

BIG DATA ANALYTICS TOOLS ON MALAYSIA'S CLIMATE CHANGE PROJECTED DATA TO REINFORCE WATER RISK MANAGEMENT

MOHD ZAKI MAT AMIN⁽¹⁾, MOHAMMAD FIKRY ABDULLAH⁽²⁾, NURUL HUDA MD ADNAN⁽³⁾, HARLISA ZULKIFLI⁽⁴⁾, MARINI MOHAMAD IDERIS⁽⁵⁾ & ZURINA ZAINOL⁽⁶⁾

^(1,2,3,5,6) Water Resources and Climate Change Research Centre, National Hydraulic Research Institute of Malaysia, Seri Kembangan, Selangor, Malaysia,

zaki@nahrim.gov.my; fikry@nahrim.gov.my; nurulhuda@nahrim.gov.my; marini@nahrim.gov.my; zurina@nahrim.gov.my

⁽⁴⁾ Information Management Division, National Hydraulic Research Institute of Malaysia, Seri Kembangan, Selangor, Malaysia, harlisa@nahrim.gov.my

ABSTRACT

Climate change and variability together with non-climatic drivers have had physical impacts on the hydrological cycle, especially in the recent decades. Evolution in the hydroinformatics discipline, with incorporating rapidly expanding body of climate change information is indispensable particularly for sustainable water resources and risk management. As such, the growth of big data analytics technology have enabled stakeholders to quickly process and analyse huge volumes of climate change data and generate accurate insights, which subsequently would open up promising disaster resilience and risk reduction approaches. The National Hydraulic Research Institute of Malaysia (NAHRIM) has been applying big data analytics over 10 billion projected hydroclimate data for the Peninsular Malaysia, downscaled and generated at 6km horizontal spatial resolution from 15 realisations based on the 4th Assessment Report (AR4) of the Intergovernmental Panel on Climate Change (IPCC AR4). Big data technology elevated processing, analysis and visualisation of hydroclimate data that assist in mining and identifying the potential future water related extreme events and their magnitudes, such as extreme rainfalls, floods, storm centers and drought. The analytics system enabled NAHRIM to highlight predictive, scientific and data-driven evidence of climate change impacts on the hydrology of Peninsular Malaysia. Therefore, the analytics and functions of big data indubitably provide support in strategies and decision making process as well as strengthening related policy on water related disaster risk management, resilience and adaptation of climate change, such as National Policy on Climate Change and National Water Resources Policy.

Keywords: Big data; climate change; water disaster risk management; disaster risk reduction; decision making tools.

1 INTRODUCTION

In recent decades, the occurrences of natural disaster events, which brought catastrophic impacts to the environment and socio-economics, are increasing as compared to the 20th century. Meteorological and hydrological events, such as storms and floods, are becoming more frequent. In 2015, both categories of disaster events reported about 41% and 42% of the total natural disasters worldwide, respectively. The 2011 flood event in Thailand is recorded as one of the event with the highest losses of USD 43 billion and 813 fatalities (Munich Re, 2016). As in Malaysia, continuous heavy monsoonal rainfall from 14th to 25th December 2014 over the east coast of Peninsular Malaysia had caused widespread floods, especially in the state of Kelantan. This flood event is considered as the worst flood in the history of the state, which thirteen (13) deaths were reported and around 340,000 flood victims were evacuated with an estimated total loss of MYR1.7billion (DID, 2015).

Furthermore, both climatic and non-climatic drivers, such as rapid population growth, increased urbanization, industrialization and pollution, factored the increasing number of water excess as well as water stress events that threatened the sustainability of our water resources. Changes in the climate system are expected to have a wide range of impacts on ecosystems, infrastructure, health systems, the economy, and particularly on natural resources. Subsequently, water-related risks are possibly magnified in the future, thus, early identification and investigation of potential hydro-meteorological extreme events throughout future climate scenarios are essential for providing and disseminating scientific evidence to develop and strengthen water related disaster risk management, resilience and adaptation strategies.

The global climate change agenda has become a key issue in water management all over the world. According to Ishida and Kavvas (2017), climate change impacts on the water-related sector and management needs to be reflected in order to enhance its reliability, and thus, watershed-scale climate change assessment is essential particularly in flood control and water resources management. Climate change projections at very coarse scales are available from various Global Climate Models (GCMs), however, they are unable to resolve significant subgrid scale features essential for climate change impact assessment to hydrologic regimes (Fowler et al., 2007; Kavvas et al., 2007; Ishida and Kavvas, 2017). In general, two (2) fundamental

downscaling approaches for bridging the resolution gaps between GCMs and regional and local scale, the statistical and dynamical methods. The statistical method established fixed empirical relationships across spatial scales between the large-scale climate and local climate. However, the latter makes use of limited area models with progressively higher spatial resolution than the GCM and is able to produce finer resolution information that can resolve atmospheric processes on a smaller scale, but requires intensive computation (Fowler et al., 2007; IPCC, 2013; Ishida and Kavvas, 2017).

Hence, the application of big data technology and predictive analytics can be important tools in mining and processing massive volumes of hydroclimate data generated from intensive and complex computation, particularly in the context of climate change modeling and its relation to climate resilience. As the impacts of climate change are accelerating and the urgent need for effective solutions is required, therefore, emerging role of big data technology in understanding and mitigating climate change risk are considered innovative and effective solutions for climate change (Namrata, 2017). Data-Pop Alliance (2015) and Emmanouil and Nikolaos (2015) have explored the opportunities, challenges and required steps for leveraging this new ecosystem of big data to monitor and detect hazards, mitigate their effects, and assist in relief efforts, which is ultimately to build resilience and maintain hazard equilibrium. Furthermore, the utility and potential of big data for disaster management is growing as the number and access to datasets are expanding rapidly (Beth et al., 2015). Although it is still regarded as an emerging technology (Frey et al., 2016), it has been recognised as a promising approach in order to harness data science and big data for climate action by means of identifying revolutionary new approaches to climate mitigation and adaptation (UN Global Pulse, 2017).

In view of high potential effectiveness and high impact of big data analytics (BDA) technology on socio-economy activities particularly to address the current challenges faced by government agencies, the government of Malaysia has announced the Big Data Analytics (BDA) initiatives in November 2013. Subsequently, in mid-2015, four (4) government agencies have been selected, which includes NAHRIM to participate in a strategic BDA initiatives project entitled "The BDA-Digital Government Open Innovation Network (BDA-DGOIN) and Proof-of-Concept (POC)", which were co-organised by Malaysian Administrative Modernisation and Management Planning Unit (MAMPU), Malaysia Digital Economy Corporation (MDEC) and MIMOS Berhad.

Therefore, this paper emphasised NAHRIM's first attempt in utilising the technology of big data to analyse our own downscaled hydroclimate data over the Peninsular Malaysia. Further development works are carried out after the POC project in order to provide scientific insights and support resources and water engineering, planning and risk management are also highlighted, particularly in the context of hydro-meteorological potential impacts and consequences to the safety level of water risk under the climate change conditions.

2 DATA USED AND METHODOLOGY

2.1 Data used - projected hydroclimate data for peninsular Malaysia

A comprehensive study conducted to assess the impact of climate change on the hydrologic conditions of Peninsular Malaysia (NAHRIM, 2014). Fifteen (15) climate projections for the 21st century by three (3) different coupled land-atmosphere-ocean Global Climate Models (ECHAM5 of the Max Planck Institute of Meteorology of Germany, CCSM3 of the National Center for Atmospheric Research (NCAR) of the United States, and MRI-CGCM2.3.2 of the Meteorological Research Institute of Japan) under four (4) different greenhouse gas emission scenarios (B1, A1B, A2, A1FI) based on IPCC AR4 were dynamically downscaled onto 3888 grids at 6 x 6km spatial resolution by means of a Regional Hydroclimate Model of Peninsular Malaysia (RegHCM-PM) (Amin et al., 2017). The model includes a mesoscale atmospheric component and a physically-based hydrology model component, which the atmospheric model component used is the MM5 (the Fifth Generation Mesoscale Model) from NCAR (National Centre for Atmospheric Research) – a non-hydrostatic model, which can be downscaled even to 0.5km spatial resolution (Kavvas et al., 2007; Shaaban et al., 2010; Amin et al., 2017). By coupling MM5 with the physically-based hydrology model known as Watershed Environmental Hydrology Model (WEHY), a realistic estimation and interactions between the atmospheric and land surface hydrologic process can be modelled (Amin et al., 2017).

There are five (5) main projected hydroclimate parameters readily available at these grids, *i.e.*, precipitation, air surface temperature, runoff, evapotranspiration and soil water storage, which are produced at up to hourly temporal increment for 30 years of simulated historical period (1970 – 2000) and 90 years future period of 2010 – 2100 for each fifteen (15) climate change realisations. These data, together with projected streamflow data are also available at basin scale for thirteen (13) river basins in the Peninsular Malaysia. The size of the total projected hydroclimate data are over 10 billion records, which is hugely challenging to be analysed using traditional approaches and methodologies, especially in order to predict the future impacts of these climatic changes to hydrologic conditions and water resources.

2.2 Big data analytics - NAHRIM Hydroclimate Data Analysis Accelerator (N-HyDAA)

The term 'big data' is generally described and characterised by the amounts (volume), velocity and variety of data, exhaustive in scope, fine-resolution, relational in nature as well as flexibility in terms of data size extensionality and scalability (Data-Pop Alliance, 2015; Emmanouil and Nikolaos, 2015; Kitchin, 2013). The projected hydroclimate data in NAHRIM is massive in volumes, but not in terms of velocity and variety, as the dataset is already pre-processed and produced in structured and flat file format. In order to utilise, visualize and analyse about 1450 simulation years of projected hydroclimate data, the BDA system in NAHRIM known as NAHRIM Hydroclimate Data Analysis Accelerator (N-HyDAA) was constructed essentially for tracing and identifying specific data pattern, magnitude and extends of extreme hydro-meteorological events throughout the century.

The NAHRIM's BDA system was established based on Mi-Galactica, a high performance query accelerator that further advances data processing speeds by using Central Processing Unit (CPU) or Graphic Processing Units (GPUs) for massive parallel processing of unpredictable, complex and long-running query workloads developed by MIMOS Berhad (MIMOS, 2016). The developed data processing accelerator system, N-HyDAA addresses the high volume data processing challenges by utilising the following features:

- Columnar storage for parallel data processing;
- Heterogeneous structure query engine using GPU technology;
- Geo-spatial accelerated processing;
- Data visualisation system of multi-billion record datasets.

For instance, the effectiveness of data processing accelerator was capable to perform quick analytics and visualisation in only 14 seconds based on one scenario table for the whole 3888 grids, and about 3.5 minutes for all tables and scenarios. Comparatively, the identification of 127 million points for polygon of river basins and regions, and transformation into visualisation in web server takes 4.3 seconds, which is 77 times than other post-GIS systems.

In general, there are currently eight (8) analytics features developed, which four (4) modules are from POC: drought, temperature, rainfall, storm center and streamflow, while three (3) new analytics features: climate change factor (CCF), water stress index (WSI) and WSI simulation are developed to assist engineers, development and utility planners, and main stakeholders in making timely decision and strategies in water-related risk management and development projects. The overall infrastructure of the developed system is generalised in Figure 1.

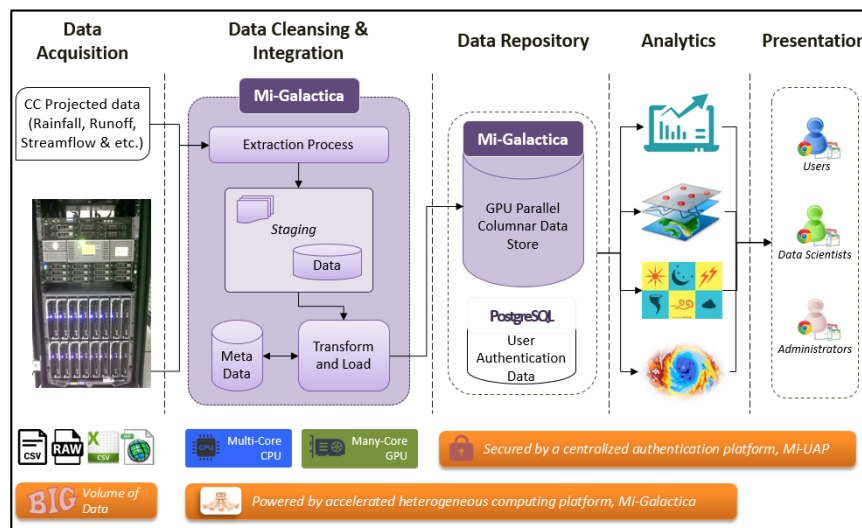


Figure 1. NAHRIM's BDA N-HyDAA data warehouse infrastructure.

3 ANALYTICS 1: POTENTIAL IMPACTS

Big data can help significantly in the prevention and preparation of water-related disaster and crisis management. Information derived from big data analysis can help to anticipate crisis or reduce the risks that would arise (Emmanouil and Nikolaos, 2015). The findings from NAHRIM (2014) indicate that future rainstorms will change in temporal and spatial variability, which are mirrored in the formation of runoff and streamflow, leading to intensification of floods and droughts.

Predictive analytics in N-HyDAA functions enable data mining and extractions of possible future rainfall trends, patterns and magnitude either yearly, monthly or even weekly basis, and based on each climate change scenarios concerned. For instance, Figure 2 shows a sample of visualised changes in the projected yearly gridded rainfall for year 2040, 2070 and 2100 extracted from the system. The figure shows that the rainfall magnitude is projected to increase towards the end of the century. At the same time, daily-based

temporal resolution for projected temperature and corresponding projected runoff to the rainfall distribution also can be investigated and visualised. The BDA system accelerated these analytics and visualisations, providing quick insights and identification of potential impacted areas, degree of severity and planning of strategic approaches in short or long-term mitigation and adaptation.

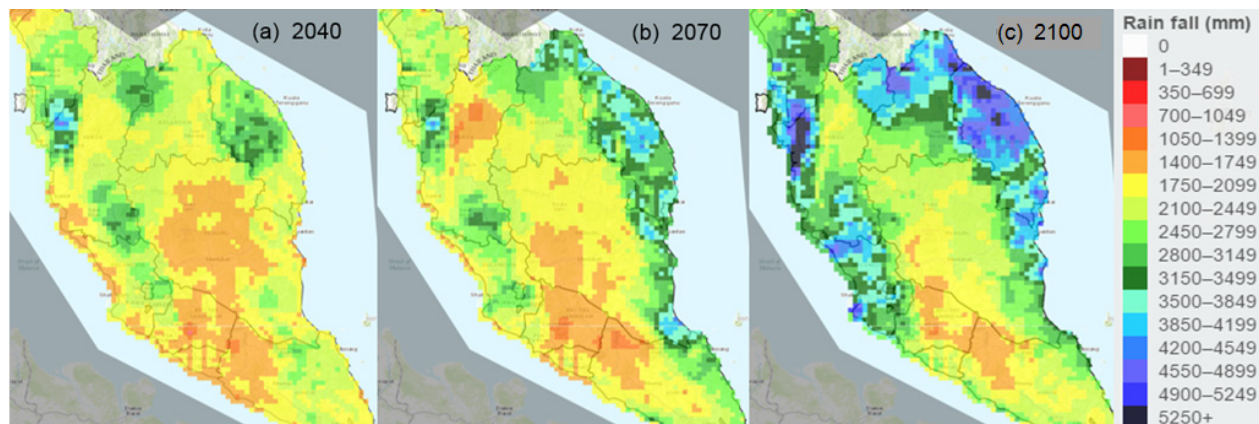


Figure 2. Projected changes in future yearly rainfall depth (in mm) and spatial distribution for year 2040, 2070 and 2100.

Further example, analysis conducted on future rainfall and river flow pattern in Kelantan River basin has detected possible rainfall and flood event in projected time horizon 2030 – 2040, with nearly similar pattern, magnitude and even storm centres with the disastrous flood that hit Kelantan and east coast states in 2014. The histogram in Figure 3 shows the projected basin-averaged daily rainfall (in mm) for the flood period in time horizon 2030 – 2040 (based on average of A1B scenario) that resembles the pattern of 2014 flood, that was identified and generated instantly by the N-HyDAA system. There are two episodes of highest daily basin rainfall identified, 168 mm and 160 mm, during the event in the mentioned period. Concurrently, the system is also able to visualise the rainfall distribution and magnitude during the event as also shown in Figure 3.

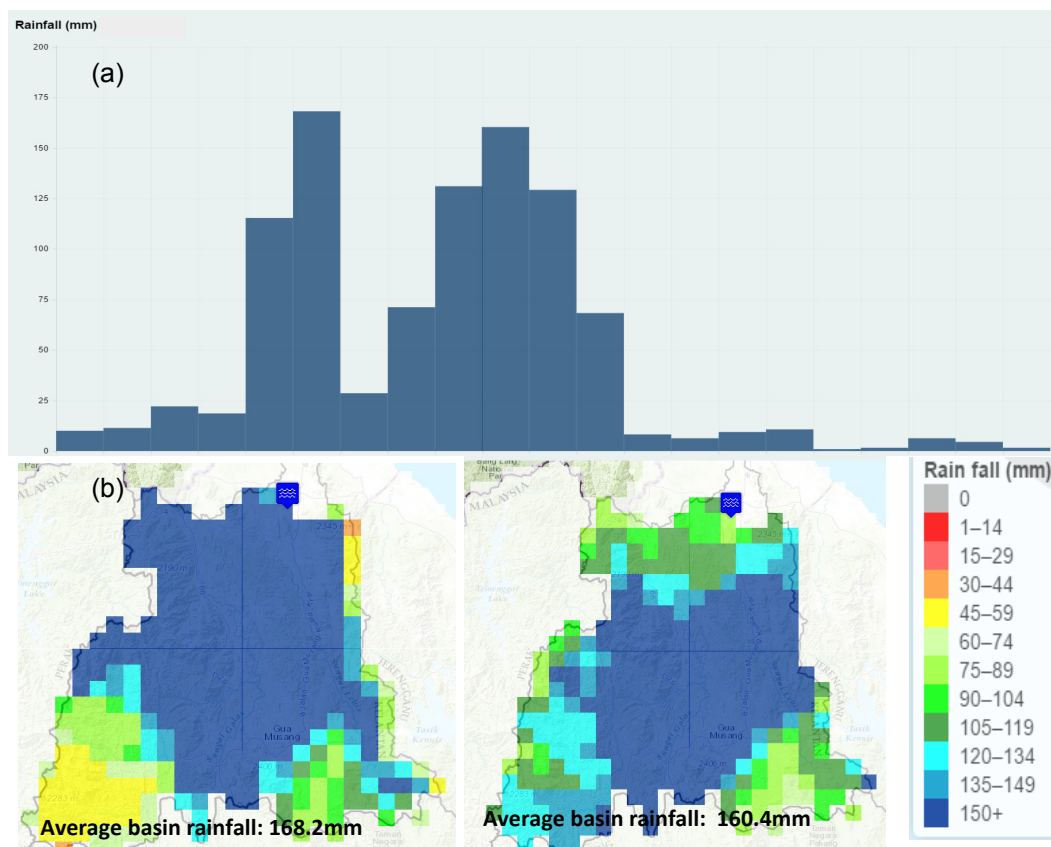


Figure 3. (a) Graph of basin averaged projected daily rainfall during the identified flood event in time horizon 2030 – 2040, and (b) the spatial distribution of the two highest rainfall depths during the future projected flood event in Kelantan River basin.

The system also enables us to query and predict the area and extents of extreme, prolonged dry periods that may poses threats to future water resources and supply. Previously, through manual hydrological assessments, NAHRIM has discovered potential future drought years based on drought intensity and recurrences. With BDA system, the areas of possible drought, reduction and changes in rainfall pattern and magnitude can be easily identified and visualised, and thus, proper mitigation and adaptation strategies can be carried out to adhere the impacts. Figure 4 shows the projected three-monthly gridded rainfall in early century for time horizon 2020 – 2030, which has been identified as one of the extreme drought event that may affected the whole peninsular. Almost all states are projected to receive very low rainfall of below 700mm during the first six months of the period mentioned above.

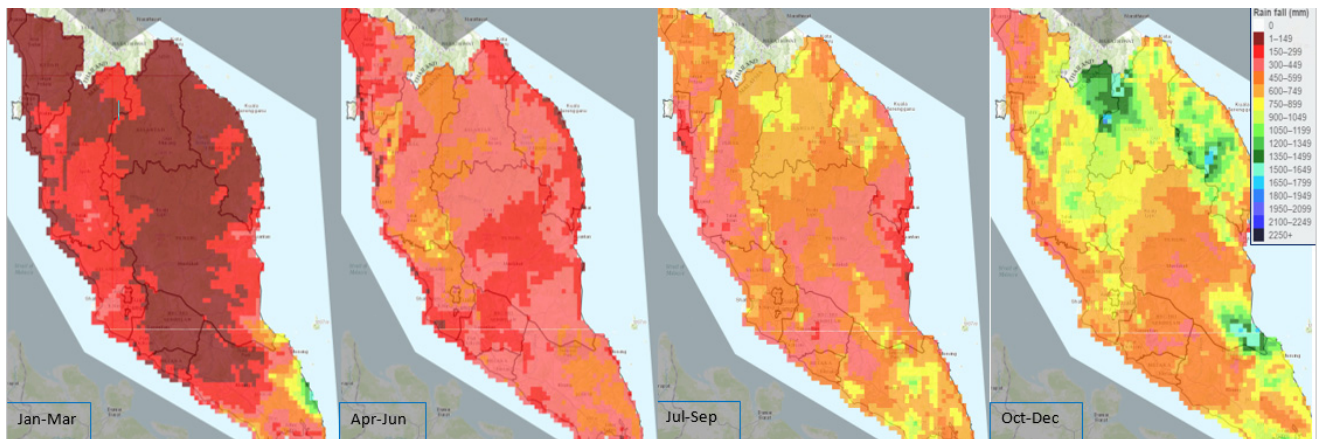


Figure 4. Changes in projected three-month rainfall (in mm) in time horizon 2020 – 2030.

Hence, the extraction and assessment of the future hydroclimate information influences longer term evaluations of hydro-meteorological hazard and risk reductions in addition of resource management strategies through assumptions about possible precipitation and runoff conditions as these physical variables are translated into assumed variability in future water supplies, demands, and/or operational constraints.

4 ANALYTICS 2: EXPLORING THE SAFETY LEVEL OF WATER RISK

4.1 Climate Change Factor (CCF)

The BDA system is then further developed to analyse degree of vulnerability of water excess and stress due to changes in rain depth and intensity, and its consequence to river flows. NAHRIM (2013) has introduced a method to estimate degree of changes and impacts of future rainfall through derivation of a climate change loading factor or Climate Change Factor (CCF). CCF is generally defined as a ratio of the projected future hydrological data, such as rainfall, to the simulated historical data. By adopting the methodologies derived in NAHRIM (2013), the Generalized Extreme Value (GEV) and Extreme Value Type 1 (EV1) approaches are used to calculate the return periods of maximum daily rainfall events with return periods of 2, 5, 10, 20, 25, 50, 100 and 200-year. The same fundamental probability distributions are also applied in developing 1-day CCF for high flows, while estimation of low flow CCFs are based on GEV and Weibull distribution.

These equations /methodologies are embedded or incorporated into N-HyDAA analytics algorithm, which then are based by custom selection of climate change scenarios, by grid or region, and future 30-year time slices (2010 – 2040, 2040 – 2070 and 2070 – 2100). Figure 5 shows the estimated 1-day maximum rainfall under the average 14 realisations from three (3) emission scenarios (A1B, A2 and B1) for 50-year and 100-year Average Recurrence Interval (ARI) during middle of the century (time horizon 2040 – 2070). The maximum CCF values for both ARI years reach 1.90. The CCF value indicates that there might be an increase of 90% in rainfall depth as compared to baseline/simulated historical years (1970 – 2000). However, it can be seen from the 100-year ARI map that the extents/areas with the maximum CCF values are increasing especially at the northern states, Selangor and Johor Bahru, as compared to 50-year ARI's. However, at the same time, there are places that may have reduction in rainfall amount, as indicated by CCF values lower than 1 (Figure 5).

Whilst analysis based on GEV third quartile distribution yet shows a possibility of higher rainfall intensity and magnitude in future period 2040 – 2070, as depicted in Figure 6. Almost the whole Peninsular Malaysia is analysed to have CCF values of more than 1 in both 50 and 100-year ARI. The areas that are projected with extreme CCF values of 1.9 and above in the return period 100-year are distinctly wider, affecting most of Kedah and Perlis in north, Terengganu and Selangor, as compared to the same ARI in Figure 5. This estimation information is an approach of quantifying the scale of climatic change to surface water systems, which can act as an alarm and indication of future hydro-meteorological condition, and thus, should be integrated into future water-related planning and risk management through strategic adaptive capacity.

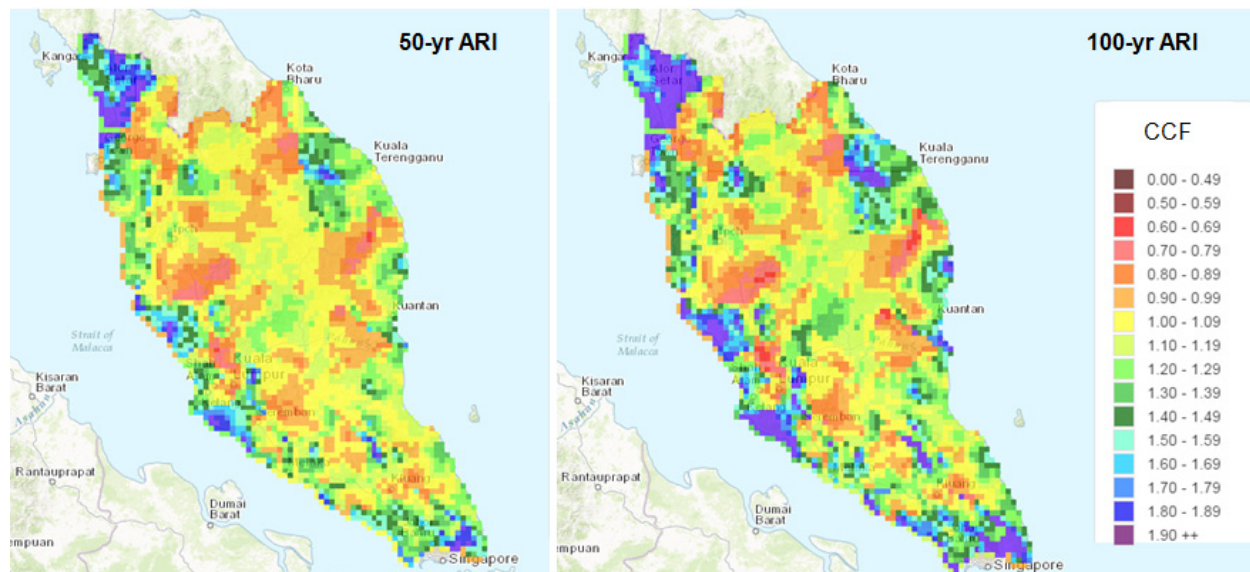


Figure 5. Comparison between 50-yr and 100-yr ARI of gridded rainfall CCF for time horizon 2040 – 2070.

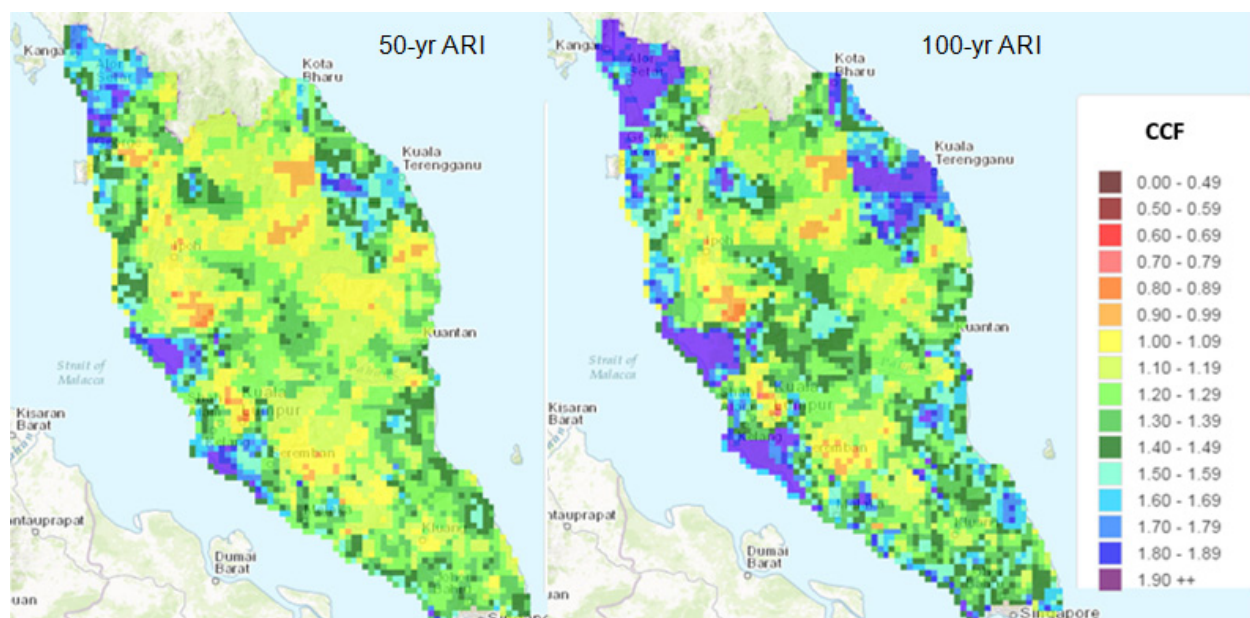


Figure 6. Comparison between 50-yr and 100-yr ARI rainfall CCF based on third quartile of GEV distribution for time horizon 2040 – 2070.

4.2 Water yield and Water Stress Index (WSI)

Drought, changing patterns of precipitation and increased evaporation will affect water availability. Through big data analytic tools, federal and state government agencies, water authorities and stakeholders can have a quick but very informative input on future water yield as compared to historical availability, as an example shown in Figure 7. The figure shows the calculated water yield for historical and future period for 2030, for 80 districts of the entire Peninsular Malaysia, calculated based on simulated future runoff under four (4) different gas emission scenarios (B1, A1B, A2 and A1FI) by means of an average of fourteen (14) projections (A1B, A2 and B1), average A1B, average A2 and average B1. It can be seen from the figure that water yield for year 2030 under average of 14 scenarios, A1B and A2 are expected to increase as compared to historical period (2 billion cubic meters (BCM) to 8 BCM) except for B1, which is expected to decrease by 3 BCM. In terms of temporal variations throughout the whole Peninsular Malaysia, the monthly trends of future water yield based on each scenario are similar to the simulated historical water yield, as depicted in Figure 8. Some monthly future show decreasing yields, though six months, *i.e.*, January, February, March, April, July and October might have increase water yield (from 1% – 13%, 19% – 26%, 27% – 31%, 15% – 27%, 4% – 9% and up to 8% respectively) in the future. In December, the yields are projected to decrease about 3% to 17% (925 million cubic meters (MCM) to 6,196 MCM) in the future scenarios.

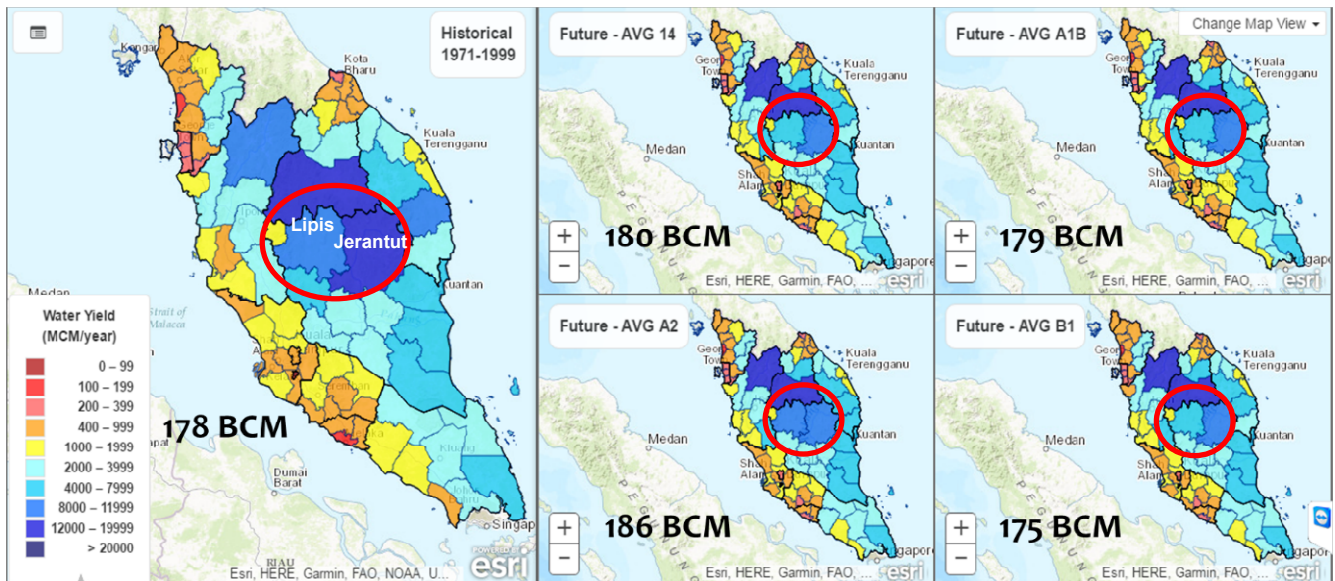


Figure 7. Comparison between simulated historical and future water yield (in billion cubic meter, BCM) in 2030 based on average scenarios of A1B, A2, B1 and fourteen scenarios.

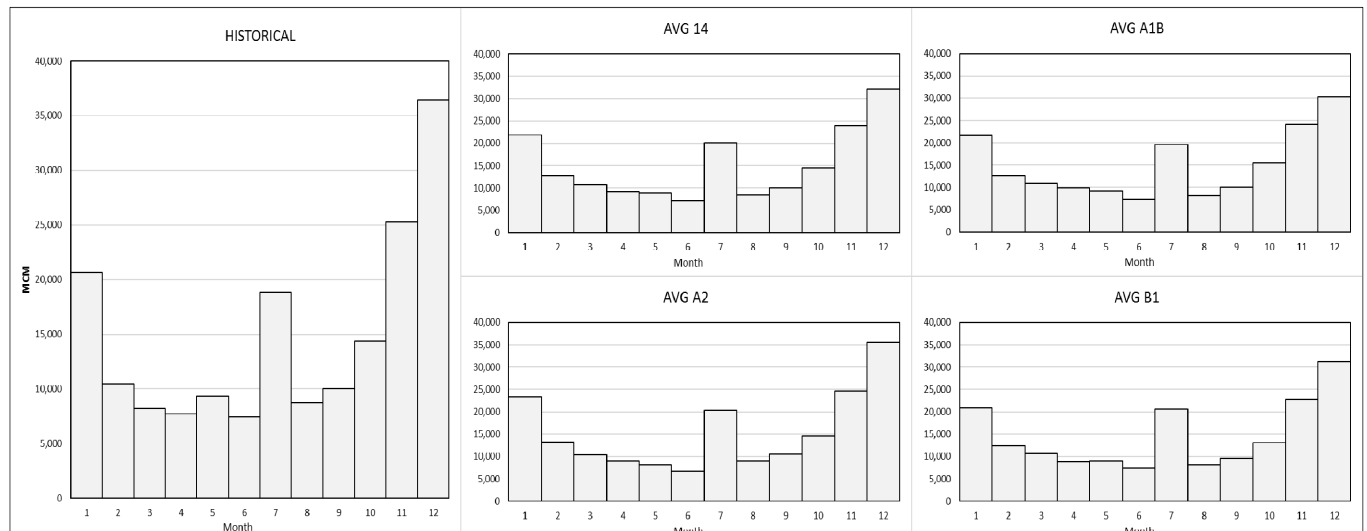


Figure 8. Comparison between monthly simulated historical and future water yield (in million cubic meter, MCM) in 2030 based on average scenarios of A1B, A2, B1 and fourteen scenarios.

Generally, all districts show an increase in future water yields, however a few districts in Pahang (circled in Figure 7) have relatively decreasing yield, which translates to possibly lesser water. The district of Lipis is projected to have 20% less of water under the scenario of an average of 14 projections as compared to the simulated historical value (from 9,645 MCM to 7,698 MCM in the future), -20% based on average A1B, -16% (average A2) and -23% under the average B1 projections. In Jerantut, future water yields under the four (4) groups of average scenarios are also estimated to decrease about 23% – 30% as compared to the simulated historical condition of 13,539 MCM.

The decreasing trend of future water yields in the two districts of Pahang can be correlated with the decreasing volume of projected monthly rainfall in 2030 compared to simulated historical rainfall, as shown in Figure 9a and 9b. In Lipis (Figure 9a), the total rainfall based on the scenarios are projected to decrease ranging from 16% – 21%, which monthly rainfall for six (6) months are projected to be significantly lower than the historical values. The same rainfall pattern is projected for Jerantut (Figure 9b), where the future rainfall for each scenario decrease for about 2,880 MCM to 3,635 MCM from the simulated historical value of 17,600 MCM.

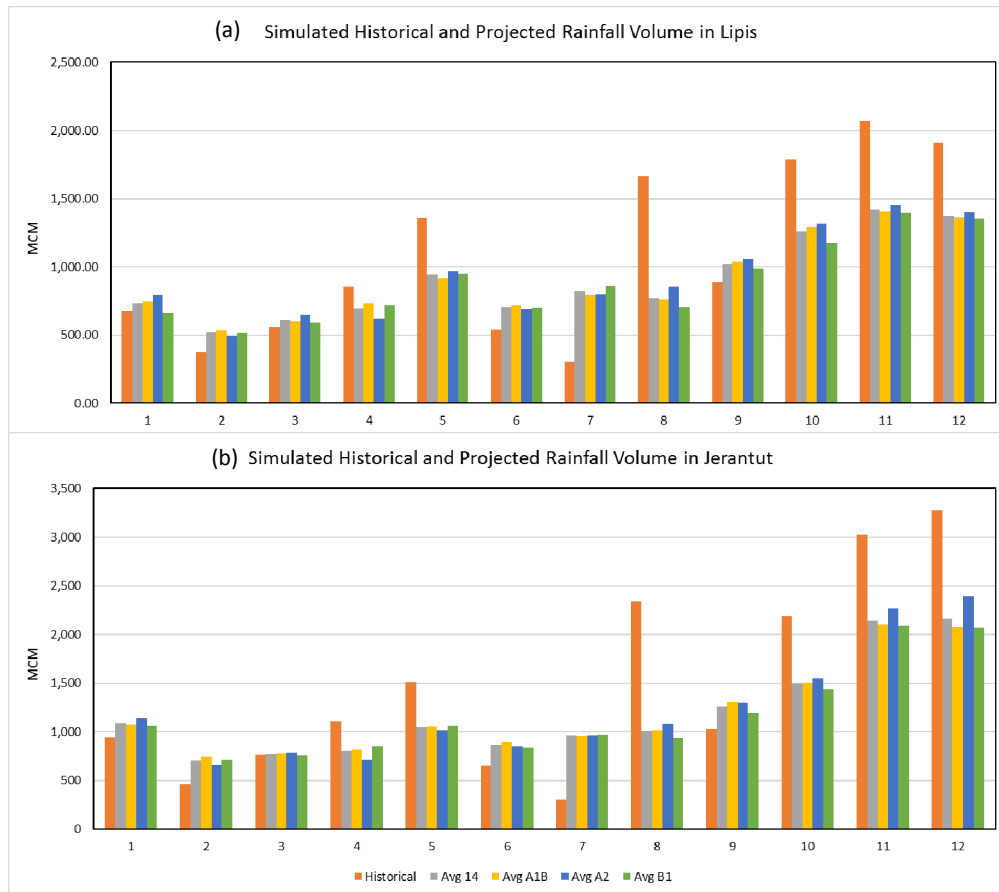


Figure 9. Monthly simulated historical and future rainfall (in million cubic meter, MCM) in 2030 based on average scenarios of A1B, A2, B1 and fourteen scenarios for district (a) Lipis, and (b) Jerantut.

Furthermore, the Water Stress Index (WSI) approach developed by Pfister et al. (2009) is used to construct water stress indices by means of projected water yield and water demand that is possibly impacted by climate change conditions. WSI is defined as an index calculated based on Stephen Pfister's model to represent the level of water stress in specific area by means of a ratio of total water demand or consumption against water yield or availability (Brown et al., 2011).

The estimation of WSI is determined by means of water withdrawal to water availability (yield) calculation as given in Eq. [1]. Water to water availability (WTA) is the ratio of total annual freshwater withdrawal (WU) made up of domestic, industrial and agricultural sector, to hydrological (water) availability (WA). The variation factor (VF) is derived to consider variation in precipitation. VF is calculated using Eq. [2] by means of the standard deviations of average monthly (s°_{month}) and annual (s°_{year}) rainfall, in order to calculate a modified WTA (WTA° as in Eq. 3), which is used to differentiate watersheds with strongly regulated flows (SRF) or non-strongly regulated flows (non SRF) (Pfister et al., 2009). Finally, WSI is calculated based on a logistic function as in Eq. [4].

$$WTA = \frac{\sum WU}{WA} \quad [1]$$

$$VF = e^{\sqrt{\ln(s^{\circ}_{\text{month}})^2 + \ln(s^{\circ}_{\text{year}})^2}} \quad [2]$$

$$WTA^{\circ} = \begin{cases} \sqrt{VF} \times WTA & \text{for SRF} \\ VF \times WTA & \text{for non SRF} \end{cases} \quad [3]$$

$$WSI = \frac{1}{1 + e^{-6.4 \cdot WTA^{\circ} \left(\frac{1}{0.01} - 1 \right)}} \quad [4]$$

WSI analysis through BDA system in N-HyDAA is carried out for all 80 districts in the Peninsular Malaysia. Consequently, the constructed district-based WSI are mapped for the respective districts and time horizon for the Peninsular Malaysia, as shown in Figure 10. As to determine the safety level of water yield-water demand, WSI is divided into five (5) stress categories, as shown in Table 1.

Table 1. Water Stress Index (WSI) stress categories.

Stress Category	WSI
Low	< 0.1
Medium low	0.1 – 0.2
Moderate	0.2 – 0.5
High	0.5 – 0.8
Extremely high	> 0.8

As shown in Figure 10, it can be observed that most WSI in high and extremely high categories are located in the West Coast of Peninsular Malaysia by year 2030, especially in urbanised and highly populated areas, such as in Penang, Klang Valley and Johor Bahru, as well in irrigation schemes areas, such as Muda Irrigation Scheme (MADA) in Kedah, Kemubu Irrigation Scheme (KADA) in Kelantan and Barat Laut Selangor Integrated Agricultural Development Project (IADA BLS). Under the average of 14 scenarios, for example, the high WSI in Sabak Bernam, Selangor would affect the water supply sustainability at IADA BLS, while the extremely high WSI in Johor Bahru is associated with high water demand due to highly populated and rapid development that would undoubtedly bring challenge in supplying sufficient treated water to consumers.

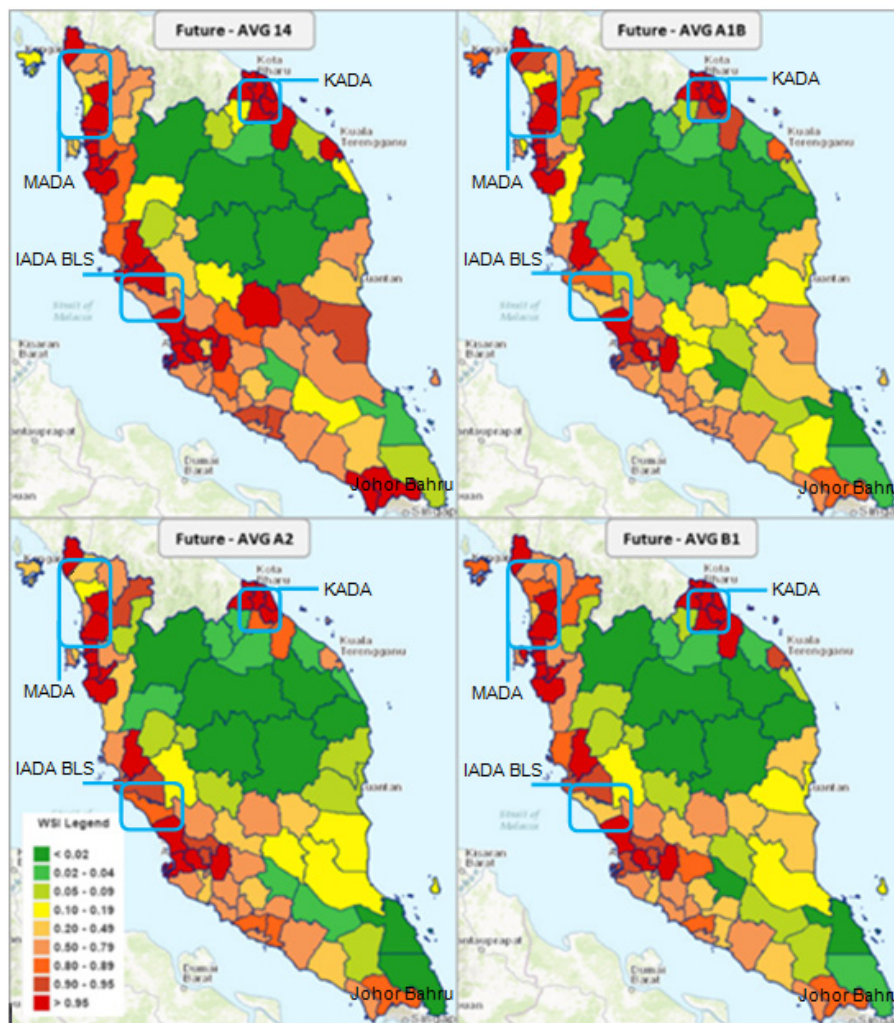


Figure 10. Water Stress Index (WSI) results for each scenario based on the districts for year 2030.

5 CONCLUSIONS

Climate change studies by various agencies all around the world are in fact, applying the big data technology analytics in assessing and projecting climate variability in all its complexity and dynamic processes. Concurrently, the field of hydro-climate informatics, sciences and computational sustainability are rapidly growing and continuously changing. The uncertainties of future climate, its impacts to water resources and environment can be quickly processed, analysed and projected as key information in addressing and managing future water-related risks. Through data linkages, data mining and analysis, and predictive analytics, decision making processes are improved and thus, can increase situational awareness amongst

high level decision makers, implementing agencies and local stakeholders in reinforcing our national policy on climate change, water management and disaster risk reduction for sustainable and resilient water resources, such as National Policy on Climate Change and National Water Resources Policy. Possible changes, intensification and impacts to future water resources vulnerability and risk of extreme flood and drought events are identified and visualized through analytics on the hydroclimate projections and development of CCF and WSI. More occurrences of water related events with higher magnitudes are expected in the future, as indicated by CCF values more than 1.9 at wider areas, decreasing trend in water yield and alarming WSI categories of high and extremely high water stress in urbanised and highly populated areas, such as in Klang Valley and Johor Bahru.

Application and enhancements of N-HyDAA in the future, such as incorporating crowd-sourcing inputs, are hugely beneficial, particularly in providing comprehensive monitoring and evaluation system regarding to climate and water related risk management. Besides, with the automated and systematic process, big data technology and analytics have reduced the current manual process by humans and improves quality and efficiency in mainstreaming climate change for a sustainable and resilient future.

ACKNOWLEDGEMENTS

The authors would like to thank Malaysian Administrative Modernisation and Management Planning Unit (MAMPU) and Malaysia Digital Economy Corporation (MDEC) for the opportunity to involve in the BDA Proof of Concept Project, MIMOS Berhad for providing the accelerating computing platform, Mi-Galactica, and the Ministry of Natural Resources and Environment for funding the development of N-HyDAA.

REFERENCES

- Amin, M.Z.M., Shaaban, A.J., Ercan, A., Ishida, K. Kavvas, M.L., Chen, Z.Q. & Jang, S. (2017). Future Climate Change Impact Assessment of Watershed Scale Hydrologic Processes in Peninsular Malaysia by a Regional Climate Model Coupled with a Physically-Based Hydrology Model. *Science of the Total Environment*, 575, 12-22.
- Armbruster, W. & MacDonell, M. (2015). Big Data for Big Problems - Climate Change, Water Availability and Food Safety, In *EnvirolInfo and ICT for Sustainability 2015*, Atlantis Press, DOI:10.2991/ict4s-env-15.2015.22.
- Beth, T., Bessie, S., Ryan, B. & Christiaan, A. (2015). *Chapter Disaster Risk Reduction: Big Data in the Disaster Cycle: Overview of Use of Big Data and Satellite Imaging in Monitoring Risk and Impact of Disasters*, UN Development Report 2015.
- Brown, A. & Matlock, M.D. (2011). *A Review of Water Scarcity Indices and Methodologies*, White Paper #106, The Sustainability Consortium, University of Arkansas.
- Data-Pop Alliance (2015). Big Data for Resilience: Realising the Benefits for Developing Countries. *Synthesis Report*.
- Department of Irrigation and Drainage (DID). (2015). *December 2014 Kelantan's Flood Report*, Department of Irrigation and Drainage Malaysia.
- Emmanouil, D. & Nikolaos, D. (2015). Big Data Analytics in Prevention, Preparedness, Response and Recovery in Crisis and Disaster Management, *Proceedings of the 29th International Conference on Computers Series: Recent Advances in Computer Engineering Series*, 32, 476-482.
- Frey, J.G., Brewer, S. & Bird, C.L. (2016). *Internet of Food Things, IT as a Utility Network+*, UK Food Standards Agency for England.
- Fowler, H.J., Blenkinsop, S. & Tebaldi, C. (2007). Review – Linking Climate Change Modelling to Impact Studies: Recent Advances in Downscaling Techniques for Hydrological Modelling. *International Journal of Climatology*, 27(12), 1547-1578
- Intergovernmental Panel on Climate Change (IPCC). (2013). *Climate Change 2013: The Physical Science Basis*, Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge Univ. Press, New York.
- Ishida, K. & Kavvas, M.L. (2017). Climate Change Analysis on Historical Watershed-Scale Precipitation by Means of Long-Term Dynamical Downscaling. *Hydrological Process*, 31(1), 35-50.
- Kavvas, M.L., Chen, Z.Q., Ohara, N., Shaaban, A.J. & Amin, M.Z.M. (2007). Impact of Climate Change on the Hydrology and Water Resources of Peninsular Malaysia, *Proceedings of International Congress on River Basin Management 2007*, 528-537.
- Kitchin, R. (2013). Big Data and Human Geography: Opportunities, Challenges and Risks. *Dialogues in Human Geography*, 3(3), 262-267, DOI: 10.1177/2043820613513388.
- MIMOS (2016). MIMOS Query Accelerator (Mi-Galactica). Technology Fact Sheet, MIMOS Berhad. (Online) Available: http://www.mimos.my/wp-content/uploads/2015/07/Fact_Sheet_Mi-Galactica_001-0421A.pdf
- Munich Re. (2016). Website of Münchener Rückversicherungs-Gesellschaft, Geo Risks Research, NatCatSERVICE. (Online). Available: <https://www.munichre.com/en/reinsurance/business/non-life/natcatservice/>.

- National Hydraulic Research Institute of Malaysia (NAHRIM). (2014). *Extension Study of the Impacts of Climate Change on the Hydrologic Regime and Water Resources of Peninsular Malaysia*, National Hydraulic Research Institute of Malaysia, 429.
- National Hydraulic Research Institute of Malaysia (NAHRIM). (2013). *Technical Guide No.1: Estimation of Future Design Rainstorm under the Climate Change Scenario in Peninsular Malaysia*, National Hydraulic Research Institute of Malaysia, 59.
- Namrata, B. (2017). Emerging Role of Big Data for Understanding and Mitigating Climate Change Risks. *12th Annual MAFSM Conference 2016*, Maryland Association of Floodplain and Stormwater Managers. (Online) Available: <http://www.mafsm.org/MAFSM/wp-content/uploads/2017/01>.
- Pal, K. (2015). How Big Data and Predictive Analytics Can Help Manage Climate Change. (Online). Available: <http://www.kdnuggets.com/2015/12/big-data-predictive-analytics-climate-change.html>.
- Pfister, S., Koehler, A. & Hellweg, S. (2009). Assessing the Environmental Impacts of Freshwater Consumption in LCA. *Environmental Science & Technology*, 43, 4098-4104.
- Shaaban, A.J., Amin, M.Z.M., Chen, Z.Q. & Ohara, N. (2010). Regional Modeling of Climate Change Impact on Peninsular Malaysia Water Resources. *Journal of Hydrologic Engineering*, 16(12), 1040-1049.
- United Nations Global Pulse (2017). Data for Climate Action – An Open Innovation Challenge to Channel Big Data for Climate Solutions. (Online) Available: <http://www.unglobalpulse.org/data-for-climate-action>.