UTILISATION OF SYNTHETIC APERTURE RADAR DATA IN BIG DATA ANALYTICS

Hazri Hassan
Department of Survey and Mapping Malaysia (JUPEM)

(This article has not been published)

Abstract

Major applications of future generation in geospatial data revolution is the integration with Big Data Analytics (BDA) that enable to cater the abundance rich dynamic information's of geographic location, time and theme axes known as 5 Dimension (5D). Based on the survey and mapping perspective, the collective of large volume and varieties of spatial data require a specific mining algorithm towards exploration of the potential benefits from each dataset. Recent advances in mapping technology such as Synthetic-Aperture Radar (SAR) has enable for environmental monitoring and earth-resource mapping require broad-area imaging at high resolutions imagery. This paper intends to highlight JUPEM's experience on utilisation of SAR data in year 2008 and 2017 which is Radar - IFSAR Orthorectified Radar Image (ORI), Digital Surface Model (DSM) and Digital Terrain Model (DTM) in Sungai Tebrau, Malaysia. The statistical and probabilistic methods were used in fusion process together with statistical data from related agencies such as Department of Irrigation and Drainages and Department of Environment for the analytic outcomes. The land use changes, modelling of demography, hydrology network and prediction of hydro pollutant area have led to amazing discoveries, and proved that the development really impact the environment from a significant distance.

Keyword: Big Data Analytic, SAR, Geospatial, Hydrography

1.0 INTRODUCTION

Synthetic aperture radar (SAR) technology is widely used in earth observations due to its illumination and weather-independence capability. In the big data era, advanced hardware and high-performance computing technologies are being invented rapidly to tackle the data challenge. Recently, deep learning is showing its self-learning power, and successfully applied to variant fields including image understanding. These will no doubt provide chances and even lead to fundamental changes in SAR remote sensing. Synthetic Aperture Radar (SAR) is a unique technology that widely used to measure ground subsidence and has already shown its ability to map such phenomena on a large spatial scale. Consequently, the processing of the Big Data is challenging for SAR analysis techniques. This paper intends to highlight the potential of SAR in Big Data Analytics application for instance hydrography pollution detection and mitigation.

2.0 SYNTHETIC APERTURE RADAR (SAR)

RADAR is an acronym for Radio Detection and Ranging. Its ability to determine range and motion make it suitable for many applications, such as air traffic control, vehicle speed detection on roadways, and storm tracking. Radar is also used as an active imaging technology in which pulses of microwave energy are emitted from an antenna and the resulting reflections are used to create images. Radar's most important attribute for imaging applications is that its relatively long wavelengths penetrate clouds, dust, and even volcanic ash, and it can image independent of most weather conditions. Considering that this is an ocean planet, with vast amounts of water vapor continuously condensing into clouds over large regions of Earth's surface, radar is a core remote-sensing technology and it becomes the primary imaging source when cloud cover prevents other means of data collection.

Table 1 shows the distribution of frequencies (f) and wave-lengths (λ) of the established radar bands (IEE Standard 521-1984).

Table 1: Designations for radar frequency bands (Ager, 2013)

Band	Frequency Range (GHz)		Wavelength Range (cm)	
UHF	0.3	1	100	30
L	1	2	30	15
S	2	4	15	7.5
C	4	8	7.5	3.75
X	8	12	3.75	2.5
Ku	12	18	2.5	1.67
K	18	27	1.67	1.11
Ka	27	40	1.11	0.75
V	40	75	0.75	0.40
W	75	110	0.40	0.27
mm	110	300	0.27	0.10

Aside from cloud-cover penetration, radar has other notable characteristics that make it valuable for remote sensing applications as follows:

i. Radar provides its own illumination, data can be collected independent of sunlight during the day or at night.

ii. The technique called synthetic aperture radar (SAR) provides high resolution with the remarkable characteristic that its resolution does not degrade with distance. Distance weakens the strength of the radar reflections and can increase image noise

iii. Resolution cell size does not increase as distance increases.

iv. Because radars do not have a fixed lens, as do optical systems, they are flexible regarding resolution and ground

v. Coverage, so that a single system can collect data from wide areas at low resolution, medium areas at medium resolution, and small areas at high resolution

2.1 Radar Echo Measurement System

A radar antenna emits individual pulses of microwave radiation. The pulses are sent at a pulse repetition frequency (PRF) in the range of 2,000 per second or more. As Figure 1 shows, these pulses are scattered in every direction, and a small portion, called radar backscatter, is returned to the antenna. The radar measures the characteristics of the echoes, including the roundtrip time for the pulse to travel from the antenna to the ground and back to the antenna, the strength of the reflection, and the phase of the return wave. That is, the radar can determine if the wave returns at its peak, trough, or somewhere in between.

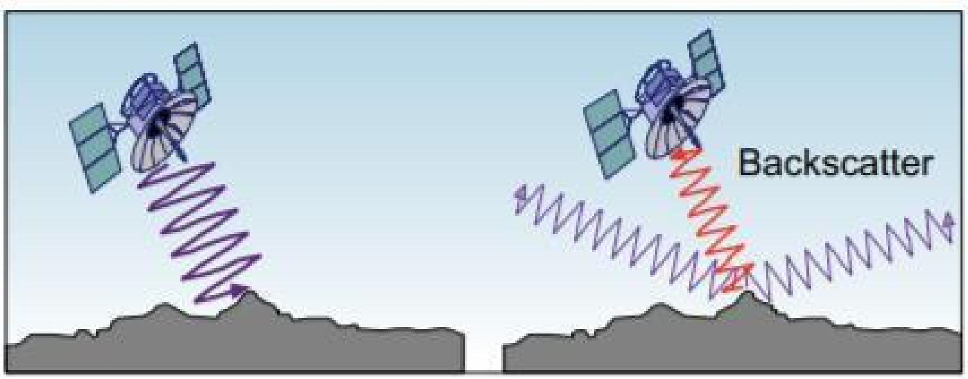


Figure 1: Radar Backscatter

The pulse travel time is used to determine the range from the antenna to the ground. Range is simple to calculate and is merely equal to the pulse speed, which is the known speed of light, multiplied by the round-trip time divided by two

$$Range = \frac{c\Delta T}{2}. \quad (1)$$

Where;

c : speed of light

ΔT: round trip time

Radar data associated with these pulse measurements are often referred to as the fast-time dimension, while the motion of the sensor along its flight path is called slow-time.

Figure 2 shows the measurement of the strength of the backscatter as a function of time. In this example, the pulse reflects first in the near-range area. Most of the associated reflection from the flat terrain travels away from the antenna, and thus the measured backscatter is relatively weak. This is followed by a stronger backscatter, caused by the texture of the trees, and then a very weak return from the calm river. The reflection of the front side of the hill is very strong, which would result in bright pixels on the radar image.

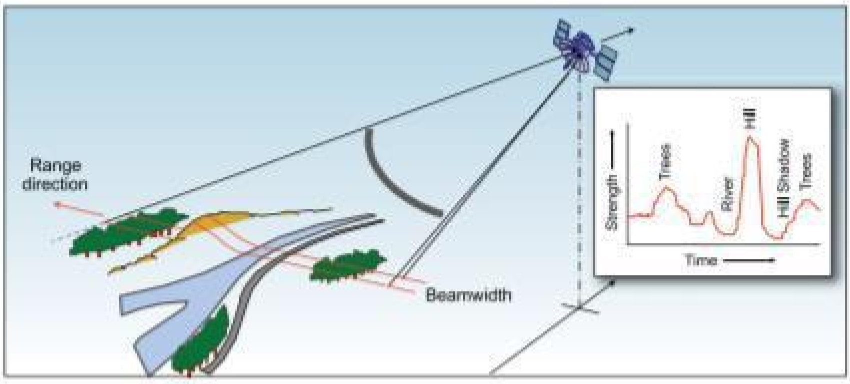


Figure 2: Measuring the Backscatter

2.2 SAR Product

The radar has no lens to focus light energy onto digital detectors from which image pixels are derived. A SAR knows only about its raw measurements of time, amplitude, and phase, and they are manipulated with signal processing algorithms into images and other product. Figure 3 (Ager,2013) shows a simple diagram of the processing flow for SAR products. The raw radar measurements are processed into phase history data, which contains the echo measurements and timing data needed for further processing. A phase history data file is not an image product and cannot be viewed, but it is processed into a pixel form called a “complex image.” Notice that the complex image is a core radar product used to create amplitude images and many other radar products that cannot be generated from simple amplitude data alone (Figure 4).

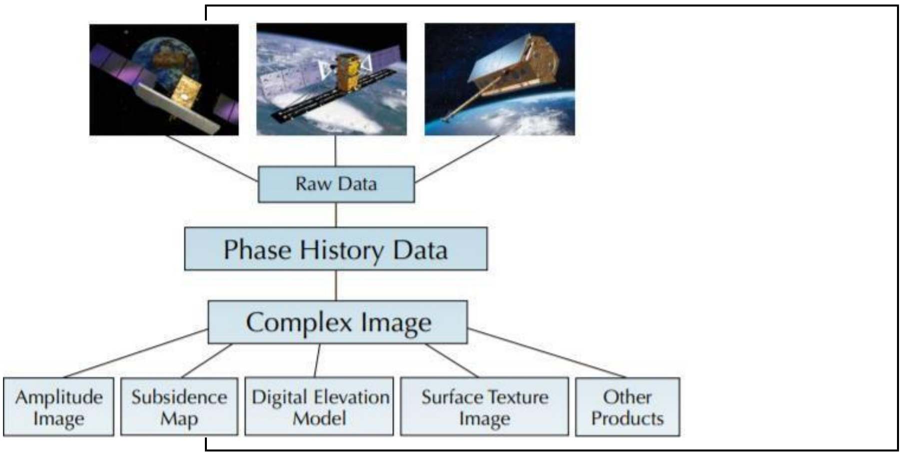
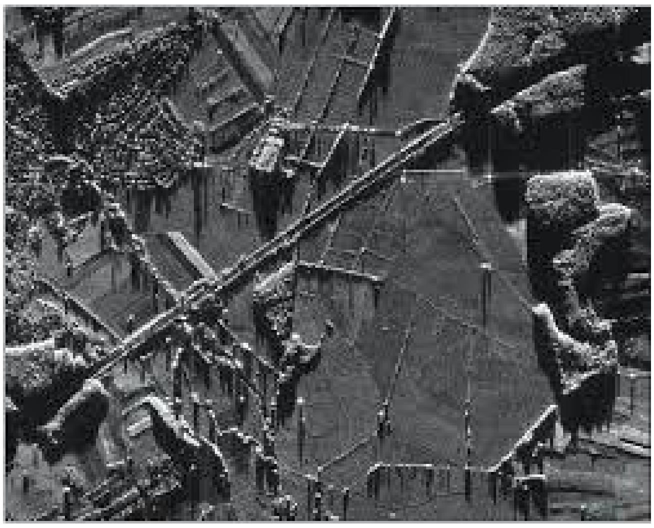
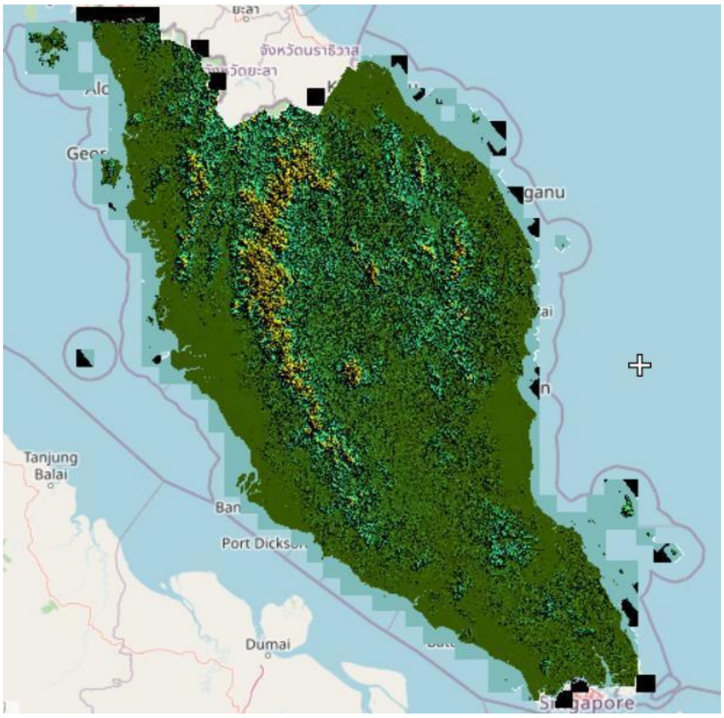


Figure 3: The SAR product formation chain.



(a)

(b)



(c)

Figure 4: (a) Orthorectified Radar Imagery, (b) Color Orthorectified Radar Imagery and (c) Digital Terrain Model (JUPEM, 2017)

3.0 BIG DATA ANALYTICS

The concept of big data has been developed for years. Big Data refers to the evolution and use of technologies that provide the right user at the right time with the right information from a mass of data that has been growing exponentially for a long time in our society. The challenge is not only to deal with rapidly increasing volumes of data but also the difficulty of managing increasingly heterogeneous formats as well as increasingly complex and interconnected data.

Big Data refers to gigantic larger datasets, more diversified, including structured, semi-structured, and unstructured as represented in terms Big Data characteristics that comprise of the following (Figure 5); velocity, variety, value, veracity and volume (Hadi, 2015).

Figure 5: Big Data Characteristic

Big data analytics refers to the process of collecting, organizing, analyzing large data sets to discover different patterns and other useful information. Big data analytics is a set of technologies and techniques that require new forms of integration to disclose large hidden values from large datasets that are different from the usual ones, more complex, and of a large enormous scale. It mainly focuses on solving new problems or old problems in better and effective ways. Types of Big data analytics can be categorized as follows:

i. **Descriptive Analytics**

It consists of asking the question: What is happening? It is a preliminary stage of the processing that creates a set of historical data. Data mining methods organize data and help uncover patterns that offer insight. Descriptive analytics provides future probabilities and trends and gives an idea about what might happen in the future.

ii. **Diagnostic Analytics**

It consists of asking the question: Why did it happen? Diagnostic analytics looks for the root cause of a problem. It is used to determine why something happened. This type attempts to find and understand the causes of events and behaviors.

iii. **Predictive Analytics**

It consists of asking the question: What is likely to happen? It uses past data in order to predict the future. It is all about forecasting. Predictive analytics uses many techniques like data mining and artificial intelligence to analyze current data and make scenarios of what might happen.

iv. **Prescriptive Analytics**

It consists of asking the question: What should be done? It is dedicated to finding the right action to be taken. Descriptive analytics provides a historical data, and predictive analytics helps forecast what might happen. Prescriptive analytics uses these parameters to find the best solution.

This paper highlights the utilization of SAR data elevation information in providing an important element in big data analytic for hydrography management.

4.0 SYNTHETIC APERTURE RADAR DATA IN BIG DATA ANALYTICS

Department of Survey and Mapping Malaysia (JUPEM) as the main National Mapping Agency has traditionally been the custodians of authoritative geospatial data, where it has been confirmed that JUPEM data is spatially accurate, reliable, credible and relevance for all the areas updated by JUPEM in Malaysia but the lack of currency of some authoritative data sets has been questioned. To this end, JUPEM is transitioning from inwardly focused and closed agencies to outwardly looking and accessible infrastructures of spatial data. The ability to inter-connect and link data provides the opportunity to leverage the vast data, information and knowledge sources across the globe. Hence JUPEM is considering the drivers of the big data phenomena and look to identify how authoritative and big data may co-exist. Existing data available in JUPEM include structured and unstructured data. Data collected from field work such as LiDAR, SAR and Orthophoto are considered as unstructured data.

However, after processing work is done, the data will be extracted into structured data such as vector data, raster as well as scheduled data. One of the business case in JUPEM BDA is natural disaster (Big Data Analytic for Geospatial Data Management- Mapping of Hydrography Theme)

4.1 Business Case: Big Data Analytic for Geospatial Data Management- Mapping of Hydrography Theme

The objective of the business case is to analyze and visualize various factors that predicts the river quality in the next 10 years in order to assist JUPEM officers to update the maps accordingly which in turn helps all other relevant agencies who needs updated and accurate maps to mitigate the issue of river pollution and its surrounding areas. The related prescriptive analytics is prediction of river speed and river flow direction using terrain data from SAR. Figure 6 shows the slope map and aspect map for particular area. These maps are critical for hydrology modelling in river pollution analysis.

Figure 6: Slope and Aspect Map from SAR Data

Figure 7 indicates that there are 3 categories of river flow speed which is fast, moderate and slow. BDA will predict the river pollution prone area based on elevation information for 5 years period.

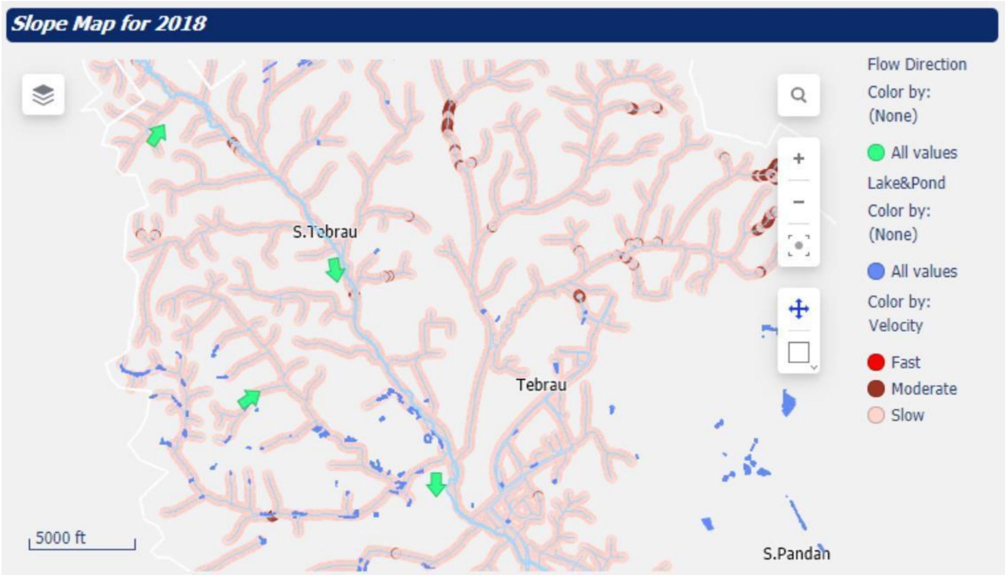


Figure 7: River Flow and River Speed

Data fusion with JUPEM topography datasets and secondary data from other agencies facilitates BDA to predict the potential mapping planning requirements in 5 years period as shown in Figure 8 (JUPEM , 2019).

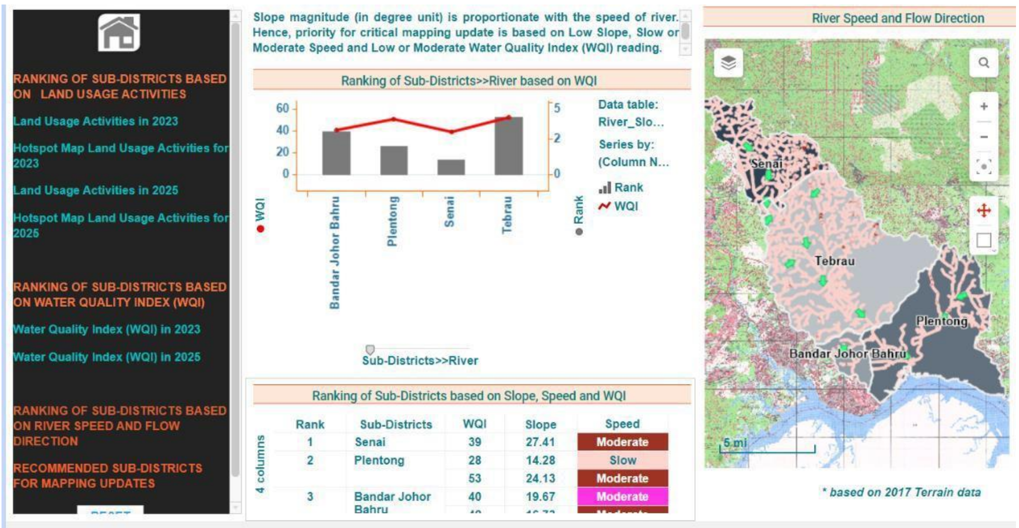


Figure 8: Predicted Planning requirement based on JUPEM Fundamental Datasets

5.0 CONCLUSION

Big Data is new and requires investigation and understanding of both technical and business requirements. Big data is not a stand-alone technology but it is a combination of the period of technological evolution including geospatial data. Applications of SAR include topography, oceanography, glaciology, geology (for example, terrain discrimination and subsurface imaging), and forestry, including forest height, biomass, deforestation. On the other hand, based on previous work, Big data initiatives have shown significant promise for policy and decision-making, as well as fostering collaboration between governments and citizens and businesses, and for ushering in a new era of digital government services.

REFERENCES

Ager, T.P, 2013. An introduction to synthetic aperture radar imaging. Oceanography 26(2):20–33, http://dx.doi.org/10.5670/oceanog.2013.28. USA

Ahamad Zakaria, 2017. Application of IFSAR Technology In Topographic Mapping: JUPEM's Experience. Proceedings of the International Cartographic Association. https://doi.org/10.5194/ica-proc-1-125-2017. Switzerland

Institute of Electrical and Electronics Engineers.1984. IEE Standard. New York.

Hiba Jasim Hadi, Ammar Hameed Shnain, Sarah Hadishaheed, Azizah Ahmad, 2015. Big Data And Five V's Characteristics. International Journal of Advances in Electronics and Computer Science, Volume-2, Issue-1.

Department of Survey and Mapping, Malaysia, 2019. Development Of Big Data Analytics Framework. Kuala Lumpur