Rapid Extreme Tropical Precipitation and Flood Inundation Mapping (Flood-Tropical) Framework: Initial Testing for the 2021-2022 Malaysia Flood

Yi Lin Tew 1, Mou Leong Tan 1,*, Liew Juneng 2, Kwok Pan Chun 3, Mohamad Hafiz bin Hassan 4, Sazali bin Osman 4, Narimah Samat 1, Chun Kiat Chang 5, Muhammad Humayun Kabir 6

1 GeoInformatic Unit, Geography Section, School of Humanities, Universiti Sains Malaysia, 11800 USM, Pulau Pinang, Malaysia; tevyilin@student.usm.my, mouleong@usm.my, narimah@usm.my
2 Department of Earth Sciences and Environment, Faculty of Science and Technology, Universiti Kebangsaan Malaysia, Bangi 43600, Selangor, Malaysia; juneng@ukm.edu.my
3 Department of Geography and Environmental Management, University of the West of England, Bristol, UK; kwok.chun@uwe.ac.uk
4 National Flood Forecasting and Warning Centre, Department of Irrigation and Drainage, Malaysia. hafizhassan@water.gov.my; sazali@water.gov.my
5 River Engineering and Urban Drainage Research Centre (REDAC), Universiti Sains Malaysia, Engineering Campus, Nibong Tebal 14300, Penang, Malaysia. redac10@usm.my
6 Department of Agricultural Extension and Information System, Sher-e-Bangla Agricultural University, Dhaka, Bangladesh; mhumayunsau@gmail.com
* Correspondence: mouleong@usm.my

Abstract: The 2021-2022 flood is one of the most serious flood events in Malaysian history, with approximately 70,000 victims evacuated daily, 54 killed and total losses were up to MYR 6.1 billion. From this devastating event, we realized the lack of extreme precipitation and flood inundation information, which is a common problem in tropical regions. Therefore, we have developed Flood-Tropical framework to provide a rapid extreme precipitation information and flood inundation mapping by utilizing (1) a cloud-computing platform, the Google Earth Engine (GEE); (2) open-sources satellite images such as Global Precipitation Mission (GPM), Sentinel-1 SAR and Sentinel-2 optical; and (3) flood victim information. The preliminary results on the 2021-2022 Malaysia flood were satisfactory, as the accuracy of inundated floods was up to 70%. Overall, two precipitation peaks resulted 60000 to 70000 people mostly in Selangor and Pahang evacuated on 21 – 24 December 2021, and 10000 to 15000 people from southern Peninsular evacuated on 2 – 6 Jan 2022. Extreme daily precipitation of up to 230 mm/day was observed and resulted in an inundated area of 77.43 km² in Peninsular Malaysia. This framework can act as a useful tool for local authorities to visualize extreme precipitation and floods for rescue planning and flood management.

Keywords: flood; inundation mapping; Sentinel-1 SAR; Malaysia; Climate Change

1. Introduction

As reported in the Sixth Assessment Report of the Intergovernmental Panel Climate Change [1], the number and intensity of extreme precipitation have increased significantly in the past few decades. Based on the Emergency Event International Disaster Database [2], a total of 5608 extreme flood events have been recorded around the world from 1906 to 2021, causing about 7 million deaths, 3.9 billion people affected and US$ 953 billion total damage losses. The EM-DAT also reported the increases of flood events over the years, particularly in the past two decades. The actual flood losses may be higher as not all flood events were recorded in the EM-DAT database and the costs of damage from floods are difficult to measure. Flood inundation mapping is therefore important to effec-
tively organize flood rescue operations, monitoring, management, mitigation and modeling for achieving rapid and effective recovery. In addition, flooded area information could also be used to measure flood damage in cities and agricultural industries.

In Malaysia, major floods are normally found during the early phase of the Northeast Monsoon (NEM) from November to January, i.e. 1971, 2006-2007, 2014-2015, 2017 and 2018 [3]. Unexpected extreme precipitation events from mid-December 2021 to early 2022 have also caused devastating floods in seven states of Malaysia, which killed 54 people, affected more than 125,000 and nearly 70,000 evacuated on a single day and resulted in losses up to MYR 6.1 billion or USD 1.46 billion USD [4-6]. Surprisingly, the extreme events hit not only the east coast region of Peninsular Malaysia, but also the west coast Peninsular Malaysia that typically experienced a dry condition during this period where it was noticeable from the direction of the 29W depression [7]. The water levels of many rivers and reservoirs in Selangor have surpassed their dangerous level. More than 120 landslides have occurred a few days after the extreme rains, causing the rescue operations to become more difficult [8]. For instance, the floods and landslides have caused the Kuala Lumpur-Karak and the East Coast Expressway phase 1 (LPT1) that connected the east coast, and the west coast became impassable [9].

The existing flood maps available in Malaysia were categorized into three types, namely flood inundation maps, flood hazard maps and flood risk maps [10]. The common practice of producing flood inundation maps in Malaysia were through coupled digital elevation model (DEM) and flood event records [11], or hydrologic and hydraulic modeling such as the InfoWorks Integrated Catchment Modeling (ICM) [12]. The production of flood maps via such methods requires longer time for data collection and model calibration. Despite the importance of near real-time precipitation and flood inundation information, there is still a lack of a comprehensive framework to provide a rapid precipitation-flood information during extreme precipitation in tropical regions.

Satellite technologies provide a cost-effective and timely resource to capture the precipitation and flood information in a large area during the flood events. Satellite precipitation products (SPPs) have emerged as a major source to monitoring global precipitation in the past few decades [13, 14]. The Global Precipitation Measurement (GPM), extended from the Tropical Rainfall Measuring Mission (TRMM) [15], is one of the most accurate and finest resolution SPPs to estimate precipitation in tropical regions [16]. The Integrated Multi-Satellite Retrievals for GPM (IMERG) products are able to provide near real-time precipitation data for the regions from 65°N to 65°S at a 30 min time-scale. Tapiador et al. (2021) have utilized IMERG to study the hydro-meteorological characteristic of the September 2019 floods in Spain and concluded that the IMERG compares well with observations [17]. In China, Qi et al. (2021) found the IMERG late run (IMERG-Late) product performs well in monitoring the extreme heavy precipitation of the super typhoon LeBron [18]. In Malaysia, Tan and Santo [19] reported the IMERG products show the lowest bias in capturing precipitation in the 2014-2015 flood events of Malaysia relative to other SPPs. Therefore, the IMERG near real-time products can provide useful information to monitor flood precipitation across the globe.

The European Space Agency’s Sentinel-1 Synthetic Aperture Radar (SAR) and Sentinel-2 optical multispectral satellites are becoming popular data sources for effective flood monitoring [20]. Flood mapping using the Sentinel data are mostly conducted in northern temperate latitudes such as Europe, the UK and Canada [21]. In tropical areas, cloud cover is a prevalent issue in optical satellite imagery [22]. Hence, the Sentinel-1 SAR is a preferred satellite sensor for tropical flood monitoring due to cloud-free and allows users to extract time-critical disaster images from the ESA Sentinel Data Hub within 45 min of data capturing [23]. Based on the literature available, the common methodology of analyzing flood via Sentinel-1 data are through classification [24] and thresholding [25]. Thresholding is the most common SAR-based technique for flood detection but is only perform well in simple flood situation and homogenous land surfaces [26] as it minimized classification error between water and land features without differentiating different land classes [25]. However, identifying an optimal threshold in the binarization and feature elevation
Google Earth Engine (GEE) is a free cloud-based computing platform that enables users to process massive geospatial datasets rapidly using high-performance computing resources [29]. Utilization of GEE in processing Sentinel datasets can achieve a rapid flood inundation mapping [26, 30]. Therefore, the main aim of this study is to develop a framework, called Flood-Tropical, for rapid precipitation information extraction and flood inundation mapping using the latest satellite technologies and GEE. The Flood-Tropical framework is applied to study the characteristics of precipitation and floods for the Dec 2021-Jan 2022 unexpected precipitation events in Peninsular Malaysia. Three specific objectives of the study are: (1) to investigate the spatio-temporal characteristics of daily precipitation over Peninsular Malaysia using the GPM IMERG product; and (2) to identify an optimal threshold for a Sentinel-based flood inundation mapping in Malaysia. This study helps local authorities to identify flooded areas, which are vital for the flood rescue and management in this region.

2. Materials and Methods

2.1. The Malaysia 2021-2022 Flood Event

The 2021-2022 flood was among the most serious floods recorded in Malaysia’s history, which involved eight states in the Peninsular: Perak, Selangor, Kuala Lumpur, Negeri Sembilan, Melaka, Kelantan, Terengganu, and Pahang [31]. Hence, the present study focused on the flooded areas of these largely affected states during the 2021-2022 flood event (Figure 1). Peninsular Malaysia, which also known as West Malaysia, is located at the Southeast Asia that shared land border with Thailand at the north and the sea border with Singapore at the south. Peninsular Malaysia accounted majority of Malaysia’s population and economy, with a total land area of 132.265 km², and total population of 25.9 million (as of year 2021) [32].

Situated within the equatorial region and exerts the tropical climate with low atmospheric pressure. The climate of Malaysia is uniform with three major characteristics: constant temperature, high humidity, and abundant rainfall throughout the year [35]. The country is situated in a strategic location where it is free from natural catastrophes, however, there are two major hydro-climatic related disaster affecting the livelihood: too much of rainfall that caused floods and water shortage during the dry period [34]. Among the two hydro-climatic related disasters, flood is the more devastating natural disaster for its duration and frequency of occurrence, extent of affected area and the impact towards the socioeconomic development [36, 37].

According to the Department of Irrigation and Drainage (DID) of Malaysia, no formal categorization was made to distinguish floods in Malaysia, but generally floods in Malaysia can be categorized as monsoonal, flash, or tidal floods [38], where the differentiation between the monsoonal and flash floods depend on the period of the flood water to recede. Flash floods are sudden caused by unpredictable heavy rainfall whereas monsoonal floods occurred during the monsoon seasons [39]. The Dec 2021 flood can be categorized as a monsoonal flood, that caused by the Northeast monsoon. In fact, monsoonal floods are an regular annual natural disaster in Malaysia, with major events occurred in the years of 1926, 1963, 1965, 1967, 1969, 1971, 1973, 1979, 1983, 1993, 1998, 2005, 2006, 2007, 2010, 2014, 2017, and 2021 [39, 40]. The Northwest monsoonal floods happened for the past decade in Malaysia was summarized in Table 1.

<table>
<thead>
<tr>
<th>Year</th>
<th>Date</th>
<th>States Involved</th>
<th>Outcome</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>2010</td>
<td>10 October – 19</td>
<td>Kedah, Perlis</td>
<td>Triggered by tropical depression and later aggravated by La Nina monsoon</td>
<td>[37, 39-41]</td>
</tr>
<tr>
<td></td>
<td>November</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
45,000 hectares of rice fields were damaged. Government pledged USD $6.5 billion to help the farmers.

<table>
<thead>
<tr>
<th>Year</th>
<th>Date</th>
<th>Location</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>2014</td>
<td>15 December</td>
<td>Johor, Kelantan, Kedah, Negeri Sembilan, Pahang, Perak, Perlis, Sabah, Sarawak, Selangor, Terengganu</td>
<td>Heavy rainfall as part of the north-east monsoon. The worst flood in Kelantan in history with 202,000 individuals evacuated. Property damage of USD $560 million.</td>
</tr>
<tr>
<td>2017</td>
<td>4 - 5 November</td>
<td>Pulau Pinang</td>
<td>Caused by the Tropical Depression 29W on 3 November. Flash flood in Pulau Pinang with maximum flood level 3.7 meters.</td>
</tr>
<tr>
<td>2021</td>
<td>16 - 31 December</td>
<td>Perak, Selangor, Kuala Lumpur, Negeri Sembilan, Melaka, Johor, Pahang, Terengganu, Kelantan, Sabah</td>
<td>Caused by the Tropical Depression 29W on 14 - 17 December. Heavy flood in four states and minor flood in four other states. Government estimates total of USD $1.55 billion of property damage.</td>
</tr>
</tbody>
</table>

National Disaster Management Agency (NADMA) reported that at least 89,723 victims were affected by the flood in December 2021 [43]. For example, the number of flood victims evacuated from their houses has been increased significantly to about 70,000 people per day for three consecutive days since 21 Dec 2021 (Figure 1c). The second peak of flood victim evacuation can be found from 3 to 5 Jan 2022 due to the second wave of flooding in the state of Johor. As the flood affected areas are getting wider, thus a rapid flood mapping is much needed for the disaster management in Malaysia.
2.2. Global Precipitation Mission (GPM)

The Global Precipitation Measurement (GPM) mission is a global rainfall and snowfall observation at 3 hours temporal resolution [16]. The mission was launched by the National Aeronautics and Space Administration (NASA) and the Japan Aerospace Exploration Agency (JAXA) via the Core Observatory (GPM CO) satellite on 28 February 2014, where it has designed a dual frequency precipitation radar for precise rainfall measurements. With the sophisticated satellite instrumentation and integrated user applications, the public release of the precipitation product can be required in near-real time at 1-5 hours post processing of downlinked observations to ground stations [44].

The Integrated Multi-Satellite Retrievals for GPM (IMERG) is a US GPM Science Team precipitation product that applied inter-calibrated estimates over various international constellation of precipitation satellites as well as the monthly surface precipitation gauge analyses to compute higher temporal (half-hourly) and spatial (0.1˚X0.1˚) resolution precipitation product [45, 46]. These characteristics of IMERG precipitation products are advantageous in precipitation extreme studies globally [21].

Basically, IMERG data can be divided into the early run, late run and final products, which are available ~4 hours, ~12 hours and ~2 months after capturing the earth surface. The IMERG late run product downloaded from the NASA Giovanni online tool (https://giovanni.gsfc.nasa.gov/giovanni/) was used to analyze the spatio-temporal changes of the daily precipitation for Peninsular Malaysia from 15 December 2021 to 7 January 2022. Basically, the quality of IMERG late run product is slightly better than the early run version due to the calibration scheme in the algorithm, so it is more appropriate for the flood precipitation analysis. This precipitation analysis is done to identify the volume of the daily precipitation that has led to the flood in each state of Peninsular Malaysia.

2.3. Sentinel Satellites

The Sentinel programme is a mission under the European Space Agency [47]. The Sentinel mission operates based on radar and super-spectral imaging for land, ocean, and atmospheric monitoring. The data processor included the Shuttle Radar Topography Mission (SRTM) 1-arc second data for terrain and radiometric corrections, that increases the ease and readiness of Sentinel data in analyzation. The strengths of the data are the ability for it to produce global, continuous, and wide coverage satellite imaging products.
Comprising a constellation of two polar orbiting satellites, the Sentinel-1 mission operates day and night, capturing C-band SAR imagery in all weathers. The C-band imaging operating in Sentinel-1 mission captures the earth with coverage up to 400km and spatial resolution of 10m. The foremost advantage of utilizing the Sentinel-1 SAR data in flood mapping is that the sensor can capture images at wavelengths beyond the cloud cover and thick moist vegetation covers, regardless of day and night-time [48]. The dataset available in the GEE platform that matches with the flood incident in each state for this study are listed in Table 2.

<table>
<thead>
<tr>
<th>State</th>
<th>Flood Date</th>
<th>Available Datasets</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pahang</td>
<td>17 December – 31 December</td>
<td>868</td>
</tr>
<tr>
<td>Selangor</td>
<td>17 December – 31 December</td>
<td>534</td>
</tr>
<tr>
<td>Negeri Sembilan</td>
<td>18 December – 31 December</td>
<td>204</td>
</tr>
<tr>
<td>Melaka</td>
<td>17 December – 25 December</td>
<td>199</td>
</tr>
<tr>
<td>Johor</td>
<td>05 December – 31 December</td>
<td>708</td>
</tr>
<tr>
<td>Terengganu</td>
<td>02 December – 31 December</td>
<td>651</td>
</tr>
<tr>
<td>Kelantan</td>
<td>17 December – 31 December</td>
<td>589</td>
</tr>
<tr>
<td>Perak</td>
<td>19 December – 21 December</td>
<td>852</td>
</tr>
</tbody>
</table>

The Level-1 Ground Range Detected (GRD) product level data were utilized for the water surface detection under the Interferometric Wide Swath IW scanning mode, where the phase information was not included. The acquired Sentinel-1 images are 10m x 10m spatial resolution, with polarization configuration types Vertical-Horizontal (VH) and Vertical-Vertical (VV) [49]. The Level-1 GRD products is a package of dataset that comprises focused SAR data that have been detected, multi-looked and projected to the ground range Earth ellipsoid model [23].

Backscattering of the GRD products has been widely used in land cover studies and water body monitoring [50]. Theoretically, VH has a stronger return over areas with volume scattering, while VV has a stronger return over specular scattering. Thus, both polarization configurations are used to enhance the band information [51]. SAR imageries are often used to map calm and open water surfaces because water surface will be recorded as reflected specular incident microwave radiation [52].

2.4. RETRACE

The author would like to present the Rapid Extreme Tropical Precipitation and Flood Inundation Mapping (RETRACE) framework in this study, where earth observation data is used to map the inundation extent of the Malaysia 2021-2022 flood event via automated thresholding method in the GEE platform, where the results are accessible by the public. Earth observation satellite data provide a precise and cost-effective tool for assessing and monitoring hydrological dynamics of the earth’s surface. As such, multispectral and synthetic aperture radar (SAR) satellite imageries are advantageous for precipitation monitoring and flood mapping due to their high spatial and temporal resolution characteristics [53]. In the present study, two satellite data are used to map the flood inundation area throughout the study area: Sentinel-1 SAR image and the Sentinel-2 MSI image, where the data are freely accessible via the Google Earth Engine (GEE). Meanwhile, the Integrated Multi-Satellite Retrievals for GPM (IMERG) was used to study the spatial and temporal changes of daily precipitation over Malaysia from 17 Dec 2021 to 10 Jan 2022 as the explanatory variable to further illustrate the relationship between the climate induced surface water and flooding extent [54].

Besides Sentinel data, the Global Administrative Unit Layers 2015 (GAUL), JRC Global Surface Water Mapping Layers, and the WWF HydroSHEDS Void-Filled DEM
data in the GEE Data Catalog are also utilized in this study. GAUL is a reliable global first level administrative boundary layer made available by the United Nation (UN) Food and Agriculture Organization (FAO), disseminated based on the year 2015 global data [55]. The Global Surface Water Mapping Layers developed by the Joint Research Centre (JRC) of European Commission (EC) is a global surface water map generated with 4,453,989 Landsat 5, 7, and 8 scenes acquired on 16 March 1984 to 31 December 2020 [56]. Lastly, the HydroSHEDS developed by the World Wildlife Fund (WWF) is an elevation hydrographic mapping product at global scale generated by the Shuttle Radar Topography Mission (SRTM) at a spatial resolution of 3 arc-second [57].

2.5. SAR-based Flood Inundation Mapping

The general framework for rapid precipitation analysis and flood inundation mapping specifically designed for tropical regions, particularly Malaysia, is illustrated in Figure 2. The Sentinel-1 SAR GRD data is used to generate the flood inundation areas, while the Sentinel-2 MSI data is used to collect the ground truth flooded area for the validation purpose. The GPM IMERG late-run data is used to perform rainfall analysis throughout the Peninsular Malaysia during the flood event. The workflow of the processing can be explained in four major steps: (1) pre-processing of the Sentinel-1 and 2 data including subset of area of interest, mosaicking, geometric correction, noise removal, etc; (2) identification of seriously affected areas based on the GPM IMERG precipitation data and flood victim evacuation data; (3) Sentinal-1 SAR flood inundation maps generation and observed floods data collection from both sky (Sentinel-2 optical) and ground (site visit); and (4) validation of the flood inundation maps with the observed data.

Figure 2 General framework of rapid flood inundation mapping with Sentinel SAR.
2.5.1. Pre-processing

The preprocessing of the Sentinel 1 and 2 datasets including the geometric and radiometric corrections. The Sentinel-1 GRD dataset was retrieved in the GEE platform with instrument mode and receives polarization VH, where the date range is set on the driest month of the state for the before flood period and the flood dates as stated in Table 1 for the during flood period.

Then, the dataset is subset with the FAO GAUL data to the extent of each state in Peninsular Malaysia. Radiometric calibration and speckle filtering is conducted at this point as it is essential to obtain the backscatter values [58]. The radiometrically calibrated images were computed with radar backscattering coefficient ($\sigma_v$) associated with SAR image brightness ($\beta_v$) with the formula:

$$\sigma_v = \beta_v \sin \alpha$$  \hspace{1cm} (1)

where $\alpha$ is the local incidence angle [28, 59]. The return scattering coefficient, $\sigma_v$ in decibels [59]:

$$\sigma_v (dB) = 10 \log_{10}(\sigma_v)$$  \hspace{1cm} (2)

With this, the radar backscattering values can be evaluated by Laur et al. (2004) [59]:

$$\sigma_v (dB) = \beta_v (dB) + 10 \log_{10}[\sin(i_p)/\sin(i_c)]$$  \hspace{1cm} (3)

$$\beta_v (dB) = 20 \log_{10}(DN) - K(dB)$$  \hspace{1cm} (4)

where

'${i_p}'$ denotes the pixel’s angle of incidence

'${i_c}'$ denotes the image center’s angle of incidence

'K' denotes the calibration constant of SAR image

Then, smoothing technique is performed to filter out the speckle by eliminating the granular noise that would increase the clarity of the image through application of a speckle filtering window [60]. The process of improving the texture of the SAR image requires computation of Haralick features [61], thus we tested different values to observe the dissimilarity. However, the dissimilarity was not presented, thus the 5x5 window is adopted in this study which is optimal to optimize the tradeoff between computational time, robustness to outliers, and edge preservation [62]. Therefore, the speckle filtering smoothing radius is set at 25 for this study.

2.5.2. Thresholding

One of the most efficient and simple approach for image binarization is through thresholding [63]. In this study, the thresholding is applied via the GEE function that was built from the Otsu’s method for image segmentation [64]. Otsu’s method is the most widely applied thresholding approach, where it detects the optimum threshold automatically based on the observed distribution of pixel values [65]. In the case of this study, the concept threshold is set in the image processing codes to determine the optimum threshold from the maximization of the between-class variance of the water pixels and the non-water pixels.

The JRC Global Surface Water Mapping Layers is incorporate into the image search result to identify the permanent water areas (sea, rivers, and lakes) that exist in the study area which would be classified as the flooded surface. A threshold of 1.25 is set to extract the Sentinel-1 classified waterbodies that outside the range of permanent waterbody are
considered inundated areas as we were trying to facilitate the detection of flooded surfaces to the pixels with difference 0.25 ratio [63]. Then, the WWF HydroSHEDS Void-Filled DEM is utilized to filter the inundation result where if the area has larger than 0.5% slope it will be eliminated from the final inundation result due to the extended nature of flooded area were mainly flat surfaces.

2.5.3. Change Detection

In flood mapping, the change detection process is done to analyze the difference between the images for the pre and post flood event [63]. As for this study, the processed Sentinel-1 image captured during the flood event is compared with the image captured during the driest month over the region (15-26 February 2021). The pre-flood image is set to obtain any permanent water surface over the region which this information is later being used to mask the overlapping water pixels found on the post-flood image, where the unmasked water pixels are the final inundated extent.

2.6. Accuracy Assessment

For the accuracy assessment of the inundation map produced by the Sentinel-1 images, the inundated areas generated by both Sentinel images are cross validated by using the Combine tool in the ArcMap 10.4. This process is done for the Pahang and Selangor state only due to the availability of the cloudless data of the Sentinel-2 sensor on the flood period.

The accuracy was assessed on the pixel-by-pixel basis of the “flooded” pixel (“0” for non-flooded pixels and “1” for flooded pixels [54]). However, the background class “0” is the majority in a flood mapping and it might cause overestimation in the accuracy of the correctly mapped non-flooded pixels [63], thus only the flooded pixels are considered in the accuracy assessment.

2.6.1. Sentinel-2

As for the validation purposes, flooded area ground truth points were obtained from the Sentinel-2 MSI data as it is the best-fit dataset available in terms of temporal and spatial resolution during the flood event. Sentinel-2 is a high spatial resolution optical sensor under the Global Monitoring for Environment and Security initiative by ESA, where it provides the global earth surface information with 13 multispectral bands, that covers visible, near infrared and short wave infrared [66].

The Sentinel-2 Level-1B product is used in this study by retrieving the Copernicus Sentinel-2 Surface Reflectance image collection in the GEE platform, where it is radiometric and geometric corrected [67]. The images on 15 – 31 December 2021 are retrieved for a 50% cloud-masked mosaic image over the whole Peninsular Malaysia, where a total of 123 images were composited and mosaic to form the image scene. The combination of Bands 8 (visible and near infrared), 3 (green) and 2 (blue) is selected to visualize the water content in the image to extract the flooded area.

Unfortunately, due to the limited availability of the Sentinel-2 image during the period, only Pahang and Selangor states are managed to collect the flooded area points. A total of 75 points are collected from the Sentinel-2 image based on the random sampling method onto the flooded surfaces.
2.6.2. Observed Flood

Besides the flood location collected from the Sentinel-2 optical data, we also utilized the observed flood data collected by the National Flood Forecasting and Warning Center.
(PRABN), DID. The locations of which flood occurred during the 2021-2022 flood event, mostly in the residential areas that were used for the validation purpose. The observed flood data was collected by the DID officers on-site when the flood receded. The real-time rainfall water level data published by PRABN is accessible at https://publicinfobanjir.water.gov.my/?lang=en. The data provided by PRABN is then filtered to select only the data that has the common record time as the Sentinel images, where in this study, a total of 275 matched flood record were adopted for the validation purpose over the Pahang (142 points) and Selangor (133 points) states.

2.6.3. Critical Success Index (CSI)

The CSI is used as the indicator to assess the performance of the flood inundation mapping of the SAR images as in this study we would like to assess only the correctly mapped flood pixels rather than the correctly classified non-flooded pixels [63]. CSI is a suitable performance indicator in this study as we were trying to compare the performance of different classification algorithms on the same image rather than classification of flood between different catchments nor difference of flood magnitudes [68]. The latter is computed as [69]:

$$CSI = \frac{tp}{(S2 + PRABN)}$$  (5)

Where $tp$ is the correctly classified flooded pixels; $S2$ is the number of flood points collected from the Sentinel-2 image and PRABN is the flood location that were provided by the DID.

3. Results

3.1. Flood Victim Analysis

According to the flood victim evacuation reported by NADMA [43], the victims started to be evacuated to the relief centers began on 17 December 2021 (Figure 5). The spatial distribution of the evacuated flood victims shows the victims evacuated in Pahang increases drastically from less than 500 on 17 December 2021 to more than 30000 people on 23 December 2021. Meanwhile, the number of evacuees in Selangor also rises from less than 10000 people on 19 December 2021 to more than 30000 people on 23 December 2021 due to the subsiding floods [5]. Based on the flood report, there were a total of 685 temporary relief centers set up across the Peninsular Malaysia during the flood event, with the most in Pahang (382 centers). In Malaysia, the flood victim relief centers were managed by the Department of Social Welfare [70]. The relocation of the victims to the relief centers was organized by the governmental and non-governmental agencies, based on the conditional analysis done by the authorities according to the physical factors i.e., current water level, number of victims for each relief centers, and capacity and quantity of the aids and materials could be supplied.

The statistics of the evacuated flood victims collected is used to validate the reliability of the flood inundation map generated in this study temporally and spatially. The flood victim analysis has narrowed the data search for Sentinel-1 and 2 for the flood inundation extent generation as only data on the peak of the flood event day is adopted. Flood victim analysis is an important component in flood management, as it is the first statistics to refer for analyzing the severity of the flood event spatially [71].

Referring to the Figure 4, it can be seen that the flood victims in Pahang, Terengganu and Kelantan states began to evacuate on 17 December 2021 with less than 500 victims each state. The evacuation continues to involve the neighbouring states and Pahang and Selangor reaches more than 30000 victims evacuated on 23 December 2021. This information is captured for the SAR based flood inundation mapping.
3.2. Extreme Precipitation Analysis

Extreme precipitation is the flood leading factor in the tropics [72]. The unexpected extreme precipitation of the Peninsular Malaysia from 15 December 2021 to 7 January 2022 was analyzed to identify the number of days and the volume of rainfall across the region (refer Figure 6 and Figure 7). There is a significant sprout of the precipitation beginning from 16 December 2021 over the states of Pahang, Selangor, and Kelantan. Particularly in the plot F of the Figure 5, it shows that daily precipitation in Selangor on the 17 December 2021 was quintuple of the amount on the day before. The copious amount of the rainfall for these two states continued to 21 December 2021, resulting in the increase of number of victims evacuated. Besides that, another extreme was observed on 31 December 2021 due to thunderstorms and heavy rain, but there is no flood victim evacuation reported in Selangor since the event mostly concentrated in the east coast part of Peninsular Malaysia (Figure 6).

The past precipitation trend and precipitation-related extremes in the east coast of Peninsular Malaysia (Pahang, Terengganu, and Kelantan) study done by Mayowa et al. [47] has revealed that there is increasing trend for the precipitation and precipitation-related extremes in the region for the past 40 years (1971 to 2010). In Pahang (refer plots C and G), the current study period shows that the intensified precipitation began on 15 December 2021, with the coastal part of Pahang having 50mm higher daily precipitation (plot C) than the central part of Pahang (plot G). Another notable precipitation sprout in Pahang was on 30 December 2021. This event had caused another wave of flood in the state due to the overflown water levels in the Lipis River and Serting River at the central Pahang [73].

The above mentioned second wave of flood also hit the Johor state (plot D). A precipitation extreme was recorded on 31 December 2021 to 2 January 2022, with daily precipitation above 100mm. This event caused flood in the northern Johor, leading to displacement of 1,646 victims [74]. However, Johor is the latest state to be affected by the flood, after Negeri Sembilan, Kelantan, and Terengganu during the same day [75]. Meanwhile, the precipitation extreme observed on 31 December 2021 for Kelantan (plot A) have resulted 1,129 victims to be evacuated due to the overflown in the Kelantan River.
Referring to the spatial changes of the daily precipitation in Figure 6 and the evaluation statistics in Figure 4, it can be indicated that flood events happened when the precipitation exceeded 100 mm per day. The finding was matched with the study done by Tan et al. (2015) whereby the extreme flood events on the year 2006/2007 in the Kelantan, Pahang and Johor states was due to the daily precipitation of over 100 mm [76].

Figure 6: Temporal changes of daily precipitation over Malaysia from 15 Dec 2021 to 7 Jan 2022.
3.3. Accuracy Assessment of Flood Inundation Maps

Due to the availability of cloud-free data from Sentinel-2 imageries during the flood event, accuracy assessment was done only onto the Pahang and Selangor state. The accuracy assessment for the inundation mapping of Pahang and Selangor at threshold of 3, 4, and 5 pixel inclusion for distance between inundated area and the preliminary water surface are shown in the Table 3. Both states achieves the best inundation mapping performance when a threshold of 5 is set. Therefore, the same processing approach is adopted for the other states for the inundation mapping purposes. The observed flood locations provided by PRABN mostly in the residential areas, while the flooded areas collected from Sentinel-2 are larger which fall on open space and vegetated area likes rubber estates.

Table 3 Accuracy for each threshold value for Pahang and Selangor state.

<table>
<thead>
<tr>
<th>Threshold</th>
<th>Pahang (%)</th>
<th>Selangor (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>68.75</td>
<td>61.25</td>
</tr>
<tr>
<td>4</td>
<td>69.45</td>
<td>61.88</td>
</tr>
<tr>
<td>5</td>
<td>70.20</td>
<td>62.80</td>
</tr>
</tbody>
</table>

3.4. Flood Inundation Mapping for the Malaysia 2021-2022 Flood Event

Figure 7 Spatial changes of daily precipitation over Peninsular Malaysia from 17 Dec 2021 to 10 Jan 2022.
The analysis of the Sentinel-1 SAR data shows that the 2021-2022 flood has led to flooding of eight states at a total of 77.435 km². It can be observed that the inundated area is located at the downstream of the major rivers, especially in the low-lying areas. Kelantan state has the greatest area flooded (32.32 km²), followed by Pahang (28.87 km²), Johor (8.25 km²), and Selangor (7.24 km²). The spatial distribution of the inundated area as per each state and district are discussed in the next section. The total flooded area for each state is tabulated in Table 4 below. Based on the Sentinel-1 SAR extracted inundated area, the most flooded states during the Peninsular Malaysia 2021-2022 Flood are Pahang, Selangor, Kelantan, and Johor, as compared to the other states with inundated area of < 1 km².

For Pahang state, the inundated area is distributed around the Pahang River, with Pekan district as the most severe flooded district (16.32 km²), followed by Rompin (5.09 km²) and Kuantan (4.21 km²). The extracted inundated area for Temerloh is 0.77 km². Majority of the flooded area in Pahang are residential in low-lying region, that are situated along the riverside. However, although Pahang has the highest number of evacuees, it is not the most flooded state by the total area flooded.

As for Selangor, the most severe flooded district is the Kuala Selangor (4.28 km²), where the affected area is mostly residential and industry. The extracted inundated area for Sabak Bernam is 0.77 km², Kuala Lumpur 0.014 km², Kuala Langat 0.95 km², and Ulu Selangor 0.1 km².

In Kelantan, the inundated area is clustered at the northern part where the downstream of the Kelantan River located, with Pasir Mas as the most flooded district (11.76 km²). Kota Bharu, the capital city of Kelantan is flooded 11.29 km², followed by Pasir Puteh 4.4 km², Tumpat 1.46 km², Machang 0.73 km², and Tanah Merah 0.43 km².

The most inundated district in Johor is Segamat (2.61 km²) that is located at the southern part of Johor. Besides that, Mersing is recorded with 0.38 km², Batu Pahat 1.6 km², Kluang 1.5 km², and Johor Bahru, the capital city of Johor is flooded 0.23 km².

Table 4 Flood inundated areas in the selected states of Peninsular Malaysia in Dec 2021 – Jan 2022.

<table>
<thead>
<tr>
<th>State</th>
<th>Flooded Area (km²)</th>
<th>District</th>
<th>Flooded Area by District (km²)</th>
<th>Figure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pahang</td>
<td>28.87</td>
<td>Kuantan</td>
<td>4.21</td>
<td>8 (b)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Temerloh</td>
<td>0.77</td>
<td>8 (d)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Rompin</td>
<td>5.09</td>
<td>8 (f)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Raub</td>
<td>0.02</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Pekan</td>
<td>16.32</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Maran</td>
<td>1.14</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Lipis</td>
<td>0.02</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Jerantut</td>
<td>0.72</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Cameron</td>
<td>0.01</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Bera</td>
<td>0.54</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Bentong</td>
<td>0.03</td>
<td></td>
</tr>
<tr>
<td>Selangor</td>
<td>7.24</td>
<td>Gombak</td>
<td>0.06</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Klang</td>
<td>0.52</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Kuala Selangor</td>
<td>4.28</td>
<td>9 (c)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Petaling</td>
<td>0.12</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sabak Bernam</td>
<td>0.77</td>
<td>9 (b)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sepang</td>
<td>0.43</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Ulu Langat</td>
<td>0.01</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Ulu Selangor</td>
<td>0.1</td>
<td>9 (f)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Kuala Langat</td>
<td>0.95</td>
<td>9 (e)</td>
</tr>
<tr>
<td>Kelantan</td>
<td>32.32</td>
<td>Bachok</td>
<td>2.24</td>
<td></td>
</tr>
<tr>
<td>Location</td>
<td>Flood Inundation (m)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---------------</td>
<td>----------------------</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kota Bharu</td>
<td>11.29</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kuala Krai</td>
<td>0.01</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pasir Mas</td>
<td>11.76</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pasir Puteh</td>
<td>4.4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tanah Merah</td>
<td>0.43</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tumpat</td>
<td>1.46</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Machang</td>
<td>0.73</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kuala Lumpur</td>
<td>0.014</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Melaka</td>
<td>0.63</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jasin</td>
<td>0.27</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Melaka Tengah</td>
<td>0.1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alor Gajah</td>
<td>0.26</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Negeri Sembilan</td>
<td>0.031</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kuala Pilah</td>
<td>0.005</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Seremban</td>
<td>0.004</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tampin</td>
<td>0.01</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Perak</td>
<td>0.08</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bagan Datuk</td>
<td>0.02</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Batang Padang</td>
<td>0.01</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Muallim</td>
<td>0.05</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Figure 4** Flood inundation maps of Pahang during the Dec 2021 flood.
Figure 5 Flood inundation maps of Selangor during the Malaysia 2021-2022 flood.

Figure 6 Flood inundation maps of Kelantan during the Malaysia 2021-2022 flood.
4. Discussion

4.1. GPM Precipitation Analysis

GPM IMERG products have been applied to study flash flood and monsoonal flood around the world. Based on the precipitation changes observed by the GPM satellites, extremely heavy daily precipitation of up to 230 mm/day was recorded over Selangor during the period of 17 to 19 December 2022 (Figure 5). Please note that GPM IMERG near real time products tended to underestimate daily precipitation of Malaysia during the 2014-2015 flood by 3.09 to 24.19% [19]. Therefore, actual precipitation amount reaches ground surface during the 2021-2022 flood may be higher than the ones observed by the GPM IMERG late run product. A comprehensive validation of the GPM IMERG and other satellite precipitation products in capturing the extreme precipitation (up to hourly assessment) during flood periods over Malaysia should be conducted in the future.

4.2. The 2021-2022 Flood

During the past flood events, the Mineral and Geosciences Department had reported a total of 121 landslides occurred nationwide, following the yellow alert continuous rain by the Malaysian Meteorological Department [4]. The topography of Peninsular Malaysia is hilly in the central spine and flat towards the coastal lines, where landslides happened along the Pahang and Langat rivers believes to sedate the cause of the flood event [8]. The landslides carried debris formed debris streams across few major rivers in the Peninsular Malaysia and thus leading to the flood. In this study, we identified the extent and total area flooded for each state through analyzing Sentinel-1 SAR images. The results shown the flooded area are mostly saturated along the major river, whereby the areas are lower in terms of elevation. Kelantan had the largest area flooded among the other states, while Pahang had the greatest number of victims evacuated during the flood. The idea of mapping of flood inundation with SAR imageries can solve the challenges of producing real-time flood extent information when it is hard to access to the flooded area. In fact, many tropical countries do not own remote sensing satellites that allow them to fully operate or control the satellite for monitoring disasters in own land. This study can provide useful suggestion for the authorities and decision makers in Malaysia for the future flood hazard management with remote sensing methods. The ability of SAR images in capturing real time flood area in large scale add value in flood mapping, especially for
the remote areas [60]. Sentinel-1 SAR is a free and open-source SAR dataset for various applications on the GEE platform is benefit for the public to assess geographical issues.

4.3. Limitations of the Study

However, there are few limitations of this study should be addressed in the future with better solutions, especially in terms of data availability and processing considerations. Sentinel-1 SAR have a revisit time of 6-12 days on average, where in some case we are not able to get the data exactly on the day when flood hit the area. Thus, alternatively in this study, we stacked a period of five days before the flood and five days after the flood to get a differencing image that the flooded area can be seen. However, this might not be a practical method for precise inundation mapping as the real time flood situation was not captured. The same issue applies for Sentinel-2 MSI data and the flood region estimated on Sentinel-2 images might be overestimated as clouds are intense during flood event [77]. In addition, based on the analysis of the inundated area results, we found that the flood in the built-up area might not be as complete as what we expected. This might be due to finer information lost during the speckle filtering or elevation delineation. This work may be improved in the future with more polarization considered [78] or data fusion, as suggested by Tulbure et al. (2022) [54]. The study done by Tulbure et al. (2022) demonstrated the improvement of flood mapping via fusion of different datasets but the study found that mapping of open water surfaces in terms of flood are still challenging as the spectral signatures may be affected by by sediment load, turbidity, dissolved matters, algae content depth and bottom reflectance signal [79]. Mason et al. (2021) suggested that merging very high resolution SAR digital slope model (DSM) data could map urban flood more accurately [69].

5. Conclusions

The 2021-2022 flood is the most deliberate flood in Malaysia’s history in terms of property damage, however, lack of flood inundation information resulted difficulties in rescue planning and resources management during and after the flood event. Hence, this study has developed a framework for rapid precipitation information extraction and flood inundation mapping for helping local authorities and the public through utilization of open source data. Open-source satellite images such as GPM IMERG late run product for precipitation and Sentinel-1 SAR for flood inundation mapping, were fully utilized in this framework. The flood victim statistics which can be obtained from social media is an important input for any flood related studies. In addition, most of the processes have been conducted using the cloud computing system, the Google Earth Engine, to save the processing time and cost. This framework was applied to study the characteristics of precipitation and floods for the Dec 2021-Jan 2022 unexpected precipitation events in Peninsular Malaysia.

Overall, two peaks of the number of flood victims evacuated from their house were observed: (1) 21 – 24 December 2022 – 60000 to 70000 people flood victims replaced in relief centers mostly in Pahang and Selangor; and (2) 2 – 6 Jan 2022 – 10000 to 15000 people relocated to relief centers of Johor. Intense daily rainfall of up to 230 mm/day over the Peninsular Malaysia during the period had triggered extreme flooding in eight states, primarily in Pahang, Selangor, Kelantan, and Johor. This analysis is important for the future hazard management so that early warning can be released when the forecasted rainfall exceeded the danger volume, as well as preparedness of the flood relief operations during monsoon seasons.

Sentinel-1 SAR images are able to produce flood inundation maps for the 2021-2022 flood. The result is validated with the flood location extracted from Sentinel-2 MSI images and observed floods collected from site visits. The total flooded area in Peninsular Malaysia during the flood is 77.43 km², and the spatial distribution of the inundated area shows that the flood is saturated along the major rivers. The result shows that the presented framework could produce flood inundated maps at 62 ~ 70% accuracy, with the threshold
values suggested in this study. The maps are useful for local authorities such as the DID in flood management and comparison with flooded areas simulated by their flood model.

**Supplementary Materials:** The following supporting information can be downloaded at: www.mdpi.com/xxx/s1, Figure S1: title; Table S1: title; Video S1: title.

**Author Contributions:** Conceptualization, Tan, M. L.; Tew, Y. L., and Juneng, L.; methodology, Tan, M. L. and Tew, Y. L.; validation, Tew, Y. L., Hassan, M. H. and Osmani, S.; formal analysis, Tan, M. L.; Tew, Y. L. and Chun, K. P.; resources, Tan, M. L. and Samat, N.; writing—original preparation, Tew, Y. L. and Tan, M. L.; writing—review and editing, Tan, M. L., Juneng, L., Chun, K.P., Chun, K.C., Kabir, M.H.; supervision, Tan, M. L. and Samat, N.; funding acquisition, Tan, M. L. and Samat, N. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by Ministry of Higher Education Malaysia with project code: (NEWTON/1/2018/SS07/USM/) (203.PHUMANITI.6780001) and Newton Fund of NERC, grant number NE/5002707/1 under the IMpacts of PRecipitation from Extreme StormS, Malaysia (IMPRESSMALAYSIA) project.

**Data Availability Statement:** Sentinel data presented in this study are available in Google Earth Engine (https://earthengine.google.com/). The IMERG late run product is available from the NASA Giovanni online tool (https://giovanni.gsfc.nasa.gov/giovanni/). Observed flood data is available on request from the Department of Irrigation and Drainage Malaysia. The codes of the processing is available at (https://code.earthengine.google.com/?scriptPath=users%2Ftewyilin%2FSentinel_flood2021%3APahang_flood).

**Acknowledgments:** The authors would also like to express gratitude to the developers of GEE, GPM, and Sentinel for providing free satellite data and cloud computing platform to the public. The authors also acknowledge editors and reviewers for their valuable suggestions to improve this manuscript.

**Conflicts of Interest:** The authors declare no conflict of interest.

**References**


31. AHA. "Floodin in 8 States Malaysia Flash Update #1." edited by AHA Center, 2021.


47. ESA. "The Sentinel Missions." The European Space Agency, https://www.esa.int/Applications/Observing_the_Earth/Copernicus/The_Sentinel_missions?text=ESA%20is%20developing%20a%20new,robust%20datasets%20for%20Copernicus%20services, (accessed 2 March 2022).
75. BERNAMA. "Johor Latest to Be Hit by Floods, Number of Evacuees up Elsewhere." *Free Malaysia Today* 2022.