cover and the water content, respectively. The lower NDVI values with higher water content from the high value of NDWI could be associated with the presence of urban areas. Both factors could indirectly affect the land subsidence occurrence as the lower vegetation cover could reduce the soil stability, and the higher water content shows that those areas were more prone to land subsidence due to the abundance of water bodies that will be more occupied by the populations that lead to the increase in urbanization that could affect the higher possibility of land subsidence occurrence [29], [53].

The precipitation could affect the land subsidence conditions over an area because the precipitation could contribute to groundwater recharge. In the area with poor drainage, the precipitation could not be infiltrated into the soil due to the higher surface runoff, and this condition could lead to land subsidence occurrences due to the lower infiltration from the precipitation [29], [54]. The land use map depicts the areas most exploited and occupied by residents, and the lithology map shows the geological conditions of an area, which have an appropriate relationship with land use. The utilization of groundwater in Pekalongan due to the scarcity of surface water could be shown in settlement areas. Another potential issue due to the scarcity of surface water in Pekalongan may also indicate that groundwater is also used for agricultural purposes. Meanwhile, the lithological features of Pekalongan that consists of an alluvium landform could affect the land subsidence due to unconsolidated sediment that could be compacted due to the gaps in soil pores that exist due to the groundwater extraction. The distance to the river map and drainage density describes the groundwater condition and could indirectly affect the land subsidence [4], [29], [35], [47], [55].

E. Susceptibility Methods

Land subsidence susceptibility mapping is the process of correlating a current incident with its related factors to predict the future possible areas that will be affected by land subsidence considering these factors [29], [56]. The whole workflow of the land subsidence susceptibility mapping is shown in Fig. 5.

1) Data Preparation: The current incidence of land subsidence in Pekalongan generated by the ICOPS measurement and validated by the GPS measurement was used in this study as the land subsidence inventory map that represents the current land subsidence condition in Pekalongan. The correlation between the inventory map and the GIS-related factors in this study was calculated using frequency ratio (FR) analysis. The FR method in this study is defined as the proportion of area affected by land subsidence in the total area of interest [57]. The result from the FR analysis was used to measure the density of the land subsidence occurrence in each class of every land subsidence-related factor.

2) Factor Analysis: The analysis of each GIS-related factor with the land subsidence location was further analyzed through factor analysis before the model was trained using the CNN algorithm. The methods to analyze the land subsidence-related factors are FR, information gain, Pearson correlation, and multicollinearity analysis. The FR method quantifies the density of the land subsidence occurrence in each class for each land subsidence-related factor. The ratio higher than 1 in the FR value describes the correspondence class in each factor that was more correlated with the land subsidence occurrences [47], [58], [59]. The information gain method evaluates the contribution of each land subsidence-related factor to the prediction process. The information gain method will rank the most influential factors for land subsidence with respect to the class of each factor [50]. The Pearson correlation analysis is a statistical method that measures the linear relationship between two variables. In this case, the Pearson correlation will quantify the linear relationship between each land subsidence-related factor. A correlation coefficient value close to 1 or -1 indicates a strong relationship, while close to 0 indicates a weak relationship or no linear relationship. Strong relationships between each variable can cause distortions during the prediction process, which generates susceptibility maps [59], [60]. Another statistical method to measure the multicollinearity between each land subsidence-related factor that could contribute to the distortion of the prediction result was the variance inflation factor (VIF) and tolerance measurement. The VIF was calculated by dividing the variance of the estimated regression coefficient that is expected to increase due to multicollinearity. Meanwhile, the tolerance value was reciprocal to the VIF value, the higher the VIF value could lower the tolerance value. The tolerance value could be acquired by subtracting 1 from the VIF value. A higher VIF value than 10 with a lower tolerance than 0.1 could indicate that each variable was dependent on each other and lead to an increase in the standard error in the predictive model [36], [59], [60], [61].

3) Training and Testing Data Preparation: Before the land subsidence susceptibility model was generated, the land
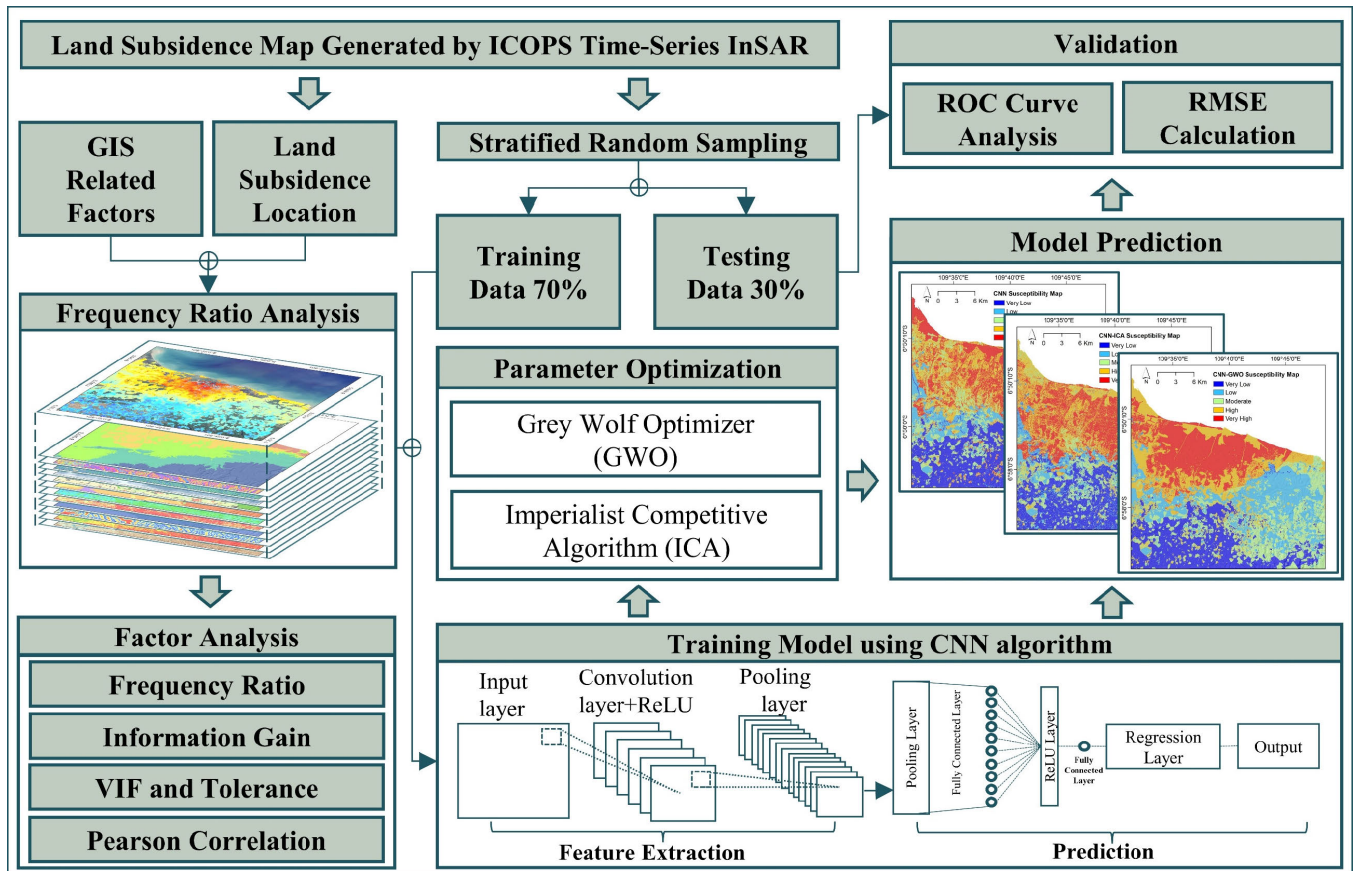
Fig. 5.　Workflow for generating land subsidence susceptibility maps from the result of ICOPS time-series InSAR.

subsidence inventory map was randomly divided using stratified random sampling between the land subsidence and nonland subsidence data. In this study, the nonland subsidence data are defined as zero data in the time-series processing result that contains stable measurements. We divided 70% of the data for training and 30% for testing. The training data and the result from the FR analysis were used to train the model. In this study, we used CNN algorithms to generate the susceptibility map. We also used optimization algorithms based on GWO and the imperialist competitive algorithm (ICA) to compare with the standalone CNN algorithm. Thus, the generated results from the CNN, CNN-GWO, and CNN-ICA algorithms were compared by assessing their accuracy values based on the receiver operating characteristic (ROC) curve and root mean square error (RMSE) calculation.

4) Model Description: The model used in this study is the deep learning algorithm based on the CNN algorithm with the optimization algorithms used were GWO and ICA. Both optimization algorithms will optimize the parameters from the CNN algorithm to generate land subsidence susceptibility maps.

1) The CNN algorithm was a deep learning algorithm based on a neural network designed to imitate the function of the cerebral cortex in the human brain [62], [63]. The main difference between CNN and conventional neural networks is that the CNN algorithm did not utilize the weighting method to learn and instead applied multiple layers [64]. Those multiple layers consisted of convolutional, pooling, and fully connected layers

[65]. The first layer, which convolutional layer, had a role in extracting the input with filters to generate the feature maps and wrap around the information to be fed to the other layer [66]. The filters used in the convolutional layer were used as the activation function that has a role in processing the nonlinear results from the convolutional process [67]. The ReLU activation layer was often used due to its popularity to manage the activation of all neurons in the process [68], [69]. The second layer, which is the pooling layer, was to simplify the convolutional layer's output by reducing the size of feature maps [66]. The last layer was the fully connected layer, which connected each pooling layer to enhance the output of nonlinear mapping [63].

2) The GWO algorithm was a metaheuristic algorithm based on swarm intelligence (SI) that tends to imitate the gray wolves' social structures and their hunting behavior [70], [71]. Gray wolves, as apex predators, roamed in packs and had a strict social structure with a leader and positional hierarchy that defined their role in the groups [70], [72], [73]. The first role in leading the packs came to the alpha wolves. The alpha has the highest authority among the packs, and the orders of the alpha wolves are absolute for the group to follow [74]. The second tier that had a role as a subordinate of the alpha wolves was occupied by the beta wolves. The wolves that followed the instructions from the dominant wolves were the delta and omega wolves. The difference between delta and omega wolves is that the delta dominated the omega

wolves, and the delta was placed in the middle class of their social hierarchy [74], [75]. The hunting mechanism of the gray wolves was initiated by approaching and encircling the prey, and then, the gray wolves started pursuing the prey and harassing the target. After the prey stopped moving, the gray wolves attacked the prey to finish the hunting. In this algorithm, the best candidate solution will be obtained from the alpha wolves due to their knowledge of the potential location of the prey [67], [71].

3) ICA was a metaheuristic algorithm that was inspired by the mechanisms of competition and colonization in human history [76]. This algorithm combines evolutionary concepts of the individual populations that formed an empire based on the colonies and the imperialists [77]. The competition between each empire will destroy weak empires, and the remaining strong empires will dominate each other until there is only one empire that rules over all the colonies, which, from this algorithm, becomes the solution to optimize the parameter during the prediction process [67], [76], [77], [78].

5) Model Performance Evaluation: The model performance evaluation of the land subsidence susceptibility maps was conducted by the ROC curve analysis and RMSE calculations. The ROC curve analysis will produce the area under the curve (AUC), which ranges from 0 to 1 that defines the performance of the land subsidence susceptibility model [79], [80]. The higher AUC values close to 1 indicate that the land subsidence susceptibility models were more accurate, and the AUC values lower than 0.5 were unacceptably inaccurate [47], [81], [82]. The RMSE calculation in this study also involves calculating the mean square error (mse), mean absolute error (MAE), and StD. Those metrics were statistical measurements that represented the difference between the predicted models with their actual values. Those model performances were statistical metrics commonly used in regression analysis and did not have any units due to its measurement to test the prediction result between the continuous regression data that was valued to be the index of land subsidence susceptibility maps. The lower mse, RMSE, MAE, and StD values indicate higher model performance [83], [84].

IV. RESULTS

A. Time-Series InSAR Map

The results of time-series InSAR processing based on the ICOPS algorithm were applied to Sentinel-1 data, resulting in combined PS and DS points over the Pekalongan area. We overlaid the result in Fig. 6(a) with a combination of natural colors from band 4 (red), band 3 (green), and band 2 (blue) of Landsat 8 OLI/TIRS imagery. The Landsat image was provided by the United States Geological Survey (USGS) and taken on October 1, 2020. The model performance of the CNN algorithm is shown in the RMSE metric [see Fig. 6(b)] and frequency error [see Fig. 6(c)]. The low RMSE (0.167) has a relatively small error compared to the variability of the data (StD = 0.168), suggesting a high level of accuracy in the prediction of the optimum result. Further validation was also conducted with the ground-truth data from the

GPS measurement. The comparison between ICOPS (P1) and GPS measurements is shown in Fig. 6(d). Other points that were also selected for the analysis are shown in southern Pekalongan (P2) and northern Pekalongan (P3). Both southern (P2) and northern (P3) time-series data are shown in Fig. 6(e) and (f), respectively.

The result from the combined PS and DS measurements (CSI) before point optimization was approximately 307 451 MPs, as shown in Fig. 6(a), with 50% transparency measurements. These combined PS and DS points consist of 7488 PS points and 299 963 DS points. After the optimization process by the CNN algorithm, the number of MPs was reduced by 85 777 points, and these points were considered to be noisy and outlier-affected data. Thus, more precise MPs of approximately 221 674 points were obtained using the optimization method based on the CNN algorithm. The statistical clustering process based on the OHSA separated MPs into several categories, such as insignificant points and clusters of deformation points. This separation was obtained by selecting the MPs with a p-value of less than 0.01, which is associated with a very high absolute value of the Z-score. After the clustering process, the number of MPs was 146 377 points, resulting in 75 297 MPs that were considered to be insignificant.

The results from the optimization process using the CNN algorithm and OHSA method were validated by comparison with the GPS measurements. We selected the MPs in a 1 km radius near the GPS station before and after the optimization process. The comparison was conducted to obtain the RMSE value. The result shows that a lower RMSE value of the time-series result was obtained after the optimization process than before the optimization process. Thus, the removal of noisy MPs could increase the reliability of the MPs. Furthermore, we attempted to plot the available data from the groundwater level in Pekalongan city with the time-series deformation results. The time-series measurements from ICOPS processing were selected from several locations in Pekalongan city, specifically the point near the GPS station and monitoring well. The time-series deformation result was compared with the time-series GPS measurement from the CORS station in Pekalongan. The daily GPS measurement was conducted using a Continuously Operating Reference Station (CORS) provided by the Indonesian Geospatial Information Agency [85]. The data from the GPS station in Pekalongan only recorded information related to deformation from 2017 to 2020, and the data have not been updated recently. The comparison between time-series InSAR measurements at the selected point within the area of the GPS station location displayed in Fig. 6(d) shows that the deformation pattern between the time-series InSAR measurement and the GPS measurements shows good agreement in terms of the deformation pattern. This is supported by the comparison between the time-series InSAR result and the GPS measurement. The RMSE results show that the RMSE value between ICOPS and GPS measurement shows a better result (1.79 cm/year) after the optimization process. A slight variation in the time-series InSAR was suggested to be caused by the seasonal effect from the groundwater recharge that
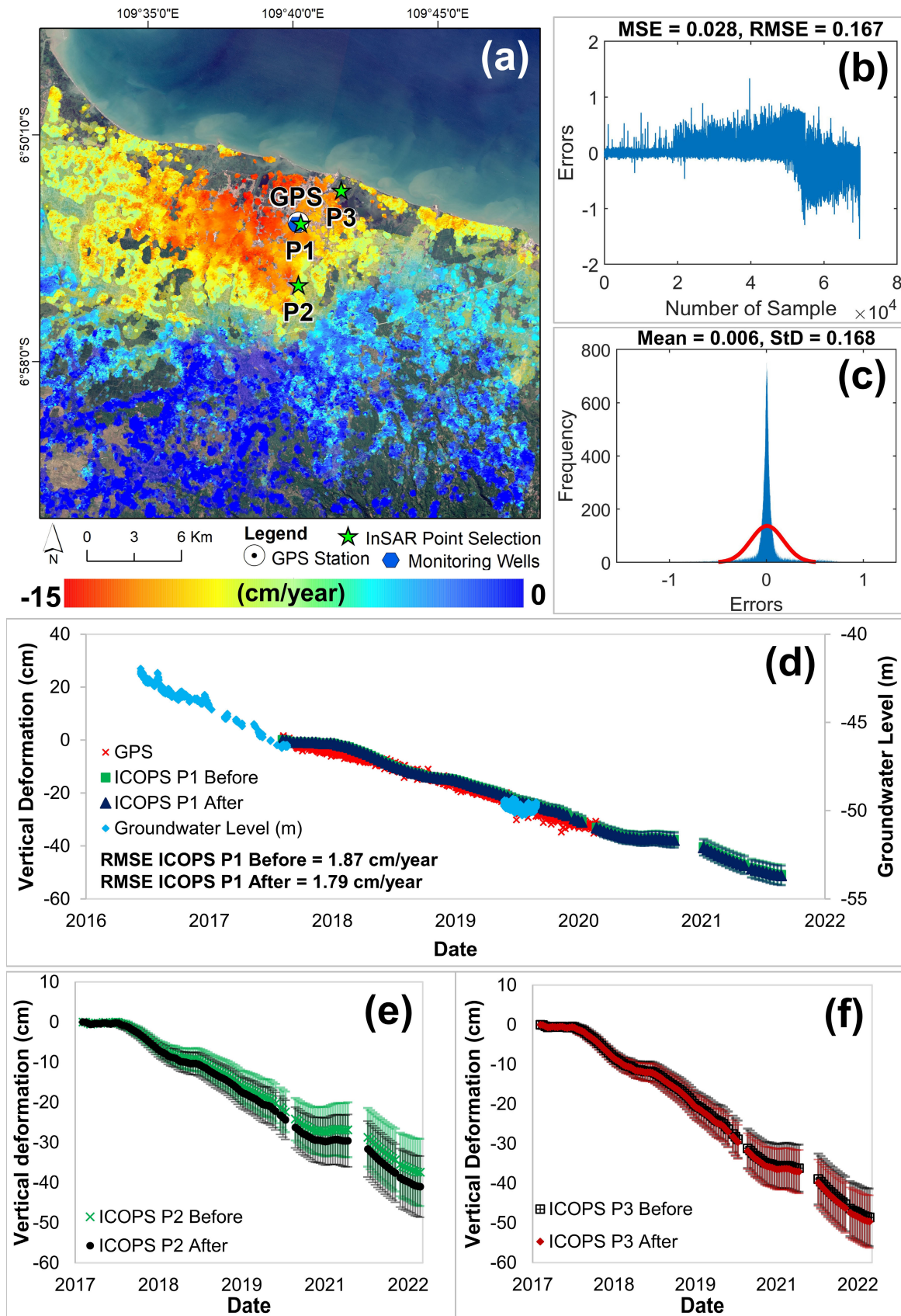
Fig. 6. (a) Mean vertical deformation rate map from before (50% transparency) and after (0% transparency) the ICOPS optimization process. (b) and (c) RMSE value and frequency errors of the testing data. (d) Comparison between time-series GPS, P1 ICOPS before and ICOPS after optimization, and groundwater level data. (e) P2 and (f) P3 of ICOPS before and after optimization.

decreased the accumulated deformation due to the compaction of pore pressure within the soil [86].

The cause of the land subsidence in Pekalongan was known to be groundwater exploitation for domestic use and agriculture. The groundwater utilization in Pekalongan was because of the scarcity of the water source, especially the surface water [35], [87]. The groundwater extraction activities in Pekalongan in this study were shown by the groundwater level data in a monitoring well near the GPS station in Fig. 6(a). The available groundwater level data were acquired from the groundwater monitoring wells provided by the Ministry of Energy and Mineral Resources of Central Java, Indonesia. The relationship between land subsidence and the groundwater level in Fig. 6(d) shows that the deformation in Pekalongan is in good agreement with the decrease in the groundwater level. However, the availability of groundwater level data has also been a problem in this city due to the lack of updated measurements that are suggested to be caused by local site effects. The data in daily measurements were collected from 2016 to 2017, data loss occurred in early 2018, and several measurements were recorded in late 2019. The gaps between the groundwater level data were suggested to be caused by the possibility of damage to the measuring device. Nevertheless, the measurement from ICOPS could allow an additional interpretation of the event that occurred regarding the groundwater level data in Pekalongan. The comparison between time-series deformation maps could further interpret the extent of the gaps in groundwater level data. Despite the possibility of inferring groundwater information from deformation data, the absence of available groundwater data at these MPs alone is insufficient for concluding groundwater conditions in Pekalongan city. Therefore, an additional analysis will be conducted to assess the spatial extent of GWS by utilizing the GRACE satellite and GLDAS data.

B. GWS and Precipitation

The cause of land subsidence in Pekalongan was known to be caused by the excessive groundwater extraction activities [35], [87]. The analysis of the groundwater usage was conducted by utilizing the groundwater level data provided by the Ministry of Energy and Mineral Resources of Central Java, Indonesia. The comparison between the land subsidence and the groundwater usage was conducted by extracting the time-series deformation data from the ICOPS measurement after the optimization process in Fig. 6(d) and the time-series of groundwater level data from the monitoring wells in Fig. 6(a) to focus on the comparison between both of the data. The activity of the groundwater usage in Pekalongan could be seen during the daily measurement from 2016 to 2017 that there is a decline pattern in groundwater level data during that period. However, the groundwater level measurement in the following year could not be acquired due to the absence of data available to be accessed for the analysis. The absence of groundwater level data was suggested to be caused by the damage to the monitoring device. Despite that the analysis of land subsidence and groundwater level cannot be conducted thoroughly due to the absence of the groundwater level data,

there is a fairly similar pattern between the two comparisons made in Fig. 7(a) that the gaps data in groundwater level data could be suggested to have the similar interpretation as the land subsidence data. The mutual interpretation between the groundwater level and the InSAR data in this study is consistent with several studies that compared groundwater level data with the land subsidence data from the time-series InSAR measurement [88], [89]. Thus, the relationship between InSAR data and groundwater levels can be mutually interpreted as the land subsidence occurrence in Pekalongan was caused by the increase in groundwater exploitation that affected the decline of the groundwater level from 2016 to 2019. A similar decline in groundwater level data was also suggested to occur between 2020 and 2021 with a similar interpretation to the deformation pattern in ICOPS data. Therefore, the groundwater extraction activities in Pekalongan, which are shown as the decline pattern in groundwater level data, could describe the causes of land subsidence in Pekalongan.

Another relationship between land subsidence and precipitation was also conducted to understand the indirect relationship between rainfall and land subsidence in Pekalongan, as shown in Fig. 7(b). Monthly total precipitation data from a subdistrict in Pekalongan were provided by the Central Bureau of Statistics of Pekalongan city and Regency from the estimated weather station distributed in Fig. 8(a). The precipitation data that are available from 2016 to 2021 were used to analyze the monthly precipitation pattern that occurred during the period of monitoring land subsidence in Pekalongan. The monthly average of precipitation data from each subdistrict was averaged. The comparison between the land subsidence and the precipitation shows a nonlinear pattern between the land subsidence and the rainfall in Pekalongan. The comparative analysis was shown by the boundaries between each rainy season from 2017 to 2021. The result shows that the land subsidence in each rainy season showed a stable deformation with an increase in deformation rate shown only during the dry season. The important mechanism behind this was suggested to be caused by the infiltration process from the precipitation to the pore soil and led to the decrease in the rate of land subsidence that was caused by the groundwater recharge. A slightly different pattern in 2019 and 2020 was suggested to be caused by the soil being saturated before recharging the groundwater table [86], [90]. A direct relationship could occur between groundwater level and precipitation due to the decrease of groundwater recharge when the precipitation is less in the northern area of Pekalongan. The decrease in the groundwater recharge with the overexploitation of groundwater usage could increase the deformation rate due to the compaction as pore spaces within the aquifers decrease in size. The relationship between this phenomenon is not linear due to the several aspects that can affect the groundwater recharge from the precipitation, such as infiltration, surface runoff, and evapotranspiration [91].

The results from GRACE processing results are shown in Fig. 8(b). Those results were produced by averaging GRACE data between 2017 and 2021 for each different institution. The mean TWS data from each different institution will be averaged to produce the final TWS result with higher
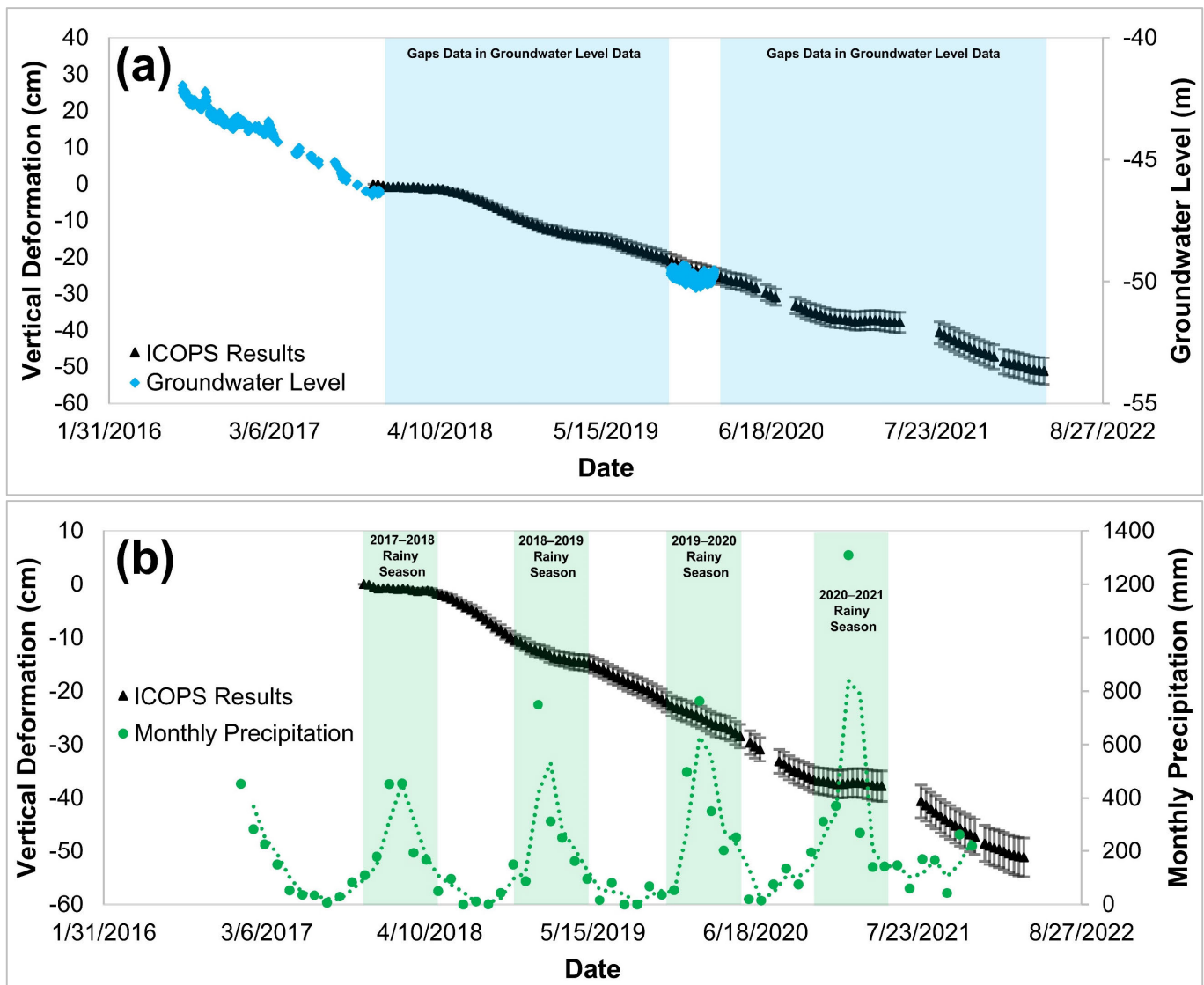
Fig. 7. Comparison between land subsidence results from ICOPS measurement with (a) groundwater level data and (b) monthly precipitation data.

accuracy of the GRACE measurement [1], [42], [43], [44], [45]. After that, the final TWS result will be subtracted from the parameters from GLDAS data used in this study, which are SM, CWS, and QS. Those parameters were averaged for each year and converted to the anomaly data to reduce the uncertainty with the GRACE results. The scale of the Pekalongan area in this study, which only has an area of approximately 716 km^2, was insufficient to properly display the result for Pekalongan as the GRACE data were processed in 1-pixel degree (approximately 12 368 km^2). Because of that, the GRACE result was displayed in a larger area; in this case, the result of GRACE measurement was shown over Central Java province with the preview of GRACE results from each institution that covers Java Island. The results of GWS from the subtraction of GRACE and GLDAS data processing are used in Fig. 8(c) to be compared with the monthly precipitation data.

The result from GRACE processing was compared with the spatially averaged monthly precipitation in the Pekalongan area from the estimated weather station distributed in Fig. 8(a). We processed the precipitation data by averaging all of the data acquired from each subdistrict to represent the area that covers the study areas that were equal to the resolution of the GWS anomaly data from the GRACE measurement. The comparison between the GWS and monthly precipitation data shows a similar seasonal pattern between the two data. This suggests that the precipitation over the Pekalongan area affects the GWS in Pekalongan. The difference in recharge time during the increase in GWS, which caused a phase shift between the peaks of the two data, was suggested to be caused by the delay of the infiltration process during the groundwater recharge. The process that caused the time lag was suggested to have occurred during the infiltration and percolation process of rainfall, which did not happen instantaneously after the precipitation occurred. This time lag can result in a phase shift between the peak of precipitation and the peak of groundwater recharge. The reason that could affect the time lag is because of the lower permeability on the aquifer characteristics that
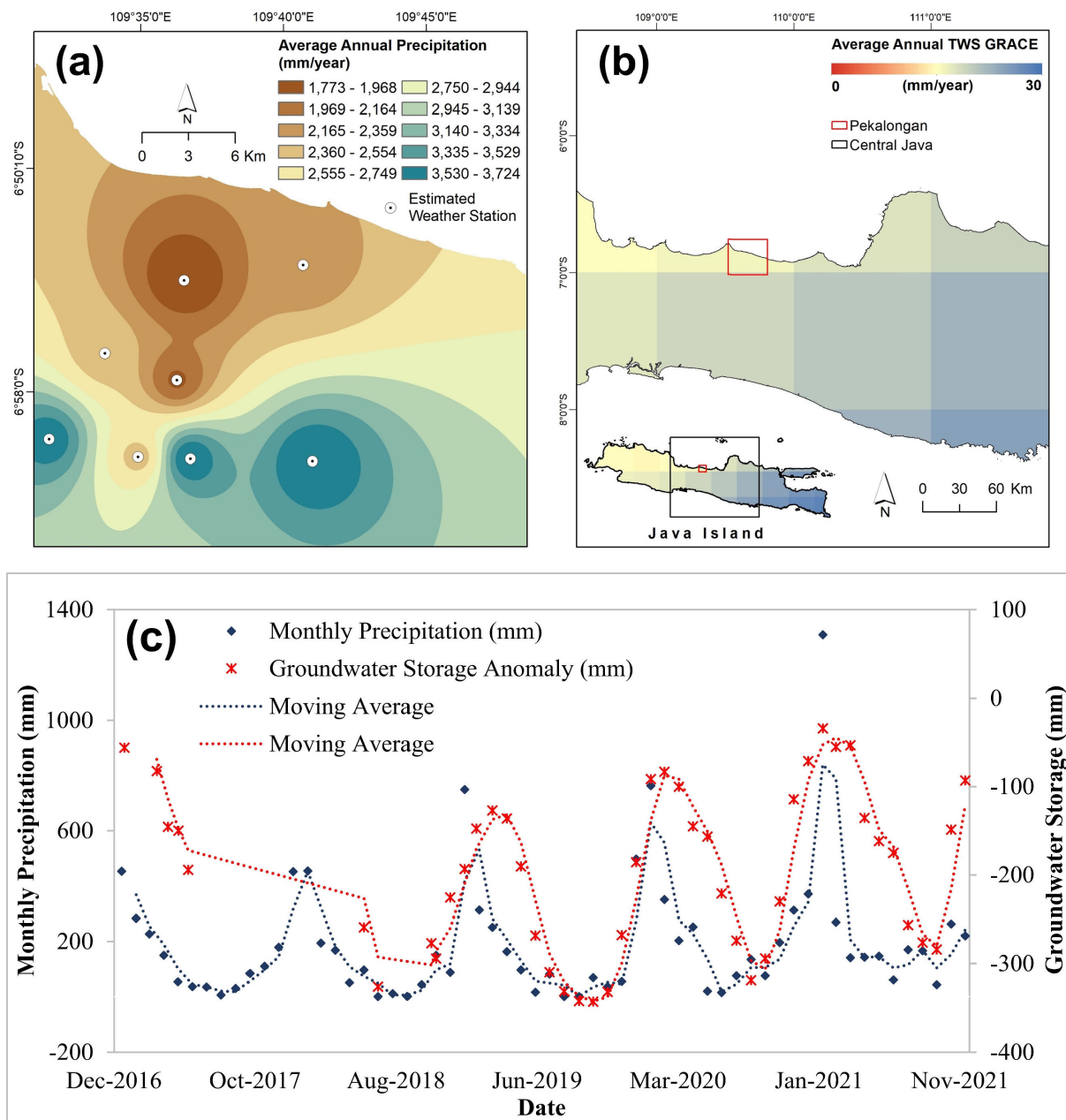
Fig. 8. (a) Average annual precipitation over the study area from 2017 to 2021. (b) Mean TWS GRACE results from 2017 to 2021 provided by averaging annual TWS data from CSR, GFZ, and JPL. (c) GWS changes in Pekalongan (represented by the red points) were analyzed using GRACE and GLDAS data processing techniques. The monthly precipitation data (represented by the blue points) were obtained from local authorities for the period spanning 2017–2021. Both datasets were smoothed using a moving average for better analysis.

affect the water to transport with a longer time. Another reason that could contribute to the phase shift between the two types of data the GWS anomaly and the monthly precipitation data is suggested to be because of the groundwater extraction activity for daily needs, such as irrigation or domestic use. The groundwater extraction activity could lead to changes in the GWS anomaly pattern that might not directly correlate with the monthly precipitation data. The time delay during the groundwater recharge proves that the soil infiltration from the high rainfall intensity could vary with time [92], [93].

C. Land Subsidence Susceptibility Maps

The result from the land subsidence susceptibility map was generated by utilizing the time-series ICOPS result as the

land subsidence inventory map. The land subsidence inventory map that is randomly divided into 70% training and 30% testing data will be used to generate land subsidence susceptibility maps; 70% of training data will be used for the training process using the CNN algorithm and its optimization algorithm. Three different combinations between CNN and its optimization were utilized to generate the land subsidence susceptibility maps. Those algorithms are standalone CNN, CNN with GWO optimization algorithm, and CNN with ICA optimization algorithm. The results of the land subsidence susceptibility maps are shown in Fig. 9(a)–(c). The prediction result of each model was classified using a quantile method into five classes: very low, low, moderate, high, and very high susceptibility indices. The quantile method in the classification
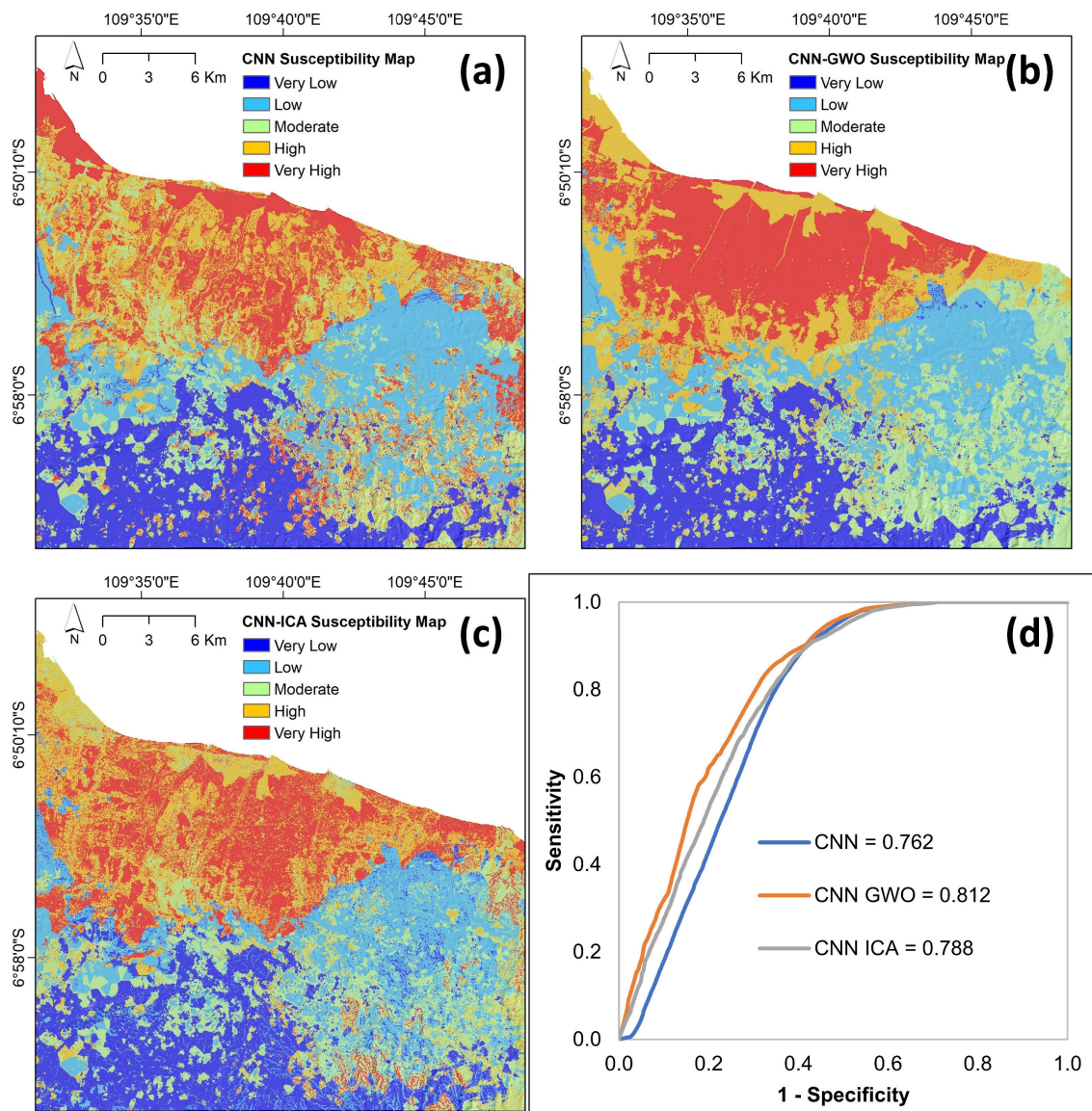
Fig. 9. Land subsidence susceptibility map from (a) CNN algorithm, (b) CNN-GWO algorithm, and (c) CNN-ICA algorithm. (d) Comparison of the ROC curves for each model.

process was used to maintain consistency across the different datasets. The land subsidence susceptibility maps from each model need to be evaluated to assess its model performance accuracy. The accuracy of each prediction model was calculated by ROC curve analysis [see Fig. 9(d)]. The ROC curve analysis was conducted by comparing the land subsidence susceptibility maps and 30% of testing data. The higher AUC value of 1 indicated that the model has an excellent accuracy (AUC between 0.9 and 1.0), and the lower AUC value than 0.5 indicated that the model is not acceptable (0 as the lowest accuracy) [50]. The results show that the CNN-GWO algorithm has a higher accuracy (0.812) than that of the CNN-ICA (0.788) and the CNN (0.762). The highest accuracy obtained by CNN-GWO in this study was in the 0.8–0.9 range, indicating very good accuracy. Meanwhile, the other models can be interpreted as having good accuracy, with an AUC value of approximately 0.7–0.8. The land subsidence susceptibility map generated in this study provides valuable information for

urban development, particularly when considering the superior performance of the CNN-GWO map compared to that of other methods. In addition, the accuracy of the land subsidence susceptibility map is higher than the unsatisfactory accuracy, which is between 0.5 and 0.6 [50].

In the study area, approximately 20% of the area is indexed by very high and low probabilities of land subsidence occurrence based on all models (see Fig. 10). The probable area of land subsidence is located along the coastal area and extends between urban centers or densely populated areas surrounded by rice fields. The overall distribution of the pixel percentage of the susceptibility models showed mostly an equal distribution with slight differences from the CNN algorithm that overestimated the high susceptibility classes while underestimating the moderate susceptibility classes. According to the distribution of the index of land subsidence susceptibility maps, most of the areas with very high probability class distributions are characterized by low altitudes (under 9 m).
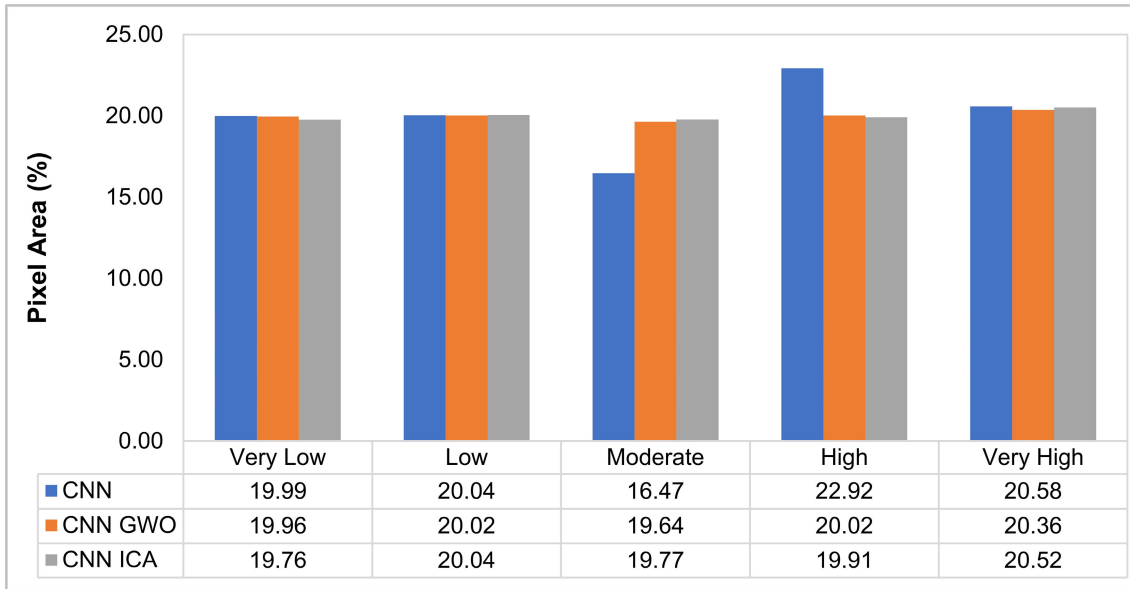
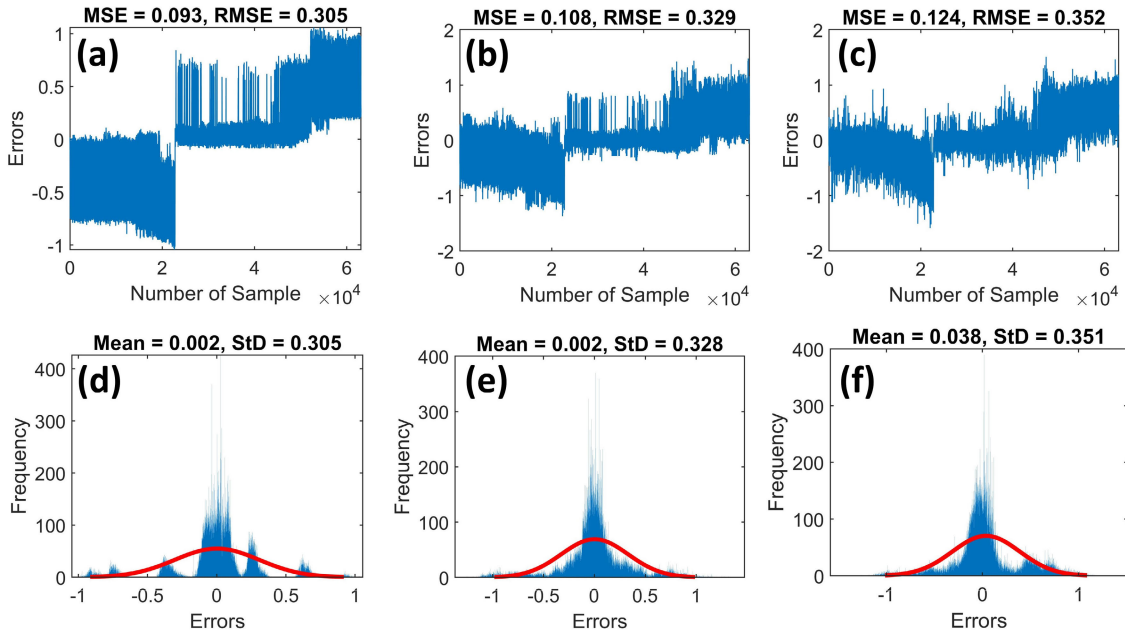Fig. 10. Percentage of area comparison between the susceptibility classes for each model.

Fig. 11. MSE, RMSE, and StD of the frequency error from testing data for (a) and (d) CNN-GWO, (b) and (e) CNN-ICA, and (c) and (f) CNN models.

Further analysis regarding the explanations for each factor's influential class will be provided using the FR method in Section IV-B.

The accuracy of the land subsidence susceptibility maps in this study needs to be verified with another accuracy assessment. In this study, we implemented the RMSE metric for evaluating the prediction model to be compared with the testing data. Both inputs for accuracy assessment between the ROC curve and RMSE metric were similar to each other. The result from the RMSE metric is shown as the mean squared error and the root mean squared error. CNN-GWO has the lowest RMSE (0.305), followed by CNN-ICA (0.329) and the standalone CNN algorithm (0.352) [see Fig. 11(a)–(c)]. In addition, the frequency error is shown by the MAE and

StD values in Fig. 11(d)–(f). The red line explains that the frequency error from the CNN-GWO model is narrower ($-0.9 <$ error $< +0.9$) than that of the other model, which has a wider range of frequency errors ($-1 <$ error $< +1$), which means that the result from the CNN-GWO algorithm was the best compared to the other algorithms.

V. DISCUSSION

A. Surface Geology and Land Use Data

We present the distribution maps of land subsidence, which are overlaid by geological formation [see Fig. 12(a)] and land use [see Fig. 12(b)]. A cross section analysis was conducted from point A to point A' in a 28-km line [see Fig. 12(c)]
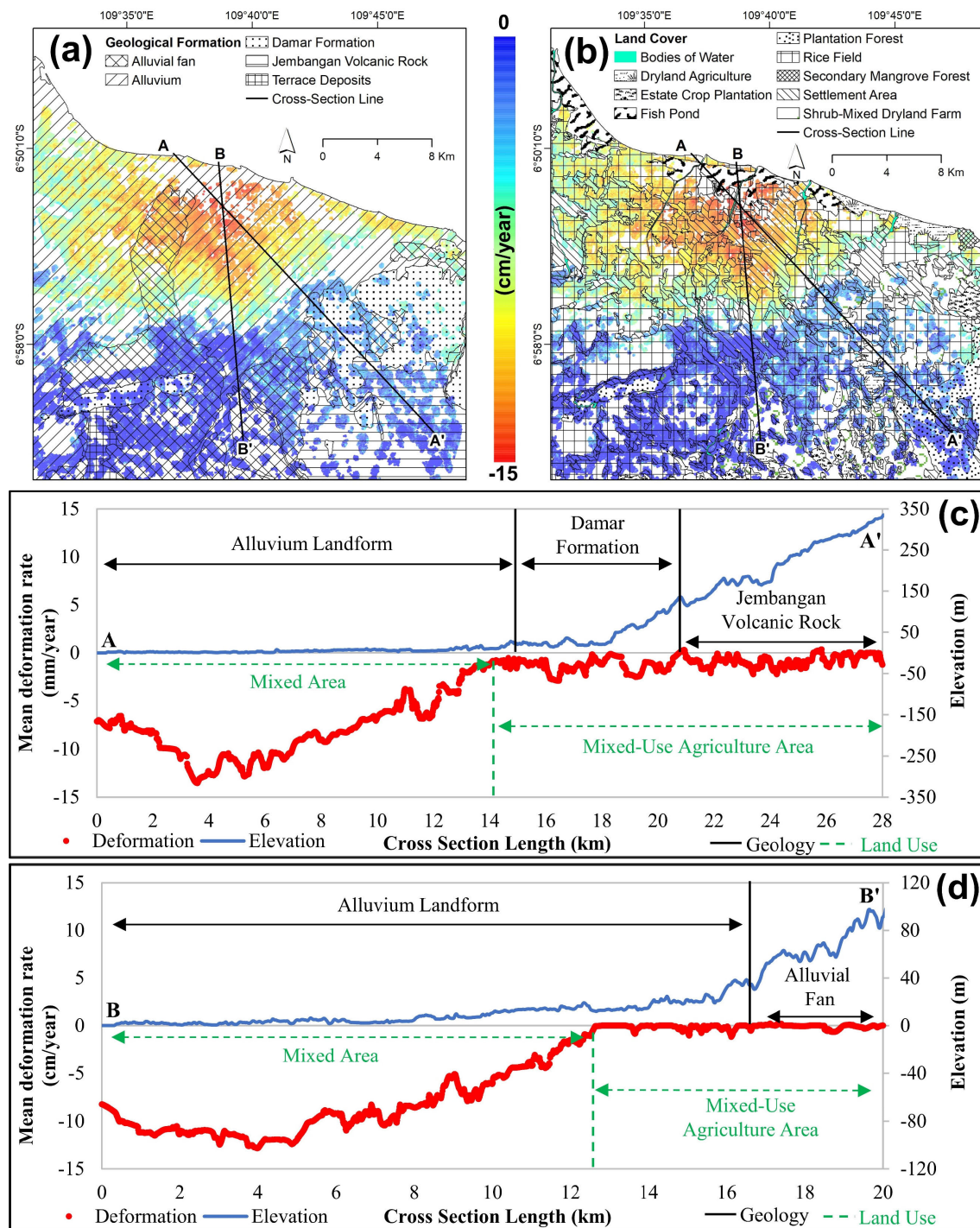
Fig. 12. Mean vertical deformation rate map from (a) land subsidence distribution map overlaid by the geological formation map and (b) land subsidence distribution map overlaid by the land use map. A cross section line to show the relationship between land subsidence, geological formation, land use, and elevation for (c) A–A' and (d) B–B' cross section lines.

and a 20-km line from point B to point B' [see Fig. 12(d)] to show the relationship between the deformation and the elevation along the cross section line in Pekalongan. This map reveals that the greatest subsidence in Pekalongan has mostly occurred in lowland mixed areas located on the alluvium landform. The alluvium landform in Pekalongan is composed of unconsolidated sediment deposited from the Java Sea. The worst events can occur in areas that are close to the coastline because the excessive use of groundwater in this area can cause

compaction of the sediment layer beneath human settlements. This can cause land subsidence, which can lead to further disasters such as high tide flooding and seawater intrusion.

The mixed area in Fig. 12(c) and (d) refers to a cluster of populated residential areas surrounded by rice fields. This mixed area relies heavily on excessive groundwater use for various daily needs, such as living and agricultural purposes. As groundwater is pumped out of an aquifer, the water pressure that supports the soil is reduced, and the soil particles
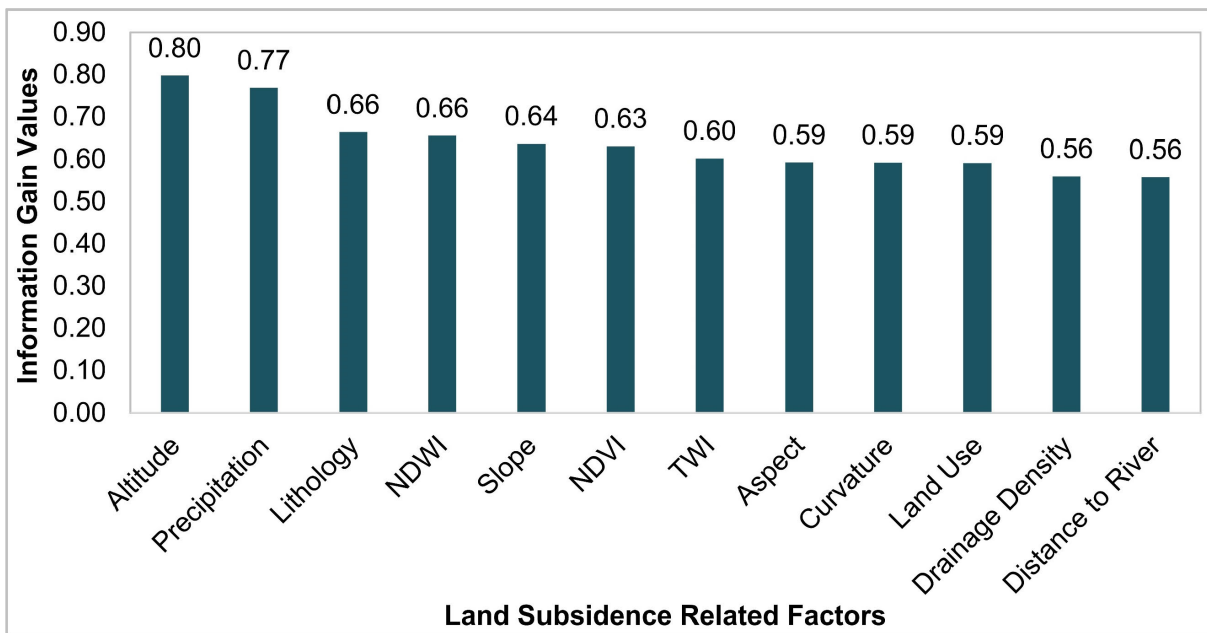
Fig. 13.   Information gain value between land subsidence-related factors.

can then settle and become more compact. This can cause the ground surface to sink, leading to infrastructure damage, increased flood risk, and other problems. In addition to subsidence, the excessive use of groundwater can lead to other problems, such as seawater intrusion, where seawater from the ocean can seep into the groundwater supply, contaminate it, and decrease soil fertility. This can further exacerbate soil compaction and reduce the quality of the groundwater available for use. These conditions can also be worsened by young alluvium soil that has the potential to subside due to natural consolidation. In addition, the load of buildings around the alluvium soil can accelerate compaction, which is usually only caused by the influence of gravity loads [2], [4], [6], [11]. The areas that are not affected by land subsidence are indicated by areas located in the highlands, which generally have geological landforms other than alluvium landforms. In this case, the southeast area was mostly occupied by the Damar formation and Jembangan volcanic rock (line A–A'). Meanwhile, line B–B' shows that the decline in the subsidence area was mostly related to the decrease in the residential area and groundwater recharge in the southern area due to the high precipitation in higher elevation areas.

B. Importance of GIS Factors and Multicollinearity Analysis

The relative importance of the land subsidence-related factors was analyzed using an information gain attribute evaluator [50], as shown in Fig. 13. This method calculates the significance of each land subsidence-related factor from the training dataset [94]. The correlation between altitude and land subsidence occurred because the most populated area is located in a lower elevation area. Precipitation could indirectly influence land subsidence during a decrease in precipitation, which could decrease groundwater recharge and lead to soil compaction. The precipitation factor is also influenced by

TABLE I
MULTICOLLINEARITY ANALYSIS USING VIF AND TOLERANCE

Factor	VIF	Tolerance
Altitude	2.4	0.4
Aspect	1.1	0.9
Curvature	1.2	0.8
Distance to River	1.1	0.9
Drainage Density	1.1	0.9
Land Use	1.2	0.9
Lithology	1.5	0.7
NDVI	3.4	0.3
NDWI	3.7	0.3
Precipitation	2.2	0.5
Slope	2.3	0.4
TWI	2.0	0.5

altitude; as a result, higher elevation areas are more likely to receive more precipitation than lower areas. This is a series of events that can lead to land subsidence for areas still dependent on groundwater use. Lithology and NDWI factors (0.66) were the third most influential factor in land subsidence occurrence, where the alluvium landform played an important role in land subsidence due to the presence of natural compaction, and this process could be worsened by groundwater extraction. The NDWI factor can provide valuable information that is indirectly related to land subsidence, as the NDWI calculates the amount of water content across an area. Further explanation regarding each influential class from each factor is shown by the FR result in Fig. 14.

The FR analysis in Fig. 14 shows the spatial relationship between land subsidence and each land subsidence-related factor. An FR value higher than 1 means that the land subsidence is mostly affected by the specific class, an FR value below 1 means that the specific class is suggested to
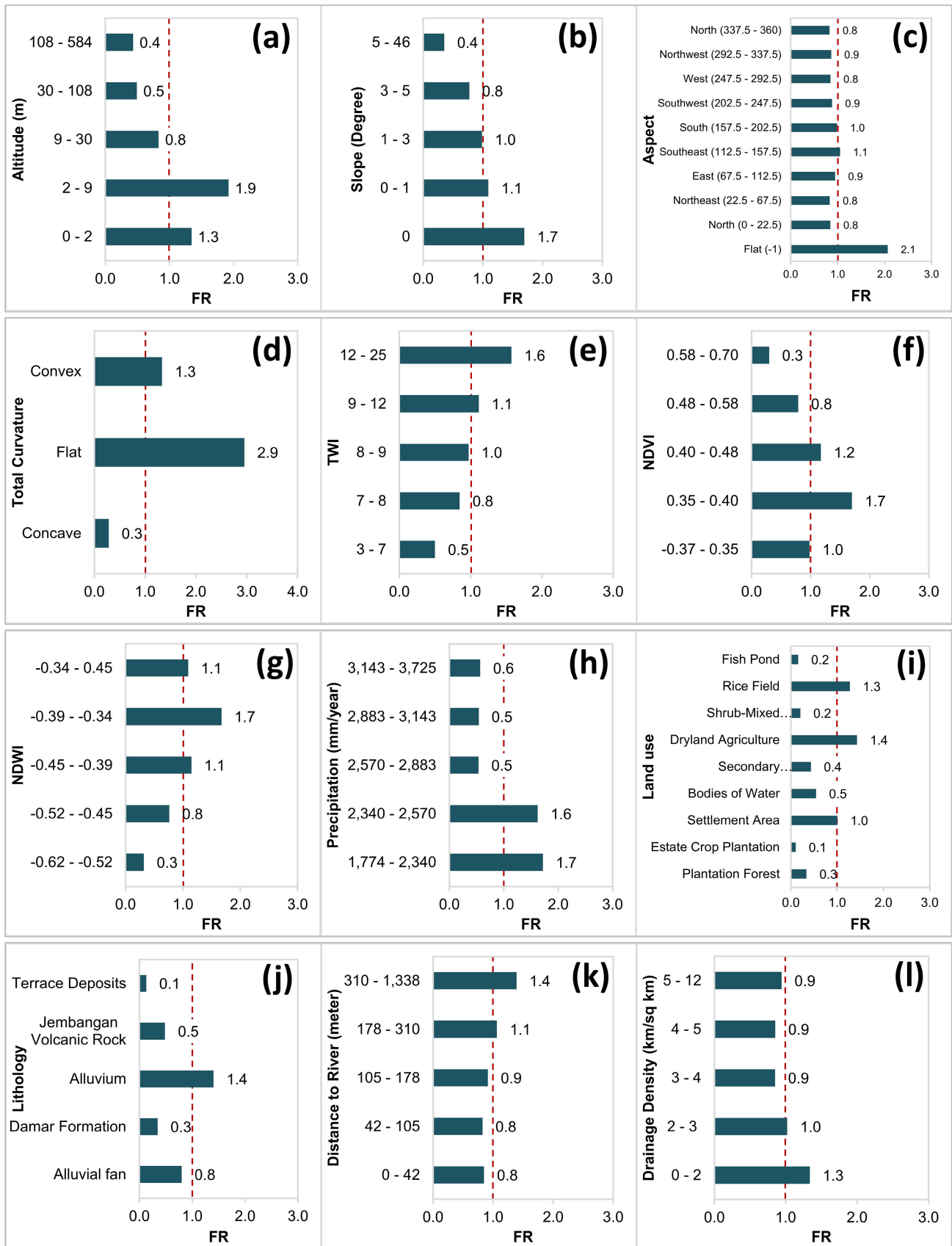
Fig. 14. Results from the FR method show the effectiveness of each class from the land subsidence-related factor: (a) altitude, (b) slope, (c) aspect, (d) total curvature, (e) TWI, (f) NDVI, (g) NDWI, (h) precipitation, (i) land use, (j) lithology, (k) distance to the river, and (l) drainage density.

TABLE II
PEARSON CORRELATION ANALYSIS BETWEEN LAND SUBSIDENCE-RELATED FACTORS

	Altitude	Aspect	Curvature	Distance to River	Drainage Density	Land Use	Litho.	NDVI	NDWI	Precip.	Slope	TWI
Altitude	1.0											
Aspect	0.1	1.0										
Curvature	0.0	0.1	1.0									
Distance to River	0.1	0.0	0.0	1.0								
Drainage Density	0.0	0.0	0.0	0.3	1.0							
Land Use	0.2	0.0	0.0	0.0	0.0	1.0						
Litho.	0.5	0.1	0.0	0.1	0.0	0.2	1.0					
NDVI	0.4	0.1	-0.1	0.1	0.0	0.3	0.3	1.0				
NDWI	0.5	0.1	-0.1	0.1	0.1	0.3	0.4	0.6	1.0			
Precip.	0.6	0.1	0.0	0.1	0.1	0.2	0.5	0.3	0.4	1.0		
Slope	0.4	0.3	0.0	0.1	0.0	0.3	0.3	0.4	0.4	0.3	1.0	
TWI	0.3	0.2	-0.3	0.0	0.0	0.2	0.2	0.3	0.3	0.2	0.6	1.0

have a low correlation, and land subsidence is less affected by that class [50]. The altitude [see Fig. 14(a)] classes show that the land subsidence mostly occurred in areas with an altitude between 0 and 2 and 2 and 9 m, and FR values of approximately 1.3 and 1.9, respectively. Thus, lower elevation areas in Pekalongan are among the most populated areas. The slope class [see Fig. 14(b)] shows that higher FR values of approximately 1.7, 1.1, and 1.0 are found within the flat slopes, with values of approximately 0, 0–1, and 1–3, respectively. This suggests that the land subsidence mostly occurred in an area with a relatively small slope degree, which is mostly habituated by the population in Pekalongan. The aspect classes [see Fig. 14(c)] show the direction of the slope faces, and the flat direction shows the highest correlation of 2.1, followed by the southeast (1.1) and south (1.0) directions. The flat direction class also shows a higher correlation with land subsidence in the total curvature factor (2.9), followed by the convex curvature, with a value of 1.3 [see Fig. 14(d)]. The slope, aspect, and total curvature factors show that land subsidence mostly occurred on flat areas with gentle slopes or plain-like surfaces where the slope degree was 0 or between 0 and 1 with a flat aspect and flat total curvature factor.

The TWI factors [see Fig. 14(e)] show that land subsidence mostly occurred in areas with higher TWI values. TWI values of approximately 8–9, 9–12, and 12–25 were shown to have a higher correlation with land subsidence, with FR values of approximately 1.0, 1.1, and 1.6, respectively. This is because higher TWI values indicate a higher potential for water accumulation, which can increase soil erosion and sediment transport. The lower NDVI area [see Fig. 14(f)], namely, −0.37 to 0.35, 0.35–0.40, and 0.40–0.48, and higher NDWI area [see Fig. 14(g)], namely, −0.45 to −0.39, −0.39 to −0.34, and −0.34 to 0.45, were associated with the presence of land subsidence because of the lower NDVI indicated with areas with limited vegetation cover, i.e., urban areas. Higher NDWI values indicate the higher presence of water content that could indirectly indicate areas that are prone to land subsidence. Areas with higher water content are often associated with higher population densities, as abundant

water resources support populated areas. The FR value from the precipitation factor [see Fig. 14(h)] shows that the land subsidence location was correlated with a lower precipitation area. The precipitation values of approximately 1774–2340 and 2340–2570 mm/year have high correlation values of approximately 1.7 and 1.6, respectively. The land use [see Fig. 14(i)] areas in Pekalongan, which were mostly related to land subsidence, were settlement areas (1.0), rice fields (1.3), and dryland agriculture areas (1.4). Meanwhile, in terms of lithological factors [see Fig. 14(j)], alluvium (1.4) has the highest probability of land subsidence occurrence. The trend for distance to the river [see Fig. 14(k)] shows that the farther the distance to the river is, the smaller the correlation to the land subsidence by its FR value. This resulted in distances to the river of approximately 178–310 and 310–1338 m having higher FR values of approximately 1.1 and 1.4, respectively. Meanwhile, the opposite trend was shown in drainage density [see Fig. 14(l)], where the lowest drainage density has the highest land subsidence occurrence compared to the highest drainage density area. This result shows that drainage densities of approximately 0–2 and 2–3 km/km^2 have higher FR values of approximately 1.3 and 1.0, respectively.

Another important analysis that should be performed when generating susceptibility maps is the identification of redundant variables by conducting a multicollinearity analysis. In this study, we analyzed the degree of correlation between each land subsidence-related factor using VIFs, tolerance, and the Pearson correlation analysis. A VIF value > 10 and tolerance value < 0.1 are considered to not correlate with land subsidence occurrence and have high multicollinearity between each factor [95], [96]. In this study, the results from the VIF and tolerance are shown in Table I. The VIF value from all the factors was lower than 4, and all tolerance values were higher than 0.2. Meanwhile, from the Pearson correlation analysis (see Table II), a moderate correlation was shown with a Pearson correlation of 0.6 between altitude and precipitation, between NDVI and NDWI, and between slope and TWI. Nevertheless, most of the land subsidence-related factors show a weak correlation, and not having a very

strong correlation with each other indicates that most of the land subsidence-related factors are independent of each other and would not cause distortion during the training process [59], [60]. This is shown by the fact that all correlation values are within the safe threshold, which is lower than 0.7 [59], [96]. Thus, the multicollinearity analysis from Pearson correlation was supported by the results of multicollinearity using the VIF and tolerance method; none of which has a value of more than 5, a tolerance of less than 0.1 shows that there is no multicollinearity found between land subsidence-related factors, and these factors are suitable for generating susceptibility maps.

VI. Conclusion

Land subsidence in Pekalongan was monitored using Sentinel-1 SAR data from 2017 to 2022. A mean vertical deformation map was generated by implementing the ICOPS algorithm. The integration of the ICOPS algorithm with the optimization of MP using a CNN algorithm could reduce the outlier data between the MPs. Spatial clustering was conducted on the MP to remove the nonsignificant points among the MPs by selecting MPs with a p-value < 0.01. The final land subsidence map was presented by the MPs with a 99% significance level from OHSA clustering. The analysis was performed to find the relationship between land subsidence and geological formation, land use, and precipitation data. The cross section analysis results between the land subsidence map, geological formation map, land use map, and elevation map found that land subsidence occurred in lowland areas that consist of alluvium soils. A comparison between the deformation and precipitation data showed that lower precipitation occurred in the deformation area. The analysis of the correlation between GWS data and monthly precipitation data in the Pekalongan area revealed a similar pattern of fluctuation between the two datasets. This observation suggests that GWS data are correlated with monthly precipitation data. Furthermore, the locations of land subsidence in Pekalongan, which coincide with areas of lower precipitation, indirectly support the correlation between land subsidence and GWS changes in Pekalongan. The land subsidence result from the ICOPS measurement was used as an inventory map for generating land subsidence susceptibility maps. Twelve land subsidence-related factors were selected and validated using multicollinearity analysis based on VIF, tolerance, and the Pearson correlation analysis. The factor analysis based on the information gain and FR analysis was also utilized to identify the most influential factor and the highest class correlation with land subsidence occurrence. The land subsidence susceptibility results were generated using three different algorithms, namely, CNN, CNN-GWO, and CNN-ICA. The accuracy assessment based on the ROC curve and RMSE metric analysis between these models was calculated, and it was found that the land subsidence susceptibility model from the CNN-GWO algorithm has the highest accuracy, with an AUC value of approximately 0.812 and RMSE of 0.305. The results of land subsidence in Pekalongan in this study can be further analyzed for multihazard risk analysis in the Pekalongan area, which experiences flooding due to high tides and monsoon rainfall. Another hazard that may arise from flooding is landslides. The correlation between land subsidence events and the increased impact of flood events can be an important factor for understanding multihazard risk analysis in Pekalongan because a series of natural disasters that can occur due to land subsidence can affect flooding up to landslides. Therefore, a multihazard risk map that includes land subsidence, flooding, and landslides in Pekalongan could be conducted to analyze more deeply all potential natural disasters that might occur in Pekalongan.

References

[1] C. Ojha, S. Werth, and M. Shirzaei, "Recovery of aquifer-systems in southwest US following 2012–2015 drought: Evidence from InSAR, GRACE and groundwater level data," *J. Hydrol.*, vol. 587, Aug. 2020, Art. no. 124943, doi: 10.1016/j.jhydrol.2020.124943.

[2] T. A. Meckel, U. S. T. Brink, and S. J. Williams, "Current subsidence rates due to compaction of Holocene sediments in Southern Louisiana," *Geophys. Res. Lett.*, vol. 33, no. 11, pp. 1–5, Jun. 2006, doi: 10.1029/2006GL026300.

[3] L. Sui, F. Ma, and N. Chen, "Mining subsidence prediction by combining support vector machine regression and interferometric synthetic aperture radar data," *ISPRS Int. J. Geo-Inf.*, vol. 9, no. 6, p. 390, Jun. 2020, doi: 10.3390/ijgi9060390.

[4] E. Chaussard, F. Amelung, H. Abidin, and S.-H. Hong, "Sinking cities in Indonesia: ALOS PALSAR detects rapid subsidence due to groundwater and gas extraction," *Remote Sens. Environ.*, vol. 128, pp. 150–161, Jan. 2013, doi: 10.1016/j.rse.2012.10.015.

[5] M. A. Marfai and L. King, "Monitoring land subsidence in Semarang, Indonesia," *Environ. Geol.*, vol. 53, no. 3, pp. 651–659, Oct. 2007, doi: 10.1007/s00254-007-0680-3.

[6] L. E. Erban, S. M. Gorelick, and H. A. Zebker, "Groundwater extraction, land subsidence, and sea-level rise in the Mekong Delta, Vietnam," *Environ. Res. Lett.*, vol. 9, no. 8, Aug. 2014, Art. no. 084010, doi: 10.1088/1748-9326/9/8/084010.

[7] H. Z. Abidin, H. Andreas, I. Gumilar, Y. Fukuda, Y. E. Pohan, and T. Deguchi, "Land subsidence of Jakarta (Indonesia) and its relation with urban development," *Natural Hazards*, vol. 59, no. 3, pp. 1753–1771, Dec. 2011, doi: 10.1007/s11069-011-9866-9.

[8] S. Mazzotti, A. Lambert, M. Van der Kooij, and A. Mainville, "Impact of anthropogenic subsidence on relative sea-level rise in the fraser river delta," *Geology*, vol. 37, no. 9, pp. 771–774, Sep. 2009, doi: 10.1130/G25640A.1.

[9] Y. Budiyono, J. C. J. H. Aerts, D. Tollenaar, and P. J. Ward, "River flood risk in Jakarta under scenarios of future change," *Natural Hazards Earth Syst. Sci.*, vol. 16, no. 3, pp. 757–774, Mar. 2016, doi: 10.5194/nhess-16-757-2016.

[10] D. Sengupta, R. Chen, M. E. Meadows, and A. Banerjee, "Gaining or losing ground? Tracking Asia's hunger for 'new' coastal land in the era of sea level rise," *Sci. Total Environ.*, vol. 732, Aug. 2020, Art. no. 139290, doi: 10.1016/j.scitotenv.2020.139290.

[11] P. E. Yastika, N. Shimizu, and H. Z. Abidin, "Monitoring of long-term land subsidence from 2003 to 2017 in coastal area of Semarang, Indonesia by SBAS DInSAR analyses using ENVISAT-ASAR, ALOS-PALSAR, and Sentinel-1A SAR data," *Adv. Space Res.*, vol. 63, no. 5, pp. 1719–1736, Mar. 2019, doi: 10.1016/j.asr.2018.11.008.

[12] M. A. Marfai and L. King, "Coastal flood management in Semarang, Indonesia," *Environ. Geol.*, vol. 55, no. 7, pp. 1507–1518, Oct. 2008, doi: 10.1007/s00254-007-1101-3.

[13] R. Dewi, W. Bijker, A. Stein, and M. Marfai, "Fuzzy classification for shoreline change monitoring in a part of the northern coastal area of Java, Indonesia," *Remote Sens.*, vol. 8, no. 3, p. 190, Feb. 2016, doi: 10.3390/rs8030190.

[14] M. A. Marfai et al., "Natural hazards in central Java Province, Indonesia: An overview," *Environ. Geol.*, vol. 56, no. 2, pp. 335–351, Nov. 2008, doi: 10.1007/s00254-007-1169-9.

[15] S. Fiaschi et al., "From ERS-1/2 to Sentinel-1: Two decades of subsidence monitored through A-DInSAR techniques in the Ravenna area (Italy)," *GIScience Remote Sens.*, vol. 54, no. 3, pp. 305–328, May 2017, doi: 10.1080/15481603.2016.1269404.

[16] C. W. Lee, H. S. Jung, and S. W. Kim, "Measurement of a small-scale subsidence in a reclaimed land using RADARSAT-1 SAR," in *Proc. Fringe Workshop*, vol. 610, Feb. 2006, pp. 1–6.

[17] Q. Zhao et al., "Generation of long-term InSAR ground displacement time-series through a novel multi-sensor data merging technique: The case study of the Shanghai coastal area," *ISPRS J. Photogramm. Remote Sens.*, vol. 154, pp. 10–27, Aug. 2019, doi: 10.1016/j.isprsjprs.2019.05.005.

[18] W.-C. Hung, C. Hwang, C.-P. Chang, J.-Y. Yen, C.-H. Liu, and W.-H. Yang, "Monitoring severe aquifer-system compaction and land subsidence in Taiwan using multiple sensors: Yunlin, The Southern Choushui River Alluvial Fan," *Environ. Earth Sci.*, vol. 59, no. 7, pp. 1535–1548, Feb. 2010, doi: 10.1007/s12665-009-0139-9.

[19] A. Ferretti et al., "Submillimeter accuracy of InSAR time series: Experimental validation," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 5, pp. 1142–1153, May 2007, doi: 10.1109/TGRS.2007.894440.

[20] Z. Lu, O. Kwoun, and R. Rykhus, "Interferometric synthetic aperture radar (InSAR): Its past, present and future," *Photogramm. Eng. Remote Sens.*, vol. 73, no. 3, pp. 217–221, Mar. 2007.

[21] J. Dong et al., "Mapping landslide surface displacements with time series SAR interferometry by combining persistent and distributed scatterers: A case study of Jiaju landslide in Danba, China," *Remote Sens. Environ.*, vol. 205, pp. 180–198, Feb. 2018, doi: 10.1016/j.rse.2017.11.022.

[22] P. Lu, S. Bai, V. Tofani, and N. Casagli, "Landslides detection through optimized hot spot analysis on persistent scatterers and distributed scatterers," *ISPRS J. Photogramm. Remote Sens.*, vol. 156, pp. 147–159, Oct. 2019, doi: 10.1016/j.isprsjprs.2019.08.004.

[23] E. Chaussard and F. Amelung, "Precursory inflation of shallow magma reservoirs at west Sunda volcanoes detected by InSAR," *Geophys. Res. Lett.*, vol. 39, no. 21, pp. 1–6, Nov. 2012, doi: 10.1029/2012GL053817.

[24] C.-W. Lee, Z. Lu, and J. W. Kim, "Monitoring Mount Sinabung in Indonesia using multi-temporal InSAR," *Korean J. Remote Sens.*, vol. 33, no. 1, pp. 37–46, Feb. 2017, doi: 10.7780/kjrs.2017.33.1.4.

[25] A. Hooper, P. Segall, and H. Zebker, "Persistent scatterer interferometric synthetic aperture radar for crustal deformation analysis, with application to Volcán Alcedo, Galápagos," *J. Geophys. Res., Solid Earth*, vol. 112, Jul. 2007, Art. no. B07407, doi: 10.1029/2006JB004763.

[26] S. Samsonov, N. d'Oreye, and B. Smets, "Ground deformation associated with post-mining activity at the French–German border revealed by novel InSAR time series method," *Int. J. Appl. Earth Observ. Geoinf.*, vol. 23, pp. 142–154, Aug. 2013, doi: 10.1016/j.jag.2012.12.008.

[27] R. Bonì, G. Pilla, and C. Meisina, "Methodology for detection and interpretation of ground motion areas with the A-DInSAR time series analysis," *Remote Sens.*, vol. 8, no. 8, p. 686, Aug. 2016, doi: 10.3390/rs8080686.

[28] P. Berardino, G. Fornaro, R. Lanari, and E. Sansosti, "A new algorithm for surface deformation monitoring based on small baseline differential SAR interferograms," *IEEE Trans. Geosci. Remote Sens.*, vol. 40, no. 11, pp. 2375–2383, Nov. 2002, doi: 10.1109/TGRS.2002.803792.

[29] W. L. Hakim, M. F. Fadhillah, S. Park, B. Pradhan, J.-S. Won, and C.-W. Lee, "InSAR time-series analysis and susceptibility mapping for land subsidence in Semarang, Indonesia using convolutional neural network and support vector regression," *Remote Sens. Environ.*, vol. 287, Mar. 2023, Art. no. 113453, doi: 10.1016/j.rse.2023.113453.

[30] *BPS-Statistic of Pekalongan Municipality, Pekalongan Municipality In Figures*, BPS-Statistic of Pekalongan Municipality, Pekalongan, Indonesia, 2020.

[31] R. W. van Bemmelen, *The Geology of Indonesia*, vol. 1. The Hague, The Netherlands: Martinus Nijhoff, 1949.

[32] D. S. Sudjono, U. Harmoko, and G. Yuliyanto, "Delineation of geothermal manifestation in Sangubanyu area based on microtremor HVSR method," in *Proc. E3S Web Conf.*, 2019, vol. 125, no. 201, pp. 10–14, doi: 10.1051/e3sconf/201912514012.

[33] D. Sarah, L. Hutasoit, R. Delinom, I. Sadisun, and T. Wirabuana, "A physical study of the effect of groundwater salinity on the compressibility of the Semarang-DemakAquitard, Java Island," *Geosciences*, vol. 8, no. 4, p. 130, Apr. 2018, doi: 10.3390/geosciences8040130.

[34] S. Habibi, A. Pribadi, and J. Sitorus, "The concept design for adaptation of climate change through integrated and sustainable flood infrastructure in the coastal area of Pekalongan, Indonesia," *Geographica Pannonica*, vol. 25, no. 2, pp. 121–135, 2021, doi: 10.5937/gp25-30852.

[35] M. Khairulbahri, "The qualitative analysis of the Nexus dynamics in the Pekalongan coastal area, Indonesia," *Sci. Rep.*, vol. 12, no. 1, pp. 1–16, Jul. 2022, doi: 10.1038/s41598-022-15683-9.

[36] M. F. Fadhillah, W. L. Hakim, M. Panahi, F. Rezaie, C.-W. Lee, and S. Lee, "Mapping of landslide potential in Pyeongchang-gun, South Korea, using machine learning meta-based optimization algorithms," *Egyptian J. Remote Sens. Space Sci.*, vol. 25, no. 2, pp. 463–472, Aug. 2022, doi: 10.1016/j.ejrs.2022.03.008.

[37] M. F. Fadhillah, A. R. Achmad, and C.-W. Lee, "Improved combined scatterers interferometry with optimized point scatterers (ICOPS) for interferometric synthetic aperture radar (InSAR) time-series analysis," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022, Art. no. 5220014, doi: 10.1109/TGRS.2021.3138763.

[38] D. H. T. Minh, R. Hanssen, and F. Rocca, "Radar interferometry: 20 years of development in time series techniques and future perspectives," *Remote Sens.*, vol. 12, no. 9, p. 1364, 2020, doi: 10.3390/RS12091364.

[39] J. J. Sousa, A. J. Hooper, R. F. Hanssen, L. C. Bastos, and A. M. Ruiz, "Persistent scatterer InSAR: A comparison of methodologies based on a model of temporal deformation vs. spatial correlation selection criteria," *Remote Sens. Environ.*, vol. 115, no. 10, pp. 2652–2663, Oct. 2011, doi: 10.1016/j.rse.2011.05.021.

[40] A. Hooper, "Persistent scatter radar interferometry for crustal deformation studies and modeling of volcanic deformation," Ph.D. dissertation, Stanford Univ., Stanford, CA, USA, 2006.

[41] A. Hooper, H. Zebker, P. Segall, and B. Kampes, "A new method for measuring deformation on volcanoes and other natural terrains using InSAR persistent scatterers," *Geophys. Res. Lett.*, vol. 31, no. 23, pp. 1–5, Dec. 2004, doi: 10.1029/2004GL021737.

[42] F. Alshehri and A. Mohamed, "Analysis of groundwater storage fluctuations using GRACE and remote sensing data in wadi As-Sirhan, northern Saudi Arabia," *Water*, vol. 15, no. 2, p. 282, Jan. 2023, doi: 10.3390/w15020282.

[43] T. H. Syed, J. S. Famiglietti, M. Rodell, J. Chen, and C. R. Wilson, "Analysis of terrestrial water storage changes from GRACE and GLDAS," *Water Resour. Res.*, vol. 44, no. 2, pp. 1–15, Feb. 2008, doi: 10.1029/2006WR005779.

[44] D. Jiang, J. Wang, Y. Huang, K. Zhou, X. Ding, and J. Fu, "The review of GRACE data applications in terrestrial hydrology monitoring," *Adv. Meteorol.*, vol. 2014, pp. 1–9, Jan. 2014, doi: 10.1155/2014/725131.

[45] S. Ali et al., "Improving the resolution of GRACE data for spatio-temporal groundwater storage assessment," *Remote Sens.*, vol. 13, no. 17, p. 3513, Sep. 2021, doi: 10.3390/rs13173513.

[46] M. Rahaman, B. Thakur, A. Kalra, and S. Ahmad, "Modeling of GRACE-derived groundwater information in the Colorado river basin," *Hydrology*, vol. 6, no. 1, p. 19, Feb. 2019, doi: 10.3390/hydrology6010019.

[47] W. Hakim, A. Achmad, and C.-W. Lee, "Land subsidence susceptibility mapping in Jakarta using functional and meta-ensemble machine learning algorithm based on time-series InSAR data," *Remote Sens.*, vol. 12, no. 21, p. 3627, Nov. 2020, doi: 10.3390/rs12213627.

[48] B. Ranjgar, S. V. Razavi-Termeh, F. Foroughnia, A. Sadeghi-Niaraki, and D. Perissin, "Land subsidence susceptibility mapping using persistent scatterer SAR interferometry technique and optimized hybrid machine learning algorithms," *Remote Sens.*, vol. 13, no. 7, p. 1326, Mar. 2021, doi: 10.3390/rs13071326.

[49] H. Gharechaee et al., "Land subsidence susceptibility mapping using interferometric synthetic aperture radar (InSAR) and machine learning models in a semiarid region of Iran," *Land*, vol. 12, no. 4, p. 843, 2023, doi: 10.3390/land12040843.

[50] K. Khosravi, F. Rezaie, J. R. Cooper, Z. Kalantari, S. Abolfathi, and J. Hatamiafkoueieh, "Soil water erosion susceptibility assessment using deep learning algorithms," *J. Hydrol.*, vol. 618, Mar. 2023, Art. no. 129229, doi: 10.1016/j.jhydrol.2023.129229.

[51] C. Pu et al., "Characterizing the topographic changes and land subsidence associated with the mountain excavation and city construction on the Chinese Loess Plateau," *Remote Sens.*, vol. 13, no. 8, p. 1556, Apr. 2021, doi: 10.3390/rs13081556.

[52] H. R. Pourghasemi and M. M. Saravi, "Land-subsidence spatial modeling using the random forest data-mining technique," in *Spatial Modeling in GIS and R for Earth and Environmental Sciences*. Amsterdam, The Netherlands: Elsevier, 2019, pp. 147–159.

[53] Y. Zheng, L. Tang, and H. Wang, "An improved approach for monitoring urban built-up areas by combining NPP-VIIRS nighttime light, NDVI, NDWI, and NDBI," *J. Cleaner Prod.*, vol. 328, Dec. 2021, Art. no. 129488, doi: 10.1016/j.jclepro.2021.129488.

[54] B. Chen et al., "Characterization and causes of land subsidence in Beijing, China," *Int. J. Remote Sens.*, vol. 38, no. 3, pp. 808–826, Feb. 2017, doi: 10.1080/01431161.2016.1259674.

[55] A. Arabameri et al., "Application of novel ensemble models and k-fold CV approaches for land subsidence susceptibility modelling," *Stochastic Environ. Res. Risk Assessment*, vol. 36, no. 1, pp. 201–223, May 2021, doi: 10.1007/s00477-021-02036-7.

[56] W. L. Hakim and C.-W. Lee, "A review on remote sensing and GIS applications to monitor natural disasters in Indonesia," *Korean J. Remote Sens.*, vol. 36, pp. 1303–1322, May 2020, doi: 10.7780/kjrs.2020.36.6.1.3.

[57] S. Samanta, D. K. Pal, and B. Palsamanta, "Flood susceptibility analysis through remote sensing, GIS and frequency ratio model," *Appl. Water Sci.*, vol. 8, no. 2, pp. 1–14, May 2018, doi: 10.1007/s13201-018-0710-1.

[58] B. Pradhan, M. H. Abokharima, M. N. Jebur, and M. S. Tehrany, "Land subsidence susceptibility mapping at Kinta Valley (Malaysia) using the evidential belief function model in GIS," *Natural Hazards*, vol. 73, no. 2, pp. 1019–1042, Sep. 2014, doi: 10.1007/s11069-014-1128-1.

[59] A. S. Nur, Y. J. Kim, and C.-W. Lee, "Creation of wildfire susceptibility maps in Plumas national forest using InSAR coherence, deep learning, and metaheuristic optimization approaches," *Remote Sens.*, vol. 14, no. 17, p. 4416, Sep. 2022, doi: 10.3390/rs14174416.

[60] A. E. M. Al-Juaidi, A. M. Nassar, and O. E. M. Al-Juaidi, "Evaluation of flood susceptibility mapping using logistic regression and GIS conditioning factors," *Arabian J. Geosci.*, vol. 11, no. 24, Dec. 2018, doi: 10.1007/s12517-018-4095-0.

[61] A. Arabameri, B. Pradhan, K. Rezaei, S. Lee, and M. Sohrabi, "An ensemble model for landslide susceptibility mapping in a forested area," *Geocarto Int.*, vol. 35, no. 15, pp. 1680–1705, 2019, doi: 10.1080/10106049.2019.1585484.

[62] S. Albawi, T. A. Mohammed, and S. Al-Zawi, "Understanding of a convolutional neural network," in *Proc. Int. Conf. Eng. Technol. (ICET)*, Aug. 2017, pp. 1–6, doi: 10.1109/ICEngTechnol.2017.8308186.

[63] R. Chauhan, K. K. Ghanshala, and R. C. Joshi, "Convolutional neural network (CNN) for image detection and recognition," in *Proc. 1st Int. Conf. Secure Cyber Comput. Commun. (ICSCCC)*, Dec. 2018, pp. 278–282, doi: 10.1109/ICSCCC.2018.8703316.

[64] M. I. Sameen, B. Pradhan, and S. Lee, "Application of convolutional neural networks featuring Bayesian optimization for landslide susceptibility assessment," *CATENA*, vol. 186, Mar. 2020, Art. no. 104249, doi: 10.1016/j.catena.2019.104249.

[65] W. L. Hakim, A. S. Nur, F. Rezaie, M. Panahi, C.-W. Lee, and S. Lee, "Convolutional neural network and long short-term memory algorithms for groundwater potential mapping in Anseong, South Korea," *J. Hydrol., Regional Stud.*, vol. 39, Feb. 2022, Art. no. 100990, doi: 10.1016/j.ejrh.2022.100990.

[66] G. Zhang, M. Wang, and K. Liu, "Forest fire susceptibility modeling using a convolutional neural network for Yunnan Province of China," *Int. J. Disaster Risk Sci.*, vol. 10, no. 3, pp. 386–403, Sep. 2019, doi: 10.1007/s13753-019-00233-1.

[67] W. L. Hakim et al., "Convolutional neural network (CNN) with metaheuristic optimization algorithms for landslide susceptibility mapping in Icheon, South Korea," *J. Environ. Manag.*, vol. 305, Mar. 2022, Art. no. 114367, doi: 10.1016/j.jenvman.2021.114367.

[68] K. Khosravi et al., "Convolutional neural network approach for spatial prediction of flood hazard at national scale of Iran," *J. Hydrol.*, vol. 591, Dec. 2020, Art. no. 125552, doi: 10.1016/j.jhydrol.2020.125552.

[69] Y. Wang, Z. Fang, and H. Hong, "Comparison of convolutional neural networks for landslide susceptibility mapping in Yanshan County, China," *Sci. Total Environ.*, vol. 666, pp. 975–993, May 2019, doi: 10.1016/j.scitotenv.2019.02.263.

[70] S. Mirjalili, S. M. Mirjalili, and A. Lewis, "Grey wolf optimizer," *Adv. Eng. Softw.*, vol. 69, pp. 46–61, Mar. 2014, doi: 10.1016/j.advengsoft.2013.12.007.

[71] H. Faris, I. Aljarah, M. A. Al-Betar, and S. Mirjalili, "Grey wolf optimizer: A review of recent variants and applications," *Neural Comput. Appl.*, vol. 30, no. 2, pp. 413–435, Jul. 2018, doi: 10.1007/s00521-017-3272-5.

[72] A.-L. Balogun et al., "Spatial prediction of landslide susceptibility in western Serbia using hybrid support vector regression (SVR) with GWO, BAT and COA algorithms," *Geosci. Frontiers*, vol. 12, no. 3, May 2021, Art. no. 101104, doi: 10.1016/j.gsf.2020.10.009.

[73] S. Nosratabadi, K. Szell, B. Beszedes, F. Imre, S. Ardabili, and A. Mosavi, "Comparative analysis of ANN-ICA and ANN-GWO for crop yield prediction," in *Proc. RIVF Int. Conf. Comput. Commun. Technol. (RIVF)*, Oct. 2020, pp. 1–5, doi: 10.1109/RIVF48685.2020.9140786.

[74] M. Panda and B. Das, "Grey wolf optimizer and its applications: A survey," in *Proc. 3rd Int. Conf. Adv. Informat. Comput. Res.*, New York, NY, USA, 2019, pp. 179–194.

[75] M. Pradhan, P. K. Roy, and T. Pal, "Oppositional based grey wolf optimization algorithm for economic dispatch problem of power system," *Ain Shams Eng. J.*, vol. 9, no. 4, pp. 2015–2025, Dec. 2018, doi: 10.1016/j.asej.2016.08.023.

[76] E. Atashpaz-Gargari and C. Lucas, "Imperialist competitive algorithm: An algorithm for optimization inspired by imperialistic competition," in *Proc. IEEE Congr. Evol. Comput.*, Sep. 2007, pp. 4661–4667, doi: 10.1109/CEC.2007.4425083.

[77] M. Abdollahi, A. Isazadeh, and D. Abdollahi, "Imperialist competitive algorithm for solving systems of nonlinear equations," *Comput. Math. with Appl.*, vol. 65, no. 12, pp. 1894–1908, Aug. 2013, doi: 10.1016/j.camwa.2013.04.018.

[78] D. Tien Bui et al., "New ensemble models for shallow landslide susceptibility modeling in a semi-arid watershed," *Forests*, vol. 10, no. 9, p. 743, Aug. 2019, doi: 10.3390/f10090743.

[79] M. Panahi, A. Gayen, H. R. Pourghasemi, F. Rezaie, and S. Lee, "Spatial prediction of landslide susceptibility using hybrid support vector regression (SVR) and the adaptive neuro-fuzzy inference system (ANFIS) with various metaheuristic algorithms," *Sci. Total Environ.*, vol. 741, Nov. 2020, Art. no. 139937, doi: 10.1016/j.scitotenv.2020.139937.

[80] T. Fawcett, "An introduction to ROC analysis," *Pattern Recognit. Lett.*, vol. 27, no. 8, pp. 861–874, Jun. 2006, doi: 10.1016/j.patrec.2005.10.010.

[81] M. F. Fadhillah et al., "Surface deformation simulation for InSAR detection using a machine learning approach on the Hantangang River volcanic field: A case study on the Orisan Mountain," *Frontiers Environ. Sci.*, vol. 10, Aug. 2022, Art. no. 968120, doi: 10.3389/fenvs.2022.968120.

[82] A. Linden, "Measuring diagnostic and predictive accuracy in disease management: An introduction to receiver operating characteristic (ROC) analysis," *J. Eval. Clin. Pract.*, vol. 12, no. 2, pp. 132–139, Apr. 2006, doi: 10.1111/j.1365-2753.2005.00598.x.

[83] Q. H. Nguyen et al., "Influence of data splitting on performance of machine learning models in prediction of shear strength of soil," *Math. Problems Eng.*, vol. 2021, pp. 1–15, Feb. 2021, doi: 10.1155/2021/4832864.

[84] M. Panahi et al., "Spatial modeling of radon potential mapping using deep learning algorithms," *Geocarto Int.*, vol. 37, no. 25, pp. 9560–9582, 2021, doi: 10.1080/10106049.2021.2022011.

[85] M. Masykur, "Analysis of accuracy the InaCORS BIG online postprocessing service," *Appl. Geomatics*, vol. 13, no. 2, pp. 227–233, Jun. 2021, doi: 10.1007/s12518-020-00343-2.

[86] D. O. V. Kotchoni, J.-M. Vouillamoz, F. M. A. Lawson, P. Adjomayi, M. Boukari, and R. G. Taylor, "Relationships between rainfall and groundwater recharge in seasonally humid Benin: A comparative analysis of long-term hydrographs in sedimentary and crystalline aquifers," *Hydrogeol. J.*, vol. 27, no. 2, pp. 447–457, Mar. 2019, doi: 10.1007/s10040-018-1806-2.

[87] D. Sarah, E. Soebowo, and N. A. Satriyo, "Review of the land subsidence hazard in Pekalongan Delta, Central Java: Insights from the subsurface," *Rudarsko-Geolosko-Naftni Zbornik*, vol. 36, no. 4, pp. 163–176, 2021, doi: 10.17794/rgn.2021.4.13.

[88] Z. Ghorbani, A. Khosravi, Y. Maghsoudi, F. F. Mojtahedi, E. Javadnia, and A. Nazari, "Use of InSAR data for measuring land subsidence induced by groundwater withdrawal and climate change in Ardabil Plain, Iran," *Sci. Rep.*, vol. 12, no. 1, pp. 1–22, Aug. 2022, doi: 10.1038/s41598-022-17438-y.

[89] M. Béjar-Pizarro et al., "Mapping groundwater level and aquifer storage variations from InSAR measurements in the Madrid aquifer, Central Spain," *J. Hydrol.*, vol. 547, pp. 678–689, Apr. 2017, doi: 10.1016/j.jhydrol.2017.02.011.

[90] Y. Zhao et al., "Analysis of the spatial and temporal evolution of land subsidence in Wuhan, China from 2017 to 2021," *Remote Sens.*, vol. 14, no. 13, p. 3142, Jun. 2022, doi: 10.3390/rs14133142.

[91] H.-F. Yeh, C.-H. Lee, J.-F. Chen, and W.-P. Chen, "Estimation of groundwater recharge using water balance model," *Water Resour.*, vol. 34, no. 2, pp. 153–162, Apr. 2007, doi: 10.1134/S0097807807020054.

[92] D. Han and G. Cao, "Phase difference between groundwater storage changes and groundwater level fluctuations due to compaction of an aquifer-aquitard system," *J. Hydrol.*, vol. 566, pp. 89–98, Nov. 2018, doi: 10.1016/j.jhydrol.2018.09.010.

[93] R. E. Horton, "An approach toward a physical interpretation of infiltration-capacity," *Soil Sci. Soc. Amer. J.*, vol. 5, pp. 399–417, Jan. 1941, doi: 10.2136/sssaj1941.036159950005000C0075x.

[94] M. Trabelsi, N. Meddouri, and M. Maddouri, "A new feature selection method for nominal classifier based on formal concept analysis," *Proc. Comput. Sci.*, vol. 112, pp. 186–194, Jan. 2017, doi: 10.1016/j.procs.2017.08.227.

[95] W. Chen, S. Zhang, R. Li, and H. Shahabi, "Performance evaluation of the GIS-based data mining techniques of best-first decision tree, random forest, and naïve Bayes tree for landslide susceptibility modeling," *Sci. Total Environ.*, vol. 644, pp. 1006–1018, Dec. 2018, doi: 10.1016/j.scitotenv.2018.06.389.

[96] D. Tien Bui, T. A. Tuan, H. Klempe, B. Pradhan, and I. Revhaug, "Spatial prediction models for shallow landslide hazards: A comparative assessment of the efficacy of support vector machines, artificial neural networks, kernel logistic regression, and logistic model tree," *Landslides*, vol. 13, no. 2, pp. 361–378, Apr. 2016, doi: 10.1007/s10346-015-0557-6.

Wahyu Luqmanul Hakim (Student Member, IEEE) was born in Jakarta, Indonesia, in 1997. He received the B.Sc. degree in physics science from the Indonesia University of Education, Bandung, Indonesia, in 2019. He is currently pursuing the combined M.Sc. and Ph.D. degree in integrated science with Kangwon National University, Chuncheon, South Korea.

He is currently a member of the Kangwon National University Earth Observation Laboratory (KEOL), Kangwon National University. His research interests focus on remote sensing and geographic information system (GIS) applications, including synthetic aperture radar (SAR) interferometry, time-series interferometry SAR (InSAR), GIS, and machine learning.

Muhammad Fulki Fadhillah (Student Member, IEEE) was born in Tasikmalaya, Indonesia, in 1996. He received the B.S. degree in physics from the Indonesia University of Education, Bandung, Indonesia, in 2019. He is currently pursuing the combined M.Sc. and Ph.D. degree with the Department of Science Education, Kangwon National University, Chuncheon, South Korea.

From 2017 to 2018, he was a Research Assistant with the Basic Physics Laboratory, Department of Physics Education, Indonesia University of Education. Since 2020, he has been a Research Assistant with the Kangwon National University Earth Observation Laboratory, Kangwon National University. His research interests synthetic aperture radar (SAR) interferometry, time-series interferometry SAR (InSAR), geographic information systems (GISs), and machine learning.

Kwang-Jae Lee received the M.S. degree in geodetic engineering from Kyungil University, Daegu, South Korea, in 2002, and the Ph.D. degree in civil engineering from Chungnam National University, Daejeon, South Korea, in 2015.

He has been working at the Korea Aerospace Research Institute (KARI), Daejeon, since 2002, where he is currently conducting research on the utilization of the Korea Multi-Purpose Satellite (KOMPSAT) series at the Satellite Operation & Application Center as a Principal Researcher. His research interests include image processing, target detection and change detection, and geospatial information.

Seung-Jae Lee received the M.S. and Ph.D. degrees in electrical engineering from the Pohang University of Science and Technology (POSTECH), Pohang, South Korea, in 2014 and 2018, respectively.

Since 2018, he has been a Senior Researcher with the Korea Aerospace Research Institute (KARI), Daejeon, South Korea. His research interests include synthetic aperture radar (SAR) automatic target recognition (ATR), SAR super-resolution, SAR signal processing, interferometry SAR (InSAR), and polarimetric SAR (PolSAR).

Sung-Ho Chae is currently pursuing the Ph.D. degree in geoinformatics (remote sensing) with the University of Seoul, Seoul, South Korea.

He is currently a Senior Researcher with the Korea Aerospace Research Institute, Daejeon, South Korea. Recently, he has been focusing on the development of algorithms and systems for the utilization of Korea's synthetic aperture radar (SAR) satellites, KOMPSAT-5 and 6, and the constellation of microsatellites. His research interests include improving the measurement performance of the conventional SAR offset tracking method and advanced 3-D surface deformation mapping by the integration of interferometry SAR (InSAR) and multiple aperture InSAR (MAI) measurements.

Chang-Wook Lee (Member, IEEE) received the B.S. degree in geology from Kangwon National University, Chuncheon, South Korea, in 2000, and the M.S. and Ph.D. degrees in earth system sciences (Remote Sensing) from Yonsei University, Seoul, South Korea, in 2002 and 2009, respectively.

He had filled the post-doctoral position in interferometry SAR (InSAR) for the Arctic Slope Regional Corporation (ASRC) Research and Technology Solutions (ARTS) contract with the U.S. Geological Survey EROS Data Center, Sioux Falls, SD, USA, and all work was to be performed as determined by the United States Geological Survey (USGS) Project Manager working in coordination with Science Support Services Contract (SSSC) Management. The work was performed at the USGS Cascades Volcano Observatory, Vancouver, WA, USA, by National Aeronautics and Space Administration (NASA) Project supports, from 2009 to 2011. He is currently an Associate Professor with the Department of Science of Education, Kangwon National University. His research interests include synthetic aperture radar (SAR), interferometry SAR (InSAR), and time-series processing technique development on natural disaster monitoring and resource characterization. He has authored more than 100 articles in these research fields.