

FORECASTING FLOOD LEVELS AND AREAL INUNDATION IN DOWNSTREAM  
MEKONG USING REMOTELY SENSED DATA AND MODELING

by  
Chi-Hung Chang

A dissertation submitted to the Department of Civil and Environmental Engineering,  
Cullen College of Engineering  
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in Geosensing Systems Engineering and Sciences

Chair of Committee: Professor Hyongki Lee, Ph.D.

Committee Member: Professor Craig L. Glennie, Ph.D.

Committee Member: Professor Faisal Hossain, Ph.D.

Committee Member: Professor Hong-Yi Li, Ph.D.

Committee Member: Professor Saurabh Prasad, Ph.D.

University of Houston  
August 2021

Copyright 2021, Chi-Hung Chang

## **DEDICATION**

*To my family.*

## ABSTRACT

Inhabitants in the Mekong River Basin (MRB) floodplain, mainly inside Cambodia and Vietnam, have been extensively relying on resources from the river as their major food and income source by developing floodplain agriculture and freshwater fishery. However, despite the potential impact of increasing extents of climate and anthropogenic changes on the MR hydrology, which can particularly influence the livelihoods of people in the downstream MRB floodplain, there is no publicly and routinely issued water level forecast inside the Mekong Delta (MD), neither inundation extent forecast for the whole MRB floodplain. This may be because of (1) the concerns of heavy computational burden and limited accuracy of conventional approaches due to the complex hydraulic conditions and flat terrain in the region, and (2) less effective data exchange between countries due to geopolitical barriers. This dissertation introduces a question: How can we build skillful, computationally efficient, and sustainable water level and inundation extent forecasting systems for the MRB, specifically for the downstream floodplain?

To answer the question, we have proposed: Chapter 3 - A freely accessible, computationally efficient daily water level forecasting system for the MR, which particularly addresses the challenges in the MD region; Chapter 4 - A satellite imagery-based inundation extent forecasting framework, called Forecasting Inundation Extents using Rotated empirical orthogonal function analysis (FIER), with the Tonle Sap Lake Floodplain as a test bed. It allows quick and continuous estimation of inundation extents of any time with available hydrological data and also addresses the concerns of heavy computational burden and extreme overestimation issues in the conventional inundation

forecasting approaches. In Chapter 5, we further implemented FIER to the whole MRB floodplain, where conventional inundation extent forecasting approaches are quite challenging to be applied. The FIER pseudo-forecasted inundation extents were then applied to spatially predict flood hazard levels and rice paddies at risk which can serve as references for local stakeholders to do more efficient decision making for better flood damage containment. The systems we developed utilize remote sensing data and are based on computationally efficient methods and can be easily implemented on cloud-based platform with enhanced scalability and accessibility.

## TABLE OF CONTENTS

<b>DEDICATION .....</b>	<b>iii</b>
<b>ABSTRACT .....</b>	<b>iv</b>
<b>TABLE OF CONTENTS .....</b>	<b>vi</b>
<b>LIST OF TABLES.....</b>	<b>ix</b>
<b>LIST OF FIGURES.....</b>	<b>xi</b>
<b>1. INTRODUCTION.....</b>	<b>1</b>
1.1 Statement of Problem .....	1
1.2 Measuring Inland Water from Space.....	3
1.2.1 Satellite Altimetry for Inland Water.....	3
1.2.2 SAR Observations for Inland Water Detection .....	5
1.3 Motivation and Objectives.....	6
1.4 Contributions .....	7
1.5 Document Structure.....	10
<b>2. BACKGROUND.....</b>	<b>11</b>
2.1 Satellite Altimetry.....	11
2.1.1 History of Satellite Altimetry Missions.....	11
2.1.2 Principle of Satellite Altimetry Measurements .....	12
2.1.3 Waveform and Retracking Methods.....	13
2.2 SAR Imaging .....	17
2.2.1 The RAR System.....	17
2.2.2 The SAR System and Processing .....	19
<b>3. A MODEL-AIDED SATELLITE-ALTIMETRY-BASED FLOOD FORECASTING SYSTEM FOR MEKONG RIVER.....</b>	<b>25</b>
3.1 Introduction .....	25
3.2 Data.....	27
3.2.1 Jason-2 Altimetry Derived River Levels.....	27
3.2.2 Modeled Discharges at Virtual Stations .....	29
3.2.3 In-situ Water Levels .....	31
3.3 Methods .....	31

3.3.1	Daily Water Level Reconstruction at Virtual Stations .....	32
3.3.2	Forecasting Models.....	35
3.3.3	Model Evaluation Statistics .....	38
3.4	Results and discussions .....	40
3.4.1	Cross-comparison with Current System .....	40
3.4.2	Flood Forecasting in the Middle Reach of the MR and CF.....	43
3.4.3	Flood Forecasting in the MD.....	50
3.5	Conclusions and Perspectives.....	56
<b>4.</b>	<b>HINDCAST AND FORECAST OF DAILY INUNDATION EXTENTS USING SATELLITE SAR AND ALTIMETRY DATA WITH ROTATED EMPIRICAL ORTHOGONAL FUNCTION ANALYSIS: CASE STUDY IN TONLE SAP LAKE FLOODPLAIN</b>	<b>60</b>
4.1	Introduction .....	60
4.2	Data.....	63
4.2.1	Sentinel-1 SAR Data .....	63
4.2.2	MODIS Surface Reflectance Data and Yearly Water Mask .....	65
4.2.3	MERIT DEM.....	67
4.2.4	Jason Altimetry-derived and In-situ Water Levels at TSL.....	67
4.2.5	ENSO Index – MEI .....	69
4.3	Methods .....	70
4.3.1	REOF-based Daily Inundation Extent Estimation Framework .....	70
4.3.2	Long-term Forecasting of TSL Levels Using ENSO Index .....	83
4.3.3	MODIS-derived Inundation Maps for Cross-comparison.....	85
4.3.4	Framework Skill Evaluation Statistics .....	89
4.4	Results and Discussions.....	91
4.4.1	Evaluation and Analysis of Framework Skills Using Long-term Historical Data .....	91
4.4.2	Evaluation of Skills of Forecasted Inundation Extents .....	108
4.5	Conclusions .....	115
<b>5.</b>	<b>FORECASTING INUNDATION EXTENTS USING ROTATED EMPIRICAL ORTHOGONAL FUNCTION ANALYSIS IN MEKONG RIVER BASIN FLOODPLAIN .....</b>	<b>119</b>

5.1	Introduction .....	119
5.2	Data.....	122
5.2.1	Sentinel-1 SAR Data .....	122
5.2.2	MODIS Surface Reflectance Data and Yearly Water Mask .....	124
5.2.3	NASA GSFC NRT Global Flood Maps .....	126
5.2.4	MRC In-situ Water Levels .....	127
5.2.5	Jason-3 Altimetry-derived Water Levels.....	128
5.2.6	Modeled Discharges at Virtual Station.....	130
5.2.7	MERIT DEM.....	131
5.3	Methods .....	132
5.3.1	FIER Framework .....	132
5.3.2	Iterative Threshold Calibration and Water Classification .....	137
5.3.3	Daily Water Level Forecasting.....	140
5.3.4	Framework Skill Evaluation Statistics .....	146
5.3.5	Inundation Depth Estimation.....	147
5.4	Results and Discussions.....	148
5.4.1	FIER-hindcasted Inundation Extents in 2020.....	148
5.4.2	FIER Pseudo-forecasted Inundation Extents in 2020.....	154
5.5	Application: Pseudo-forecasted Flood Risk in 2020 .....	158
5.6	Conclusions .....	164
<b>6.</b>	<b>CONCLUSION AND FUTURE WORK .....</b>	<b>168</b>
6.1	Conclusion .....	168
6.2	Future Works .....	171
	<b>REFERENCES .....</b>	<b>180</b>

## LIST OF TABLES

<b>Table 3.1</b> The in-situ gauges used in this study. Data were collected at 6 locations in the middle reach of the MR, 3 locations in the CF, and 13 locations in the MD. ....	31
<b>Table 3.2</b> Dominant ocean tide frequencies. ....	37
<b>Table 3.3</b> Cross-comparison with the RFMMC’s 5-day forecasting in the wet season. ....	42
<b>Table 3.4</b> Average of years (2011 - 2015) of ARs (%) of our forecasting system at locations from the middle reach of MR, and CF upstream of MD, and the corresponding RFMMC satisfactory benchmarks (cm).....	48
<b>Table 4.1</b> Climatological monthly averages and STDs of CSI, omission and commission error, overall accuracy, and statistics of altimetry-derived TSL levels including average level and RMSE. In-situ average TSL levels and STDs are shown as well. ....	94
<b>Table 4.2</b> Climatological monthly averages and STDs of CSI, omission and commission error (OE, and CE, respectively) and overall accuracy (OA) with high-elevation inundated pixels being excluded. Change of framework skills are also listed.....	100
<b>Table 4.3</b> Temporal correlation coefficients and corresponding P-values (in the bracket) of the number of high-elevation inundated pixels with the change of CSI, omission error, commission error, and overall accuracy.....	102
<b>Table 4.4</b> Climatological monthly averages and STDs of the number of high-elevation inundated pixels, and change of CSI, omission, commission error and overall accuracy. Correlation coefficients and corresponding P-values (in the bracket) between STDs are also listed. ....	103
<b>Table 4.5</b> Climatological monthly averages and STDs of the number of high-elevation inundated pixels, differences between elevations of such inundated pixels and contemporary TSL levels and their products, called extents of influence. ....	106
<b>Table 4.6</b> Climatological monthly averages and STDs of CSI, omission and commission error, and overall accuracy of our forecasting from January to July 2019. ....	112

<b>Table 4.7</b> Climatological monthly averages and STDs of CSI, omission and commission error (OE, and CE, respectively), and overall accuracy (OA) of our forecasting from January to July 2019 with high-elevation inundated pixels excluded. Differences with original framework skills are also listed.....	112
<b>Table 4.8</b> Climatological monthly averages and STDs of CSI, omission and commission error, and overall accuracy of our forecasting from January to July 2019, which were obtained by cross-comparing with inundation extents directly estimated from updated Sentinel-1A/-1B SAR imagery with K-means clustering algorithm.....	114
<b>Table 5.1</b> Dominant ocean tide frequencies. ....	143
<b>Table 5.2</b> Skills of the FIER hindcast in October 2011 cross-compared with MODIS-based inundations maps. ....	153
<b>Table 5.3</b> Definition of hazard levels categorized by inundation depths (modified from Balica et al., 2014).....	162
<b>Table 6.1</b> Implementation of FIER that have been tested in areas other than the MRB. ....	172

## LIST OF FIGURES

<b>Figure 2.1</b> Illustration of the principle of satellite altimetry for inland hydrology studies (AVISO+, <a href="https://www.aviso.altimetry.fr/en/applications/hydrology-and-land.html">https://www.aviso.altimetry.fr/en/applications/hydrology-and-land.html</a> ).....	13
<b>Figure 2.2</b> Illustration of the evolution of emitted radar pulse, illuminated area, and recorded waveform (AVISO+, <a href="https://www.aviso.altimetry.fr/en/techniques/altimetry/principle/pulses-and-waveforms.html">https://www.aviso.altimetry.fr/en/techniques/altimetry/principle/pulses-and-waveforms.html</a> ).....	14
<b>Figure 2.3</b> Schematic plot of the structure of OCOG retracker (Lee, 2008). .....	16
<b>Figure 2.4</b> Illustration of the geometry of a RAR system. ....	18
<b>Figure 3.1</b> Distribution of VSs (blue triangle) and in-situ gauges (red triangle) in (a) the middle reach of the MR and CF, and (b) the MD (within colored area) used in this study. ....	28
<b>Figure 3.2</b> Jason-2 derived water level time series and validation. (a) and (b) are results at Jason-2 VSs in the MR mainstem and (c) is for the TSL. ....	29
<b>Figure 3.3</b> Rating curves in logarithmic scale (top panel) and original scale (bottom panel). ....	34
<b>Figure 3.4</b> Reconstructed daily water levels at three VSs. JVS P179-01 and JVS P001-01 are located in the upstream MR mainstem, while JVS P001-02 is located in the TSL.....	34
<b>Figure 3.5</b> Flowcharts of (a) building forecasting model considering water flow from the MR (and the TSL if necessary, depends on forecasting location), and (b) building tidal influence model. The reconstructed daily water levels at upstream VSs and reconstructed daily water levels at the TSL are the same as those in (a). ....	37
<b>Figure 3.6</b> Flowchart of conducting water level forecast.....	38
<b>Figure 3.7</b> The MAEs and NSEs of our system at 6 locations in the middle reach of the MR and 3 locations in the CF. ....	47
<b>Figure 3.8</b> The yearly averages and STDs of MAEs and NSEs of 1-day to 5-day forecasting at 6 locations in the middle reach of the MR and 3 locations in the CF. ....	47
<b>Figure 3.9</b> The correlation coefficients of our forecasted water levels with in situ observations at 6 locations in the middle reach of the MR and 3 locations in the CF. ....	48

<b>Figure 3.10</b> Time-series of the 5-day forecasted and in situ water levels at 6 locations in the middle reach of the MR and 3 locations in the CF. ....	49
<b>Figure 3.11</b> The ARs of our forecasting at 6 locations in the middle reach of the MR and 3 locations in the CF. Different lines indicate the ARs in the case of different benchmarks. Note that the benchmarks are different in two regions since they have different levels of forecasting skills). ....	49
<b>Figure 3.12</b> The MAEs and NSEs of our system at 13 locations in the MD. ....	51
<b>Figure 3.13</b> Yearly averages and STDs of the MAEs and NSEs of 1-day to 20-day forecasting at 13 locations in the MD. ....	52
<b>Figure 3.14</b> The correlation coefficients of our forecasted water levels with in situ observations at 13 locations in the MD. ....	53
<b>Figure 3.15</b> The 10-day forecasted (orange) and in situ (blue) water levels from 2011 to 2015 at Chau Doc/Tan Chau and from 2011 to 2012 at the other 11 locations in the MD. The extreme peak water levels (in red boxes) at Chau Doc, Tan Chau, Vam Nao and Cho Moi in 2011 wet season and the strong oscillations (green boxes) at Chau Doc and Tan Chau in June 2015 can be clearly seen in the plots right next to full time series. ....	54
<b>Figure 3.16</b> The ARs of our forecasting at 13 locations in the MD. Different lines indicate the AR in the case of different benchmarks. ....	56
<b>Figure 4.1</b> Frames of Sentinel-1 GRDH intensity images used in this study. The image acquired on April 3 <sup>rd</sup> , 2016 was used as an example. ....	65
<b>Figure 4.2</b> (a) Geographical locations of Jason altimetry satellite ground track passing through the TSL and the corresponding VS and in situ gauge at Kampong Luong, (b) Jason-1/-2/-3 concatenated altimetry-derived water levels at the TSL and comparison with in situ data. ....	69
<b>Figure 4.3</b> Eigenvalues and corresponding sampling errors calculated by the rule of thumb of North et al. (1982) and cumulative percentages of explained variances of significant modes in the cases of using 500 m spatial resolutions of Sentinel-1A GRDH images as input. ....	73
<b>Figure 4.4</b> (a) to (c) are examples of Sentinel-1A GRDH images within the TSW with MERIT DEM elevation below 23 m at	

	different acquisition times. (d) is an example of images with permanent water bodies masked out. ....	75
<b>Figure 4.5</b>	Mode-1 to mode-4 RSMs (top), corresponding RTPCs, and percentages of explained variance (bottom) of input multi-temporal stack of Sentinel-1 GRDH intensity images at 500 m spatial resolution. ....	76
<b>Figure 4.6</b>	Polynomial regression models between altimetry-derived water levels w.r.t. WGS84 ellipsoid at TSL and the mode-1 to mode-4 RTPCs of Sentinel-1 GRDH intensity images at 500 m spatial resolution. ....	77
<b>Figure 4.7</b>	Percentages of pixels of each SAR image whose original and synthesized intensity are both negative. ....	79
<b>Figure 4.8</b>	Example of (a) original SAR intensity, (b) synthesized SAR intensity, and (c) difference by subtracting (a) from (b) where there are large positive difference values in high-elevation areas (red circles in (c)). ....	80
<b>Figure 4.9</b>	Example of (a) original SAR intensity, (b) synthesized SAR intensity, and (c) difference by subtracting (a) from (b) where there are large positive difference values in the areas around the boundary of TSLF (red circles in (c)). ....	80
<b>Figure 4.10</b>	Example of (a) original SAR intensity, (b) synthesized SAR intensity, and (c) difference by subtracting (a) from (b). Red circles mark areas where there are really small negative difference values in (c). ....	80
<b>Figure 4.11</b>	Illustration of how data were clustered into non-inundated (blue) and inundated (red) clusters by K-means algorithm. (a) is an example in the dry season while (b) is in the wet season. ....	82
<b>Figure 4.12</b>	Flowchart of the proposed REOF-based daily inundation extent estimation framework. (*Altimetry-derived TSL levels were used to build regression models with RTPCs. #Altimetry-derived TSL levels were used for estimated RTPCs.) ....	83
<b>Figure 4.13</b>	Illustration of the data used to build regression models between MEIs and TSL levels. ....	85
<b>Figure 4.14</b>	Offsets between datums of TSL water levels, which include the WGS84 ellipsoid, local zero gauge of in situ TSL water level data at Kampong Luong, and the EGM96 geoid. ....	88
<b>Figure 4.15</b>	Flowchart for generation of MODIS-derived inundation extents (* Time span of data could be different	

depending on the time of estimated inundation extents to be validated). .....	88
<b>Figure 4.16</b> A $2 \times 2$ confusion matrix, which displays the number of pixels that is <i>hit</i> , <i>false alarm</i> , <i>miss</i> , or <i>correct negative</i> .....	89
<b>Figure 4.17</b> (a) Climatological monthly variation of CSIs and RMSEs for altimetry-derived TSL levels. Corresponding scatter plots with fitted linear regression models are in (b).....	95
<b>Figure 4.18</b> Climatological monthly variation of (a-1) CSIs and altimetry-derived TSL levels, (b-1) omission errors and altimetry-derived TSL levels, and (c-1) CSIs and omission errors. (a-2) to (c-2) are corresponding scatter plots with fitted regression models. ....	95
<b>Figure 4.19</b> Climatological monthly variation of (a-1) commission errors and altimetry-derived TSL levels, and (b-1) CSIs and commission errors. (a-2) to (b-2) are corresponding scatter plots with fitted regression models. ....	96
<b>Figure 4.20</b> Climatological monthly variation of omission errors and rapidity of changes of altimetry-derived TSL levels.....	98
<b>Figure 4.21</b> (a) Climatological monthly variation of omission errors and RMSEs of altimetry-derived TSL levels. Corresponding scatter plots with fitted linear regression models are in (b).....	98
<b>Figure 4.22</b> Examples of inundation with elevations higher than contemporary TSL levels (yellow). Areas with elevations lower than TSL levels are shown in blue, while green is the background color. ....	99
<b>Figure 4.23</b> Climatological monthly variations of difference of CSIs and difference of omission errors when excluding inundation with elevation higher than contemporary TSL levels (a). Corresponding scatter plot is in (b).....	100
<b>Figure 4.24</b> Climatological monthly variations of difference of CSIs and difference of commission errors when excluding inundation with elevation higher than contemporary TSL levels (a). Corresponding scatter plot is in (b).....	101
<b>Figure 4.25</b> Time series of the number of high-elevation inundated pixels for each month from 2003 to 2015. ....	104
<b>Figure 4.26</b> Climatological monthly variations of difference of CSIs, omission errors, and commission errors with extents of influence of high-elevation inundated pixels with order of (a-1) to (c-1), respectively. Corresponding scatter	

plots with fitted linear regression models are in (a-2) to (c-2).....	106
<b>Figure 4.27</b> Evolution of inundation extents for the hydrological year 2011 (May 2011 to April 2012). Inundation extents on the 15 <sup>th</sup> of each month are shown.....	107
<b>Figure 4.28</b> Evolution of inundation extents for the hydrological year 2015 (May 2015 to April 2016). Inundation extents on the 15 <sup>th</sup> of each month are shown.....	107
<b>Figure 4.29</b> Correlation coefficients between TSL levels of each date and monthly MEIs in the past 12 months. ....	110
<b>Figure 4.30</b> (a) Months of MEIs which have the highest adjusted R <sup>2</sup> for linear regression models with TSL levels on the date when forecasting is performed and the corresponding lead time. The highest adjusted R <sup>2</sup> and p-value are shown in (b). ....	110
<b>Figure 4.31</b> (a) MEI-forecasted TSL levels with months of lead time which were validated by in situ data up to July 2019, and (b) MEI-forecasted TSL levels for all of 2019.....	111
<b>Figure 4.32</b> Forecasted inundation extents on the 15 <sup>th</sup> day of each month in 2019. ....	115
<b>Figure 5.1</b> Example of Sentinel-1A VV-polarized GRDH intensity image used in this study acquired on March 13 <sup>th</sup> , 2017, covering the TCF and MD. The basemap is World Imagery provided by ArcMap 10.7.....	124
<b>Figure 5.2</b> Geographical locations of the MRC in-situ water level gauges used (Basemap is World Imagery provided by ArcMap 10.7).....	128
<b>Figure 5.3</b> Geographical location of VSs used and corresponding in-situ gauges for validation. ....	129
<b>Figure 5.4</b> Comparison of altimetry-derived water levels at (a) JVS P001-01 with Khong Chiam and (b) JVS P001-02 with Kompong Luong. ....	130
<b>Figure 5.5</b> Normalized RTPCs (with percentages of explained variance).....	133
<b>Figure 5.6</b> REOF-extracted RSM-01 to RSM-04 for multi-temporal Sentinel-1A imagery used in this study. Grey areas are permanent water bodies.....	134
<b>Figure 5.7</b> Zero-lag correlation between MRC in-situ water levels at different gauges and RTPCs. Note that MM means Mekong Mainstem, TCF means Tonle Sap Lake and	

Cambodian Floodplain, while MD represents Mekong Delta .....	136
<b>Figure 5.8</b> Hydro-to-RTPC models between MRC in-situ river levels at selected gauges and RTPCs. ....	136
<b>Figure 5.9</b> Examples of difference images in the wet season generated from (a) FIER-synthesized SAR-like images and (b) Sentinel-1 SAR images where lower difference values over inundated areas can be clearly seen. Grey areas are permanent water bodies.....	138
<b>Figure 5.10</b> (Top panel) CSIs from different threshold candidates for each calibration date, (bottom panel) calibrated monthly thresholds, for the case of (a) FIER-synthesized SAR-like images, and (b) Sentinel-1 SAR images. ....	139
<b>Figure 5.11</b> (a) Iterative definition of the minimum depth at JVS P001-01 by maximizing $R^2$ of fitted depth-discharge power-law relationship and (b) the final fitted rating used to reconstruct altimetry-derived water levels. ....	141
<b>Figure 5.12</b> Comparison of reconstructed daily water levels at (a) JVS P001-01 and (b) JVS P001-02 with in-situ water levels at Khong Chiam and Kompong Luong, respectively.....	141
<b>Figure 5.13</b> Illustration of the gap-filling process for daily JVS P001-02 water levels for pseudo forecasting. ....	144
<b>Figure 5.14</b> Reconstructed and gap-filled daily water levels at (a) JVS P001-01 and (b) JVS P001-02 along with in-situ water levels at Khong Chiam, and Kompong Luong, respectively. The original JVS P001-02 water levels with the in-situ levels at Kompong Luong are shown separately in (c). ....	145
<b>Figure 5.15</b> A $2 \times 2$ confusion matrix, which displays the number of pixels that is <i>hit</i> , <i>false alarm</i> , <i>miss</i> , or <i>correct negative</i> .....	146
<b>Figure 5.16</b> Examples of FIER-hindcasted inundation extents on the 15 <sup>th</sup> of each month for 2020 using MRC in-situ river levels as inputs. The black box in the background shows the boundary of Sentinel-1 frames. ....	149
<b>Figure 5.17</b> (a) Overall accuracies, and CSIs, (b) omission, and commission errors for the FIER-hindcasted inundation extents in 2020, compared with original SAR image-derived values.....	151
<b>Figure 5.18</b> (a) In-situ river level at Kratie in 2017 – 2020 where an abrupt drop in early March 2020 is marked in the red box and (b) percentages of false alarm pixels on March	

9 <sup>th</sup> , 2020, that have high positive RSM values (> 90% percentile).....	152
<b>Figure 5.19</b> Occurrence (%) of errors over 30 images used for cross-comparison. ....	152
<b>Figure 5.20</b> Cross-comparison between the FIER hindcast and MODIS-based inundation maps in October 2011 (FA: <i>false alarm</i> ; CN: <i>correct negative</i> ). ....	154
<b>Figure 5.21</b> Skills water level forecasting at selected gauges with up to 20-day lead time. ....	154
<b>Figure 5.22</b> FIER pseudo-forecasted inundation extents on the 15 <sup>th</sup> of July to December, 2020.....	158
<b>Figure 5.23</b> Averaged (a) overall accuracies, and CSIs, and (b) omission, and commission errors of FIER pseudo-forecasted inundation extents with 1-day to 20-day lead times. ....	158
<b>Figure 5.24</b> 20-day lead time pseudo-forecasted inundation depths on the 15 <sup>th</sup> of July to December. Grey areas are permanent water bodies. ....	160
<b>Figure 5.25</b> Rice paddies (light purple) in the Sentinel-1 frame with those pseudo-forecasted to be under flood risk on the 15 <sup>th</sup> of July to December being marked in purple. Grey areas are permanent water bodies. ....	161
<b>Figure 5.26</b> Rice paddies (light purple) in the Sentinel-1 frame with the start DOY of pseudo-forecasted flood risk areas marked as colorbar shows. Grey areas are permanent water bodies. ....	161
<b>Figure 5.27</b> Examples of 20-day pseudo-forecasted maps of flood hazard levels in November 2020 over the (a) TSLF, (b) CF, (c) northern MD, and (d) southern MD. Grey areas are permanent water bodies.....	163
<b>Figure 6.1</b> Area over the metropolitan area of Houston that FIER implementation has been tested. ....	174
<b>Figure 6.2</b> Water-related (a) RSMs, (b) normalized RTPCs and in-situ water levels, and (c) scatter plots between RTPCs and in-situ water levels over the metropolitan areas of Houston. ....	174
<b>Figure 6.3</b> Preliminary hindcast inundated extents along Buffalo Bayou in Houston in 2017 on the dates near the strike of Hurricane Harvey. ....	175

<b>Figure 6.4</b> Area over the Mississippi River near New Madrid, Missouri, that FIER implementation was tested. ....	176
<b>Figure 6.5</b> Water-related (a) RSMs, (b) normalized RTPCs and in-situ water levels, and (c) scatter plots between RTPCs and in-situ water levels over the Mississippi River near New Madrid, Missouri.....	177
<b>Figure 6.6</b> Preliminary hindcast inundation extents on (a) March 12 <sup>th</sup> , and (b) March 25 <sup>th</sup> , 2018.....	178

# 1. INTRODUCTION

## 1.1 Statement of Problem

The Mekong River (MR) is the twelfth-longest river in the world and the largest river in Southeast Asia and flows through six countries (China, Lao PDR, Myanmar, Thailand, Cambodia, and Vietnam). The importance of the MR on a regional and global scale is more than apparent from multiple aspects. Inhabitants in the MR Basin (MRB) have been extensively relying on resources from the river as their major food and income source through floodplain agriculture and freshwater fishery (Mekong River Commission, MRC, 2011), which also contribute to the global food supply.

In recent years, the extent of land-use change and the amount of water infrastructure has rapidly increased, which together with climate change-induced rainfall intensification, influences the MR hydrology by altering the amount of river flow and the extents of seasonal inundation. This can pose an impact on agricultural, and ecological systems especially in the downstream areas of MRB like the Cambodian Floodplain (CF) and Mekong Delta (MD), impacting the livelihoods of millions of people (Pokhrel et al., 2018; Shin et al., 2020; Try et al., 2020a, 2020b). Hence, flood forecasting in these areas is urgently needed.

Among all hydrology properties, water level is an indicator for flood warning. The Regional Flood Management and Mitigation Center (RFMMC) of the MRC has been regularly issuing water level forecasting at 22 gauges using a hydrological model (Pagano, 2014). However, the forecasting is only regularly issued down to the very upper part of MD near the Cambodia-Vietnam national border. Another critical quantity

for flood warning is inundation extent, which directly reflects the distribution of floodwater in space, and thus, is very important for establishing a fast response plan and disaster mitigation (Liu et al., 2017; Smith, 1997). Yet, to the best of our knowledge, there is currently no publicly available information about forecasted inundation extents issued by the MRC to the best of our knowledge. In fact, the complex hydraulic conditions in the MRB floodplain, especially in the MD, have posed challenges to water level forecasting, as a complicated model structure and heavy computational burden are required (Pagano, 2014). For inundation extent forecasting, conventional hydrodynamic modeling can be computationally impractical in fast and frequent operations due to a heavy computational loading, influencing the forecasting lead time (Liu et al., 2018; Zheng et al., 2018a, 2018b). The GIS-based strategy could provide extremely overestimated results over the low relief areas (Johnson et al., 2019) like the MRB floodplain despite being computationally efficient. On the other hand, in transboundary river basins like the MRB, even if international data-sharing agreements exist among the member countries, the process can still be complex in practice (Du et al., 2020; Gerlak et al., 2011), which is another challenge in flood forecasting in the downstream areas. In fact, even though the MRC sets a successful example in regulating transboundary water resources as well as coordinating the relationship between member countries, the data exchange procedure is still not very effective (Anh, 2021)

Therefore, computationally efficient water level and inundation extent forecasting systems that utilize remotely sensed measurements without geopolitical restrictions are necessary to address the needs of downstream countries like Cambodia and Vietnam, toward an independent and sustainable capacity for early risk assessment,

decision-making, and fast response. Previous studies using remote sensing techniques to measure water levels and inundation extents are reviewed below, with a focus on radar satellite altimetry and Synthetic Aperture Radar (SAR).

## **1.2 Measuring Inland Water from Space**

### **1.2.1 Satellite Altimetry for Inland Water**

Satellite altimetry was initially designed to provide accurate water level observations over oceans, as the shape of returning waveform is simple, allowing easy definition of signal time-of-flight and the resulting satellite range. Its application on inland water bodies is more complicated as the shape of the returning waveform is more likely to be influenced by river bank topography, lakeshore, or coastline, making the definition of time-of-flight and satellite range a challenging task. Fortunately, the difficulty has been mitigated by “waveform retracking” techniques, allowing retrieval of water levels of inland water bodies such as rivers and lakes or coastal sea levels with improved accuracy. This also opens the door for applying satellite altimetry to hydrology studies. Direct application examples are the retrieval of water levels over various types of inland water bodies (Biancamaria et al., 2017; Boergens et al., 2019; Da Silva et al., 2012; Lee et al., 2011; Sulistioadi et al., 2015; Tourian et al., 2016). Since satellite altimetry is a profiling tool providing one-dimensional along-track observations, many studies have combined altimetry-derived water levels with other satellite-based measurements for a wider range of applications. These include calibration of hydrodynamic model (Jiang et al., 2019), estimation of lake and reservoir

water volumes (Busker et al., 2019; Zhou et al., 2016), river discharges (Kim et al., 2019a, 2019b; Paris et al., 2016; Tarpanelli et al., 2019; Tourian et al., 2017) and bathymetry (Brêda et al., 2019). In the MRB, Frappart et al. (2006) combined satellite altimetry-derived water levels with inundation extents derived from satellite images for analysis of water volume changes over downstream floodplain areas including the CF and MD. More recently, Frappart et al. (2018) used long-term altimetry-derived Tonle Sap Lake (TSL) levels along with inundation extents from satellite imagery to study the connection between El Niño and Southern Oscillation (ENSO) and TSL Floodplain (TSLF) water volume changes. For the MR level estimation, a recent work by Pham et al. (2018) derived daily MR levels at the upstream of MRB by fitting regression models between satellite-observed land surface temperatures and altimetry-derived river levels. However, it focused on the upstream areas of MRB and did not address the challenges in the downstream areas. For water level forecasting, recent studies have applied satellite altimetry data for computationally efficient regression analysis-based water level forecasting in the Ganges-Brahmaputra-Meghna River Basin, a transboundary river basin (Biancamaria et al., 2011; Hossain et al., 2014a, 2014b). However, such application of satellite altimetry toward a low computational cost daily water level forecasting in the MRB, addressing the complex hydraulic conditions in its downstream areas is yet to be investigated.

### 1.2.2 SAR Observations for Inland Water Detection

SAR is considered the most useful sensor for detecting flooded areas under cloud cover (Yan et al., 2015). Its use of active microwave electromagnetic pulses allows it to penetrate clouds and be independent of illumination and atmospheric conditions. Thus, it can provide surface observations without spatial gaps both day and night under all weather conditions (Martinis et al., 2015; Pierdicca et al., 2013). Its application for depicting inundation extents began with the First ERS Thematic Working Group Meeting on Flood Monitoring in 1995, where numerous investigators presented inundation maps by using C-band ERS-1 SAR images (Smith, 1997). Studies such as Henry et al. (2006) used ENVISAT Advanced SAR (ASAR) to delineate the Elbe River flood of August 2002, while Bouvet et al. (2009) used ENVISAT ASAR data for rice crop monitoring over the MD. More recent studies have used inundation extents derived from SAR imagery for hydrological/hydraulic model calibration (Hostache et al., 2018; Wood et al., 2016). Other studies used SAR imagery with high resolution of up to 1 m and short revisit time, such as COSMO-SkyMed and TerraSAR, for urban flood mapping (Mason et al., 2010; Pierdicca et al., 2013). However, due to either the high cost of SAR imagery or its infrequent image acquisition (Markert et al., 2018), frequent, rapid, and publicly accessible flood mapping has not been realized until the launch of the European Space Agency (ESA)'s Sentinel-1 in 2014. Sentinel-1 has promoted the use of SAR imagery as it features consistent image acquisition, free accessibility, and shorter revisit time (12 days) compared to earlier launched SAR satellites, such as ERS-1/-2, ENVISAT, and ALOS. This allows continuous monitoring of ground features and their changes over time (Tsyganskaya et al., 2018a; White et al., 2014). Recently,

several studies have shown the ability of Sentinel-1 imagery to provide rapid and frequent areal flood mapping (Amitrano et al., 2018; Bioresita et al., 2018; Li et al., 2018; Twele et al., 2016). In the MRB, Pham-Duc et al. (2017) used Sentinel-1 SAR images for surface water detection and monitoring in Cambodia and the MD through neural network classification with Landsat-8 imagery as training targets. Dinh et al. (2019) used time series of Sentinel-1 images along with other optical imagery for a higher temporal resolution inundation mapping in the MD. Markert et al. (2020) used Sentinel-1 imagery for inundation mapping in Cambodia by utilizing Google Earth Engine. Tuan et al. (2021) used Sentinel-1 imagery along with ALOS-2 imagery toward a better inundation mapping in the MD. However, despite the rapid development and growing use of SAR imagery for inundation extent detection, there is a lack of investigations using SAR imagery toward cloud-free inundation extent forecasting.

### 1.3 **Motivation and Objectives**

This research aims to explore new approaches for flood forecasting, specifically for water level and areal inundation. The motivation is to address: (1) The challenges in conventional approaches, such as the heavy computational load for hydrological and hydrodynamic modeling for water level and inundation extent forecasting as well as the extreme overestimation when implementing non-modeling GIS-based inundation forecasting approaches in flat-terrain areas like the MRB floodplain; (2) Potential practical complex data-sharing process between the upstream and downstream countries in the MRB due to geopolitical barriers, which can hinder the operation of flood forecasting in the downstream.

The objective of this dissertation is to answer the research question: How can we build skillful, computationally efficient, and sustainable flood forecasting systems of the MR water levels and inundation extents for the MRB, specifically for downstream areas? Therefore, we propose novel computationally efficient water level and inundation extent forecasting approaches utilizing remotely sensed observations to mitigate geopolitical restrictions and help stakeholders in the downstream countries of MRB, including Cambodia and Vietnam, build independent and sustainable capacity of early risk assessment, more efficient decision-making and fast response for flood damage mitigation.

#### 1.4 Contributions

To fulfill the objectives listed above, we integrated multiple remote sensing data sources (satellite altimetry and SAR) and hydrological modeling for building novel flood forecasting systems. Detail contributions of this dissertation include:

- Proposing a freely accessible, computationally efficient daily water level forecasting system for the MR, which particularly addresses the challenges in the MD region, where there is no routinely issued forecasting from the MRC (Chang et al., 2019):
  - First, daily upstream MR levels at cross-sections of satellite altimetry ground tracks and the MR, called Virtual Stations (VSs), were reconstructed through rating curves generated between the model-estimated discharges and altimetry-derived MR levels. Daily altimetry-derived TSL levels at VS were reconstructed by interpolation.

- Second, computationally efficient forecasting models for downstream water levels were built by linear regression analysis considering temporal shifts between the upstream reconstructed MR water levels and downstream in-situ water levels. Influences of altimetry-derived daily TSL levels and sinusoid model-simulated tidal fluctuations were additionally considered as inputs for the forecasting from Phnom Penh to the MD.
- Proposing a satellite imagery-based inundation extent forecasting framework, called Forecasting Inundation Extents using Rotated empirical orthogonal function analysis (FIER), with the TSLF as a test bed (Chang et al., 2020). The framework allows quick and continuous estimation of inundation extents at any time with available hydrological data, and also addresses the concerns of heavy computational burden and extreme overestimation in conventional inundation forecasting approaches:
  - First, spatiotemporal patterns in multi-temporal Sentinel-1 SAR image stacks, covering the TSLF, were extracted by Rotated Empirical Orthogonal Function (REOF) analysis. The temporal patterns of SAR images were then coupled with altimetry-derived TSL levels to build regression models.
  - Second, the temporal patterns of SAR images can then be estimated by feeding altimetry-derived TSL levels into the generated regression models. The estimated temporal patterns can then be combined with the spatial patterns to synthesize SAR-like images at the time of input altimetry-derived TSL levels. To demonstrate the forecasting capacity of FIER, a TSL level forecasting system using ENSO index was also proposed.

- Finally, inundation extents were estimated by applying unsupervised K-means clustering with the aid of a Digital Elevation Model (DEM).
- Implementing FIER for the whole MRB floodplain, where the application of conventional inundation extent forecasting approaches is challenging. Applications of the continuous FIER-forecasted inundation extents in the spatial prediction of flood hazard levels and rice paddies at risk are also demonstrated. Such applications are extremely crucial for the MRB floodplain which is heavily populated with highly rice-dependent food intake and economy:
  - FIER of MRB floodplain was built by coupling spatiotemporal patterns of multi-temporal Sentinel-1 SAR imagery with gauges of in-situ water levels distributed over the floodplain.
  - By integrating FIER with the water level forecasting system that we developed (Chang et al., 2019), we obtained inundation extent forecasting over the MRB floodplain where conventional approaches have difficulties to be applied.
  - The FIER pseudo-forecasted inundation extents are used to generate corresponding inundation depths, which are commonly used for flood risk assessment. FIER's capacity for quickly estimating inundation extents is advantageous when practically performing continuous areal inundation forecast, allowing instantaneous spatial prediction of flood hazard levels and flood risk for rice cultivation.

## 1.5 Document Structure

The outline of this dissertation is shown as follows:

Chapter 2 reviews satellite radar altimetry and SAR imaging techniques; Chapter 3 presents a feasibility study for daily water level forecasting system for the MRB. The chapter is based on the work presented in Chang et al. (2019); Chapter 4 introduces a novel framework, named FIER, toward inundation extent hindcast and forecast by using the TSLF as a test bed; Chapter 5 presents the work of expanding the implementation of FIER to the whole MRB floodplain, encompassing TSLF, CF as well as MD, by integrating with the daily water level forecasting system presented in Chapter 3. Applications of the FIER pseudo-forecasted inundation extents are also demonstrated. Finally, Chapter 6 provides conclusions and future work.

## **2. BACKGROUND**

### **2.1 Satellite Altimetry**

#### **2.1.1 History of Satellite Altimetry Missions**

Satellite altimetry techniques started to emerge in 1960s following the development of artificial satellites. The Williamstown Symposium in 1969 revealed a plan for showing the feasibility of applying a single satellite equipped with a suitably designed radar altimeter in near-earth orbit for the observation of physical sea surface (Kaula, 1970).

On May 14th, 1973, Skylab 1, the very first National Aeronautics and Space Administration (NASA) space station, was launched for experimental purposes but due to limited accuracy, was far from practical use. The first altimetry mission that was really applied to inland hydrology studies was Topex/Poseidon (T/P) launched in 1992 under the cooperation of NASA and Centre National d'Études Spatiales (CNES). T/P was followed by the Jason series of altimetry missions including Jason-1, Jason-2, and the on-going Jason-3, launched in 2002, 2008 and 2016 respectively. T/P and the Jason-series of missions follow the same orbit and together have provided about 30 years of observations. The other major group of satellite altimetry missions which also provides measurements for inland hydrology studies consists of the European Remote Sensing-2 (ERS-2), ENVIRONMENTAL SATellite (ENVISAT), and Satellite with ARGOS and ALtiKa (SARAL). ERS-2 and ENVISAT, launched in 1995 and 2002, were led by the ESA and were followed by SARAL, launched in 2013, by the Indian Space Agency. These missions also share the same orbit and collect long-term measurements. More recently,

ESA has launched Sentinel-3A/-3B in 2016 and 2018, respectively. Both are equipped with a SAR radar altimeter with decimeter level accuracy. The Sentinel-3A/-3B satellites are on-air but with different orbits, resulting in denser ground tracks and more comprehensive measurements over inland water bodies. Other satellite altimetry missions such as Jason-CS/Sentinel-6, and Surface Water and Ocean Topography (SWOT) missions are to be launched in the near future (Calmant et al., 2016). In this study, measurements collected by Jason-1/-2/-3 satellite altimetry missions were used, with a conventional altimetry technique known as Low Resolution Mode. The fundamental of this mode is explained in the following sections.

### 2.1.2 Principle of Satellite Altimetry Measurements

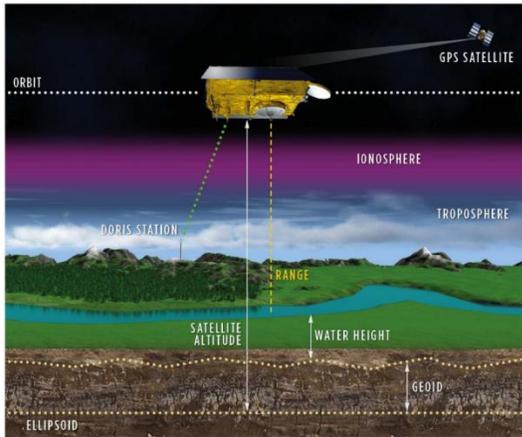
The principle of satellite altimetry measurements consists of two major components, satellite range ( $R$ ), that is the distance from the satellite to the reflective surface, and satellite altitude with respect to the reference ellipsoid ( $h_{sat}$ ). The altimeter emits a radar pulse and receives an echo from the reflective surface and records the corresponding two-way travel time ( $\Delta t$ ). By dividing  $\Delta t$  by 2 and multiplying it with the speed of light in free space ( $c = 2.997925 \times 10^8$  m/s),  $R$  can be determined (Calmant et al., 2016)

$$R = c \cdot \frac{\Delta t}{2}. \quad (2.1)$$

The height of the reflective surface with respect to the reference ellipsoid can be determined by subtracting  $R$  from  $h_{sat}$  with correction terms (Calmant et al., 2016)

$$H = h_{sat} - (R + C_p + C_g) \quad (2.2)$$

where  $C_p$  is propagation error corrections, and  $C_g$  is a geophysical correction.  $C_p$  includes dry and wet tropospheric correction and ionospheric correction. In inland hydrology studies,  $C_g$  includes solid earth tide and pole tide correction. Figure 2.1 shows an illustration of the principle of satellite altimetry.

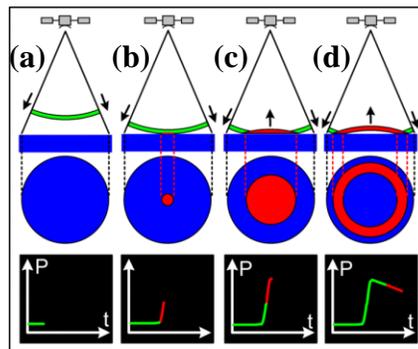


**Figure 2.1** Illustration of the principle of satellite altimetry for inland hydrology studies (AVISO+, <https://www.aviso.altimetry.fr/en/applications/hydrology-and-land.html>).

### 2.1.3 Waveform and Retracking Methods

The definition of two-way travel time ( $\Delta t$ ) is the core of the satellite altimetry technique. To determine it, the altimeter onboard the satellite records the amplitude of radar pulse echoes as a function of gates (bins). The evolution of the recorded amplitude is called a “waveform”. Since a gate (bin) corresponds to a time width, called bin width, the waveform is actually equivalent to the recorded amplitude as a function of time. Figure 2.2 is an illustration of the evolution of radar pulse transmission, illuminated area, and the corresponding waveform. Once the front edge of the radar pulse hits the

reflective surface, the amplitude of the echo starts to rise (Figure 2.2(b)) and then increases proportionally until it reaches the peak when the rear edge of the radar pulse hits the reflective surface (Figure 2.2(c)), which results in the “leading edge” of the waveform. Afterward, the amplitude of the echo starts to decay (Figure 2.2(d)), resulting in the “trailing edge” of the waveform. Hence, there should be a jump in the waveform between the epoch that the leading and trailing edge hit the reflective surface. In the ideal case, such as over the surface of oceans or open water bodies, the recorded time at the mid-point of this jump, called the Leading Edge Point (LEP), represents  $\Delta t$ , which can be accurately estimated by the on-board tracker. However, the definition of LEP by the on-board tracker does not necessarily work for inland hydrology studies as inland water bodies, such as rivers or lakes, typically have a small spatial scale. Since the illuminated area of the altimeter radar pulse on the Earth surface, called footprint, is kilometer-level wide, the recorded waveform may be interfered by not only inland water bodies but also the complex surrounding land topography. Therefore, the definition of LEP (and resulting  $\Delta t$ ) requires an additional correction called a retracking correction.



**Figure 2.2** Illustration of the evolution of emitted radar pulse, illuminated area, and recorded waveform (AVISO+, <https://www.aviso.altimetry.fr/en/techniques/altimetry/principle/pulses-and-waveforms.html>).

Since the non-ocean surfaces have different characteristics, different retracking methods have been developed to address these differences, such as threshold retracker (Davis, 1997), Offset Center Of Gravity (OCO<sub>G</sub>) retracker (Wingham et al., 1986). The ICE-1 retracker (Bamber, 1994) was developed based on the OCO<sub>G</sub> retracker, using 25% of amplitude derived from OCO<sub>G</sub> for determining  $\Delta t$ , while ICE retracker in Jason Geophysical Data Record (GDR) uses 30% of the amplitude derived from OCO<sub>G</sub> to define  $\Delta t$ . Since ICE retracker is the only retracker adopted in this study, only its details are described as below.

For OCO<sub>G</sub> retracker, the algorithm determines the Center Of Gravity (COG) position, and the corresponding width and amplitude (Wingham et al., 1986)

$$COG = \frac{\sum_{i=1+n_a}^{64-n_a} iP^2(i)}{\sum_{i=1+n_a}^{64-n_a} P^2(i)}, \quad (2.3)$$

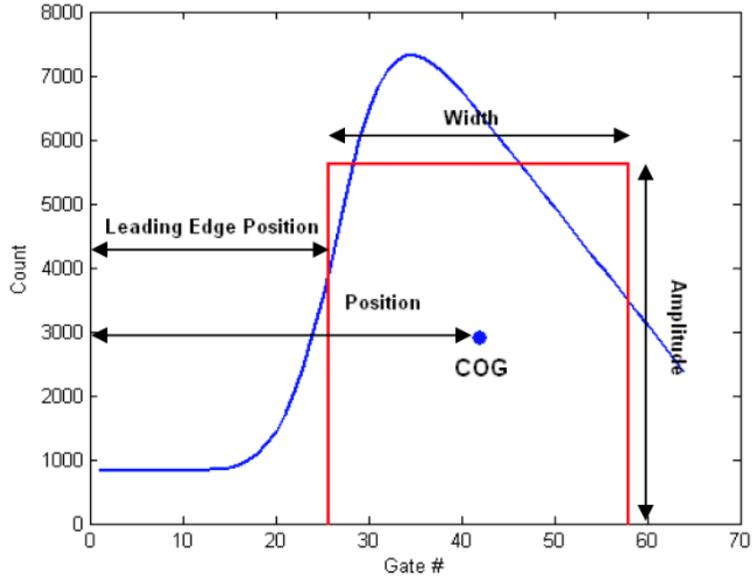
$$Amplitude = \sqrt{\frac{\sum_{i=1+n_a}^{64-n_a} iP^4(i)}{\sum_{i=1+n_a}^{64-n_a} P^2(i)}}, \quad (2.4)$$

$$\text{and } Width = \left( \frac{\sum_{i=1+n_a}^{64-n_a} P^2(i)}{\sum_{i=1+n_a}^{64-n_a} P^4(i)} \right)^2 \quad (2.5)$$

where  $P(i)$  is the value of the  $i$ -th bin from waveform samples and  $n_a$  is the number of bins. Figure 2.3 is a schematic plot of the OCO<sub>G</sub> retracker. LEP of the OCO<sub>G</sub> retracker is defined as

$$LEP = Position - \frac{1}{2} \cdot Width \quad (2.6)$$

where *Position* is the corresponding bin of the COG.



**Figure 2.3** Schematic plot of the structure of OCOG retracker (Lee, 2008).

For the ICE retracker, LEP is defined as the first bin whose corresponding recorded amplitude exceeds 30% of the amplitude of COG. It is different from the definition of OCOG retracker. The retracking correction ( $\Delta R$ ) is then determined by

$$\Delta R = (LEP - \text{Tracking bin}) \cdot c \cdot \frac{\tau}{2}. \quad (2.7)$$

Tracking bin is the bin defined by the on-board tracker, while  $\tau$  is bin width, which are 31 and 3.125 ns respectively, in the case of Jason series of satellite altimetry missions. Thus Equation 2.2 is rephrased as follows after waveform retracking process, which is more accurate for inland hydrology studies

$$H = h_{sat} - (R + C_p + C_g + \Delta R). \quad (2.8)$$

## 2.2 SAR Imaging

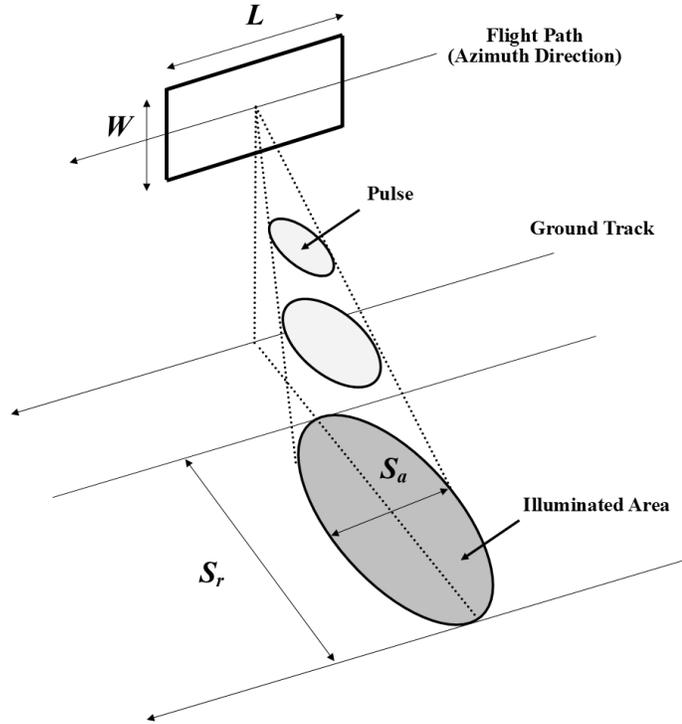
### 2.2.1 The RAR System

To explain the process of Synthetic Aperture Radar (SAR) imaging, we begin with a brief review of conventional Real Aperture Radar (RAR) systems. Radar is an abbreviation of Radio Detection And Ranging which was developed from the 1930s to 1940s for detecting ships and aircraft. In the 1950s, by mounting it on the side of an aircraft and performing “side-look” towards the ground, its application for Earth surface mapping was first discovered. The technique is called Side-Looking Airborne Radar (SLAR), which is an implementation of the RAR system (Stewart and Larson, 1999). In a RAR system, the illuminated area (swath) is defined by antenna size ( $W$ : antenna width,  $L$ : antenna length). Then the widths of swath in range ( $S_r$ ) and azimuth ( $S_a$ ) directions can be determined as (van Zyl and Kim, 2011)

$$S_r \approx \frac{\lambda h}{W \cos^2 \theta_i} \text{ and} \quad (2.9)$$

$$S_a \approx \frac{\lambda h}{L \cos \theta_i} \quad (2.10)$$

where  $\lambda$  is the radar wavelength,  $h$  is the altitude of sensor and  $\theta_i$  is the range-dependent incidence angle at the center of the swath.



**Figure 2.4** Illustration of the geometry of a RAR system.

In the range (cross-track) direction, two objects can be discriminated if the leading edge of the pulse returned from the farther object is received later than the trailing edge of the pulse returned from the closer object. Thus, the ground range resolution ( $\rho_r$ ) is

$$\rho_r = \frac{c\tau}{2 \sin\theta_i} \quad (2.11)$$

In the azimuth direction, ground azimuth resolution ( $\rho_a$ ) is given by the smallest separation between two objects that can be detected (Moreira et al., 2013), which in the RAR system is equal to the width of the swath in azimuth direction

$$\rho_a = r \cdot \Theta_a \approx \frac{\lambda h}{L \cos\theta_i} \quad (2.12)$$

where  $r$  is the slant range from sensor to target and  $\Theta_a$  is the angular beam width in the azimuth direction. Therefore, in the RAR system, ground azimuth resolution can be enhanced by either applying electromagnetic waves with a shorter wavelength, lowering sensor altitude, lengthening the antenna, or increasing the look angle. However, for a satellite-borne system, sensor altitude is always high, while the wavelength cannot be too short due to the atmospheric transmission. Moreover, the expansion of look angle and antenna length are also limited by the physical size of satellites. Such limitation in azimuth resolution is the major drawback of the RAR system and is the main motivation for the advent of the SAR system.

### 2.2.2 The SAR System and Processing

In 1951, Carl Wiley of the Goodyear Aircraft Corporation first noted that Doppler frequency analysis of signals from a moving coherent radar can be used to improve azimuth resolution. He noticed that two targets at different along-track locations will have different angles relative to the aircraft velocity vector, resulting in Doppler frequencies, which can be further adopted to distinguish targets through frequency analysis (Stewart and Larson, 1999; van Zyl and Kim, 2011). The technique is called SAR. The essence of the SAR processing algorithm is that it compresses the energy of the return signal, which is generally distributed on an image along the range and azimuth directions, to enhance the resolution in both directions. The received raw data is a two-dimensional matrix of complex numbers. In order to fulfill high range and

azimuth resolution from raw data, the Range-Doppler Algorithm (RDA) is applied. The RDA consists of several steps: (1) range compression, (2) azimuth Fast Fourier Transform (FFT), (3) Range Cell Migration Correction (RCMC), and (4) azimuth compression (Curlander and McDonough, 1991). First, the return signals are processed with matched filtering. Second, the range-compressed signals are transformed to the range-doppler domain using an azimuth FFT. Next, since the slant range from the satellite to a target varies along the azimuth direction, return signal energy from the same target may be distributed over different range cells along the azimuth direction, resulting in a quadratic shape of signal distribution called RCM. Hence, RCMC is applied to ensure these return signal energies are aligned parallel to the azimuth direction so that all return signal energy from the same target can be captured in the following azimuth compression process, which generates the Single-Look Complex (SLC) image.

To explain SAR processing, let us start with the range compression process. As described in the previous section, the ground range resolution ( $\rho_r$ ) is

$$\rho_r = \frac{c\tau}{2 \sin\theta} \quad (2.13)$$

where  $c$  is the speed of light and  $\tau$  is the effective time length of the radar pulse. Hence, the range resolution depends on the effective time length of the radar pulse. The shorter the pulse, the higher the range resolution. However, in order to keep Signal-to-Noise Ratio (SNR) high to retain the quality of the return signal, the peak power of the pulse needs to be strong enough; this is often done by increasing the pulse time length. To achieve a high range resolution without the use of a short pulse, a Linear Frequency-

Modulated (LFM) pulse, also called a “chirp”, is adopted. The chirp is then compressed to synthesize an effectively short pulse which enhances the range resolution. The process is called pulse compression or range compression as the process is done along the range direction. In fact, the range compression process is applied in both RAR and SAR systems. The difference between them is that the RAR system only applies range compression, while the SAR system applies compression in both range and azimuth directions.

The transmitted LFM pulse in the time domain  $g_t(t)$  can be expressed as

$$g_t(t) = \text{rect}\left(\frac{t}{T}\right) e^{j\pi k_r t^2} \quad (2.14)$$

where  $\text{rect}$  is a rectangular function,  $t$  is the time variable (seconds),  $T$  is transmitted pulse duration (seconds) and  $k_r$  is the LFM rate (Hz). The phase of the pulse can be computed as

$$\phi_p(t) = \pi k_r t^2. \quad (2.15)$$

The instantaneous frequency of the pulse is the derivative of phase with respect to time, it is given as

$$f_p(t) = \frac{1}{2\pi} \frac{d\phi_p(t)}{dt} = \frac{1}{2\pi} 2\pi k_r t = k_r t \quad (2.16)$$

which is a linear function of time  $t$  with slope  $k_r$ , and is the reason that it is called LFM pulse. Equation 2.16 also indicates that the transmitted pulse has a bandwidth  $B_p$  of

$$B_p = |k_r| \cdot T. \quad (2.17)$$

The return pulse which is received time  $t_0$  later is then represented as

$$g_r(t) = \text{rect}\left(\frac{t - t_0}{T}\right) e^{j\pi k_r (t - t_0)^2}. \quad (2.18)$$

The range compression process is realized by applying a matched filter to the received return LFM pulse, in which the matched filter kernel  $h(t)$  is the time-reversed, complex conjugate of the transmitted LFM pulse

$$h(t) = g^*(-t) = \text{rect}\left(\frac{t}{T}\right) e^{-j\pi k_r (-t)^2} = \text{rect}\left(\frac{t}{T}\right) e^{-j\pi k_r t^2}. \quad (2.19)$$

In the time domain, matched filtering process (pulse compression) is the convolution of the received return pulse and matched filter, which in the case of range compression is written as

$$g_{rc}(t) = g_r(t) \otimes h(t) = T \text{sinc}\{k_r T(t - t_0)\} \quad (2.20)$$

where  $\otimes$  is the convolution operator. After pulse compression, the pulse width has been shortened to  $1/B_r$ . Then, we obtain the ground range resolution ( $\rho_r$ ) (Curlander and McDonough, 1991) as

$$\rho_r = \frac{c}{2 B_r \sin\theta}. \quad (2.21)$$

To explain the Doppler effect in the SAR context, we can think about the change of frequency of the received pulse from a point target as the satellite is flying in the azimuth direction. If the slant range from the antenna to the point target is decreasing,

the frequency of the received signal increases. On the other hand, if the slant range from the antenna to the point target is increasing, the frequency of the received signal decreases. This frequency, governed by the relative speed of the sensor and the target, is called SAR Doppler frequency.

Let us make the slant range  $r$  between the sensor and the target a function of time  $t_s$  along the azimuth direction

$$r(t_s) = \sqrt{r_0^2 + v^2 t_s^2} \quad (2.22)$$

where  $r_0$  is the slant range between the sensor and the target along the perpendicular of the sensor flight line and  $v$  is the satellite velocity. Then the phase of the radar signal  $\phi_s$  would be

$$\phi_s(t_s) = -\frac{4\pi r(t_s)}{\lambda} = -\frac{4\pi}{\lambda} \sqrt{r_0^2 + v^2 t_s^2} \approx -\frac{4\pi r_0}{\lambda} - \frac{2\pi v^2}{r_0 \lambda} t_s^2. \quad (2.23)$$

The instantaneous frequency of the signal then can be derived as

$$f_s(t_s) = \frac{1}{2\pi} \frac{d\phi_s(t_s)}{dt} = -\frac{2v^2}{r_0 \lambda} t_s = k_s \cdot t_s. \quad (2.24)$$

Therefore, the radar signal in azimuth direction is also an LFM signal (chirp) with slope

$k_s = -\frac{2v^2}{r_0 \lambda}$ . To determine the bandwidth of this LFM signal in the azimuth direction,

the total illuminated time ( $T_{tot}$ ) that the target will be in the antenna beam needs to be defined first, and can be represented as

$$T_{tot} = \frac{\lambda r_0}{Lv}. \quad (2.25)$$

By substituting the right side of Equation 2.17 with  $k_s$  and  $T_{tot}$  we obtain the bandwidth of this LFM signal in the azimuth direction, called Doppler bandwidth, as

$$B_D = |k_s| \cdot T = \frac{2v^2}{r_0\lambda} \cdot \frac{\lambda r_0}{Lv} = \frac{2v}{L}. \quad (2.26)$$

This shows that if the radar signal is compressed in the azimuth direction (azimuth compression) using a matched filter similar to the case of range compression, the resulting signal will have a width in time of  $1/B_D$ . Since the sensor is moving with velocity  $v$ , it leads to an azimuth resolution of

$$\rho_a = \frac{v}{B_D} = \frac{L}{2}. \quad (2.27)$$

As Equation 2.27 shows, the azimuth resolution of SAR is finer when the antenna length decreases, which is opposite to RAR. This is a great advantage for the SAR system over conventional RAR, allowing realization of high-resolution satellite-borne SAR system. However, there is a limitation that antenna size cannot be too short as it degrades the SNR.

### **3. A MODEL-AIDED SATELLITE-ALTIMETRY-BASED FLOOD FORECASTING SYSTEM FOR MEKONG RIVER**

#### **3.1 Introduction**

This chapter is based on the content of Chang et al. (2019). In recent years, several studies have indicated increases in annual and seasonal river discharges on the Mekong River (MR) under future climate change, which results in increased flood risks during the wet season on the Cambodian and Vietnamese floodplain (Hoang et al., 2016; Hoanh et al., 2010; Västilä et al., 2010). Robust and sustainable flood forecasting can be a great aid in decision-making for flood risk reduction and response, making the Mekong River Basin (MRB) more resilient when facing an uncertain future. The Regional Flood Management and Mitigation Center (RFMMC) affiliated with the Mekong River Commission (MRC) currently issues an up to 5-day lead time of daily water level forecasting at each of 22 locations from Chiang Saen to Tan Chau/Chau Doc during the wet season, which is from June to October (Pagano, 2014). The flood forecasting system adopted by the RFMMC uses the Unified River Basin Simulator (URBS) rainfall-runoff/runoff routing model and ISIS hydrodynamic model (Pagano, 2014; Tospornsampan et al., 2009). The operation of the RFMMC's forecasting system is based on the Deltares Flood Early Warning System (Delft-FEWS) platform, which links data and models in real time to perform daily forecasting (Werner et al., 2013). At the locations of each in-situ water level gauge, rainfall forecast is first provided by the Fifth Generation Mesoscale Model (MM5) (Cox et al., 1998) and then used as input to a URBS model for discharge estimation, which is then converted to forecasted water

level via a rating curve. The RFMMC also uses the ISIS hydrodynamic model for water level forecasts downstream from Stung Treng to the ocean (Pagano, 2014; Tospornsampan et al., 2009). However, the ISIS hydrodynamic model is computationally intensive and therefore is run only for retrospective analyses or when the demand arises (Pagano, 2014). The RFMMC also does not issue any forecasting in the Mekong Delta (MD) downstream from Tan Chau/Chau Doc. Considering the increasing vulnerability of the MD under future climate change, a sustainable flood forecasting system which can be operated on a routine basis with low computational cost is necessary.

However, complex hydraulic conditions in the MD hinder such a flood forecasting system. The Tonle Sap Lake (TSL) has an effect on the MD and needs to be considered in addition to the flow from the MR (Figure 3.1) in order to perform flood forecasting. The existence of TSL leads to a unique flow reversal. The water flows from the lake to the Cambodian Floodplain (CF), the Bassac River, and the MD through the Tonle Sap River in September/October and flows back to the lake in May/June. This flow reversal poses a challenge from a modeling perspective, which necessitates the use of a high-resolution hydrodynamic model to physically mimic the flow reversal due to the hydrologic drivers and is often computationally expensive; which is also the reason why the RFMMC does not routinely operate it in the MD. Moreover, the ocean tide intrusion can reach up to Phnom Penh in the dry season (Pagano, 2014) and needs to be considered (Takagi et al., 2015).

To address this issue, we propose a model-aided altimeter-based river level forecasting system. It integrates Jason-2 altimetry derived water levels and discharge

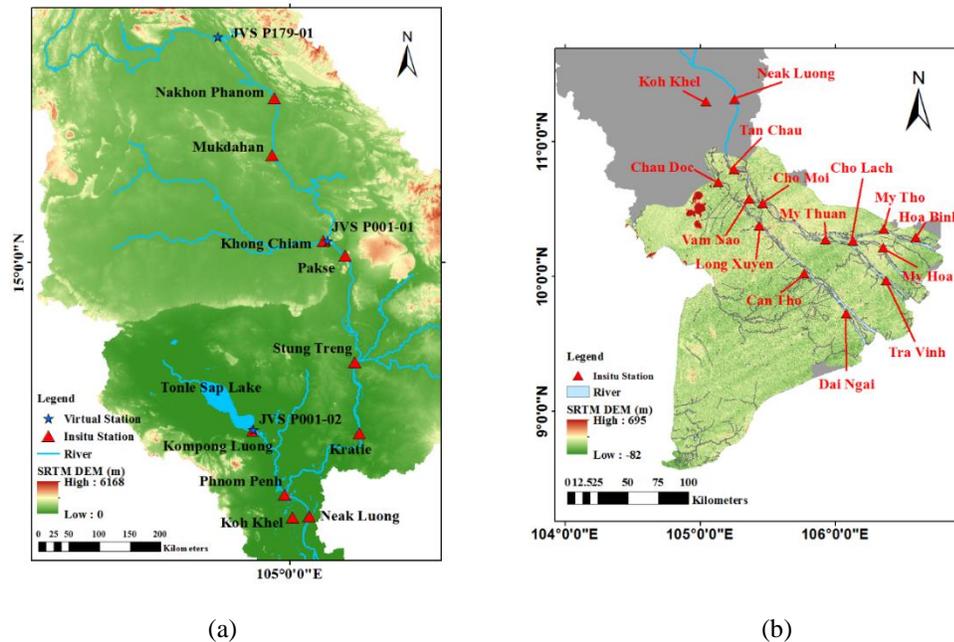
estimations from the Variable Infiltration Capacity (VIC) hydrologic model (Hossain et al., 2017; Liang et al., 1994). Here, we present a feasibility study to demonstrate its forecasting skill in the MR, especially on the MD. A multivariate regression model and the sum of sinusoidal functions are employed to address the downstream river level changes due to different causes, considering multiple variables for a better representation (Maçaira et al., 2018) of water levels in the MD. The unique features of our flood forecasting system are: (1) an easy-to-set-up hydrologic model is implemented which circumvents the need to have frequent altimeter samplings upstream; (2) the forecasting approach in the delta is based on a computationally efficient regression analysis instead of complex hydrodynamic modeling; (3) it is freely accessible for stakeholders and users with tolerable computational cost; (4) it can be easily extended for other deltas and transboundary rivers in developing countries such as the Vietnamese Red River, Ganges-Brahmaputra-Meghna River, and Niger River. With upstream altimetry Virtual Stations (VSs), a hydrologic model such as the VIC model, and historic in-situ water levels (no updated in-situ water levels are needed) in the downstream, our system can be easily implemented *as long as no major dams exist* between the upstream VSs and the downstream in-situ stations.

## 3.2 Data

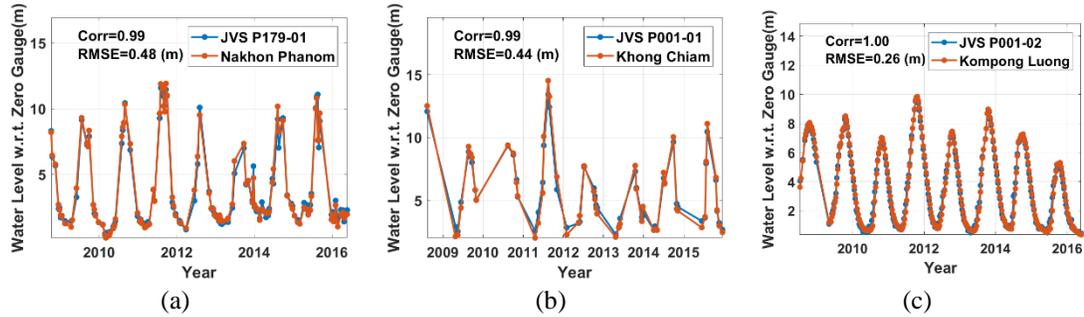
### 3.2.1 Jason-2 Altimetry Derived River Levels

Jason-2, launched on June 20<sup>th</sup>, 2008, as the follow-up mission of Jason-1, operating under the cooperation of the National Aeronautics and Space Administration

(NASA), Centre National d'Etudes Spatiales (CNES), National Oceanic and Atmospheric Administration (NOAA) and European organization for the exploitation of METeorological SATellites (EUMETSAT). It succeeds its predecessors including T/P and Jason-1 missions to continuously provide highly accurate altimetry data with a 10-day repeat cycle. In this feasibility study, we used 20-Hz ICE-retracked ranges from the Jason-2 Geophysical Data Record (GDR) to extract river levels at VSs located in the MR mainstem and TSL. The locations of the VSs are shown in Figure 3.1. The Root Mean Square Error (RMSE) of altimetry-derived water levels over the MR mainstem and TSL were about 0.4 to 0.5 m, and 0.26 m respectively when compared with the nearest in-situ gauge data (See Figure 3.2.). For the JVS P001-02 over the TSL, the altimetry water levels have a RMSE since the TSL is a large inland water body where more robust radar altimetry measurements can be collected.



**Figure 3.1** Distribution of VSs (blue triangle) and in-situ gauges (red triangle) in (a) the middle reach of the MR and CF, and (b) the MD (within colored area) used in this study.



**Figure 3.2** Jason-2 derived water level time series and validation. (a) and (b) are results at Jason-2 VSs in the MR mainstem and (c) is for the TSL.

### 3.2.2 Modeled Discharges at Virtual Stations

The river discharges at the locations of VSs are obtained by the VIC hydrological model (Liang et al., 1994) and a streamflow routing model (Lohmann et al., 1996). The VIC model is comprised of a three-layer soil column structure and considers land cover, soil type, and meteorological forcing data such as precipitation, temperature and wind speed to characterize the hydrological mechanism in the soil column of given ground cells to estimate runoffs and baseflows. The streamflow routing model (Lohmann et al., 1996) was then used to route surface runoff and baseflows estimated by the VIC model to the river channels based on the given flow direction map to simulate discharges. In our study, a 0.1-degree-resolution VIC model was set up in the MRB (Hossain et al., 2017) using the Global Land Cover Characterization (GLCC) land cover dataset and the Harmonized World Soil Database (HWSD) soil data (Siddique-E-Akbor et al., 2014). The monthly leaf area index and albedo were obtained from MODerate resolution Imaging Spectro-radiometer mission (MODIS) data. The

flow direction map was derived from the Shuttle Radar Topography Mission (SRTM) DEM over the MRB. Six sub-basin segments were selected for model calibration including Chiang Sean, Luang Prabang, Vientiane, Nakhon Phanom, Pakse, and Kampong Cham. The VIC model was calibrated using soil parameters including the variable infiltration curve parameter ( $b_{inf}$ ), maximum velocity of baseflow ( $Ds_{max}$ ), fraction of  $Ds_{max}$  ( $Ds$ ), and soil moisture ( $Ws$ ) where non-linear baseflow occurs. The cell impulse response function of the routing model was also calibrated (Hossain et al., 2017). The current VIC model output downstream of Kampong Cham, where bidirectional flow due to TSL and a complex river network exists, turns out to be less reliable. However, the VIC model is only used to estimate water discharges at VSs in the upstream of Mekong in this study. The meteorological forcings such as maximum and minimum temperatures, wind speed and precipitation were obtained from 237 weather station records archived as the Global Summary of the Day (GSOD) by the National Climatic Data Center (NCDC). The model was run at a daily time step to simulate the discharge from 2002 to 2015 at the locations of Jason-2 VSs. The output of year one (2002) was excluded to avoid spin-up error. The calibration period was set to be from 2003 to 2008. The validation period was from 2009 to 2013. The simulated water discharges at Nakhon Phanom and Khong Chiam, which are the nearest stations to JVS P179-01 and JVS P001-01, have Nash-Sutcliffe Efficiency (NSE) of 0.82 and 0.86, respectively. For more details about the VIC model set up in the MRB, readers are referred to Hossain et al. (2017). Currently, this model runs operationally for nowcast forced with satellite precipitation at several locations along the MR (See <http://depts.washington.edu/saswe>).

### 3.2.3 In-situ Water Levels

The in-situ water levels in the MR were provided by the Asian Disaster Preparedness Center (ADPC) and were used to build and test the feasibility of our model-aided altimetry-based forecasting system. In our study, the in-situ water level data collected at 6 locations in the middle reach of the MR, 3 locations in the CF, and 13 locations in the MD were used (See Table 3.1.). Figure 3.1 shows the locations of all of the in-situ gauges used (a) in the middle reach of the MR and CF, and (b) in the MD toward the river mouth.

**Table 3.1** The in-situ gauges used in this study. Data were collected at 6 locations in the middle reach of the MR, 3 locations in the CF, and 13 locations in the MD.

<b>The middle reach of the MR</b>	Nakhon Phanom / Mukdahan / Khong Chiam
	Pakse / Stung Treng / Kratie
<b>The CF</b>	Phnom Penh / Neak Luong / Koh Khel
<b>The MD</b>	Tan Chau / Chau Doc / Vam Nao
	Cho Moi / Long Xuyen / Can Tho
	My Thuan / Cho Lach / My Tho
	My Hoa / Hoa Binh / Tra Vinh / Dai Ngai

### 3.3 Methods

To obtain daily water level forecasting, Hossain et al. (2014b) used Jason-2 altimetry derived water levels with 10-day repeat cycle at upstream VSs to obtain daily forecasted downstream water discharges at downstream in situ gauges. As mentioned above, it requires a sufficient number of VSs to ensure that at least one of the VSs gets sampled in the upstream neighborhood for deriving expected water levels in the downstream (Hossain et al., 2014a). However, the MR is a north- to-south flowing river where only a few VSs can be found, hindering the direct implementation of this system.

Hence, we attempt to use the VIC hydrologic model, which can provide daily river discharges at VSs to fill the altimeter data temporal gap. At the model-building step, data from October 2008 to December 2010 were used (hereafter called “historical”). At each of the VSs in the upstream of the MR mainstem, a discharge-to-level rating curve was built using historical VIC-derived discharges and Jason-2 altimetry-derived water levels. The rating curve was then used to reconstruct historical daily water levels (See section 3.3.1). A relationship between upstream water levels at each of the VSs and downstream water levels at each of the in situ gauges was then built through a simple or multivariate linear regression, taking TSL water levels into account depending on the locations (See section 3.3.2). Tidal influence models were also added where necessary (See section 3.3.2). At the pseudo-forecasting step, data from 2011 were used. By using the rating curves and models from the model-building step, pseudo-forecasting was performed and validated to examine the skill of the system. The indices for the skill evaluation are described in section 3.3.3.

### 3.3.1 Daily Water Level Reconstruction at Virtual Stations

The conventional water level-to-discharge rating curve follows a power law curve in the form of

$$Q_{VIC} = C \cdot (h_{Alt.} - a)^b \quad (3.1)$$

where  $h_{Alt}$  is Jason-2 altimetry-derived water level,  $Q_{VIC}$  is discharge simulated from the VIC and streamflow routing model, and  $C$  and  $b$  are coefficients to be estimated.

Equation 3.1 can be transformed into a linear form by taking the logarithm of both sides and then re-arranging as

$$\log(h_{Alt.} - a) = \frac{1}{b} \cdot \log(Q_{VIC}) - \frac{1}{b} \cdot \log(C) = A \cdot \log(Q_{VIC}) + B \quad (3.2)$$

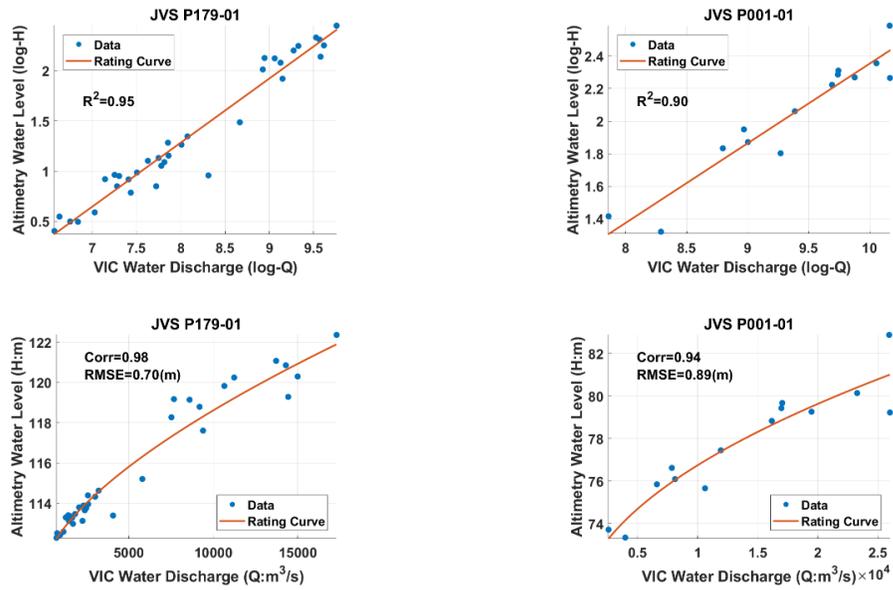
which is a linear function where coefficients  $A$  and  $B$  can be estimated by least-squares, while  $a$  can be directly solved by Johnson's method (Rantz, 1982; World Meteorological Organization, 1980). The logarithm value of daily  $Q_{VIC}$  can then be used to reconstruct daily water levels  $h_{Rec.}$

$$h_{Rec.} = e^{[A \cdot \log(Q_{VIC}) + B]} + a. \quad (3.3)$$

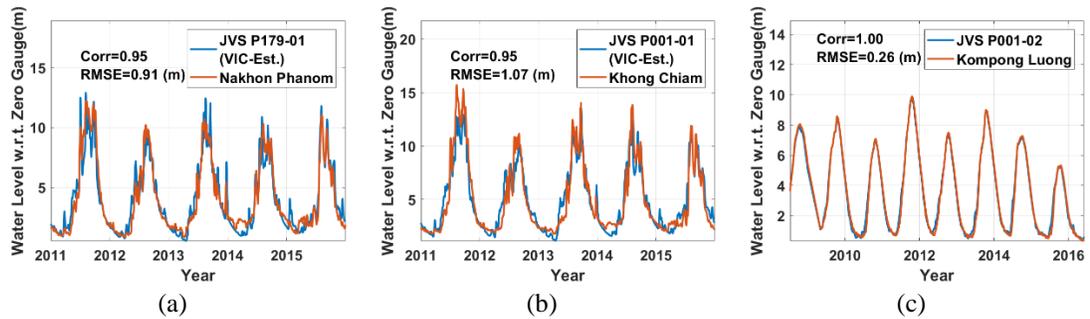
In the TSL where water levels vary smoothly (See Figure 3.2(c)), the daily water levels were reconstructed by linearly interpolating 10-day repeat Jason-2 altimetry-derived water levels.

The rating curves and reconstructed water level time series are shown in Figure 3.3 and Figure 3.4, respectively. In this study, data from October 2008 to December 2010 were used to build the rating curves for historical daily water level reconstruction at the Jason-2 VSs (Figure 3.3, also see Figure 3.5(a) for flow chart.), and data from 2011 were used for validation with in situ gauge data (Figure 3.4, also see Figure 3.6 for flow chart.). As the bottom panel of Figure 3.3 shows, the reconstructed water levels in the middle reach of the MR have RMSEs of about 1 m with temporal correlation coefficients of 0.95. The RMSEs of reconstructed water levels at the VSs in the MR may depend on the accuracies of altimetry-derived water level itself and the VIC-derived discharge. Furthermore, some "spikes" in the reconstructed water levels during

the dry season could result from errors of in situ precipitation records that are the input data for the VIC model. For the VS in TSL, the reconstructed water levels show a smaller RMSE of about 0.3 m with a temporal correlation of  $\sim 1$  due to the fact that TSL is a large water body where a larger number of 20-Hz Jason-2 measurements are available.



**Figure 3.3** Rating curves in logarithmic scale (top panel) and original scale (bottom panel).



**Figure 3.4** Reconstructed daily water levels at three VSs. JVS P179-01 and JVS P001-01 are located in the upstream MR mainstem, while JVS P001-02 is located in the TSL.

### 3.3.2 Forecasting Models

Considering that the hydraulic conditions of the MR are complex due to its large geographic coverage and various geographic features, we built two different forecasting models for (1) the middle reach of the MR, and (2) the CF and the MD. We used the reconstructed water levels at upstream VSs and downstream in situ water levels from October 2008 to December 2010 to build forecasting and tidal influence models (See Figure 3.5 for flow charts.). Data from 2011 were used to perform “pseudo forecasting” (See Figure 3.6 for flow chart.). The forecasted water levels were then validated with in situ observations.

For each of the locations from Nakhon Phanom to Kratie (See Figure 3.1(a).), in the middle reach of the mainstem MR, the hydraulic condition is relatively simple (Pagano, 2014). Thus, we performed a simple linear regression analysis using the reconstructed daily water levels at upstream VSs and observed historical daily water levels at downstream in situ gauges, considering different lead times to build the forecasting model as

$$h_{Insitu}(t + k) = E \cdot h_{Rec.}(t) + F \quad (3.4)$$

where  $h_{Insitu}(t)$  is historic water levels at downstream in situ gauges at time  $t$ ,  $h_{Rec.}(t)$  is the reconstructed water levels at upstream VSs at time  $t$  which were derived by Equation 3.3,  $k$  is the forecasting lead time (day) and  $E$  and  $F$  are coefficients to be estimated for this linear forecasting model.

For each of the locations from Phnom Penh to the MD (See Figure 3.1(b).), river flows are dominated by water exchange between river channel and floodplain and flow

reversal due to TSL. Furthermore, ocean tides can also intrude up to Phnom Penh in the dry season (Pagano, 2014). Considering these multiple sources of water level changes, we propose a two-step method to build a forecasting model for the lower reach of the MR. The first step is to use reconstructed water levels at VSs in the upstream MR mainstem and TSL to perform multivariate linear regression analysis with downstream in situ water levels

$$h_{Insitu}(t+k) = G \cdot h_{Rec.}^{MK}(t) + H \cdot h_{Rec.}^{TSL}(t) + I + v(t+k) \quad (3.5)$$

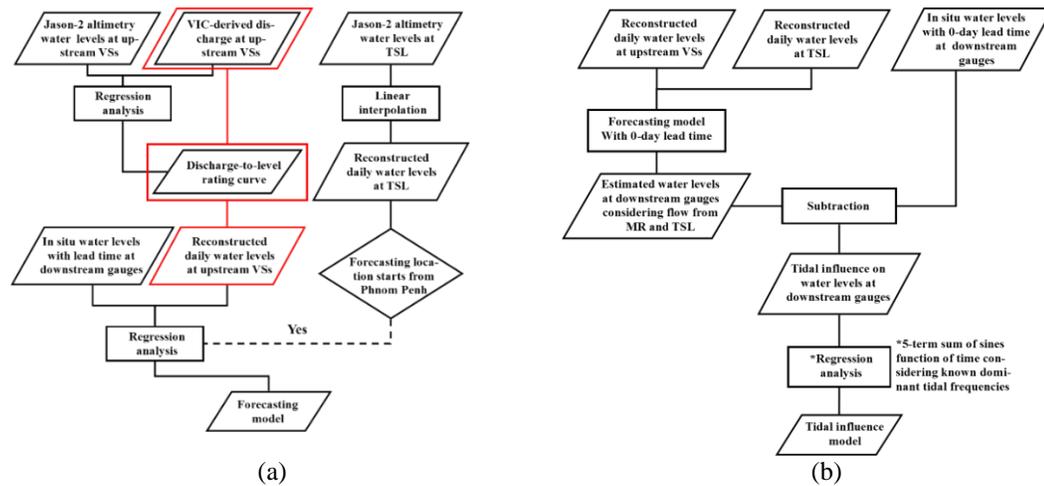
where  $h_{Rec.}^{MK}(t)$  is the reconstructed daily water levels in the MR mainstem at time  $t$  derived by Equation 3.3,  $h_{Rec.}^{TSL}(t)$  is the reconstructed daily water levels in the TSL which were obtained by linearly interpolating Jason-2 altimetry water levels as section 3.1 described,  $G$ ,  $H$ , and  $I$  are coefficients to be estimated for this multivariate linear regression model,  $k$  is the forecasting lead time (day), and  $v(t+k)$  is the residual between modeled results and  $h_{Insitu}(t+k)$ . Since the model only considers the influence of the MR mainstem and TSL, we assume that  $v$  here is dominated by ocean tide influence and hereafter rephrased as  $h_{Tide}(t)$  for the case of  $k=0$ . Next, we adopted a model of tidal influence as the sum of a 5-term sinusoidal function while considering dominant tidal frequencies including annual, semi-annual, monthly, fortnightly, and synodic fortnightly tides (See Table 3.2.) (Egbert and Ray, 2003; Zheng and Zhao, 1984).

$$h_{Tide}(t) = \sum_{i=1}^5 M_i \cdot \sin(2\pi \cdot \omega_i \cdot t + N_i) \quad (3.6)$$

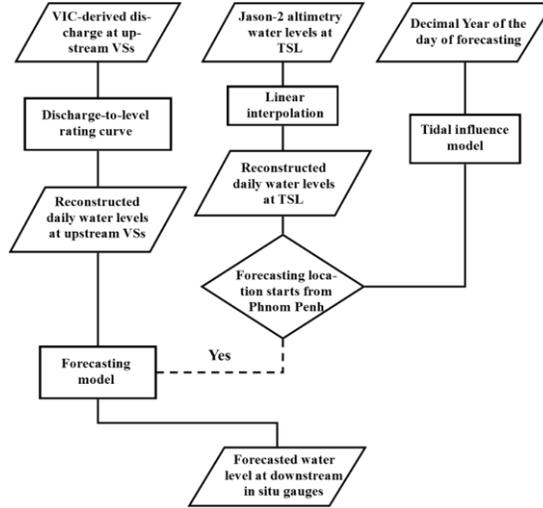
where  $\omega$  is the dominant tidal frequency (number of cycles per year), while  $M$  and  $N$  are the magnitude and phase shift to be estimated, respectively. Once  $M$  and  $N$  are fitted,  $t$  can be updated as decimal year at the time we want to perform forecasting. Subsequently, the summation of modeled results given by Equation 3.5 and Equation 3.6 provides us with forecasted water levels which consider influences from the MR mainstem, TSL, and ocean tide intrusion.

**Table 3.2** Dominant ocean tide frequencies.

Name of Tide	Period (Days)	Frequency (1/year)
Annual	365.26	1.00
Semi-annual	182.62	2.00
Monthly	27.55	13.26
Fortnightly	14.77	24.73
Synodic fortnightly	13.66	26.74



**Figure 3.5** Flowcharts of (a) building forecasting model considering water flow from the MR (and the TSL if necessary, depends on forecasting location), and (b) building tidal influence model. The reconstructed daily water levels at upstream VSs and reconstructed daily water levels at the TSL are the same as those in (a).



**Figure 3.6** Flowchart of conducting water level forecast.

### 3.3.3 Model Evaluation Statistics

To evaluate model capacity, several statistical indices including the mean absolute error (MAE), error of standard deviation (ESTD), correlation coefficient (Cor.) as well as Nash-Sutcliffe Efficiency (NSE) and achievement rate (AR) were adopted.

The MAE is a measure of error which can be calculated by

$$\text{MAE} = \frac{\sum_{i=1}^N |Y_i^{obs} - Y_i^{fct}|}{N} \quad (3.7)$$

where  $N$  is the total number of data points, with  $i = 1 \sim N$ .  $Y_i^{obs}$  is the in situ observations, and  $Y_i^{fct}$  is the forecasting results. A MAE value of 0 indicates perfect forecasting.

The ESTD is the standard deviation of differences between forecasting results and observations

$$\text{ESTD} = \sqrt{\frac{\sum_{i=1}^N (e_i - \bar{e})^2}{N}} \quad (3.8)$$

where  $e_i = Y_i^{obs} - Y_i^{fct}$ ,  $\bar{e}$  is the mean of  $e_i$ .

The Cor. is a measure of temporal agreement between forecasting and observations

$$\text{Cor.} = \frac{\sum_{i=1}^N (Y_i^{fct} - \overline{Y^{fct}})(Y_i^{obs} - \overline{Y^{obs}})}{\sqrt{\sum_{i=1}^N (Y_i^{fct} - \overline{Y^{fct}})^2} \sqrt{\sum_{i=1}^N (Y_i^{obs} - \overline{Y^{obs}})^2}} \quad (3.9)$$

where  $\overline{Y^{fct}}$  is the mean of forecasting, and  $\overline{Y^{obs}}$  is the mean of the observations.

The NSE is a normalized statistic which reflects the relative magnitude of the residual variance (noise) compared to the measured data variance (information) (Nash and Sutcliffe, 1970). It indicates how well the plot of observations versus forecasting results fits the 1:1 line and is commonly used to evaluate the performance of a forecasting model

$$\text{NSE} = 1 - \frac{\sum_{i=1}^N (Y_i^{obs} - Y_i^{fct})^2}{\sum_{i=1}^N (Y_i^{obs} - \overline{Y^{obs}})^2} \quad (3.10)$$

NSE ranges from  $-\infty$  to 1.0. An NSE of 1 indicates perfect forecasting skill, 0 means the forecasting skill is no better than adopting the average of observations as forecasting results, and a negative NSE means unacceptable skill (Moriasi et al., 2007).

The AR is the rate of days in the wet season that meets the satisfactory benchmark, in which we used absolute error (AE)

$$\text{AR} = \frac{S}{N} \quad (3.11)$$

where  $S$  is the number of days in the wet season when the difference between forecasted water levels and in situ observations is smaller than or equal to the benchmark.

### 3.4 Results and discussions

#### 3.4.1 Cross-comparison with Current System

For cross-comparison of our system with the system adopted by the RFMMC, we compared the results of 5-day forecasting, which is the longest forecasting lead time that the RFMMC performs during the wet season (June to October). The cross-comparison was conducted at 6 locations in the middle reach of the MR (Nakhon Phanom, Mukdahan, Khong Chiam, Pakse, Stung Treng, and Kratie), 3 locations in the CF (Phnom Penh, Koh Khel, and Neak Luong), and 2 locations in the upstream of the MD (Tan Chau and Chau Doc). The MAE-based annual accuracy assessment in the RFMMC's seasonal flood situation report, available at [http://ffw.mrcmekong.org/report\\_seasonal.php](http://ffw.mrcmekong.org/report_seasonal.php), is used for comparison from 2011 to 2013, year by year (RFMMC, 2011, 2012, 2013). The ESTDs of our forecasting results during the wet season from 2011 to 2015 were also compared with Pagano (2014)'s evaluation of performance of the RFMMC's forecasting system using historical data from 2002 to 2012 for a long-term evaluation. Although the time spans of the long-term ESTDs are different, Pagano (2014)'s estimation is the only long-term accuracy assessment that we can find. Forecasting at the locations down to Khong Chiam was performed using JVS P179-01, while JVS P001-01 was used for forecasting at locations from Pakse. For forecasting at locations downstream of TSL, both JVS P001-01 and JVS P001-02 were used to represent different water sources from the MR mainstem and TSL, respectively. As Table 3.3 shows, in the middle reach of the MR, our forecasted river levels were, in general, less accurate than the RFMMC's forecasts. We obtained

yearly MAEs of about 1 – 2.2 m and long-term ESTDs of about 0.8 – 1.5 m during the wet season, while the RFMMC’s yearly MAEs are about 0.4 – 0.7 m with long-term ESTDs of about 0.7 – 0.9 m. For the locations in the CF, the accuracy of our system improved. Yearly MAEs of our system are about 0.4 – 0.5 m and long-term ESTDs are about 0.5 m, while the RFMMC’s yearly MAEs are of about 0.2 – 0.25 m, with long-term ESTDs of about 0.25 – 0.3 m. Overall, in the middle reach of the MR and the CF, the RFMMC’s forecasting showed slightly better accuracy than ours. Since our system uses VIC-derived discharges to build the rating curve with altimetry-derived water levels for reconstruction of daily water levels at VVs, both of their accuracies can be contributing factors of our forecasting skill. For example, the RFMMC’s forecasting system uses a calibrated set of 51 URBS models which cover over 2,244 sub-areas (740,000 km<sup>2</sup>) (Tospornsampan et al., 2009), while the VIC model we adopted was calibrated at only 6 sub-basins. In addition, the accuracy of altimetry-derived water levels can be up to ~ half of a meter as shown in Figure 3.2. On the other hand, our forecasting system only assumes that there is a linear relationship between upstream and downstream water levels but does not consider the impact of tributaries and direct rainfall in between. Therefore, it is expected that our forecasting skill can be further improved by fine-tuning the VIC model, by performing more advanced altimetry waveform retracking and by considering those additional contributions to the downstream water levels.

However, it needs to be emphasized that our forecasting system uses freely accessible and publicly available satellite data and hydrologic model and is able to perform forecasting without updated in situ data. Our system is purely based on simple

regression analysis instead of a computationally expensive weather forecast and hydrodynamic model. For example, the RFMMC’s system uses MM5 to forecast rainfall as an input to the URBS models, which have a computational cost and skill set requirements, while our system does not. Features including strong affordability, easy-to-set-up, and negligible computational cost make our system operationally sustainable and scalable, especially for developing countries.

It is also worthwhile to mention that in Tan Chau and Chau Doc, which are close to the upstream boundary of the MD in Vietnam, our system showed yearly MAEs of about 0.2 – 0.3 m and long-term ESTDs of about 0.3 m, which are compatible with the RFMMC’s results. The smaller MAEs might be due to the lower amplitude of water levels in the MD, but it also reflects the fact that the sinusoidal model to represent the tidal influence was effective in the MD. Moreover, our system can provide forecasting further downstream of Tan Chau and Chau Doc inside the MD where the RFMMC currently does not provide any forecasting. After cross-comparison with the RFMMC’s report, we attempted to perform a “pseudo forecasting” considering 20 different forecasting lead times (from 1-day forecasting to 20-day forecasting) in each region, and then validated our forecasting results using in situ data. In the following section, the forecasting skill of our system will be discussed.

**Table 3.3** Cross-comparison with the RFMMC’s 5-day forecasting in the wet season.

Location Name	The RFMMC (Approximate)				This Study			
	MAEs (m)			ESTDs (m)	MAEs (m)			ESTDs (m)
	2011	2012	2013	Pagano (2014)	2011	2012	2013	2011-2015 Wet season
				2002-2012 Wet season				
Middle reach of the MR								
Nakhon Phanom	0.78	0.55	0.60	0.83	1.58	1.04	1.07	1.13
Mukdahan	0.80	0.60	0.66	0.83	1.67	1.04	0.91	1.09

Table 3.3 (Continued)								
Khong Chiam	0.85	0.65	0.80	0.90	2.23	1.12	1.28	1.34
Pakse	0.68	0.58	0.78	0.85	1.55	1.15	1.23	1.17
Stung Treng	0.55	0.43	0.65	0.68	1.03	0.82	0.95	0.85
Kratie	0.75	0.70	0.88	0.87	1.56	1.61	1.54	1.47
<b>Average</b>	<b>0.74</b>	<b>0.59</b>	<b>0.73</b>	<b>0.83</b>	<b>1.60</b>	<b>1.13</b>	<b>1.16</b>	<b>1.18</b>
<b>±STD</b>	<b>±0.11</b>	<b>±0.09</b>	<b>±0.11</b>	<b>±0.08</b>	<b>±0.38</b>	<b>±0.26</b>	<b>±0.24</b>	<b>±0.21</b>
<b>CF</b>								
Phnom Penh	0.23	0.25	0.25	0.33	0.55	0.44	0.47	0.62
Neak Luong	0.15	0.20	0.18	0.25	0.34	0.32	0.32	0.44
Koh Khel	0.15	0.18	0.16	0.24	0.71	0.36	0.45	0.56
<b>Average</b>	<b>0.18</b>	<b>0.21</b>	<b>0.20</b>	<b>0.27</b>	<b>0.53</b>	<b>0.37</b>	<b>0.41</b>	<b>0.54</b>
<b>±STD</b>	<b>±0.05</b>	<b>±0.04</b>	<b>±0.05</b>	<b>±0.05</b>	<b>±0.19</b>	<b>±0.06</b>	<b>±0.08</b>	<b>±0.09</b>
<b>MD</b>								
Chau Doc	0.18	0.25	0.25	0.30	0.27	0.16	0.21	0.29
Tan Chau	0.15	0.20	0.20	0.25	0.25	0.20	0.21	0.31
<b>Average</b>	<b>0.17</b>	<b>0.23</b>	<b>0.23</b>	<b>0.28</b>	<b>0.26</b>	<b>0.18</b>	<b>0.21</b>	<b>0.30</b>
<b>±STD</b>	<b>±0.02</b>	<b>±0.04</b>	<b>±0.04</b>	<b>±0.04</b>	<b>±0.01</b>	<b>±0.03</b>	<b>±0.00</b>	<b>±0.01</b>

**STD:** Standard deviation

### 3.4.2 Flood Forecasting in the Middle Reach of the MR and CF

In the middle reach of the MR and the CF, in situ data are available until 2015, and were used to validate our forecasting results. Figure 3.7 shows the MAEs and NSEs of our system during the wet season from 2011 to 2015. Up to 20-day lead time forecasting was performed. Forecasting at Nakhom Phanom, Mukdahan, and Khong Chiam was performed by using reconstructed water levels at JVS P179-01. We can see that the MAEs of 1-day forecasting at Nakhom Phanom and Mukdahan are ~ 0.9 m then gradually increase with lead time, while initially about 1.2 m at Khong Chiam. The relatively higher MAEs at Khong Chiam might be due to its longer distance from JVS P179-01, which is ~ 477 km compared to the 141 km and 248 km for Nakhom Phanom

and Mukdahan, respectively. Since our forecasting is based on a linear relationship between upstream VS water levels and downstream in situ water levels, hydrologic or hydraulic condition changes between upstream VSs and downstream in situ stations due to geographic remoteness can influence forecasting skill. On the other hand, Khong Chiam is close to the confluence of the Mun River and MR which might be another reason why forecasting at Khong Chiam has higher MAEs compared to that at Nakhom Phanom and Mukdahan. The NSEs of forecasting at these locations start from about 0.8 for 1-day forecasting and gradually decrease, but can still maintain at 0.7 up to about 5-day lead time.

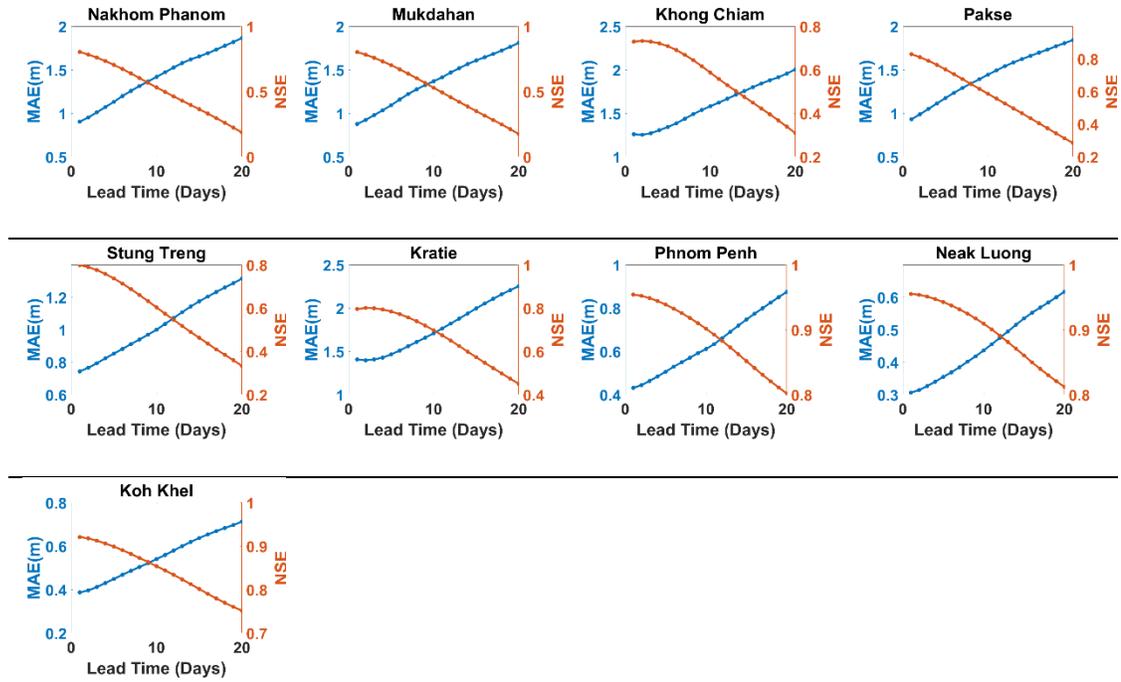
Forecasting starting from Pakse was performed by using reconstructed water levels at JVS P001-01. The MAEs of 1-day forecasting are 0.9 m and 0.8 m at Pakse and Stung Treng, respectively, while it is 1.4 m at Kratie. They increase with longer lead time. Kratie is 373 km away from JVS P001-01, which is more distant than 44 km and 240 km for Pakse and Stung Treng. This may be the reason why the MAEs at Kratie are relatively higher. On the other hand, Kratie is the location where the hydraulic condition changes from mainstream channel river flow to one where significant water exchange occurs between the river and floodplains. This fact may influence forecasting skill as well. At the locations in the CF downstream from Phnom Penh, the reconstructed water levels of TSL in addition to JVS P001-01 were used in order to consider the influence of reverse flow. At Phnom Penh and Koh Khel, the MAEs of 1-day forecasting are about 0.4 m while it is about 0.3 m at Neak Luong. They also increase with longer lead times. The NSEs of forecasting start from about 0.9 to 0.95 for 1-day forecasting

and gradually decrease with longer lead times but can still maintain NSEs higher than 0.75 to 0.8 for up to 20-day lead time.

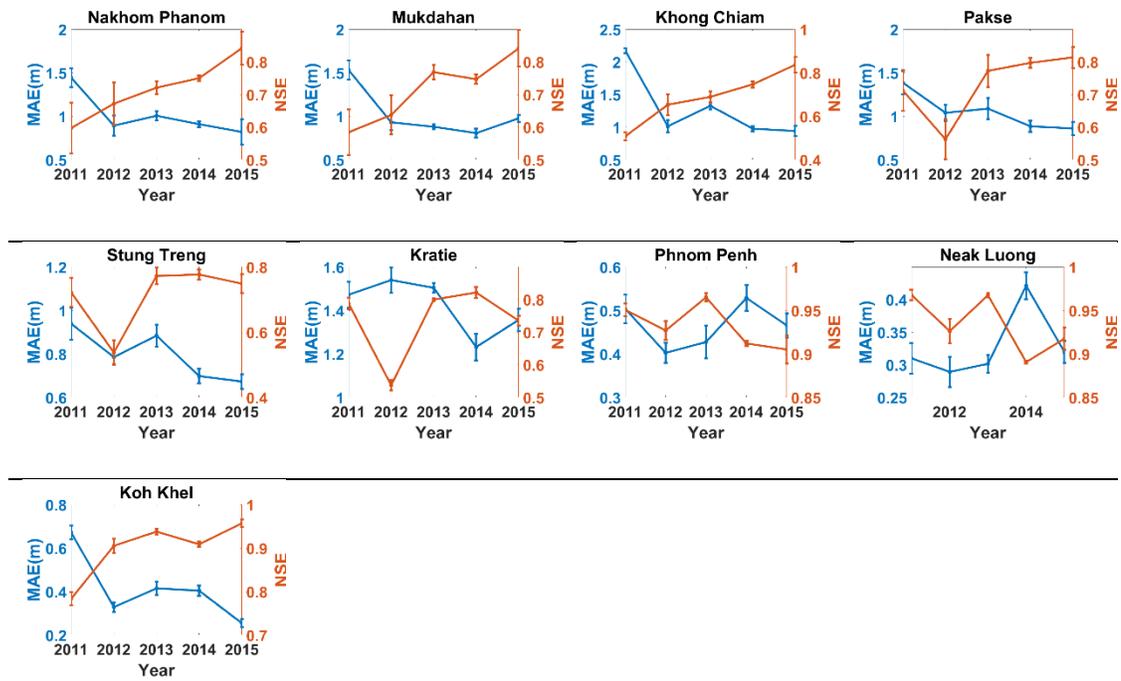
Since the change of the MAEs within 20 days of forecasting can be decimeters to around one-meter, it can be misleading to evaluate our forecasting skill if we calculate the yearly averages and STDs of the MAEs and NSEs from 20 different days of forecasting. Thus, as shown in Figure 3.8, we computed yearly averages and STDs of the MAEs and NSEs for 1-day to 5-day forecasts. At Nakhom Phanom, Mukdahan, and Khong Chiam, the average MAEs in 2011 are higher than other years with the lowest average NSEs. At Pakse, Stung Treng, and Kratie, the average NSEs in 2012 are lower than other years. At Phnom Penh, the average MAEs in 2011, 2014 and 2015 are higher. At Neak Luong and Koh Khel, the average MAEs are higher in 2014 and 2011, respectively. The higher MAEs in 2011 might be due to the strike of two tropical storms - Haima and Nokten, followed by the influence of southwest monsoons and the intertropical convergence zone (ITCZ). This led to two flood events with amplitudes over 4 m at stations from Nakhom Phanom to Pakse, and another flood event with amplitudes over 2 m at Stung Treng and over 3.5 m at Kratie (RFMMC, 2011). The average NSEs are all above 0.8 and some of them are higher than 0.95 with no significant differences. The variation of the average NSEs between years is less than 0.15. Figure 3.9 shows the correlation coefficients of our forecasting results with in situ observations. Generally speaking, the correlation coefficients of our forecasting results with in situ water levels are about 0.9 at the locations in the middle reach of the MR and gradually decrease to about 0.6 – 0.7 with longer lead times. They become larger than 0.95 at those in the CF in the case of 1-day forecasting, then gradually decrease to 0.9

as lead time extends. Figure 3.10 shows the time-series of 5-day forecasting and in situ water levels. Table 3.4 shows averages for years (2011-2015) of ARs for our forecasting system based on the benchmarks from RFMMC (RFMMC, 2011, 2012, 2013). It can be seen that in the middle reach of the MR, most ARs range around 10 – 30%, except for Khong Chiam and Kratie which range 5 – 25% and 5 – 17%, respectively. The lower ARs at these two locations are probably due to the fact that Khong Chiam is close to the confluence of the Mun River and MR, and the complicated hydraulic condition at Kratie. At Phnom Penh, Neak Luong, and Koh Khel in the CF, the values range 12 – 30%, 20 – 40%, and 18 – 33%, respectively. Therefore, although our system generally provides higher MAEs than the satisfactory benchmarks of RFMMC. Based on Figure 3.8, it can still provide satisfactory forecasting during the wet season.

Figure 3.11 shows the ARs corresponding to 1-day to 20-day forecasting at 6 locations in the middle reach of the MR and 3 locations in the CF. The benchmarks in the two regions are different (0.80 m/1.00 m/1.20 m in the middle reach of the MR, and 0.30 m/0.40 m/0.50 m in the CF) since they have different levels of forecasting skills. Generally, the ARs decrease with longer lead times. The MAEs also have such pattern as Figure 3.7 shows. During the wet season of 2011 to 2015, the ARs of the 0.8 m benchmark are about 50 – 70% at Nakhom Phanom, 55 – 70% and 55 – 65% at Mukdahan and Khong Chiam, respectively, and are 50 – 66% at Pakse, 70 – 78% at Stung Treng. At Kratie, 55 – 60% of days meet the 1 m benchmark. At Phnom Penh, 45 – 62% of days during the period meet the 0.4 m benchmark and 55 – 66% and 50 – 66% of days can meet the 0.3 m benchmark at Neak Luong, and Koh Khel, respectively.



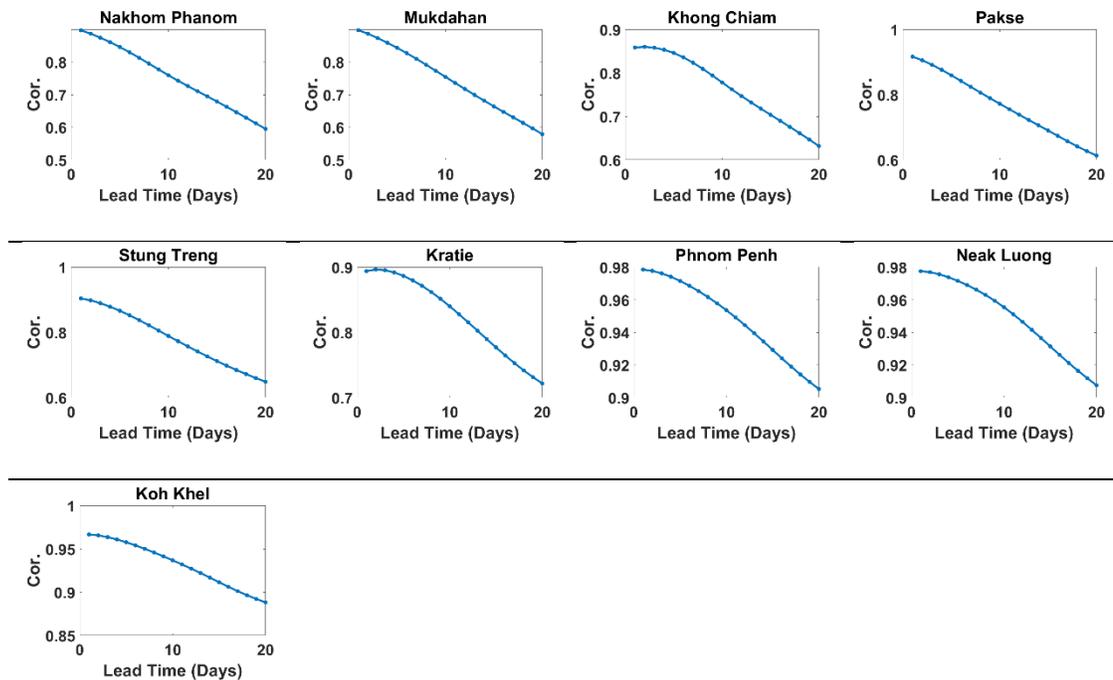
**Figure 3.7** The MAEs and NSEs of our system at 6 locations in the middle reach of the MR and 3 locations in the CF.



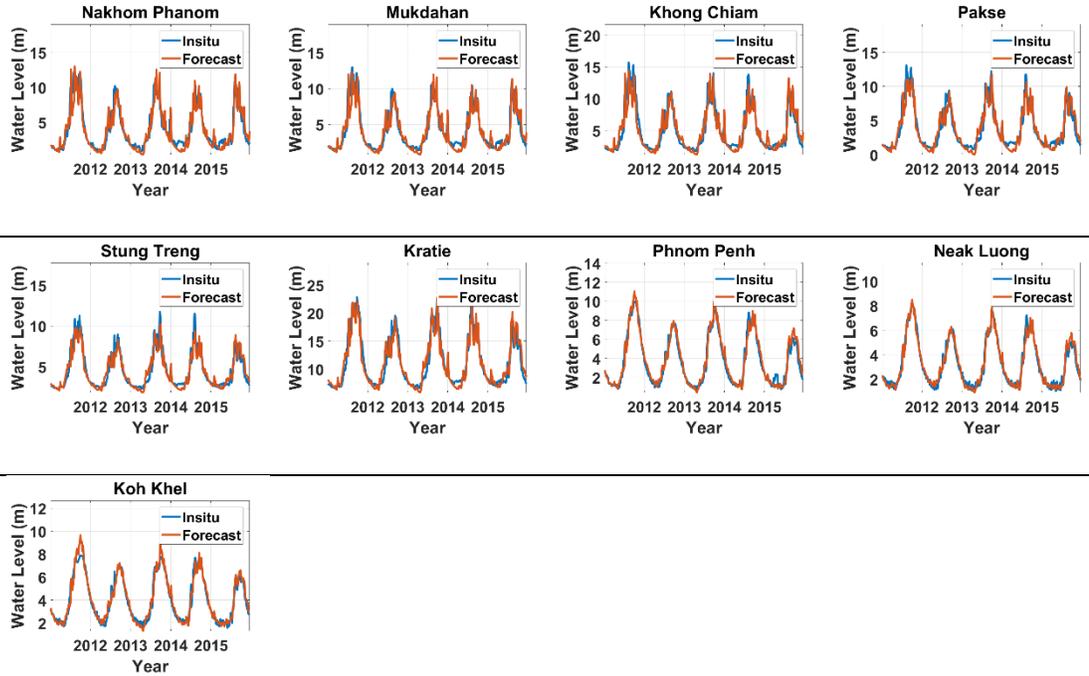
**Figure 3.8** The yearly averages and STDs of MAEs and NSEs of 1-day to 5-day forecasting at 6 locations in the middle reach of the MR and 3 locations in the CF.

**Table 3.4** Average of years (2011 - 2015) of ARs (%) of our forecasting system at locations from the middle reach of MR, and CF upstream of MD, and the corresponding RFMMC satisfactory benchmarks (cm).

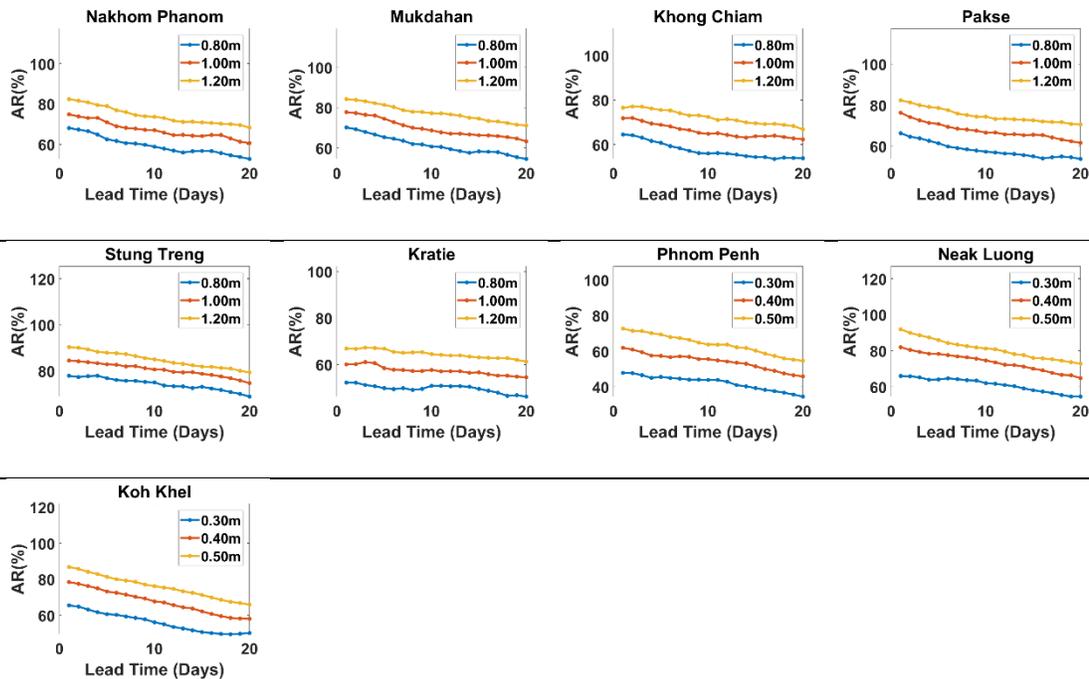
Location Name	Lead time (day)				
	1	2	3	4	5
<b>Middle reach of the MR</b>					
Nakhom Phanom	11.37	31.44	17.65	31.37	28.63
Mukdahan	10.06	21.18	17.91	33.99	32.68
Khong Chiam	5.10	14.77	15.42	26.67	24.70
Pakse	9.02	17.25	15.16	28.23	26.67
Stung Treng	9.41	19.61	18.82	35.69	33.33
Kratie	5.75	11.64	11.24	18.04	16.73
<b>RFMMC satisfactory benchmark (cm)</b>	<b>10</b>	<b>25</b>	<b>25</b>	<b>50</b>	<b>50</b>
<b>CF</b>					
Phnom Penh	15.30	13.60	12.03	29.80	29.93
Neak Luong	20.26	20.52	20.39	42.48	42.09
<b>RFMMC satisfactory benchmark (cm)</b>	<b>10</b>	<b>10</b>	<b>10</b>	<b>25</b>	<b>25</b>
Koh Khel	17.65	18.04	17.78	17.39	33.46
<b>RFMMC satisfactory benchmark (cm)</b>	<b>10</b>	<b>10</b>	<b>10</b>	<b>10</b>	<b>25</b>



**Figure 3.9** The correlation coefficients of our forecasted water levels with in situ observations at 6 locations in the middle reach of the MR and 3 locations in the CF.



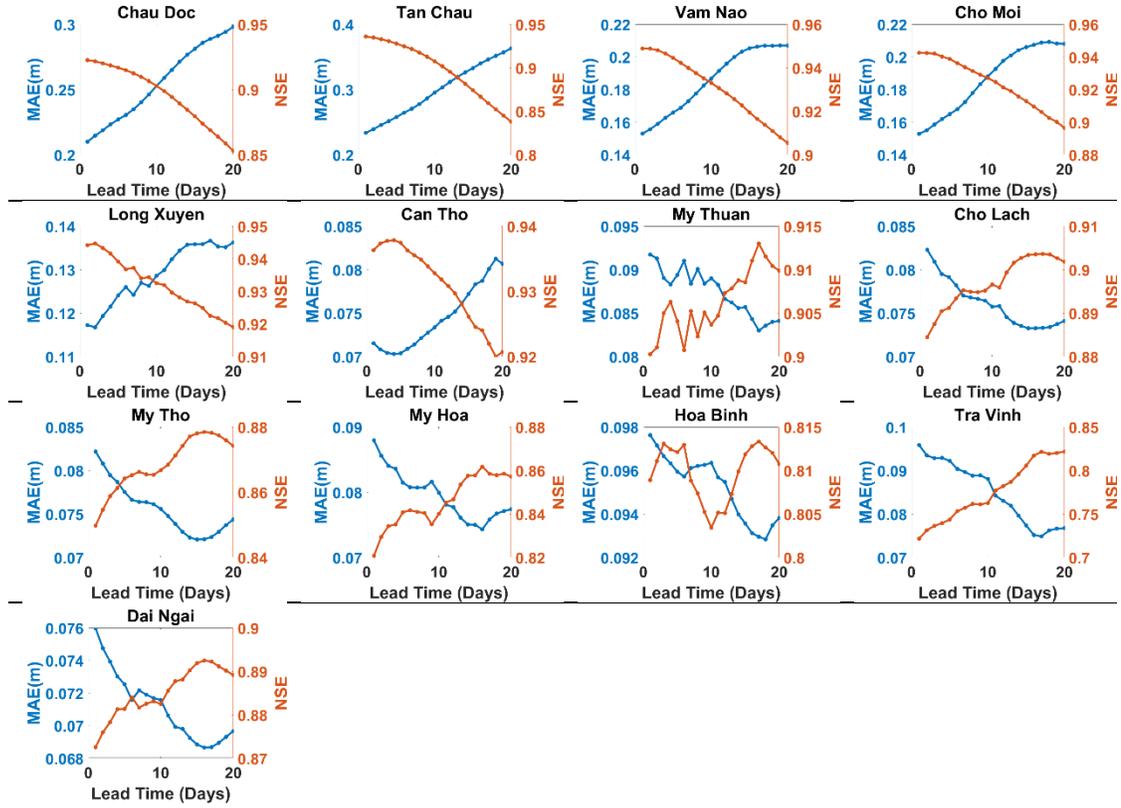
**Figure 3.10** Time-series of the 5-day forecasted and in situ water levels at 6 locations in the middle reach of the MR and 3 locations in the CF.



**Figure 3.11** The ARs of our forecasting at 6 locations in the middle reach of the MR and 3 locations in the CF. Different lines indicate the ARs in the case of different benchmarks. Note that the benchmarks are different in two regions since they have different levels of forecasting skills).

### 3.4.3 Flood Forecasting in the MD

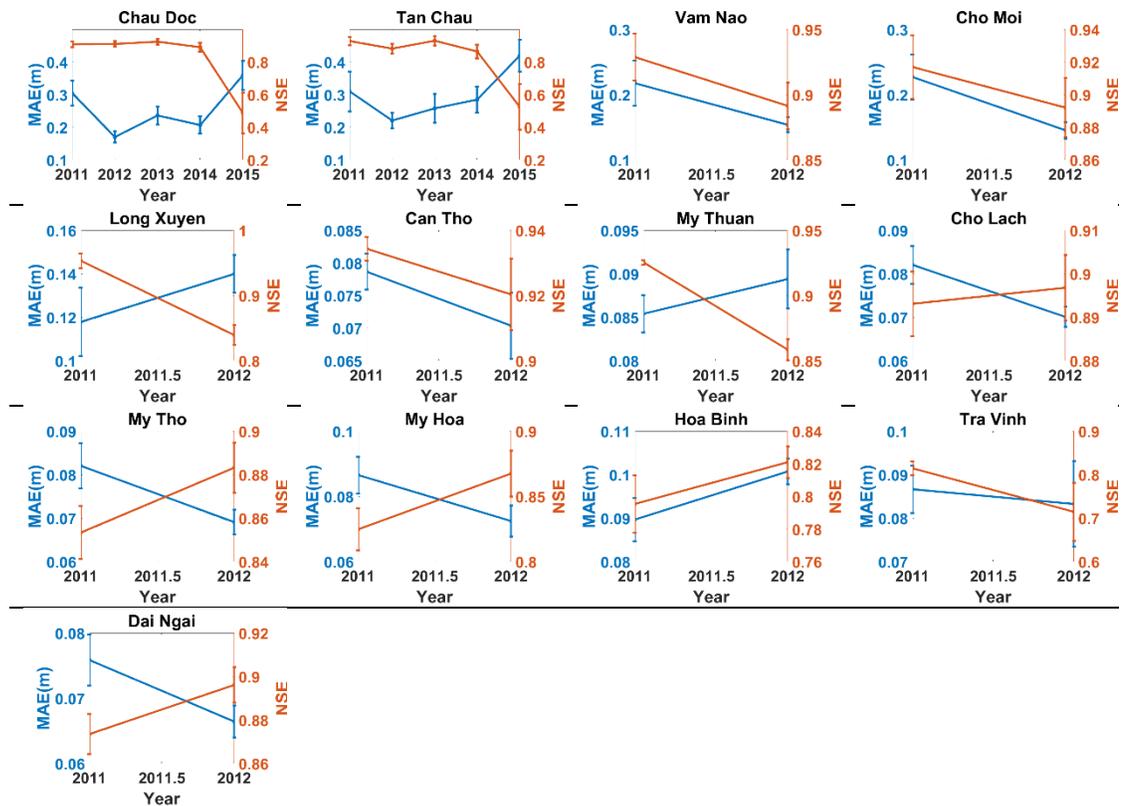
In the MD, in situ data at Tan Chau and Chau Doc span until 2015, while it spans until 2012 at the other 11 locations. Pseudo forecasting with 20 different days of forecasting lead times during the wet season from 2011 to 2015 and 2011 to 2012 was performed at Tan Chau/Chau Doc and the other 11 locations in the MD, respectively (See Figure 3.1(b) for their locations in the MD.). Figure 3.12 shows the change of the MAEs and NSEs of our forecasted water levels with respect to 1-day to 20-day lead times. From Chau Doc to Can Tho, the MAEs generally increases with longer lead time. Starting from My Thuan, the MAEs decrease with an increase of lead time until they reach a minimum at about 15- to 17-day lead time, perhaps because there are stronger fortnightly and synodic fortnightly tidal influences on the water levels at these locations. At Chau Doc and Tan Chau, the differences between the maximum and minimum of the MAEs within the 1- to 20-day lead times are about 10 cm. At Vam Nao and Cho Moi, they are about 0.04 m, and it is about 0.02 m at Long Xuyen. From Can Tho/My Thuan to the river mouth, they are mostly less than 0.01 – 0.02 m. In summary, at locations from Can Tho/My Thuan to the river mouth, the forecasted river levels have less than or about 0.10 m of MAEs even with longer than 10-day lead time. The average NSEs range from 0.8 to 0.95, which are also promising, and with less than or about 0.1 of variations within 1- to 20-day lead times.



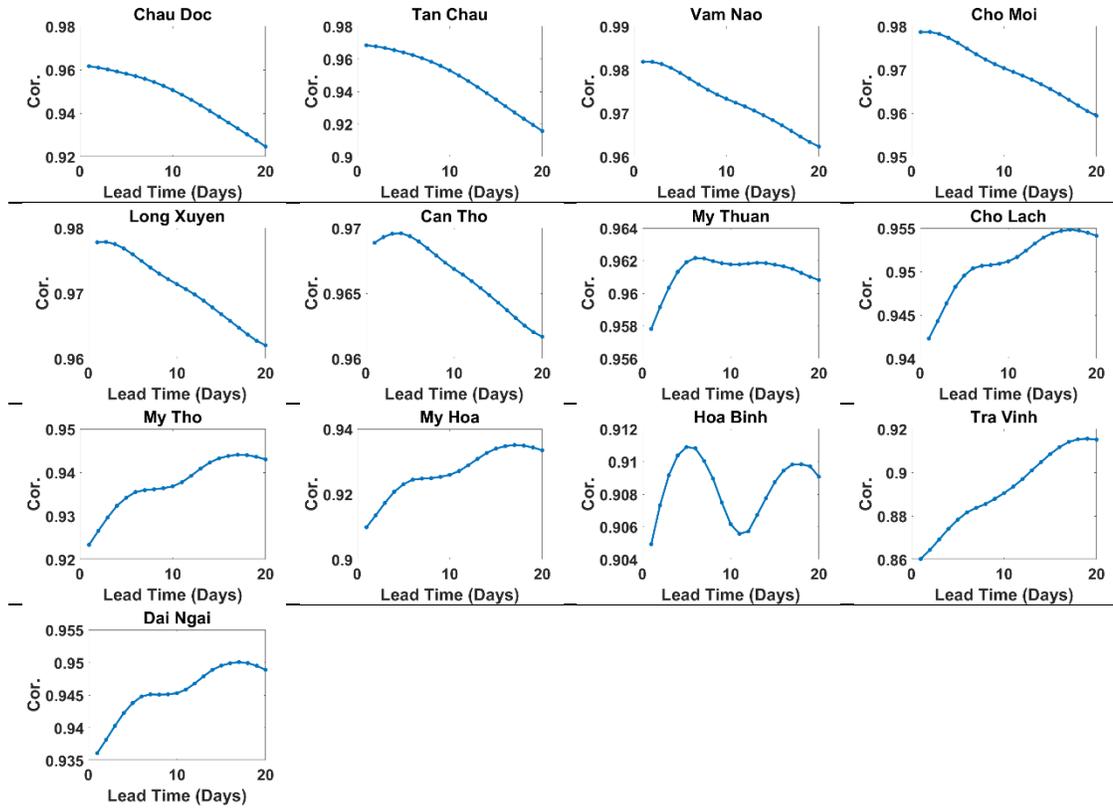
**Figure 3.12** The MAEs and NSEs of our system at 13 locations in the MD.

Figure 3.13 shows the yearly averages and STDs of the MAEs, and NSEs from 20 different days of forecasting lead time. The average MAEs range from 0.2 to about 0.4 m at Chau Doc and Tan Chau. At Vam Nao and Cho Moi, they range from 0.15 – 0.20 m while they range from 0.12 – 0.14 m at Long Xuyen. The average MAEs then decrease to mostly less than 0.10 m at the locations downstream from Can Tho/My Thuan approaching the river mouth. Note that at Chau Doc and Tan Chau, the average MAEs are 0.3 m and 0.4 m in 2011 and 2015, respectively, which are relatively higher than the other years, with the lowest average NSEs in 2015. Moreover, at Vam Nao and Cho Moi, the average MAE in 2011 are about 0.07 m higher than that in 2012, while there are only 0.01 – 0.02 m of differences between the two years at the other locations

toward the river mouth. The higher MAEs in 2011 may be due to the strike of tropical storm Haima and Nokten. According to the RFMMC’s seasonal flood situation report of 2011 (RFMMC, 2011), from the end of September to early November of 2011, water levels at Tan Chau and Chau Doc were recorded to be 3 – 4 m higher than flood levels as the result of flood water inflows from the upper and middle reaches of the MR. This resulted in extreme peak water levels at Vam Nao and Cho Moi as well and impacted the MAEs. Figure 3.14 shows the correlation coefficients of our forecasting results with in situ observations, indicating excellent temporal agreement with in situ water levels. The correlation coefficients are at or higher than 0.9.



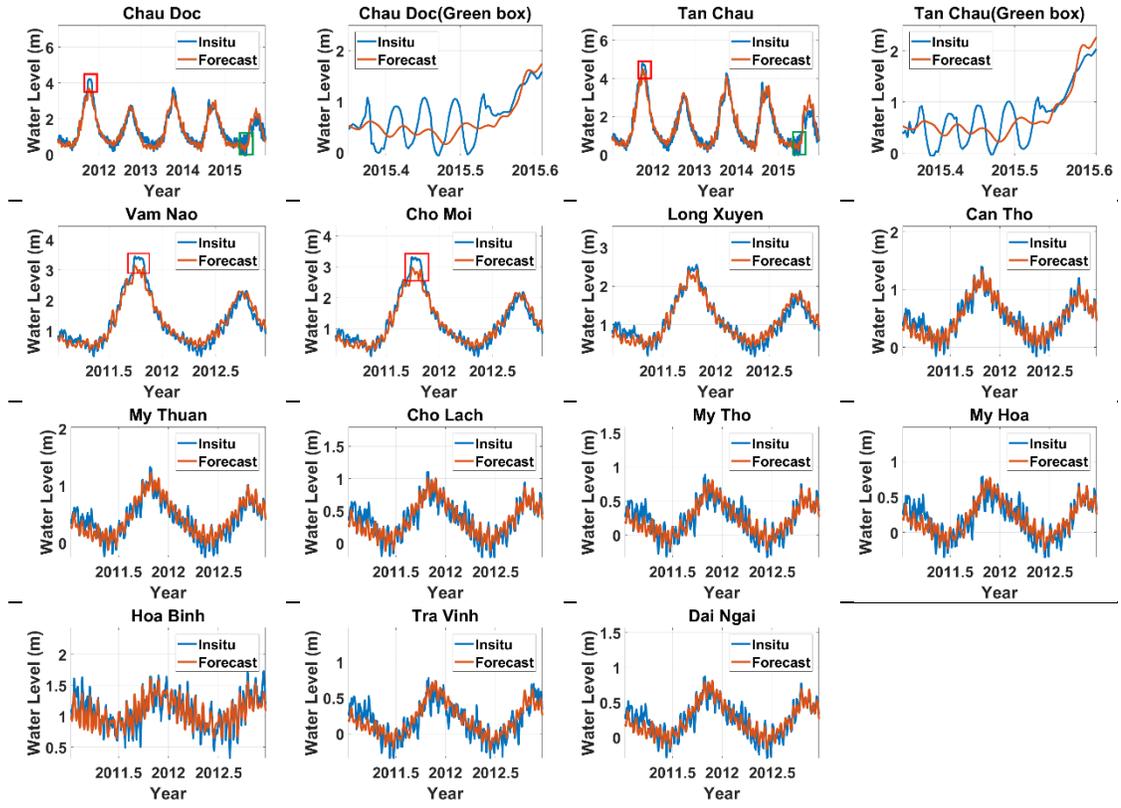
**Figure 3.13** Yearly averages and STDs of the MAEs and NSEs of 1-day to 20-day forecasting at 13 locations in the MD.



**Figure 3.14** The correlation coefficients of our forecasted water levels with in situ observations at 13 locations in the MD.

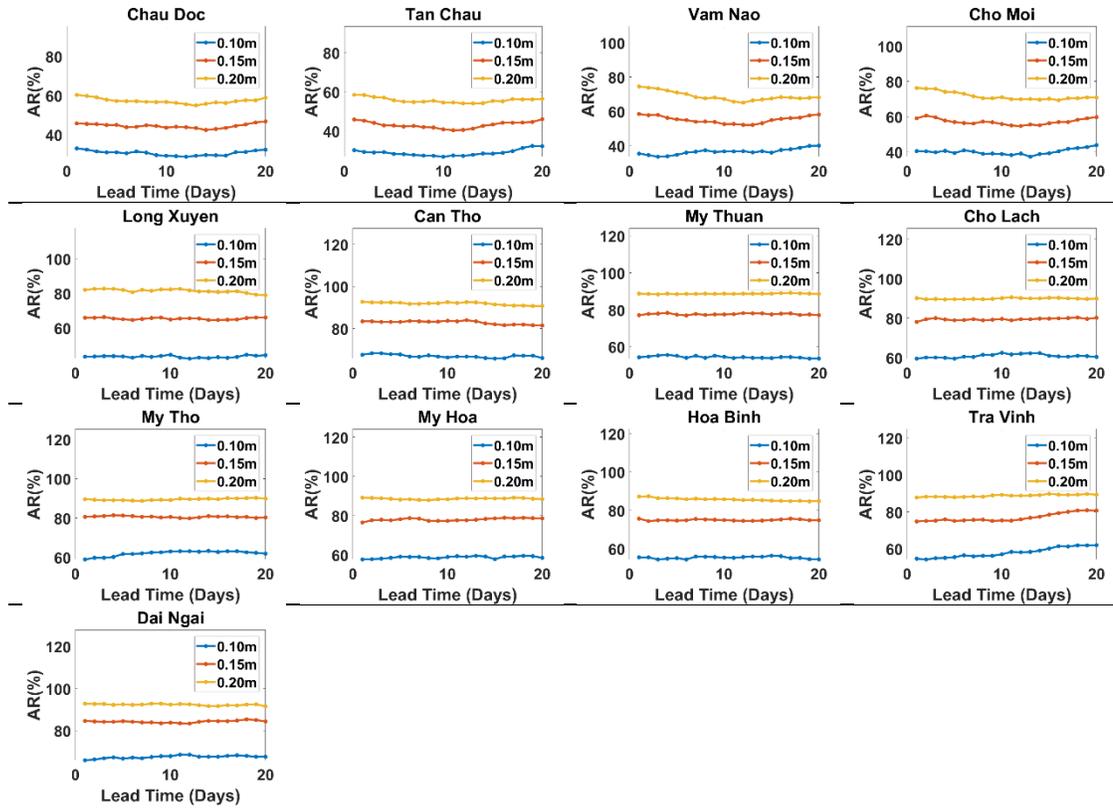
Figure 3.15 shows the time series of in situ and 10-day forecasted water levels. As marked in the red box of Figure 3.15 from Chau Doc to Cho Moi, a significant jump of water levels from the end of September to early November can be clearly observed. However, such extreme peaks did not occur at locations further downstream to the river mouth. Therefore, the forecasting skills at locations closer to the river mouth remain consistent. On the other hand, at Chau Doc and Tan Chau, the green boxes mark a period of June 2015 when the water level oscillations with a 14-day period were extremely strong. These strong water level oscillations may have resulted from a severe drought in 2015 (United Nations Country Team, 2016) caused by ENSO which may have

significantly intensified the tidal influence, leading to higher MAEs and lower NSEs in 2015 as Figure 3.13 shows. The tidal influence model built for this study are not able to accomodate the influence of ENSO on water levels, which has 3 – 6 years of period, since we used only less than three years of data to build the model. However, it can be addressed with longer time spans of data in the future, which would enable us to assimilate long-period ENSO signals into our tidal model.



**Figure 3.15** The 10-day forecasted (orange) and in situ (blue) water levels from 2011 to 2015 at Chau Doc/Tan Chau and from 2011 to 2012 at the other 11 locations in the MD. The extreme peak water levels (in red boxes) at Chau Doc, Tan Chau, Vam Nao and Cho Moi in 2011 wet season and the strong oscillations (green boxes) at Chau Doc and Tan Chau in June 2015 can be clearly seen in the plots right next to full time series.

Figure 3.16 shows the ARs in the MD with respect to 1-day to 20-day lead times. We calculated the ARs using 0.10 m, 0.15 m, and 0.20 m of AEs as the benchmark. It can be seen that at Chau Doc and Tan Chau, about 60 % of days meet the 0.20 m benchmark during 2011 to 2015 wet season. For Vam Nao and Cho Moi, about 60% of days during 2011 to 2012 wet season meet the 0.15 m benchmark and about 75 - 80% in the case of the 0.20 m benchmark. At Long Xuyen, about 66% of days reach the 0.15 m benchmark and about 81% of days in the case of the 0.20 m benchmark. At the locations downstream from Can Tho/My Thuan, there are about 55 - 70%, 75 - 85%, and 90 - 93% of days that meet the 0.10 m, 0.15 m, and 0.20 m benchmarks, respectively. As Figure 3.16 shows, the ARs become higher when performing forecasting closer to the river mouth. The patterns shown in Figure 3.12, 3.13 and 3.16 indicate that the forecasted river levels become more accurate near the river mouth. As the tidal effects on water levels in the coastal region become dominant and consistent compared to the upper MD (Dang et al., 2018), it is expected that the ocean tide model we proposed becomes more effective near the river mouth. As the acceptable accuracy of forecasting is 15 cm in the MD (Dr. Duong Du Bui, NAWAPI, personal communication), our system can provide promising forecasting for about 50 – 80% of days during wet season.



**Figure 3.16** The ARs of our forecasting at 13 locations in the MD. Different lines indicate the AR in the case of different benchmarks.

### 3.5 Conclusions and Perspectives

This study proposed a model-aided altimetry-based flood forecasting system for the MR, including the MR mainstem, CF and MD. The system integrates satellite altimetry, the VIC hydrological model and the tidal influence models. It is freely accessible, easy-to-set-up and has negligible computational cost, making it operationally sustainable and scalable to other places, especially for developing countries. Currently, the VIC hydrological model has already been set up in the Vietnamese Red River Basin (Hiep et al., 2018), Indus River Basin, Ganges River Basin, Brahmaputra River Basin, and Mangla River Basin

([http://depts.washington.edu/saswe/datavis\\_Timeseries.html](http://depts.washington.edu/saswe/datavis_Timeseries.html)). With VIC-derived discharges, altimetry-derived water levels and historic in situ data, our forecasting system can be easily implemented.

Although the cross-comparison with the MRC's forecasting shows that the forecasting skill of our system in the region outside of the MD still has room for improvement, our forecasting system is promising inside the MD which has complex hydraulic conditions. We have performed a 10-day or longer lead time of "pseudo-forecasting" at 13 locations in the MD starting from Tan Chau/Chau Doc which are the last downstream locations where the MRC provides forecasting. We obtained a promising forecasting skill, although due to the strike of tropical storms and strong water level oscillation in June 2015, the peak in situ water levels at Chau Doc to Cho Moi in 2011 were not captured in our forecasting results (See Chau Doc to Cho Moi in Figure 3.15.). Our system would be relatively easy and computationally inexpensive to be implemented by end-users and stakeholders in Cambodia and Vietnam. Johnston and Kummu (2012) summarized the MRC reports and estimated the cost of developing basin-scale models for the MRC to be about USD 20 million. Banks et al. (2014) also pointed out that the base price of a full-featured ISIS model starts at \$7,680 per year for a single-user license with an annual support and maintenance fee starting from \$1,350. Considering the expense of model building and maintenance, our freely available forecasting system could be operated as a complementary system to the RFMMC's forecasting, especially in the MD.

In our feasibility study, we linearly interpolated Jason-2 derived 10-day repeat TSL water levels and used them as "nowcast" for TSL daily water levels. However,

such interpolation cannot be done in real operational mode. Therefore, for operational purpose, up to 10 days of forecasted water levels in the TSL need to be obtained until the next revisit of the altimetry satellite. They may be obtained by time-series analysis methods such as autoregressive–moving-average model (ARMA) or autoregressive integrated moving average model (ARIMA). Alternatively, the smooth and seasonally dominant pattern of TSL water level changes are well represented by a Fourier series with known dominant frequencies by performing the Fourier analysis on historical TSL water levels.

For future studies, we would work on enhancing the forecasting skill and extending the lead time of our forecasting system with the aid of weather-based flow forecasting upstream. Sikder and Hossain (2018) proposed an operationally feasible method called bias-correction to obtain quantitative precipitation forecasting up to 15 days, and currently provides up to 7 days of weather-based VIC-derived water discharge forecasting with promising accuracy (Readers are referred to [http://depts.washington.edu/saswe/datavis\\_Timeseries.html](http://depts.washington.edu/saswe/datavis_Timeseries.html)). Since our forecasting system uses VIC-derived discharges to reconstruct daily water levels at upstream VSs, the use of forecasted VIC-derived water discharges can give us “forecasted” reconstructed daily water levels, which then can be used to perform forecast with extended lead times. For example, if the 7-day VIC-derived water discharge forecasting at an upstream VS is obtained, the 7-day forecasted water level can be reconstructed by using the discharge-to-level rating curve. This reconstructed water level forecasting with 7-day lead time can then be put into the forecasting model with 10-day lead time, for example, to forecast 17-day-later downstream water levels at the locations of

downstream in situ gauges. Although the accuracy of these “forecasted” water discharges are not as good as “nowcasted” water discharges as Sikder and Hossain (2018) indicated 0.77 – 0.57 of Cor. with 34 – 40% of normalized RMSEs (NRMSEs) with 1 to 7 days of lead times, compared with 0.86 of Cor. and 24.9% of NRMSE of nowcasted water discharges at Kampong Cham in the MR, it is expected that the capacity of the forecasting system proposed in this study still has potential to be improved with extended lead times. In addition, our system can be easily updated from year to year to keep up with climatic and anthropogenic structural and morphological changes, which are gradually altering the response of downstream water levels to the upstream discharges.

Local rainfall events at the forecasting locations also need to be addressed in the future as it may be an error source for our forecasting, since currently our system is only based on the linear relationship between upstream and downstream water levels. Moreover, it is expected that the ENSO-induced strong water level oscillations in the MD can be addressed if data with longer time spans are available in the future to build more comprehensive tidal influence models.

# **4. HINDCAST AND FORECAST OF DAILY INUNDATION EXTENTS USING SATELLITE SAR AND ALTIMETRY DATA WITH ROTATED EMPIRICAL ORTHOGONAL FUNCTION ANALYSIS: CASE STUDY IN TONLE SAP LAKE FLOODPLAIN**

## **4.1 Introduction**

This chapter is based on the content of Chang et al. (2020). Tonle Sap Lake (TSL) is the largest natural freshwater lake in Southeast Asia. The lake is well known for its unique seasonally reversed flow. In the wet season, the Mekong River (MR) level continuously rises and eventually exceeds the TSL water level. Consequently, water flows across the floodplain toward the lake. In the dry season, water flows from the lake and floodplains down the Tonle Sap River toward the sea, as the MR level recedes (Campbell et al., 2009). Such flow reversal cannot be found anywhere else in the world and has a significant impact on the surrounding TSL Floodplain (TSLF) ecosystem, making it one of the most productive ecosystems in the world (Lamberts, 2006) and extremely important for Cambodia for food security and economy (Kummu et al., 2006). In recent years, several studies have pointed out that climate change and water infrastructure development would pose stress on flood pulse and intensity of TSL (Kummu and Sarkkula, 2008; Pokhrel et al., 2018) and consequently impact the ecosystem and productivity of agriculture and aquaculture in the TSLF ecosystem (Kummu and Sarkkula, 2008; Lamberts and Koponen, 2008; Lauri et al., 2012; Lutz et al., 2014; Pokhrel et al., 2018; Västilä et al., 2010). Simulations by Västilä et al. (2010)

also show an increase of the average and maximum water levels and flood duration may cause more severe damages to roads, buildings, and other infrastructure located on the floodplain as well as other flood-related impacts such as destruction of rice crops, rise of hygiene problems and more human victims particularly those living close to the TSL (Keskinen, 2006; Nuorteva et al., 2010). Therefore, to have a better understanding of the dynamics of TSLF inundation extents and the ability to forecast future inundation extents are of great importance and urgently needed. They can also be helpful by providing more information for assessing the change of fish catches and potential flooding damages as well as corresponding relief services (Schumann and Moller, 2015).

For the purpose of detecting flooded areas, Synthetic Aperture Radar (SAR) is considered the most useful sensor (Yan et al., 2015) with the capacity to provide surface observations without spatial gaps both day and night under all weather conditions thanks to its use of active microwave signals. However, there is still no study using SAR imagery to perform daily, gap-free areal inundation extent mapping. The use of SAR imagery for performing forecasting of areal inundation extents has never been investigated before either, despite its increasing importance for improving first response and water resource management. In this study, to address the need of forecasting high temporal frequency areal inundation extents, we propose a daily areal inundation estimation framework based on synthesized SAR intensity imagery. The framework first applies Rotated Empirical Orthogonal Function (REOF) analysis (Kaiser, 1958; Lorenz, 1956), which has been shown to improve physical interpretability over conventional EOF analysis (A Hannachi et al., 2006; Hannachi et al., 2007; Lian and

Chen, 2012), on multi-temporal SAR intensity images to extract their spatiotemporal patterns. The extracted spatiotemporal patterns were then associated with satellite altimetry-derived water levels. Next, satellite altimetry-derived water levels were used as input to synthesize SAR intensity images, where the synthesized SAR intensity images are of the same date as the input satellite altimetry-derived water levels. In other words, SAR intensity at any time can be synthesized if satellite altimetry-derived water levels are available. Finally, corresponding inundation extents were estimated from the use of both synthesized SAR intensity imagery and the Multi-Error-Removed Improved-Terrain Digital Elevation Model (MERIT DEM) (Yamazaki et al., 2017) through an unsupervised K-means clustering algorithm (Arthur and Vassilvitskii, 2007; Lloyd, 1982). Note that even though EOF analysis (Lorenz, 1956) has been widely used in climate sciences for coupling different fields, reconstruction and prediction (Bracher et al., 2015; Church et al., 2004; Imani et al., 2017; Taylor et al., 2013; Yosef et al., 2017), this is the first study applying EOF analysis to SAR imagery for SAR intensity synthesis and depicting areal inundation extents by integrating with satellite altimetry data, to the best of our knowledge. Such application of satellite altimetry is also not reported in the literature.

Here, the framework was applied to TSLF for daily hindcasting and forecasting of areal inundations by using multi-temporal Sentinel-1A imagery and daily Jason series altimetry-derived TSL levels. The daily hindcasting and forecasting are fulfilled by using daily linear-interpolated historical and El Niño/Southern Oscillation (ENSO) index-forecasted altimetry-derived TSL levels. The estimated inundation extents were cross-compared with reference datasets including inundation maps derived from 8-day

composite MODerate resolution Imaging Spectro-radiometer mission (MODIS) products as well as Sentinel-1 images. The proposed framework has the following features: (1) Synthesis of SAR intensity image and estimation of areal inundation extents at any time as long as the water level data is available; (2) The framework is fully remote sensing-based in which a computationally expensive model is not required; (3) Since the framework exploits SAR imagery, the resulting estimated inundation extents are free from cloud cover with no spatial gap; (4) The framework has the potential to be applied to the floodplains of other major river basins such as the Amazon River Basin and Congo River Basins.

## 4.2 Data

### 4.2.1 Sentinel-1 SAR Data

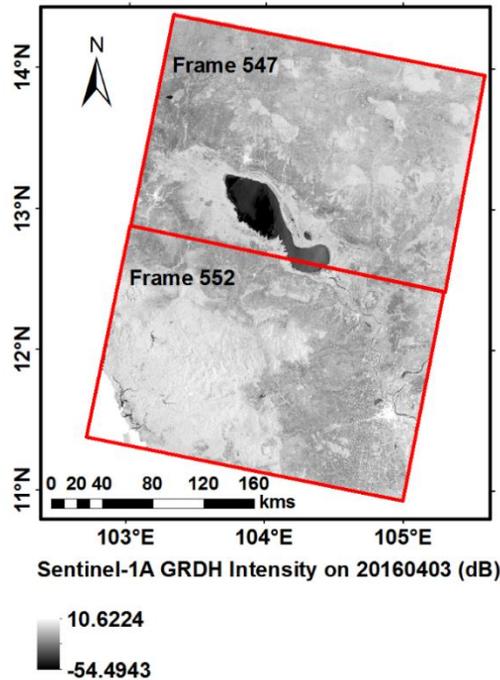
Sentinel-1 is a two-satellite-constellation mission (Sentinel-1A/-1B), equipped with C-band (5.405 GHz) SAR, under the Copernicus Earth observation program at ESA. The first satellite, Sentinel-1A, was launched on April 3<sup>rd</sup>, 2014, while Sentinel-1B was launched on April 25<sup>th</sup>, 2016. Both satellites feature free accessibility and a 12 day revisit time, which have expanded the use of SAR imagery in the study of environmental change (Markert et al., 2018; Tsyganskaya et al., 2018a; White et al., 2014). Sentinel-1A VV-polarization images in the Ground Range Detection High-resolution (GRDH) product were downloaded from the Alaska Satellite Facility (ASF). In order to have image scenes covering the whole TSL and surrounding floodplain area, two frames, frame 552 and 547 of path 91, acquired on the same date were used. The

two frames on each acquisition date were mosaicked into one single scene (See Figure 4.1).

In this study, Sentinel-1A data were used solely for building the framework. This is to avoid the influence of potential systematic differences between backscattering intensities acquired by two satellites (Sentinel-1A/-1B) on the REOF analysis, which would affect coupling the extracted temporal patterns with altimetry data, and thus the estimated inundation extents (For detail of the framework, see Section 4.3.1). The purpose was to minimize the influence of signals not related to natural phenomena. Time span of images used for building the framework are from April 3<sup>rd</sup>, 2016 to December 31<sup>st</sup>, 2018, giving us 78 images in total after mosaic. Sentinel-1A images with temporal coverage from January 6<sup>th</sup> to July 29<sup>th</sup>, 2019, and Sentinel-1B images from January 12<sup>th</sup> to July 23<sup>th</sup>, 2019, were used as reference datasets, in addition to MODIS imagery for cross-comparison purposes, giving us 54 images in total after mosaic. Since our proposed framework exploits Sentinel-1 SAR imagery, using Sentinel-1A/-1B imagery for cross-comparison avoids the influence of inherent inconsistency between radar and optical imagery in cross-comparison results. The cross-comparison using Sentinel-1 imagery as a reference dataset was conducted in the forecasting case with time span starting from January 2019 because the earlier Sentinel-1 images were all acquired within the time span of imagery used for building the framework and may affect cross-comparison independence to some degree.

Mosaicked images were pre-processed (multi-looked, radiometric terrain corrected, and geocoded) and co-registered with respect to the image acquired on April 3<sup>rd</sup>, 2016, which is the image with the earliest acquisition date. Note that in this study

we multi-looked the images to 500 m of spatial resolution to fit the MODIS reference dataset. The pre-processing procedure including image mosaic was performed using the GAMMA software (Werner et al., 2000).



**Figure 4.1** Frames of Sentinel-1 GRDH intensity images used in this study. The image acquired on April 3<sup>rd</sup>, 2016 was used as an example.

#### 4.2.2 MODIS Surface Reflectance Data and Yearly Water Mask

The MODIS spectro-radiometer onboard the Terra and Aqua satellites, launched in 1999 and 2002 respectively, acquires Earth surface radiances in 36 spectral bands. (<https://modis.gsfc.nasa.gov/>). MODIS is the only dataset that has long temporal coverage starting from 2000 and has high temporal resolution as well as high spectral resolution (36 bands). We would like to investigate the ability of the proposed framework with as much reference data as possible for more robust cross-comparison results. Therefore, MODIS data was used as reference despite its relatively coarse

spatial resolution. In fact, MODIS data has been used for monitoring and long-term analysis of inundation extents in many previous studies due to its high temporal resolution (Frappart et al., 2018; Gumma et al., 2014; Huang et al., 2014; Islam et al., 2010; Normandin et al., 2018; Sakamoto et al., 2009, 2007) in spite of the cloud cover issues (Huang et al., 2014, 2013).

In this study, tile H28V07 of MODIS products MOD09A1 and MOD44W were downloaded from the United States Geological Survey (USGS) Earthexplorer website (<https://earthexplorer.usgs.gov/>). MOD09A1 is surface reflectance data derived from the Terra satellite raw radiance measurements. It includes surface spectral reflectance for band 1 to band 7 at 500 m spatial resolution with atmospheric conditions, including gasses, aerosols, and Rayleigh scattering, corrected. For each pixel, the best surface reflectance data during an 8-day period was selected based on cloud cover and solar zenith. Here, MOD09A1 images from over a decade (2003 to 2015) and January to July in 2019 at 500 m spatial resolution were used as the cross-comparison reference dataset.

MOD44W is a 250 m resolution water mask product which was available from 2000 to 2015. Since 2015 is the year with the minimum inundation extent for the TSLF area in the last two decades (Frappart et al., 2018), the MOD44W water mask in 2015 was adopted as a permanent water body mask when building the framework. This is to avoid the influence of surface roughness change-induced intensity variation over permanent water body on the REOF analysis results (See Section 4.3.1 for detail).

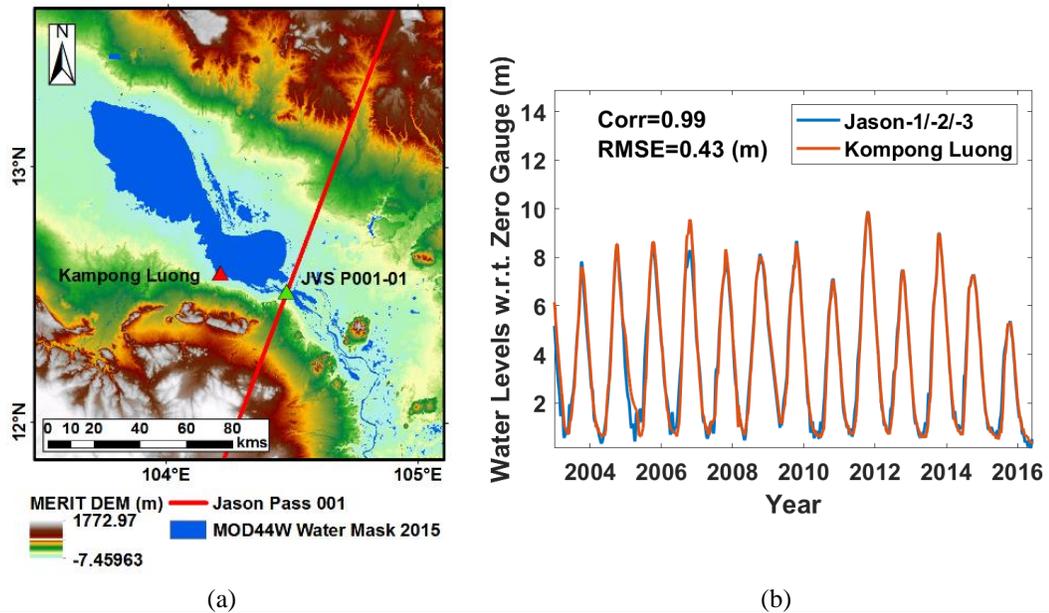
#### 4.2.3 **MERIT DEM**

MERIT DEM (Yamazaki et al., 2017) is a global DEM with respect to the Earth Gravitational Model 1996 (EGM96) with 3 arc-second spatial resolution (about 90 m at the equator). The baseline DEMs include 3 arc-second spatial resolution Shuttle Radar Topography Mission DEM (SRTM3 DEM) and the Advanced land observing satellite World 3D-30 m DEM (AW3D-30m DEM), in the region of 60° S to 60° N and 60° N to 90° N, respectively. The unobserved gaps in both SRTM3 and AW3D-30 m DEMs are filled with the Viewfinder Panoramas DEM. The NASA Ice, Cloud, and land Elevation Satellite (ICESat) laser altimetry global land surface elevation data (GLAH14) is used as the reference ground elevation for DEM bias estimation. DEM errors due to forest canopy are estimated with the use of the University of Maryland Landsat forest cover data (Hansen et al., 2013) and NASA global forest height data (Simard et al., 2011). For more details, please refer to Yamazaki et al. (2017). In this study, the DEM was multi-looked to spatial resolutions of 500 m as preprocessed Sentinel-1A GRDH images.

#### 4.2.4 **Jason Altimetry-derived and In-situ Water Levels at TSL**

In this study, we used TSL water levels from two data sources, including Jason series satellite altimetry and the in-situ gauge at Kampong Luong. Jason series altimetry are the successors of Topex/Poseidon (T/P). The satellite series consists of Jason-1, Jason-2 and Jason-3, launched on December 7th, 2001, June 20th, 2008, and January 17th, 2016, respectively. The series of missions are under the cooperation of NASA and Centre National d'Etudes Spatiales (CNES), with additional partnership from the

National Oceanic and Atmospheric Administration (NOAA), and the European organization for the exploitation of METeorological SATellites (EUMETSAT) for Jason-2 and Jason-3. As the series of satellites maintain the same orbit configuration as T/P, they continuously provide highly accurate altimetry data with an  $\sim 10$ -day repeat cycle and allow long-term inland water level monitoring. In this study, we used 20-Hz ICE-retracked ranges from the Geophysical Data Record (GDR) E and D for Jason-1 and Jason-2/-3 respectively, to extract water levels at a Virtual Station (VS) on TSL. Outliers in each cycle of measurements were removed (Okeowo et al., 2017). Biases between water level time series of different Jason missions were calculated based on the difference of mean water level time series during the overlapping period between missions and were aligned with those of Jason-1. The concatenated Jason-1/-2/-3 water levels consist of Jason-1 data from cycle 1 to cycle 238, Jason-2 data from cycle 1 to cycle 280, and Jason-3 data from cycle 1 to cycle 106 and are with respect to the World Geodetic System 1984 (WGS84) ellipsoid. In situ water levels at Kampong Luong with respect to local zero gauge up to 2016 were provided by the Asian Disaster Preparedness Center (ADPC) and were treated as in situ water levels of TSL. Locations of the Jason satellites ground track, the VS and in situ gauge are shown in Figure 4.2(a). Figure 4.2(b) shows the time series of altimetry-derived and in situ water levels at TSL up to 2016. The bias between altimetry-derived and in situ water levels is 14.53 m. Shifting the altimetry-derived time series toward the in situ time series allows us to obtain accuracy (root mean square error, RMSE) of Jason-1/-2/-3 concatenated altimetry-derived TSL water levels of 0.43 m with a high temporal correlation of 0.99. In situ TSL levels from January to July of 2019 were also used for validating forecasting results.



**Figure 4.2** (a) Geographical locations of Jason altimetry satellite ground track passing through the TSL and the corresponding VS and in situ gauge at Kampong Luong, (b) Jason-1/-2/-3 concatenated altimetry-derived water levels at the TSL and comparison with in situ data.

#### 4.2.5 ENSO Index – MEI

The second version of Multivariate ENSO Index (MEI) (Wolter and Timlin, 2011, 1998, 1993) used in this study was processed, organized, and distributed by the Physical Sciences Division of NOAA Earth System Research Laboratory (ESRL) (<https://www.esrl.noaa.gov/psd/enso/mei/>). It uses 5 variables, including sea level pressure, sea surface temperature, surface zonal and meridional winds and outgoing longwave radiation to generate time series of ENSO conditions since 1979. The sea surface temperature, sea level pressure, and surface winds are obtained from the high-quality Japanese 55-year Reanalysis (JRA-55) (Kobayashi et al., 2015). The outgoing longwave radiations are obtained from the NOAA Climate Data Record (CDR) monthly outgoing longwave radiation product version 2.2-1. All of the data fields are interpolated

to 2.5° grid size. Standardized anomalies of each field are then computed with respect to the period from 1980 – 2018. MEI is then calculated as the leading principal component time series of the EOF for the standardized anomalies of 5 combined variables within the region from 30°S – 30°N, 100°E – 70°W, excluding the Atlantic Ocean and land with latitudinal weighting. Positive MEIs represent El Niño events, while negative MEIs represent La Niña events.

### 4.3 Methods

#### 4.3.1 REOF-based Daily Inundation Extent Estimation Framework

##### 4.3.1.1 REOF Analysis

The REOF analysis starts from conventional EOF analysis (Lorenz, 1956). Consider  $X$  is an array of input data which is  $p$ -dimensional time series data with  $n$  observations

$$X_{n \times p} = \begin{bmatrix} x_{1,1} & x_{1,2} & \cdots & x_{1,p} \\ x_{2,1} & x_{2,2} & \cdots & x_{2,p} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n,1} & x_{n,2} & \cdots & x_{n,p} \end{bmatrix}. \quad (4.1)$$

We can take  $X$  as an aggregation of  $n$  maps, each map has  $p$  pixels; therefore, each row of  $X$  is a map at an acquisition time, while each column is a time series of values of a pixel from  $n$  maps. Since the essence of EOF analysis is to find the variables that can effectively represent the variability of  $X$ , a covariance matrix of  $X$  needs to be formed first. In this study, we attempt to retrieve the temporal variability of  $X$ ; therefore, a temporal anomaly array  $X'$  is first calculated by subtracting the temporal average array  $\bar{X}$  from  $X$

$$X'_{n \times p} = X_{n \times p} - \bar{X}_{n \times p} \quad (4.2)$$

where each column of  $\bar{X}$  is a column vector whose elements have the same value as the average value of elements in the corresponding column in  $X$ .

Then a  $p$ -by- $p$  covariance matrix  $C$  of the input data  $X$  can be obtained by

$$R_{p \times p} = X'^T_{p \times n} X'_{n \times p}. \quad (4.3)$$

The covariance matrix  $R$  here describes the temporal variability of  $X$  as its derivation adopts temporal anomalies of  $X$ . Then an  $p$ -by- $p$  array  $U$  whose column vectors are eigenvectors of  $R$  and a  $p$ -by- $p$  diagonal array  $\Lambda$  whose elements are the corresponding eigenvalues  $\lambda$  can be calculated by solving the eigenvalue problem

$$R_{p \times p} U_{p \times p} = U_{p \times p} \Lambda_{p \times p}. \quad (4.4)$$

Each eigenvector is a unit vector in  $p$ -dimensional space, pointing to the direction where  $X$  has significant variances. Each column  $j$  of  $U$ , where  $j$  can be from 1 to  $p$ , is also called a *mode* of an eigenvector which can be plotted as a map, representing the pattern of spatial variability of  $X$ , that is Spatial Mode (SM). The eigenvector array  $U$  is an orthogonal array, that is  $U^T U = U U^T = I$ , meaning that eigenvectors are uncorrelated (orthogonal) to each other over space.

How each mode of SM evolves in time can be determined by projecting  $X$  onto it

$$Z_{n \times p} = X_{n \times p} U_{p \times p} \quad (4.5)$$

where each column of  $Z$  is a  $n$ -dimensional vector representing the time series of evolution of the corresponding SM, that is Temporal Principal Component (TPC), which is uncorrelated (orthogonal) in time. Each row of  $Z$  represents a time epoch, while each column corresponds to a mode of SM. The explained variance of each mode is the

variances of the columns of  $Z$ . Columns of  $U$  and  $Z$ , which are SMs and TPCs, respectively, are sorted by their corresponding explained variances. Hereafter, the sorted  $U$  and  $Z$  are simply called  $U'$  and  $Z'$ . Therefore, the first column of  $U'$  and  $Z'$  is the SM that explains the maximum extent of variability in  $X$  (called mode-1), while the second-column one explains the second most variability (called mode-2), and so forth. By using  $Z'$  and  $U'$ , input data at a specific acquisition time, noted as  $X(t)$  can be synthesized through linear combination as follows if  $k = p$ :

$$X(t) = \sum_{j=1}^k z'_{t,j} \cdot u'_j + \bar{X}; t = 1, 2, \dots, n \quad (4.6)$$

where  $z'_{t,j}$  is a scalar that is the element of the mode- $j$  TPC at time epoch  $t$  and  $u'_j$  is the mode- $j$  SM. The process is called *synthesis*. If  $k = m < p$ , it is called *truncated synthesis*, which is often the primary motivation when applying EOF analysis (Wilks, 2011). Since these modes of SMs and TPCs from conventional EOF are orthogonal to each other, difficulty can occur if interested in physically interpreting SMs and the corresponding TPCs as natural phenomena which are rarely mutually independent. REOF analysis, on the other hand, is able to relieve the orthogonality constraint of EOF and improves the physical interpretability of SMs and TPCs (Hannachi et al., 2007, 2006; Lian and Chen, 2012). Since we are associating TPCs with altimetry-derived water levels, which is a way of physical interpretation (will be explained later in Section 4.3.1.2), REOF analysis is applied. Other advantages of REOF analysis such as the ability to avoid unphysical dipole-like patterns from EOF analysis and simplification of spatial structures while retaining robust patterns have been mentioned by several previous studies (Cheng et al., 1995; Dommenges and Latif, 2002; A. Hannachi et al.,

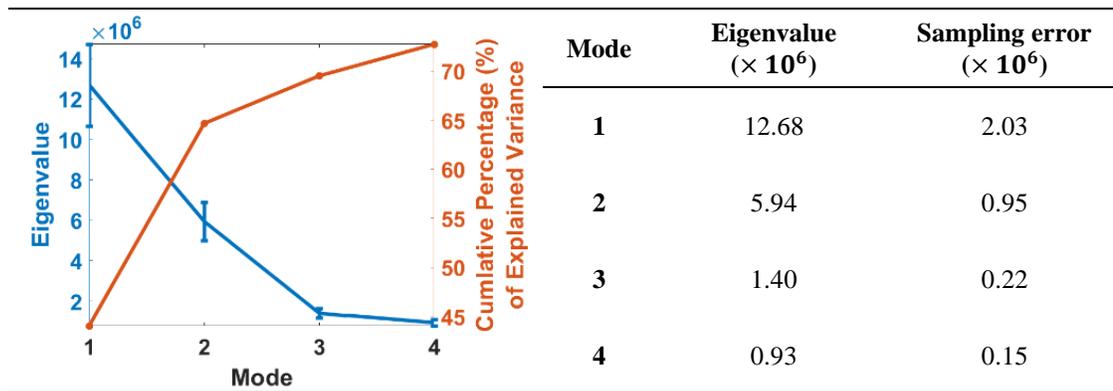
2006; Houghton and Tourre, 1992). REOF analysis is based on rotation of SMs with a linear transformation of a truncated  $m$  subset of  $U'$ :

$$\tilde{U}_{p \times m} = U'_{p \times m} T_{m \times m} \quad (4.7)$$

where  $\tilde{U}$  is an array whose column vectors are modes of Rotated SMs (RSMs) and  $T$  is a rotation matrix. The number of modes to be truncated,  $m$ , is arbitrary and is typically the number of leading EOF modes selected based on some truncation criteria (Wilks, 2011). Here, the “rule of thumb” of North et al. (1982) was used to select significant modes for truncation. The differences between the eigenvalue of a mode and its adjacent mode need to be at least the sampling error for a mode to be significant. The sampling errors of the eigenvalues were first calculated by:

$$\delta\lambda = \lambda \sqrt{\frac{2}{N}} \quad (4.8)$$

where  $\delta\lambda$  is sampling error of a specific eigenvalue  $\lambda$  and  $N$  is the number of samples which is the number of observations. The rule of thumb results in  $m = 4$ , accounting for 72.7% of the total explained variances (See Figure 4.3).



**Figure 4.3** Eigenvalues and corresponding sampling errors calculated by the rule of thumb of North et al. (1982) and cumulative percentages of explained variances of significant modes in the cases of using 500 m spatial resolutions of Sentinel-1A GRDH images as input.

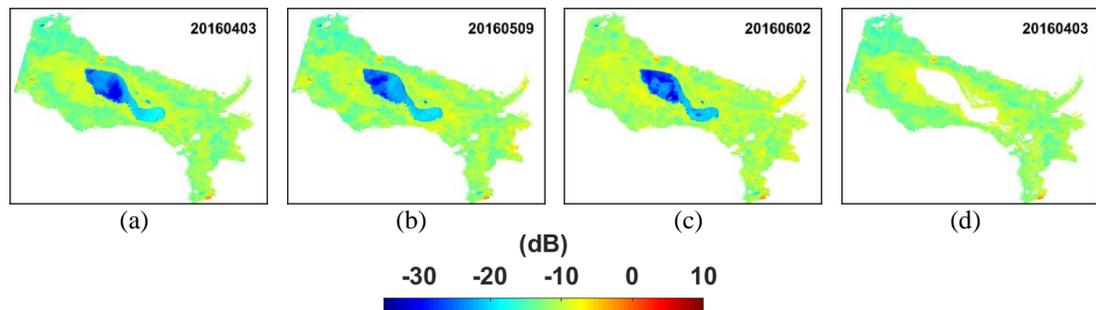
Varimax orthogonal rotation (Kaiser, 1958), the most commonly used rotating approach (Richman, 1986), was then applied to rotate  $U'$ , which is determined by choosing the elements of  $T$  that maximizes the condition:

$$\sum_{k=1}^m \left[ \frac{1}{p} \sum_{j=1}^p u_{j,k}^{*4} - \left( \frac{1}{p} \sum_{j=1}^p u_{j,k}^{*2} \right)^2 \right], \quad (4.9)$$

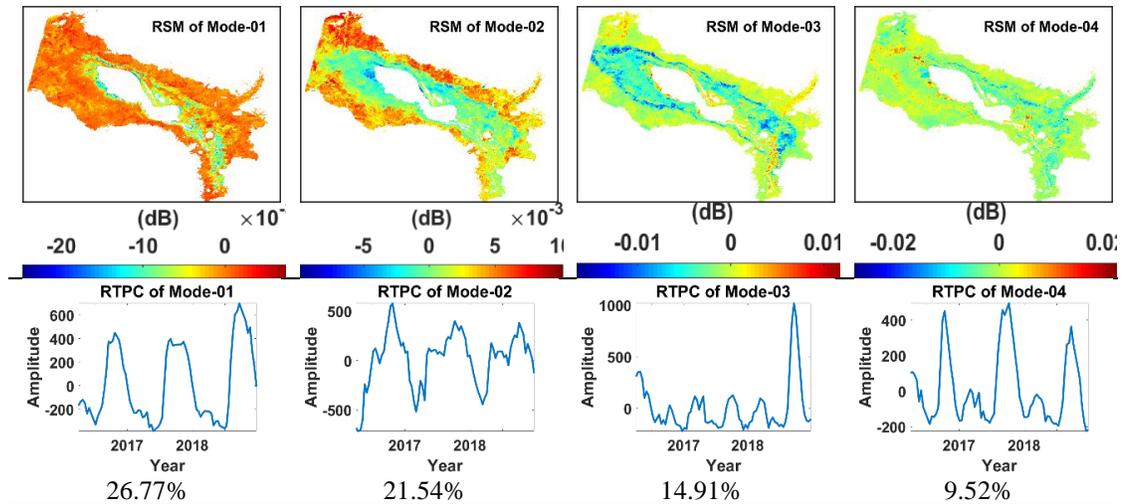
$$u_{j,k}^* = \frac{\tilde{u}_{j,k}}{\sqrt{\sum_{k=1}^m \tilde{u}_{j,k}^2}}$$

where  $\tilde{u}_{j,k}$  is element of  $\tilde{U}$  at  $j$ -th row and  $k$ -th column and  $u_{j,k}^*$  is row-normalized  $\tilde{u}_{j,k}$ . By replacing  $U$  in Equation (4.5) with  $\tilde{U}'$ , array  $\tilde{Z}'$  whose column vectors are Rotated TPCs (RTPCs) can be obtained. The varimax rotation was calculated using the National Center for Atmospheric Research (NCAR) Common Language (NCL) (NCL, 2019). Since REOF analysis redistributes the variance represented by the results from conventional EOF, columns of  $\tilde{U}'$  and  $\tilde{Z}'$  were reordered based on the explained variance of each mode, that is the variance of columns of  $\tilde{Z}'$ . The reordered  $\tilde{U}'$  and  $\tilde{Z}'$  is then noted as  $\tilde{U}''$  and  $\tilde{Z}''$  with each column representing a mode of RSM and RTPC, respectively. By plotting RSMs as maps and RTPCs as time series, spatiotemporal patterns of  $\llcorner$  input multi-temporal stack of Sentinel-1 GRDH intensity images can be seen. Since we performed REOF analysis up to mode 4, we have 4 such time series of RTPCs and corresponding RSMs. Data synthesis can be fulfilled by replacing  $z'_{t,j}$  and  $u'_j$  from Equation (4.6) with their counterparts in  $\tilde{Z}''$  and  $\tilde{U}''$ . Therefore, by estimating historical or future  $\tilde{Z}''$ , we are able to hindcast historical data or perform forecasting. This will be covered in the next section.

In this study, areas within the Tonle Sap Watershed (TSW) with elevations below 23 m were taken as the study area (Frappart et al., 2018). Considering that the change of surface roughness over the water bodies can increase the intensities of SAR images and alter the REOF analysis results, only pixels within the study area excluding those over the permanent water bodies were adopted as input data  $X$  (See Figure 4.4). This is to ensure the spatiotemporal patterns retrieved by the analysis are really from the surrounding floodplain. The extraction of pixels was conducted using the TSW boundary shapefile provided by Open Development Cambodia (<http://www.opendevdevelopmentcambodia.net/maps/downloads/>), MERIT DEM and the MOD44W water mask from 2015. The resulting RSMs and RTPCs are shown in Figure 4.5.



**Figure 4.4** (a) to (c) are examples of Sentinel-1A GRDH images within the TSW with MERIT DEM elevation below 23 m at different acquisition times. (d) is an example of images with permanent water bodies masked out.

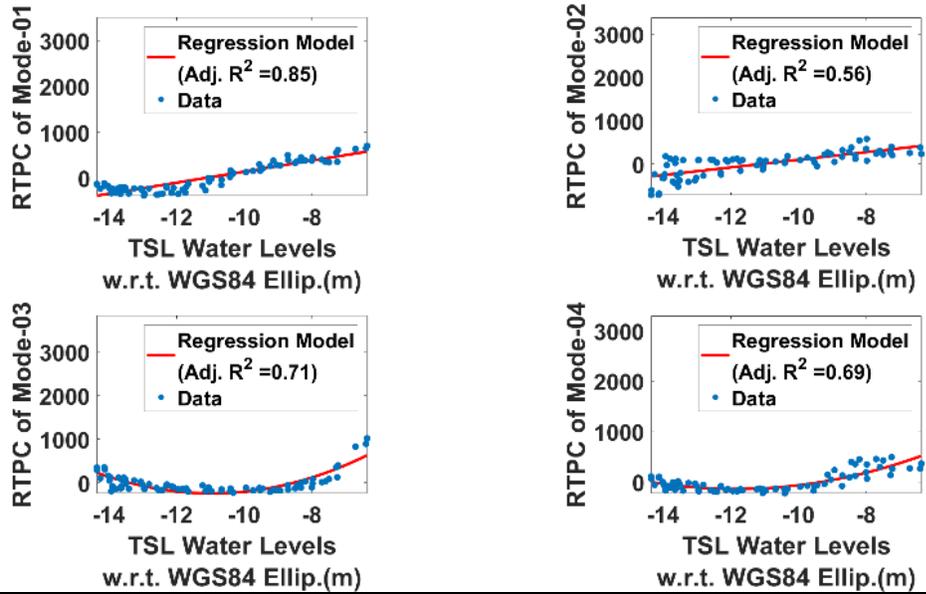


**Figure 4.5** Mode-1 to mode-4 RSMs (top), corresponding RTPCs, and percentages of explained variance (bottom) of input multi-temporal stack of Sentinel-1 GRDH intensity images at 500 m spatial resolution.

#### 4.3.1.2 Synthesis of SAR Intensity Images

Following the *synthesis formula* in Equation (4.6), Sentinel-1 GRDH intensity images, either historical or forecasted ones, can be generated by multiplying RSMs with corresponding RTPCs at past or future time. To achieve this, estimating temporally varying RTPCs at the given time is required. Here, we coupled the resulting RTPCs with TSL water levels based on polynomial regression. We applied a 1-degree (linear) polynomial model in the case of mode-1 and mode-2. For mode-3 and mode-4, 2-degree (quadratic) polynomial models were adopted. The choice of polynomial degree for different modes can be justified by the corresponding RSMs. As Figure 4.5 shows, RSMs for mode-1 and mode-2 have strong negative signals in the area around TSL, while mode-3 and mode-4 RSMs are distributed in the area farther from TSL. Therefore, RTPCs for mode-1 and mode-2 quickly respond to changes in TSL water levels, leading to a linear pattern in the scatter plots. RTPCs of mode-3 and mode-4, on the other hand, would only respond when TSL water surface rises to a certain level. The data

distribution in the scatter plot reflects this fact (See Figure 4.6). There is relatively flat and dense data distribution in the scatter plot when TSL water level is lower. Once the TSL water level reaches a certain level, the data distribution shows a rising slope. The fitted polynomial models are shown in Figure 4.6 as well. With the given TSL water levels at a specific time epoch the modes of RTPCs can be estimated. Then the data at a given time can be synthesized by using Equation (4.6).



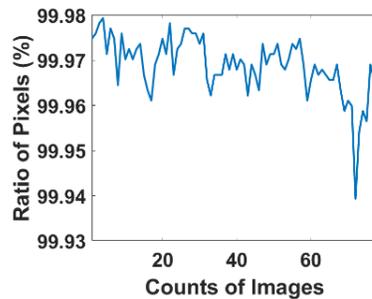
**Figure 4.6** Polynomial regression models between altimetry-derived water levels w.r.t. WGS84 ellipsoid at TSL and the mode-1 to mode-4 RTPCs of Sentinel-1 GRDH intensity images at 500 m spatial resolution.

As the error of the REOF-based synthesized SAR intensity can influence the estimation of inundation extents, we analyzed the difference between the synthesized and original SAR intensity. SAR images adopted for REOF analysis were used for comparison. The sign of both synthesized and original SAR intensity for each pixel was first investigated. Figure 4.7 shows the time series of the percentages of pixels whose original and synthesized intensities are both negative, which are nearly 100%, of each SAR image. This means that when differences between synthesized and original SAR

intensities (subtracting original ones from synthesized ones) are large positive values, the original SAR intensities would be much smaller than the synthesized ones. For example, the red circles in Figure 4.8(c) mark areas with large positive difference values. These areas match with areas where the original intensities are much smaller than the synthesized ones and may be a source of underestimated inundation extents, leading to omission errors. These areas have relatively higher elevations. Since the synthesis of SAR intensity is based on coupling temporal patterns of SAR intensity variations with TSL levels, the SAR intensity variations over areas with elevations higher than TSL levels may not be synthesized accurately as they are not necessarily caused by TSL level variations. Another example shows that there are large positive difference values in the areas around the boundary of TSLF (Figure 4.9(c)). The original intensities in these areas are much smaller than the synthesized ones as well and may lead to underestimation in our inundation extents. Since the areas are along the boundary of TSLF, the intensity variations are possibly related to TSL levels. Hence, the large positive differences in these areas may be caused by discrepancies between real temporal patterns and the altimetry-estimated ones.

On the other hand, if the differences are small negative values, the synthesized SAR intensities would be much smaller than the original ones, which may result in overestimated inundation extents and lead to commission errors. This can be seen in the red circles in Figure 4.10(c) which are along the boundary of TSL. As these areas are along the boundary of TSL, the intensity variations are also possibly related to TSL levels. The differences in these areas may also be due to errors of our altimetry-estimated temporal patterns. For a related discussion about framework skills, omission

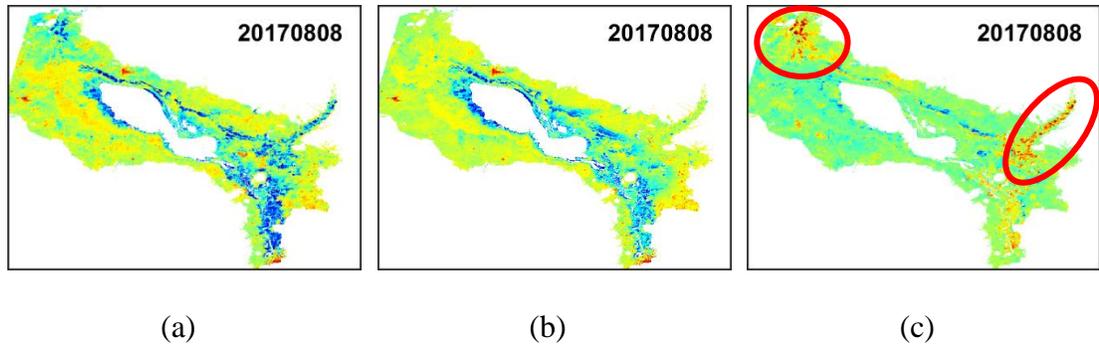
and commission errors of result and the influence of high-elevation inundation extent, please refer to Section 4.4. Finally, since pixels over permanent water bodies have been excluded before REOF analysis based on MOD44W water mask of 2015, long-term temporal averages of intensities over these permanent water body pixels were calculated and applied to the synthesized data to obtain a complete scene, which then can be used to estimate inundation extents. Note that for Sentinel-1 images used as reference datasets for cross-comparing with our estimated inundation extents, intensities over permanent water body pixels were also replaced with long-term temporal averages. This is because intensities over permanent water body can be enhanced by surface roughness change, which influences the estimation of inundation extents.



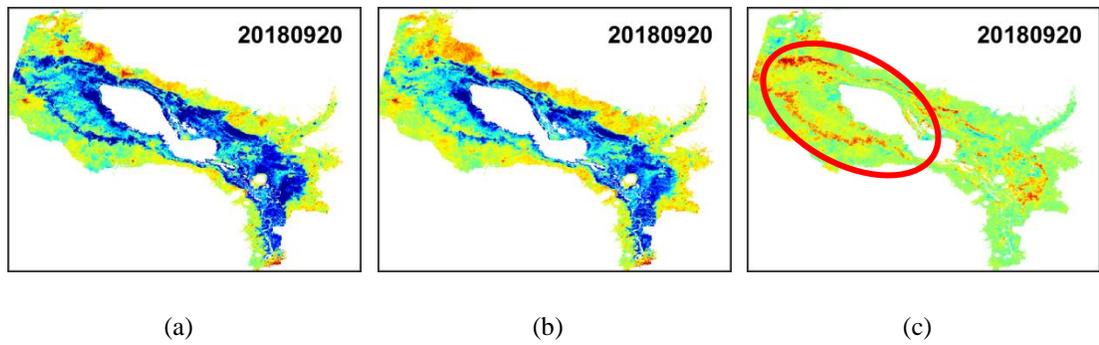
(a)

---

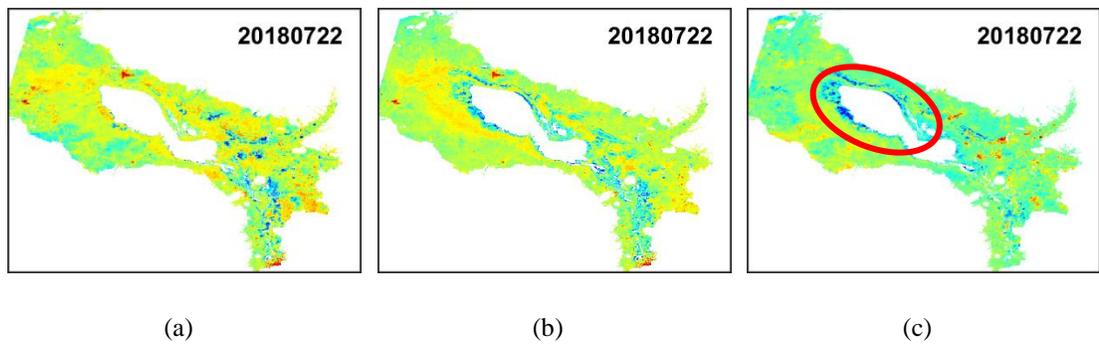
**Figure 4.7** Percentages of pixels of each SAR image whose original and synthesized intensity are both negative.



**Figure 4.8** Example of (a) original SAR intensity, (b) synthesized SAR intensity, and (c) difference by subtracting (a) from (b) where there are large positive difference values in high-elevation areas (red circles in (c)).



**Figure 4.9** Example of (a) original SAR intensity, (b) synthesized SAR intensity, and (c) difference by subtracting (a) from (b) where there are large positive difference values in the areas around the boundary of TSLF (red circles in (c)).



**Figure 4.10** Example of (a) original SAR intensity, (b) synthesized SAR intensity, and (c) difference by subtracting (a) from (b). Red circles mark areas where there are really small negative difference values in (c).

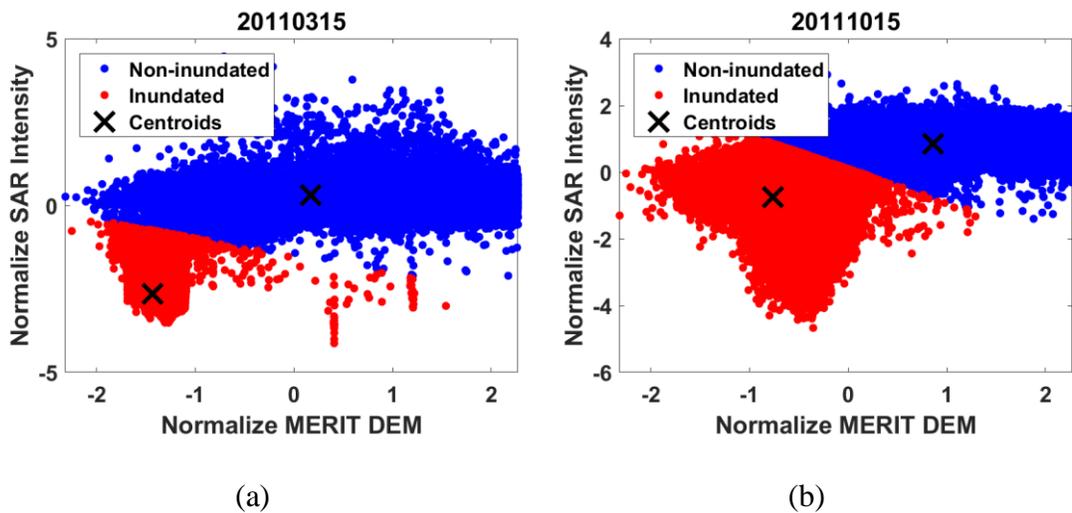
#### 4.3.1.3 K-means Clustering Inundation Extent Classification

After successfully synthesizing SAR intensity images, the K-means clustering algorithm (Lloyd, 1982) was adopted to classify pixels from synthesized intensity maps into inundated and non-inundated classes with the aid of MERIT DEM. K-means clustering is one of the most frequently used clustering techniques, can be easily implemented, provides relatively high-quality clusters with low computational effort (Chang et al., 2018; Lin et al., 2013; Tsyganskaya et al., 2018b) and has been applied on SAR images in recent studies for change detection (Celik, 2009; Zheng et al., 2014) and water pixel segmentation (Ruzza et al., 2019). The advantage of the K-means algorithm is that it is an unsupervised method that does not require additional training data as “ground truth.” The only necessary input to the algorithm is the user-predefined number of classes  $K$ . In this study, the number of classes is  $K = 2$ , representing inundated and non-inundated clusters. The algorithm first randomly selects  $K$  points as initial “centroids”. Each centroid corresponds to a class. Data are then assigned to the class whose centroid is the nearest based on squared Euclidean distance until the sum of squared distance from data to the centroid of each class has been minimized

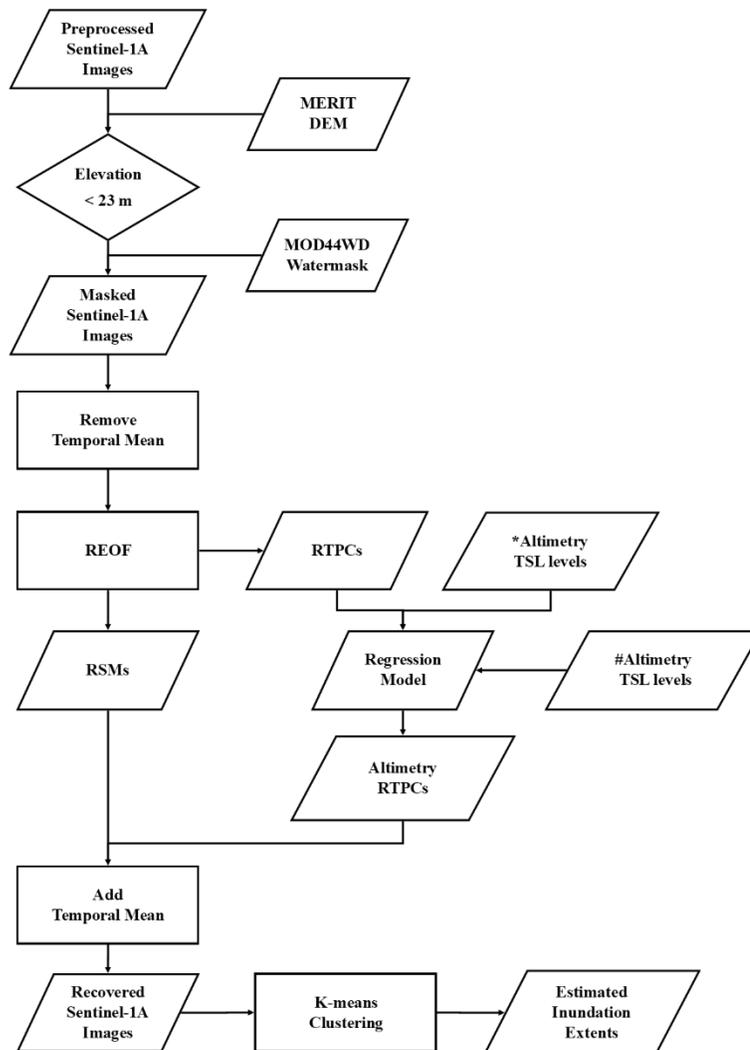
$$\min \sum_{y \in G_i} \|y - \mu_i\|^2 \quad (4.10)$$

where  $G_i$  is the class  $i$ ,  $y$  is the data which belongs to class  $i$  and  $\mu_i$  is the centroid of class  $i$ . The mean of data assigned to the same class is then taken as the new centroid. The algorithm iteratively updates centroids and data may be assigned to from one class to another until centroids stay unchanged. We used the “K-means” function in MATLAB® software R2017b. The software implements K-means++ algorithm (Arthur

and Vassilvitskii, 2007) to initialize the centroid and has been shown to have improved running time, robustness, and quality of the final solution over Lloyd's classical K-means method (Lloyd, 1982). The number of times to repeat clustering using new initial cluster centroid positions was set to be 20 (See supplementary data) to find a lower local minima and ensure the quality of the clustering results. For each of the 20 initial cluster centroid positions, the K-means clustering algorithm iterates up to 100 times to satisfy Equation (4.10). The final solution is the one among the 20 initializations that results in the minimum total sum of the distance between data points and centroids. Figure 4.11 is an illustration of how data were clustered by K-means algorithm where the x-axis is the normalized MERIT DEM and the y-axis is the normalized synthesized SAR intensities. Figure 4.12 shows a comprehensive flowchart of the proposed REOF-based daily inundation extent estimation framework. For Sentinel-1 images used as reference datasets for cross-comparison, inundated extents were directly estimated by the K-means algorithm with the same settings.



**Figure 4.11** Illustration of how data were clustered into non-inundated (blue) and inundated (red) clusters by K-means algorithm. (a) is an example in the dry season while (b) is in the wet season.

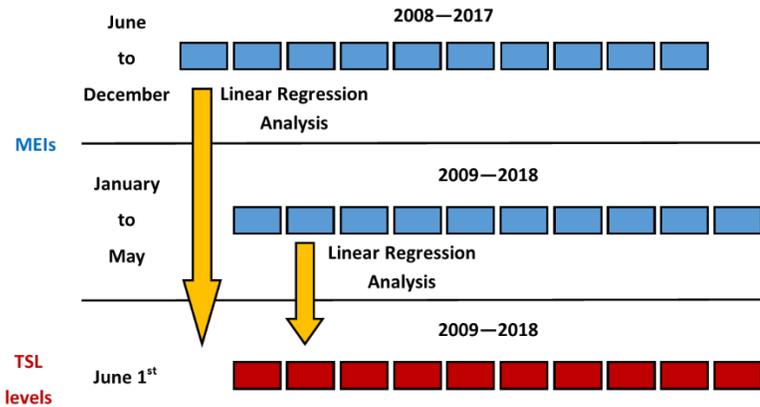


**Figure 4.12** Flowchart of the proposed REOF-based daily inundation extent estimation framework. (\*Altimetry-derived TSL levels were used to build regression models with RTPCs. #Altimetry-derived TSL levels were used for estimated RTPCs.)

#### 4.3.2 Long-term Forecasting of TSL Levels Using ENSO Index

Inspired by Frappart et al. (2018) and Räsänen and Kummu (2013), that both point out a negative correlation between MRB’s flood pulse and El Niño and La Niña events, linear regression models between MEIs (Wolter and Timlin, 2011, 1998, 1993) and TSL levels were built for forecasting TSL levels with months of lead time. Similar

work has been done by Fok et al. (2018) in which water levels in the MD were predicted. We performed regression analysis between Jason-2/-3-derived daily TSL levels from 2009 to 2018 and monthly MEIs. First, for each required forecasted TSL level (hereafter called forecasting date) the corresponding month was used as a reference month. We then used years of MEIs from each of the past 12 months with respect to the reference month to build linear regression models with years of TSL levels of the forecasting dates. For example, if we would like to forecast TSL level on June 1<sup>st</sup>, June would be the reference month. Therefore, MEIs of June to December of the previous year and January to May of the current year were used. Since we were building the regression model using Jason-2/-3 daily TSL levels from 2009 to 2018, for each month within June to December, MEIs from previous years from 2008 to 2017 were used. For each month from January to May, MEIs of current years from 2009 to 2018 were used. Figure 4.13 shows an illustration of the approach. Hence, there would be 10 MEIs and 10 Jason-2/-3-derived daily TSL levels for each forecasting date for linear regression analysis. After performing the linear regression analysis, MEIs of the month with the highest adjusted  $R^2$  with Jason-2/-3-derived daily TSL level on the forecasting dates was used as input to the regression models to forecast TSL levels with months of lead time.



**Figure 4.13** Illustration of the data used to build regression models between MEIs and TSL levels.

### 4.3.3 MODIS-derived Inundation Maps for Cross-comparison

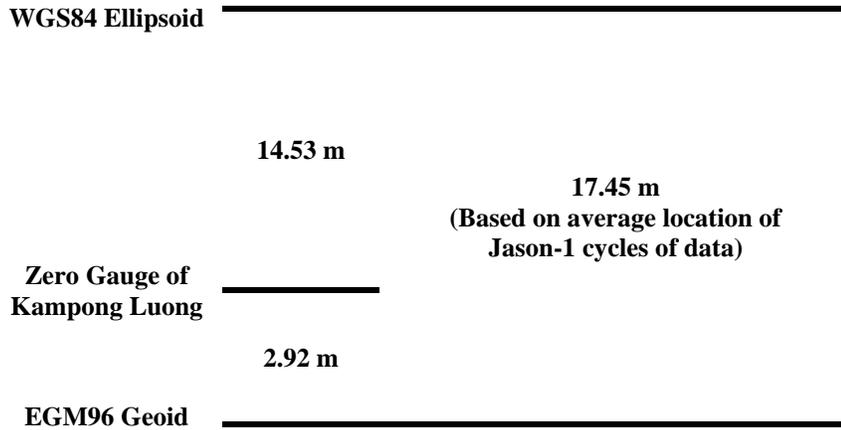
In this study, MODIS-derived inundation maps were used as one of the reference datasets for cross-comparing with our estimated inundation extents. The approach was originally proposed by Sakamoto et al. (2007) in studying change of inundation extents in the Lower Mekong. Normandin et al. (2018) simplified the approach and applied it over the Mackenzie Delta. The approach uses thresholdings on indices including the Enhanced Vegetation Index (EVI), Land Surface Water Index (LSWI), and Difference Value between EVI and LSWI (DVEL), derived from 8-day composite MODIS images of surface reflectance to classify pixels into classes of non-flooded, mixture and flooded or permanent water body. Frappart et al. (2018) also applied it for long-term MODIS-based inundation mapping and analysis over the TSL area, additionally considering a SRTM DEM and altimetry-derived water levels in TSL to judge whether a mixture class pixel is inundated or not.

According to Sakamoto et al. (2007), in order to implement the approach, EVI, LSWI, and DVEL are first calculated. The EVI (Huete et al., 1997) and LSWI (Xiao et al., 2002) are defined as

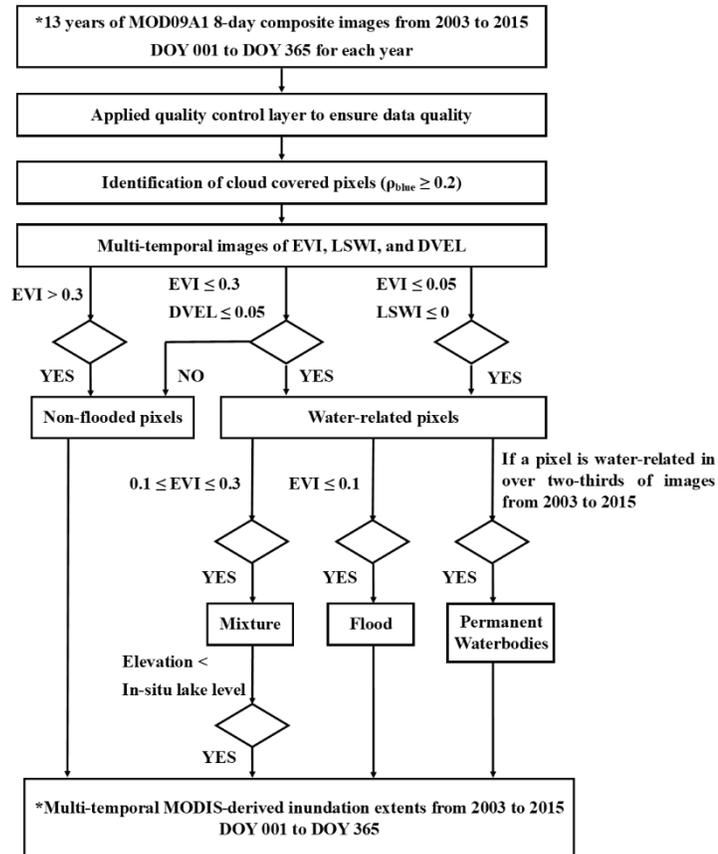
$$\begin{aligned} \text{EVI} &= 2.5 \times \frac{\rho_{NIR} - \rho_R}{\rho_{NIR} + 6 \times \rho_R - 7.5 \times \rho_B + 1} \\ \text{LSWI} &= \frac{\rho_{NIR} - \rho_{SWIR}}{\rho_{NIR} + \rho_{SWIR}} \end{aligned} \quad (4.11)$$

where  $\rho_{NIR}$  is the Near Infra-Red (NIR) surface reflectance (841-875 nm, band 2),  $\rho_R$  is the surface reflectance in the red (621-670 nm, band 1),  $\rho_B$  is the surface reflectance in the blue (459-479 nm, band 3), and  $\rho_{SWIR}$  is the surface reflectance of the Short-Wave Infra-Red (SWIR) (1628-1652 nm, band 6). DVEL is defined by subtracting LSWI from EVI. The pixels with  $\rho_B \geq 0.2$  were identified as cloud-covered and were excluded. Remaining pixels were then classified into two major classes including (1) non-flooded ( $\text{EVI} > 0.3$  or  $\text{EVI} \leq 0.3$  but  $\text{DVEL} > 0.05$ ) and (2) water-related pixels ( $\text{EVI} \leq 0.3$  and  $\text{DVEL} \leq 0.05$  or  $\text{EVI} \leq 0.05$  and  $\text{LSWI} \leq 0$ ). The water-related class is comprised of 3 sub-classes including flooded pixels when  $\text{EVI} \leq 0.1$ , mixture pixels when  $0.1 < \text{EVI} \leq 0.3$ , and permanent water bodies if the number of a pixel being classified as either flooded or mixture pixel exceeds two-thirds of the total number of MODIS images. Note that before calculating these indices, we conducted quality control upon all used bands with the reflectance band quality layer within each MOD09A1 image. Only those pixels with the highest quality in all the above-mentioned bands were kept for cross-comparison use. This was to ensure the derived reference dataset had the best quality.

To generate the final MODIS-based inundation maps, we used MERIT DEM and in situ water levels of TSL at Kampong Luong provided by the ADPC as auxiliary data. We assumed that TSL surface is parallel to the geoid surface. Hence, the TSL levels with respect to the geoid would be the same everywhere. As the datum of the MERIT DEM is EGM96 geoid, in situ water levels of TSL at Kampong Luong with respect to EGM96 geoid were generated and used as in situ water levels for TSL. This was achieved by calculating the means of the Jason-1/-2/-3 altimetry-derived water levels with respect to the EGM96 geoid and the in situ water levels at TSL with respect to local zero gauge, where the latter were then shifted toward the former by the difference between their means, which is 2.92 m. Figure 4.14 shows the offsets between datums of TSL water levels including WGS84 ellipsoid, local zero gauge of in situ water levels at Kampong Luong, and the EGM96 geoid. Finally, a pixel was considered as inundated if it was classified as (1) flooded pixel or, (2) mixture pixel with elevation lower than in situ TSL water level on that date or (3) the pixel is a permanent water body. Figure 4.15 shows the flowchart that summarizes the whole procedure for generation of MODIS-based inundation maps.



**Figure 4.14** Offsets between datums of TSL water levels, which include the WGS84 ellipsoid, local zero gauge of in situ TSL water level data at Kampong Luong, and the EGM96 geoid.



**Figure 4.15** Flowchart for generation of MODIS-derived inundation extents (\* Time span of data could be different depending on the time of estimated inundation extents to be validated).

#### 4.3.4 Framework Skill Evaluation Statistics

Evaluation of framework skill was based on a  $2 \times 2$  confusion matrix (Kohavi, R., Provost, 1998) which displays the absolute counts of combinations of framework estimated and observed event pairs. As Figure 4.16 shows,  $a$  and  $d$  are the counts of observed flood and non-flood events that the framework correctly estimates, representing *hits* and *correct negative*, respectively. Contrarily,  $b$  and  $c$  are the counts of events that are misestimated, representing *false alarms* and *misses*, respectively. In this study, these statistics were based on the number of pixels by comparing estimated inundation maps with MODIS-derived ones.

Confusion Matrix		MODIS-derived		Marginal Total
		Inundation	Non-inundation	
Framework-estimated	Inundation	$a$ (Hits)	$b$ (false alarms)	$a + b$
	Non-inundation	$c$ (missed)	$d$ (correct negative)	$c + d$
Marginal Total		$a + c$	$b + d$	Total = $a + b + c + d$

**Figure 4.16** A  $2 \times 2$  confusion matrix, which displays the number of pixels that is *hit*, *false alarm*, *miss*, or *correct negative*.

We used overall accuracy, critical success index (CSI), omission error and commission error as evaluation indices. The overall accuracy is the percentage of pixels which were estimated correctly by our framework over the total number of pixels:

$$\text{Overall accuracy} = \frac{a + d}{a + b + c + d} \times 100(\%). \quad (4.12)$$

The range of overall accuracy is from 0% to 100%, indicating zero skill to perfect skill.

CSI (Gilbert, 1884), also called *threat score*, is the number of correctly estimated inundated pixels over the number of pixels which are either real or framework-estimated inundated

$$\text{CSI} = \frac{a}{a + b + c} \times 100(\%). \quad (4.13)$$

CSI accounts for both *false alarms* and *misses* and is considered to be more complete. It avoids possible bias in the analysis results caused by correct negative (*d*) (Wing et al., 2017). Its value ranges from 0% to 100% where 0% means there is no match between observed and framework estimation, while 100% means perfect framework skill. CSI is frequently used as a standard validation measure (World Meteorological Organization, 2017).

Omission error represents the percentage of pixels which are actually inundated but are not captured by our framework over the total number of actually inundated pixels and can be determined by

$$\frac{c}{a + c} \times 100(\%). \quad (4.14)$$

Its value also ranges from 0% to 100% with 0% being perfect skill. It indicates the extent of *missed* pixels.

Commission error, on the other hand, is the percentage of framework-estimated inundated pixels which are actually non-inundated over the total number of framework-estimated inundated pixels which can be calculated as

$$\frac{b}{a + b} \times 100(\%). \quad (4.15)$$

Its value also ranges from 0% to 100% with 0% being perfect skill. It indicates the extent of *false alarm* pixels.

## 4.4 Results and Discussions

### 4.4.1 Evaluation and Analysis of Framework Skills Using Long-term Historical Data

Since 2003 is the first complete year of Jason-1 observations and 2015 is the last complete year before the acquisition time of Sentinel-1 GRDH used to build the proposed framework, we cross-compared the inundation extents in the TSLF estimated by our framework from 2003 to 2015 to provide a long-term evaluation of framework skill. The spatial resolution of our estimation is 500 m in order to fit the resolution of the MODIS inundation maps derived from the MOD09A1 8-day composite product. For each MOD09A1 image, we first counted the number of pixels within an 8-day period using day-of-year layer in the product. The dates, which have the greatest number of pixels in MOD09A1 images within corresponding 8-day periods, were considered as cross-comparison dates. The inundation extents estimated by our framework on the cross-comparison dates were evaluated against the corresponding MOD09A1-derived inundation maps. This mitigated the influence of date differences between our daily estimated inundation extents and those derived from 8-day composite MOD09A1 images. We adopted CSI, omission error and commission error and overall accuracy as evaluation indices of our framework skill. The evaluation results are listed in Table 4.1 from a climatological monthly perspective to see the performance of our framework skill in each month by a “hydrological year”, starting from May to April of next year (Kummu, M., Tes, S., Yin, S., Adamson, P., Józsa, J., Koponen, J., Richey, J., Sarkkula, 2015; Kummu and Sarkkula, 2008). Note that since the number of valid pixels in each

of the MOD09A1 images are different due to cloud cover and missing data, the climatological monthly averages and standard deviations (STDs) of CSIs, overall accuracies, omission errors, and commission errors here are weighted average and STDs, considering the total number of valid pixels of each MOD09A1 image. CSIs indicate that among pixels which were inundated either in MODIS images or estimated by our framework, 70% to 91% of pixels were both really inundated and successfully captured by our framework. CSIs are from 85% to 91% in the relatively dry period of February to June, and are about 75% to 80% in July, September, October, November, and January, but are relatively low, about 70% in August and December. Omission errors indicate that there are about 10% to 26% of pixels, which were inundated as MODIS images show, missed in our framework estimation. Commission errors indicate that 1% to 10% of our framework-estimated inundated pixels were actually not inundated in the MODIS images. The overall accuracies of our framework indicate that 90% to 99% of pixels, considering both inundated and non-inundated, are correctly classified.

We further analyzed the connection between variation of our framework skills over all months and the corresponding altimetry-derived TSL levels. Table 4.1 lists the climatological monthly variation of altimetry-derived TSL levels with corresponding RMSEs, which are 0.3 m to 0.7 m. Figure 4.17(a) shows the climatological monthly variation of CSIs, and RMSEs of altimetry-derived TSL levels and the corresponding fitted linear regression model in Figure 4.17(b), where we can see there is no correlation between CSIs, and RMSEs of altimetry-derived TSL levels with adjusted  $R^2$  of -0.09 in the fitted linear regression model. This indicates that the skills of our inundation extent

estimation framework have no significant connection with such level of errors of altimetry-derived TSL levels. However, interestingly, when pairing CSIs with TSL levels, as in Figure 4.18(a-2), we found a relationship with a convex quadratic shape between CSIs and TSL levels. The fitted convex quadratic curve between them has an adjusted  $R^2$  of 0.86. Such quadratic connection indicates that the relationship between CSIs and TSL levels changes depends on altimetry-derived TSL levels. As Figure 4.18(a-2) shows, when altimetry-derived TSL levels (with respect to WGS84 ellipsoid) are lower than about -9 m, there is a negative correlation between CSIs and TSL levels. Conversely, the correlation changes are positive when TSL levels with respect to WGS84 ellipsoid exceed -9 m. Such TSL levels, which are in August and December, seems to be a turning point in the relationship between CSIs and altimetry-derived TSL levels. In August and December, CSIs are also the lowest. A similar pattern can be seen in the case of omission errors as well. When pairing omission errors with altimetry-derived TSL levels, there is a concave quadratic relation with an adjusted  $R^2$  of 0.75 between them with the same turning point at about -9 m of TSL level (with respect to WGS84 ellipsoid), as Figure 4.18(b-2) shows. This means that when TSL levels with respect to WGS84 ellipsoid are below -9 m, there is a positive correlation between omission errors and TSL levels that changes to a negative correlation when TSL levels with respect to WGS84 ellipsoid exceeds -9 m. These facts also imply a strong connection between CSIs and omission errors. As Figure 4.18(c-2) shows, CSIs are negatively correlated with omission errors, resulting in a fitted linear model with an adjusted  $R^2$  of 0.91. Commission errors, on the other hand, do not have an obvious quadratic shape of relation with altimetry-derived TSL levels. The relation between

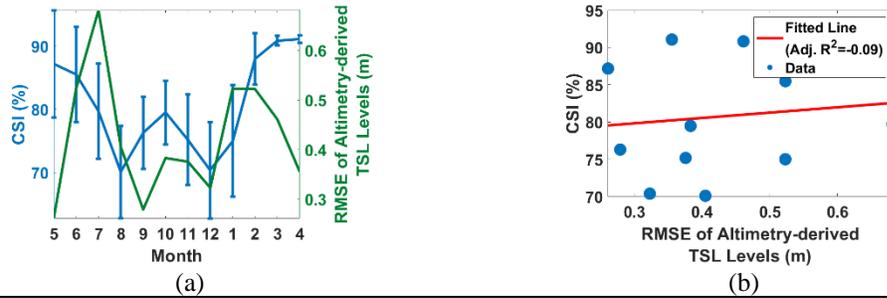
commission errors and TSL levels is more like a uniformly positive correlation, resulting in a fitted linear model with an adjusted  $R^2$  of 0.88 shown in Figure 4.19(a-2). Furthermore, the connection of CSIs with commission errors is weaker than that with omission errors as the adjusted  $R^2$  of the fitted linear model between CSIs and commission errors is 0.67 as Figure 4.19(b-2) shows, which is lower than the adjusted  $R^2$  of the fitted linear model between CSIs and omission errors. This indicates that the impact of omission errors on the variation of CSIs is more dominant than commission errors considering the results over all months.

**Table 4.1** Climatological monthly averages and STDs of CSI, omission and commission error, overall accuracy, and statistics of altimetry-derived TSL levels including average level and RMSE. In-situ average TSL levels and STDs are shown as well.

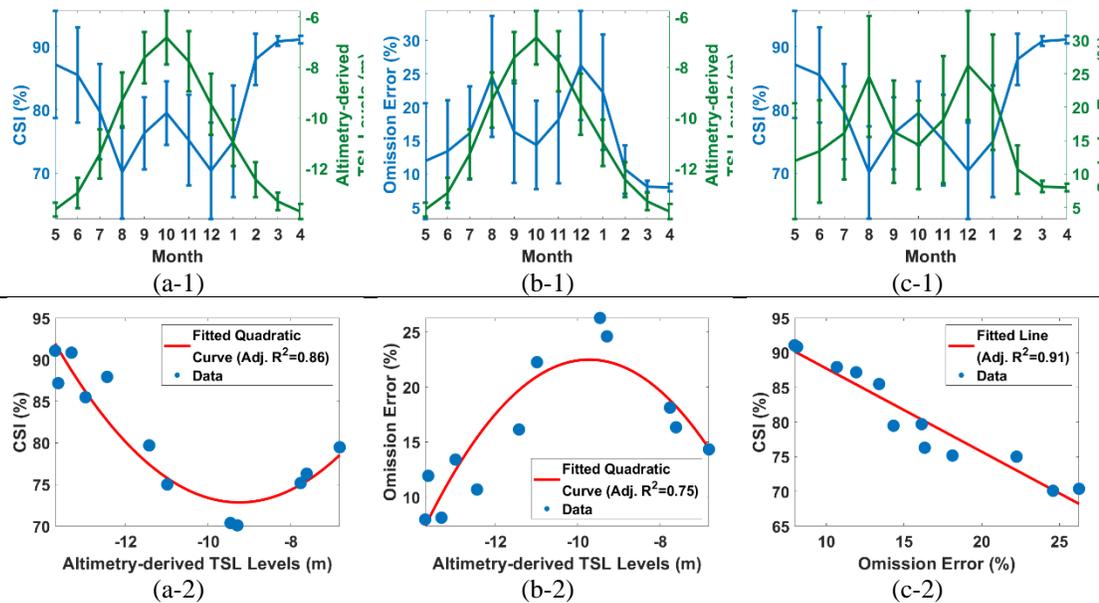
Month	CSI (%)	Omission Error (%)	Commission Error (%)	Overall Accuracy (%)	Statistics of TSL Levels		
					Altimetry-derived		In-situ
					*Average Levels (m)	RMSE (m)	#Average Levels (m)
May	87.17	11.93	1.02	98.14	-13.63	0.26	0.81
	±8.45	±8.64	±0.15	±2.28	±0.27		±0.21
Jun.	85.48	13.40	1.36	97.78	-12.97	0.52	1.17
	±7.50	±7.66	±0.74	±1.66	±0.60		±0.53
Jul.	79.69	16.13	5.99	96.29	-11.43	0.68	2.55
	±7.52	±6.98	±4.00	±2.03	±0.97		±1.02
Aug.	70.09	24.59	8.61	90.83	-9.29	0.40	4.94
	±7.27	±9.05	±3.89	±3.76	±1.10		±1.27
Sep.	76.30	16.33	9.67	88.77	-7.62	0.28	6.97
	±5.68	±7.68	±5.30	±2.86	±1.01		±1.10
Oct.	79.47	14.33	7.94	89.04	-6.82	0.38	7.86
	±5.03	±6.61	±4.88	±2.23	±1.06		±1.16
Nov.	75.17	18.12	9.07	90.07	-7.76	0.38	6.96
	±7.15	±9.51	±5.10	±2.02	±1.19		±1.24
Dec.	70.38	26.26	6.02	91.68	-9.46	0.32	5.28
	±7.61	±8.16	±1.94	±3.17	±1.20		±1.16
Jan.	75.00	22.24	4.62	95.22	-10.99	0.52	3.81
	±8.81	±8.63	±2.39	±2.78	±0.92		±0.96
Feb.	87.93	10.68	1.58	98.53	-12.45	0.52	2.34
	±4.06	±3.64	±1.06	±0.75	±0.69		±0.79
Mar.	90.83	8.12	1.02	98.95	-13.31	0.46	1.47
	±0.79	±0.86	±0.14	±0.21	±0.36		±0.50

Table 4.1 (Continued)							
<b>Apr.</b>	91.07	7.96	1.01	98.90	-13.70	0.36	0.96
	$\pm 0.58$	$\pm 0.56$	$\pm 0.20$	$\pm 0.30$	$\pm 0.30$		$\pm 0.29$

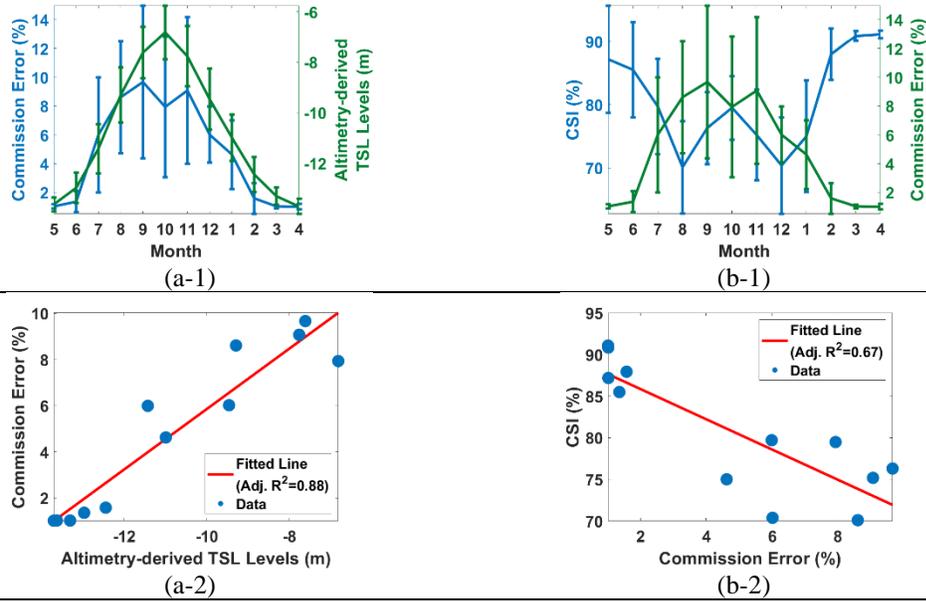
\*Altimetry-derived TSL water levels in this table are with respect to WGS84 ellipsoid. #In-situ TSL water levels are with respect to local zero gauge.



**Figure 4.17** (a) Climatological monthly variation of CSIs and RMSEs for altimetry-derived TSL levels. Corresponding scatter plots with fitted linear regression models are in (b).



**Figure 4.18** Climatological monthly variation of (a-1) CSIs and altimetry-derived TSL levels, (b-1) omission errors and altimetry-derived TSL levels, and (c-1) CSIs and omission errors. (a-2) to (c-2) are corresponding scatter plots with fitted regression models.

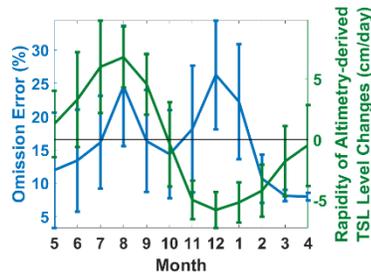


**Figure 4.19** Climatological monthly variation of (a-1) commission errors and altimetry-derived TSL levels, and (b-1) CSIs and commission errors. (a-2) to (b-2) are corresponding scatter plots with fitted regression models.

The relationship between the variation of omission errors and TSL levels has a concave quadratic shape with the highest errors occurring when TSL levels with respect to the WGS84 ellipsoid are about -9 m in August and December. This “-9 m” turning point may correspond to the influence of specific vegetation around TSL on SAR backscatter characteristics. Since Sentinel-1 is a C-band SAR satellite, which has limited vegetation penetration depth, its radar backscatter may be dominated by volume scattering if TSL levels are not high enough. Consequently, intensities over some of the inundated areas may not be low enough for the K-means clustering algorithm to be able to distinguish them properly from non-inundated areas, leading to the positive correlation between omission errors and TSL levels. By contrast, when TSL rises above this level, water surface scattering may become dominant over most of the inundated areas, leading to intensities which are low enough for K-means clustering to recognize.

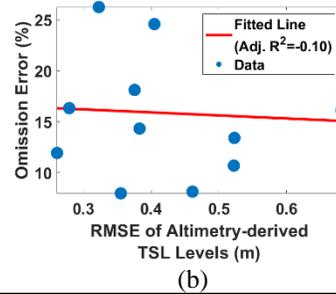
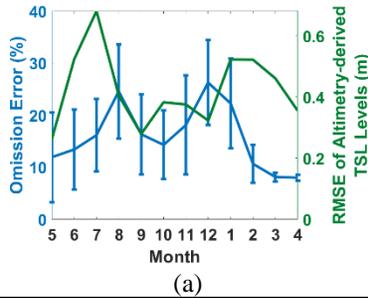
Thus, the connection between omission errors and TSL levels is a negative correlation. In fact, according to Van Trung et al. (2013), there is a rapid increase in areas which change from lowland shrubs to water surfaces in the TSLF when in-situ TSL levels with respect to local zero gauge is about 5 m, which agrees with the “-9 m” altimetry-derived TSL level with respect to WGS84 ellipsoid (See Table 4.1). Figure 4.10(b) in Van Trung et al. (2013) shows the land cover variation model in terms of flooded areal percentages and the variation of in-situ TSL levels during a flood pulse, where a “trough” can clearly be seen in the variation curve of flooded lowland area shrubs when in-situ TSL level exceeds about 5 m. Interestingly, the “trough” of variation of area of flooded lowland shrubs also crosses the curve of in-situ TSL level at the point when the latter is about 5 m. This “trough” may result from the change of land cover from flooded lowland shrubs to fully inundated water surfaces as Van Trung et al. (2013) pointed out. Hence, this indicates that the flooded lowland shrubs start to completely submerge underwater when in-situ TSL level rises above 5 m, which corresponds to the altimetry-derived TSL level of -9 m. Lowland shrub is the dominant land cover type over the TSLF as Figure 3(a) of Sáenz et al. (2016) shows. Arias et al. (2012) also pointed out that shrubland is the dominant land cover type in the TSLF which are flooded 5 to 9 months in an average year. Hence, the relationship between the variation of area for flooded lowland shrubs and TSL level found by Van Trung et al. (2013) may also explain the connection between both CSIs and omission errors of our results with TSL levels. Furthermore, August and December also have peak rapidity for changes of TSL levels at rising stage and receding stage of flood pulse, respectively, as Figure 4.20 shows. This means that August and December are months when TSL levels experience the most significant

change, which also explains why these two months are when radar backscattering characteristic changes, leading to a reverse correlation between omission errors (thus CSIs), and TSL levels. On the other hand, omission errors have no connection with the RMSEs of altimetry-derived TSL levels, which means that such amount of errors in altimetry-derived TSL levels do not have a significant impact on the skill of estimating inundation extent (See Figure 4.21).



Month	Rapidity of Changes of TSL Levels (cm/day)	Month	Rapidity of Changes of TSL Levels (cm/day)
May	1.28±2.73	Nov.	-4.98±1.59
Jun.	3.30±3.91	Dec.	-5.81±1.48
Jul.	6.01±3.82	Jan.	-5.18±1.66
Aug.	6.81±2.50	Feb.	-4.23±2.15
Sep.	4.57±2.51	Mar.	-1.78±2.90
Oct.	-0.38±3.48	Apr.	-0.48±3.34

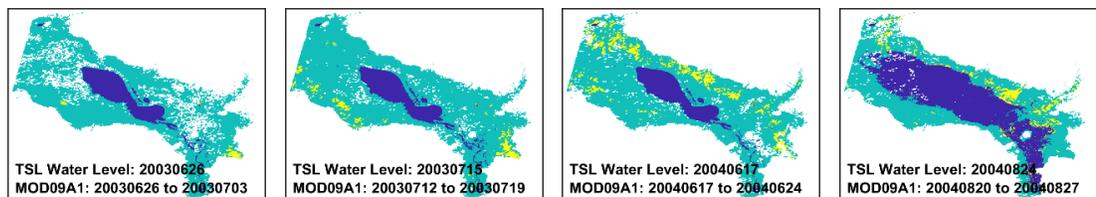
**Figure 4.20** Climatological monthly variation of omission errors and rapidity of changes of altimetry-derived TSL levels.



**Figure 4.21** (a) Climatological monthly variation of omission errors and RMSEs of altimetry-derived TSL levels. Corresponding scatter plots with fitted linear regression models are in (b).

Furthermore, since the estimation of inundation extents of our framework is based on the relationship between TSL levels and SAR backscatter intensity changes, inundation caused by regional rainfall, river overflow in the areas with higher elevation cannot be captured and leads to some degree of omission errors as well. Figure 4.22

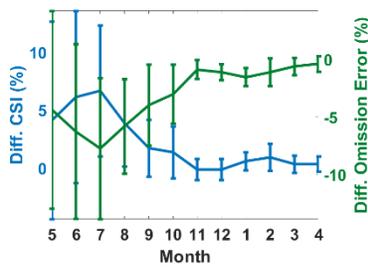
shows examples of inundation that occurred in areas higher than contemporary TSL levels in yellow (hereafter called high-elevation inundated pixels), while areas with elevations below contemporary TSL levels are in blue with the background in green. An evaluation of the framework skills excluding the influence of such high-elevation inundation, which could be unrelated to the TSL levels, was then performed. This was achieved by not removing the high-elevation inundated pixels (yellow pixels in Figure 4.22), when calculating evaluation indices as described before. The climatological monthly average and STD of the updated framework skills and the original framework skill and differences between them (hereafter called change of skills) are listed in Table 4.2 for comparison. When excluding high-elevation inundation, omission errors decrease by 1% to 8% with CSIs increase by up to 7%, while commission errors slightly increase up to 2%. The difference of CSIs is highly negatively correlated with the change of omission errors as there is a fitted linear regression model with an adjusted  $R^2$  of 0.90 between them as Figure 4.23 shows. The more the omission errors decrease, the better the CSIs, as can be seen in Figure 4.23(b). Change of commission errors, contrarily, has no connection with the changes in CSIs (Figure 4.24).



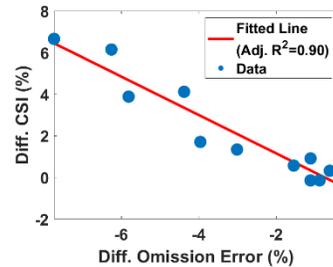
**Figure 4.22** Examples of inundation with elevations higher than contemporary TSL levels (yellow). Areas with elevations lower than TSL levels are shown in blue, while green is the background color.

**Table 4.2** Climatological monthly averages and STDs of CSI, omission and commission error (OE, and CE, respectively) and overall accuracy (OA) with high-elevation inundated pixels being excluded. Change of framework skills are also listed.

Mon.	CSI (%)	OE (%)	CE (%)	OA (%)	Change of Skills			
					CSI (%)	OE (%)	CE (%)	OA (%)
May	91.29 ±0.45	7.55 ±0.50	1.02 ±0.15	98.97 ±0.16	4.12 ±8.54	-4.38 ±8.60	0.00 ±0.00	0.84 ±2.28
Jun.	91.63 ±0.77	7.14 ±0.54	1.43 ±0.75	98.94 ±0.24	6.15 ±7.44	-6.26 ±7.59	0.07 ±0.25	1.16 ±1.64
Jul.	86.35 ±5.48	8.39 ±3.27	6.32 ±4.28	97.77 ±1.50	6.66 ±5.68	-7.74 ±6.13	0.33 ±0.57	1.48 ±1.25
Aug.	73.98 ±6.24	18.78 ±7.39	10.60 ±3.77	92.66 ±2.85	3.88 ±3.79	-5.81 ±4.11	1.99 ±1.18	1.82 ±1.89
Sep.	78.00 ±5.30	12.37 ±7.70	11.76 ±5.76	90.09 ±1.92	1.71 ±2.45	-3.96 ±3.51	2.09 ±1.07	1.32 ±1.77
Oct.	80.81 ±5.03	11.31 ±6.39	9.42 ±5.52	90.19 ±1.31	1.34 ±2.24	-3.01 ±2.55	1.48 ±1.03	1.15 ±1.72
Nov.	75.03 ±7.17	17.24 ±9.63	10.36 ±5.08	90.11 ±2.07	-0.14 ±0.90	-0.88 ±0.82	1.29 ±0.79	0.04 ±0.40
Dec.	70.24 ±7.61	25.15 ±8.25	7.87 ±2.30	91.71 ±3.16	-0.14 ±0.91	-1.11 ±0.71	1.85 ±0.93	0.03 ±0.44
Jan.	75.00 ±7.17	20.70 ±8.58	6.00 ±3.14	95.37 ±2.77	0.57 ±0.79	-1.54 ±0.79	1.38 ±1.35	0.15 ±0.36
Feb.	88.84 ±3.58	9.57 ±3.11	1.78 ±1.35	98.66 ±0.62	0.91 ±1.16	-1.11 ±1.20	0.21 ±0.46	0.13 ±0.34
Mar.	91.15 ±0.50	7.51 ±0.61	1.04 ±0.20	98.99 ±0.11	0.32 ±0.75	-0.61 ±0.76	0.02 ±0.14	0.03 ±0.18
Apr.	91.43 ±0.49	7.57 ±0.49	1.01 ±0.20	98.94 ±0.24	0.32 ±0.66	-0.39 ±0.68	0.00 ±0.00	0.04 ±0.20

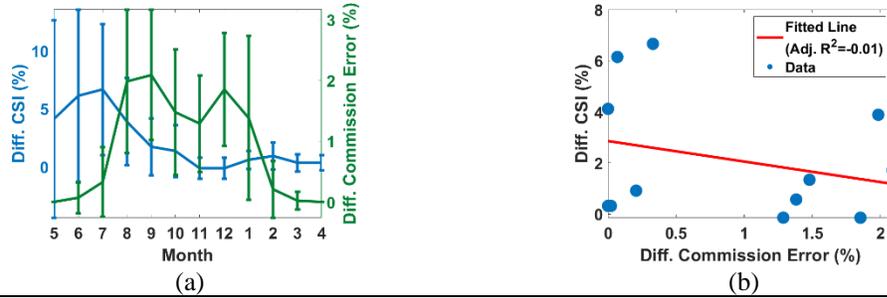


(a)



(b)

**Figure 4.23** Climatological monthly variations of difference of CSIs and difference of omission errors when excluding inundation with elevation higher than contemporary TSL levels (a). Corresponding scatter plot is in (b)



**Figure 4.24** Climatological monthly variations of difference of CSIs and difference of commission errors when excluding inundation with elevation higher than contemporary TSL levels (a). Corresponding scatter plot is in (b)

As Table 4.2 shows there are some of months with high climatological monthly STDs of skill change. Since STDs represent temporal variation, we then analyzed the temporal correlation between change of framework skills and the number of high-elevation inundated pixels (listed in Table 4.3). In most months, both changes of CSI and overall accuracy have significantly strong positive temporal correlation, while the change of omission error has a significantly strong negative correlation with the number of high-elevation inundated pixels. This means that when there are more high-elevation inundated pixels, omission error will be reduced when these pixels are excluded, leading to improvement in CSI and overall accuracy. Changes in commission error, on the other hand, have a relatively moderate positive correlation with the number of high-elevation inundated pixels. This may be caused by the increase of “false alarm” pixels when excluding high-elevation inundated pixels, but the correlation is relatively weak compared with those of other skills. Temporal correlations in April and May are both “NaN” because there are no changes of commission error in these two months when excluding high-elevation inundated pixels, indicating that these pixels have no influence on commission errors in these two months (See Table 4.2). The analysis results indicate that the temporal variations and thus STDs of change of skills including CSI, omission

error, and overall accuracy are related to the temporal variation of the number of high-elevation inundated pixels to some degree, while the change of commission error is not necessarily influenced by it.

**Table 4.3** Temporal correlation coefficients and corresponding P-values (in the bracket) of the number of high-elevation inundated pixels with the change of CSI, omission error, commission error, and overall accuracy.

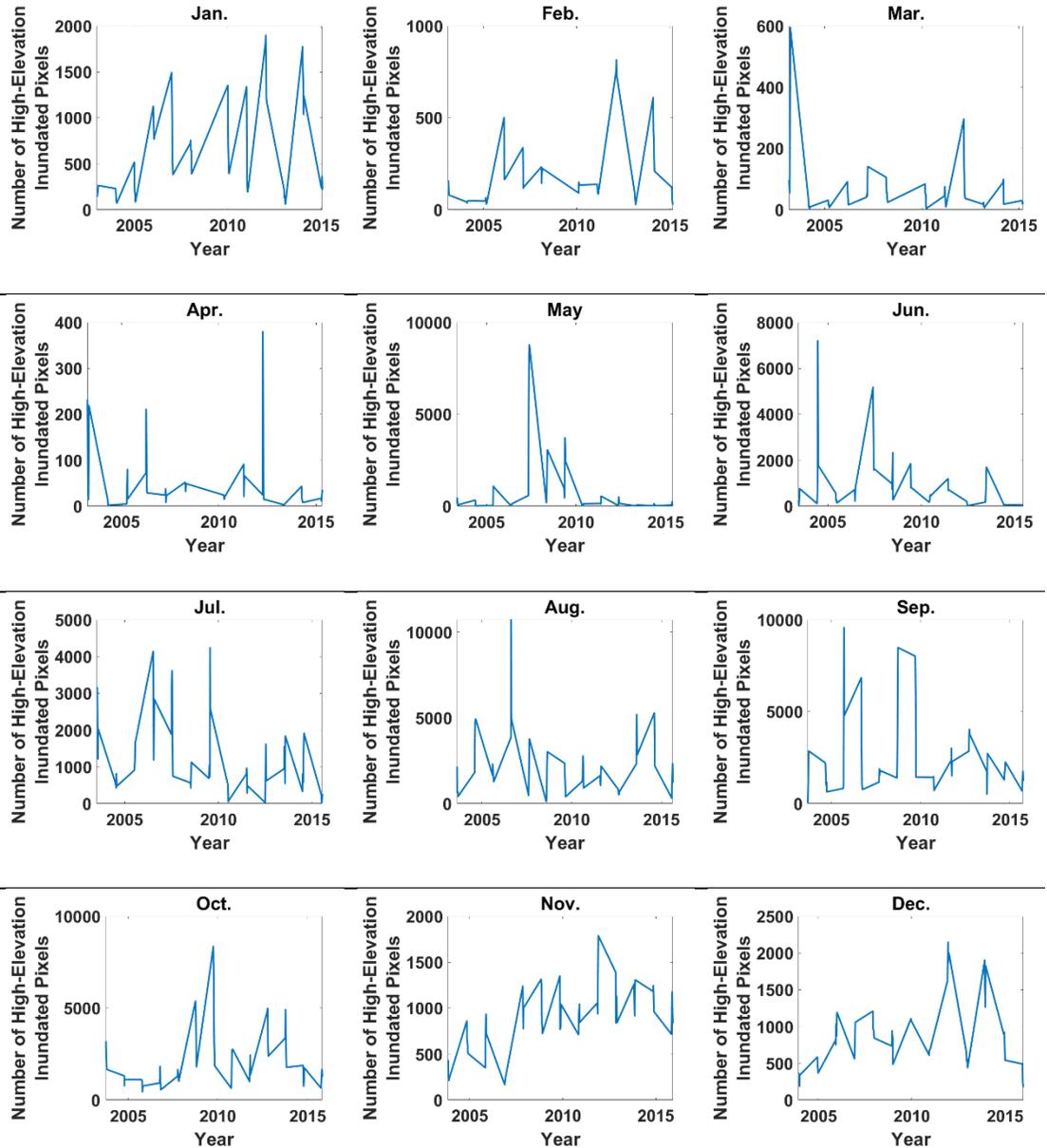
Month	Correlation coefficient and P-value of number of high-elevation inundated pixels with change of							
	CSI		Omission Error		Commission Error		Overall Accuracy	
<b>May</b>	0.99	(0.00)	-0.99	(0.00)	NaN	(NaN)	0.97	(0.00)
<b>Jun.</b>	0.99	(0.00)	-0.99	(0.00)	0.05	(0.74)	0.97	(0.00)
<b>Jul.</b>	0.92	(0.00)	-0.96	(0.00)	0.34	(0.01)	0.94	(0.00)
<b>Aug.</b>	0.66	(0.00)	-0.72	(0.00)	0.62	(0.00)	0.92	(0.00)
<b>Sep.</b>	0.81	(0.00)	-0.94	(0.00)	0.68	(0.00)	0.90	(0.00)
<b>Oct.</b>	0.83	(0.00)	-0.94	(0.00)	0.49	(0.00)	0.80	(0.00)
<b>Nov.</b>	0.00	(0.97)	-0.44	(0.00)	0.46	(0.00)	0.09	(0.56)
<b>Dec.</b>	-0.52	(0.00)	0.01	(0.93)	0.74	(0.00)	-0.35	(0.02)
<b>Jan.</b>	-0.19	(0.18)	-0.51	(0.00)	0.91	(0.00)	0.12	(0.42)
<b>Feb.</b>	0.82	(0.00)	-0.91	(0.00)	0.46	(0.00)	0.51	(0.00)
<b>Mar.</b>	0.89	(0.00)	-0.80	(0.00)	0.05	(0.72)	0.87	(0.00)
<b>Apr.</b>	0.85	(0.00)	-0.83	(0.00)	NaN	(NaN)	0.73	(0.00)

In Table 4.4, the climatological monthly averages and STDs of the number of high-elevation inundated pixels and change of skills, together with the correlation coefficients and P-values between their STDs, are shown. The STD of the number of high-elevation inundated pixels were found to have a moderate to strong temporal correlation with STDs of change of CSI, omission error, and overall accuracy, while its temporal correlation with STD of change of commission error is weak. The analysis results supported what was inferred from Table 4.3. On the other hand, we found in Table 4.4 that there are STDs of the number of high-elevation inundated pixels larger than the corresponding averages in months including February to June. This means there

were large temporal variations of the number of such pixels for the period from 2003 to 2015 in these months. Figure 4.25 shows a time series of the number of high-elevation inundated pixels for each different month where more abrupt peaks in February to June are observed. This may result from sudden heavy rainfalls near the acquisition dates of certain MODIS images, leading to large STDs in these months.

**Table 4.4** Climatological monthly averages and STDs of the number of high-elevation inundated pixels, and change of CSI, omission, commission error and overall accuracy. Correlation coefficients and corresponding P-values (in the bracket) between STDs are also listed.

Month	Number of High-elevation Inundated Pixels	Change of Skills			
		CSI (%)	Omission Error (%)	Commission Error (%)	Overall Accuracy (%)
May	752.28 ±1796.20	4.12 ±8.54	-4.38 ±8.60	0.00 ±0.00	0.84 ±2.28
Jun.	914.49 ±1347.05	6.15 ±7.44	-6.26 ±7.59	0.07 ±0.25	1.16 ±1.64
Jul.	1240.79 ±1088.77	6.66 ±5.68	-7.74 ±6.13	0.33 ±0.57	1.48 ±1.25
Aug.	2191.74 ±1882.02	3.88 ±3.79	-5.81 ±4.11	1.99 ±1.18	1.82 ±1.89
Sep.	2432.74 ±2110.79	1.71 ±2.45	-3.96 ±3.51	2.09 ±1.07	1.32 ±1.77
Oct.	2176.02 ±1604.01	1.34 ±2.24	-3.01 ±2.55	1.48 ±1.03	1.15 ±1.72
Nov.	850.96 ±369.61	-0.14 ±0.90	-0.88 ±0.82	1.29 ±0.79	0.04 ±0.40
Dec.	884.77 ±488.59	-0.14 ±0.91	-1.11 ±0.71	1.85 ±0.93	0.03 ±0.44
Jan.	650.96 ±536.17	0.57 ±0.79	-1.54 ±0.79	1.38 ±1.35	0.15 ±0.36
Feb.	199.33 ±210.77	0.91 ±1.16	-1.11 ±1.20	0.21 ±0.46	0.13 ±0.34
Mar.	58.57 ±91.15	0.32 ±0.75	-0.61 ±0.76	0.02 ±0.14	0.03 ±0.18
Apr.	49.11 ±74.51	0.32 ±0.66	-0.39 ±0.68	0.00 ±0.00	0.04 ±0.20
<b>Correlation Coefficient between STDs</b>		0.64 (0.03)	0.70 (0.01)	0.29 (0.36)	0.96 (0.03)



**Figure 4.25** Time series of the number of high-elevation inundated pixels for each month from 2003 to 2015.

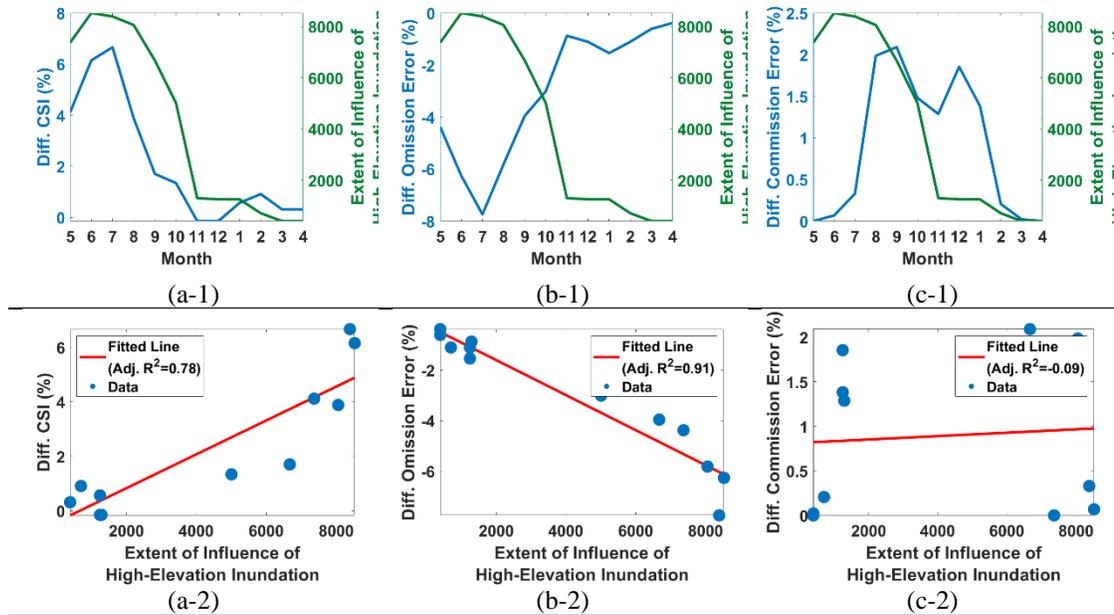
Table 4.5 listed the climatological monthly average and STD of the number of high-elevation inundated pixels, together with elevation differences of such pixels with contemporary TSL levels as well as their extent of influence. Derivation of the extent of influence of high-elevation inundated pixels assumes that if the elevation differences

are larger, the inundation was more likely to be caused by factors such as regional rainfall or river overflow. Therefore, we weighted the climatological monthly average of the number of such inundated pixels with average elevation differences to provide a reference of the extents of influence of such high-elevation inundated pixels. We found that there is a significant negative correlation between changes of omission errors when excluding high-elevation inundated pixels and the extents of influence of high-elevation inundated pixels, leading to a fitted linear regression model with adjusted  $R^2$  of 0.91. The significant negative correlation indicates the stronger the influence of high-elevation inundated pixels, the more the omission errors decrease when excluding them. On the other hand, change of commission errors has no connection with the extent of influence of high-elevation inundated pixels with negative adjusted  $R^2$ . Change of CSIs has a significant positive correlation with the extent of influence of high-elevation inundated pixels, resulting in a fitted linear regression model with adjusted  $R^2$  of 0.78 (See Figure 4.26(a-2) to Figure 4.26(c-2)). The strong positive correlation of change of CSIs with the extent of influence of high-elevation inundated pixels indicates that the stronger the influence of high-elevation inundated pixels, the more the CSIs increase when excluding them. It explains why there are different degrees of enhancement of our framework skills in different months, if high-elevation inundated pixels were not considered.

**Table 4.5** Climatological monthly averages and STDs of the number of high-elevation inundated pixels, differences between elevations of such inundated pixels and contemporary TSL levels and their products, called extents of influence.

Month	Number of High-elevation Inundated Pixels	Difference of Elevations (m)	*Extent of Influence
May	752.28±1796.20	9.79±3.56	7368.42
Jun.	914.49±1347.05	9.32±3.55	8526.73
Jul.	1240.79±1088.77	6.76±3.82	8388.84
Aug.	2191.74±1882.02	3.68±3.32	8056.12
Sep.	2432.74±2110.79	2.74±2.96	6671.90
Oct.	2176.02±1604.01	2.30±2.47	5006.12
Nov.	850.96±369.61	1.52±2.35	1293.16
Dec.	884.77±488.59	1.41±2.38	1249.11
Jan.	650.96±536.17	1.92±2.98	1249.39
Feb.	199.33±210.77	3.54±4.13	706.61
Mar.	58.57±91.15	6.80±4.58	398.49
Apr.	49.11±74.51	8.09±3.09	397.29

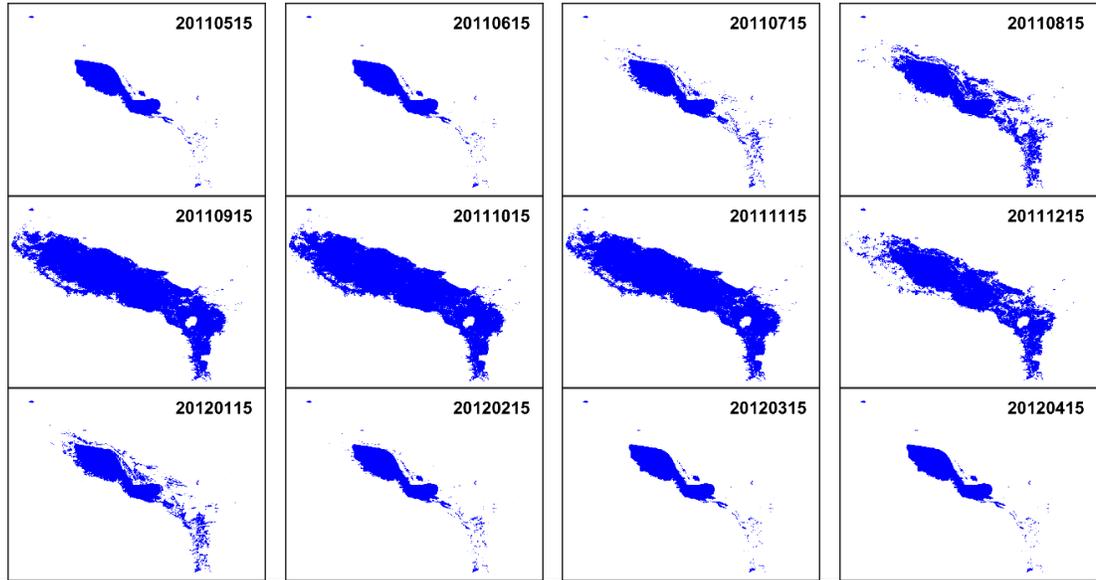
\*The extent of influence was determined as the product of the average of number of high-elevation inundated pixels and corresponding average of elevation differences with TSL levels.



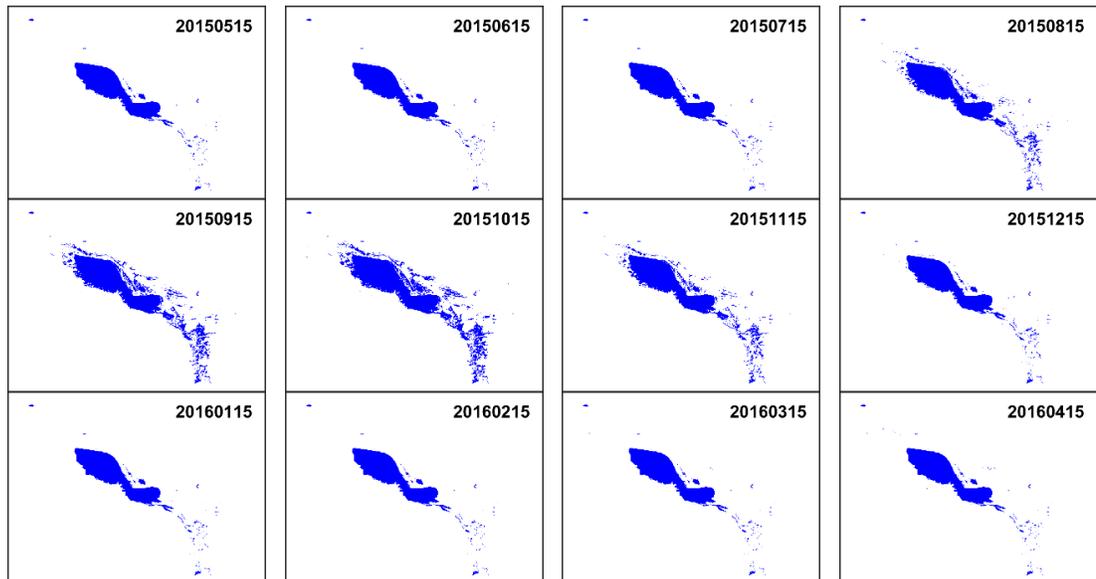
**Figure 4.26** Climatological monthly variations of difference of CSIs, omission errors, and commission errors with extents of influence of high-elevation inundated pixels with order of (a-1) to (c-1), respectively. Corresponding scatter plots with fitted linear regression models are in (a-2) to (c-2).

In Figure 4.27 and 4.28, inundation maps on the 15<sup>th</sup> of each month in the hydrological years 2011 and 2015 (May to April of next year) were chosen to display the evolutions of inundation extents, as these two years were extreme scenarios for the maximum and minimum inundation extents, respectively (Frappart et al., 2018).

Distinct differences in inundation extents between these two hydrological years, especially from July to January, can be seen.



**Figure 4.27** Evolution of inundation extents for the hydrological year 2011 (May 2011 to April 2012). Inundation extents on the 15<sup>th</sup> of each month are shown.



**Figure 4.28** Evolution of inundation extents for the hydrological year 2015 (May 2015 to April 2016). Inundation extents on the 15<sup>th</sup> of each month are shown.

#### 4.4.2 Evaluation of Skills of Forecasted Inundation Extents

In this section, the capacity of our framework to forecast daily inundation extents in the TSLF is demonstrated. The results were evaluated by cross-comparing with MODIS-derived and Sentinel-1-derived inundation maps. To forecast inundation extents, daily TSL levels were first forecasted using MEI based on the linear regression models between them. Due to the relatively poor accuracy of Jason-1 altimetry-derived water levels over inland water bodies (Ablain et al., 2010; Martin-Puig et al., 2016), only Jason-2/-3 altimetry-derived TSL levels were used to build the linear regression model with MEIs. The time span covers 2009 to 2018. Figure 4.29 shows correlation coefficients between altimetry-derived TSL levels for each date and each month of MEIs over the past 12 months. Strong negative correlations of up to -0.8 between the interannual variation of TSL levels on each date and MEIs is observed. Such negative correlation with months of lead time when flood pulses in the MRB respond to ENSO events indicate the connection between them, which has been shown and discussed in several previous studies (Fok et al., 2018; Frappart et al., 2018; Räsänen and Kummu, 2013). Here, for each date that we intend to forecast TSL levels, the month of MEI which has the highest adjusted  $R^2$  for the linear regression models with TSL levels was selected as Figure 4.30(a) shows with corresponding lead time. The months of MEI in the past 12 months which have the highest adjusted  $R^2$  for linear regression models with each date of TSL levels are from May of the previous year to June of the current year, resulting in lead times with a range of 2 to 11 months. The lead times of forecasted altimetry-derived TSL levels are also the lead times of forecasted inundation extents.

Note that the lead time here does not consider the day of month since MEI is a monthly index.

Figure 4.30(b) shows the highest adjusted  $R^2$  value for the linear regression model for each date and the corresponding p-value. The highest adjusted  $R^2$  values are from about 0.3 to 0.8 with p-values from 0 to 0.05. From late March to the mid of July, the highest adjusted  $R^2$  values have larger oscillations which may be because these months are in the dry period, hence the influence of ENSO events is not that consistent as in the wet period. Despite the lower highest adjusted  $R^2$  values within the period, the correlations are still significant, with 95% of confidence interval as p-values are only up to 0.05. We then forecasted the TSL levels in 2019. Figure 4.31(a) shows cross-comparison of our forecasted TSL levels with in-situ data from January to July 2019. We can see that the overall RMSE of our forecasting is about 0.78 m with a high positive correlation of 0.9. Note that we were performing long-term forecasting with 2 to 11 months of lead time. Figure 4.31(b) shows our forecasted TSL levels for all of 2019, with clear seasonal variation. The results demonstrate the possibility of forecasting TSL levels using the ENSO index.

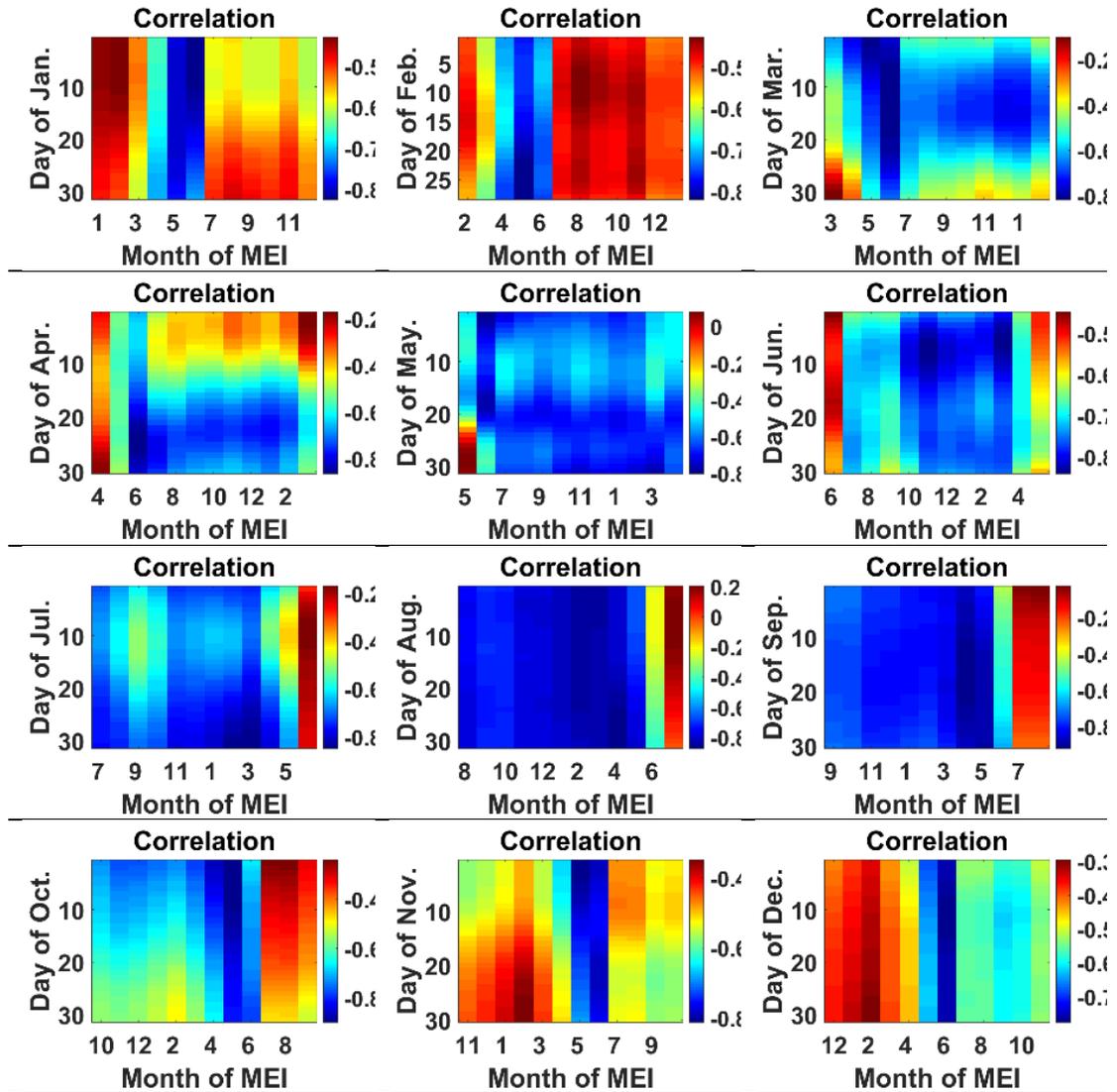


Figure 4.29 Correlation coefficients between TSL levels of each date and monthly MEIs in the past 12 months.

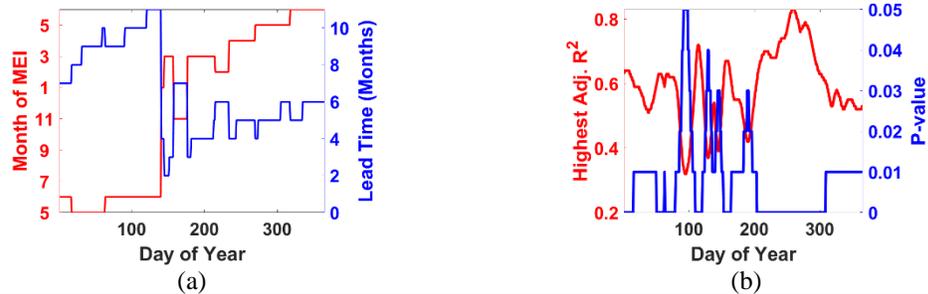
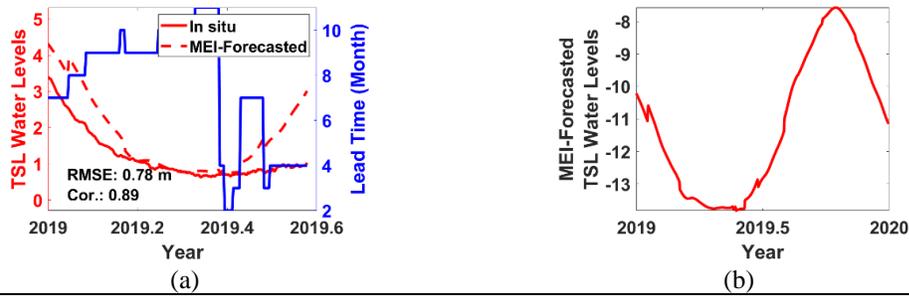


Figure 4.30 (a) Months of MEIs which have the highest adjusted  $R^2$  for linear regression models with TSL levels on the date when forecasting is performed and the corresponding lead time. The highest adjusted  $R^2$  and p-value are shown in (b).



**Figure 4.31** (a) MEI-forecasted TSL levels with months of lead time which were validated by in situ data up to July 2019, and (b) MEI-forecasted TSL levels for all of 2019.

The forecasted TSL levels were then used as input to our REOF-based daily inundation extent estimation framework to forecast the inundation extents over the TSLF. The forecasted inundation extents were validated by MOD09A1-derived inundation maps based on CSI, omission error, commission error, and overall accuracy as we did in Section 4.1. To have an understanding of the framework skill without the influence of inherent inconsistency between radar and optical imagery, inundation extents directly derived from updated Sentinel-1A/-1B SAR imagery using the K-means algorithm were used as another reference dataset in addition to MODIS-derived inundation maps.

Table 4.6 shows cross-comparison results of our forecasted inundation extents in 2019 from January to July using MODIS imagery as the reference dataset. Monthly average CSIs, omission errors, commission errors, and overall accuracies are from 80% to 91%, 7% to 11%, 1% to 10%, 97% to 99%, respectively, during the period from January to July. Table 4.7 shows monthly averages and STDs of CSI, omission and commission error, and overall accuracy of our forecasted inundation extent when excluding high-elevation inundated pixels and the difference compared with original skills. CSIs, omission errors, commission errors, and overall accuracy are from 81% to

92%, 6% to 7%, 1% to 13%, and 97% to 99%, respectively. The differences with original skill are mostly less than or around 1% except the omission error and commission error in January.

**Table 4.6** Climatological monthly averages and STDs of CSI, omission and commission error, and overall accuracy of our forecasting from January to July 2019.

Month	CSI (%)	Omission Error (%)	Commission Error (%)	Overall Accuracy (%)
<b>Jan.</b>	80.25 ±2.59	11.50 ±3.20	10.50 ±1.50	97.00 ±0.71
<b>Feb.</b>	90.28 ±0.96	7.36 ±0.48	2.72 ±0.96	99.00 ±0.00
<b>Mar.</b>	91.53 ±0.50	7.23 ±0.42	1.00 ±0.00	98.85 ±0.36
<b>Apr.</b>	91.30 ±0.46	7.30 ±0.46	1.00 ±0.00	99.00 ±0.00
<b>May</b>	91.00 ±0.00	7.70 ±0.46	1.00 ±0.00	98.81 ±0.39
<b>Jun.</b>	91.44 ±0.50	7.55 ±0.50	1.00 ±0.00	98.81 ±0.39
<b>Jul.</b>	90.91 ±0.70	7.43 ±1.30	2.66 ±0.83	99.00 ±0.00

**Table 4.7** Climatological monthly averages and STDs of CSI, omission and commission error (OE, and CE, respectively), and overall accuracy (OA) of our forecasting from January to July 2019 with high-elevation inundated pixels excluded. Differences with original framework skills are also listed.

Mon.	Change of Framework Skills							
	CSI (%)	OE (%)	CE (%)	OA (%)	CSI (%)	OE (%)	CE (%)	OA (%)
<b>Jan.</b>	81.50 ±2.18	7.00 ±1.58	13.25 ±1.48	97.25 ±0.43	1.25 ±0.83	-4.50 ±1.80	2.75 ±0.83	0.25 ±0.43
<b>Feb.</b>	90.91 ±0.79	6.27 ±0.45	3.44 ±1.23	99.00 ±0.00	0.63 ±0.48	-1.09 ±0.79	0.73 ±0.45	0.00 ±0.00
<b>Mar.</b>	91.53 ±0.50	7.00 ±0.00	1.00 ±0.00	98.85 ±0.36	0.00 ±0.00	-0.23 ±0.42	0.00 ±0.00	0.00 ±0.00
<b>Apr.</b>	91.30 ±0.46	7.00 ±0.00	1.00 ±0.00	99.00 ±0.00	0.00 ±0.00	-0.30 ±0.46	0.00 ±0.00	0.00 ±0.00
<b>May</b>	91.19 ±0.34	7.22 ±0.42	1.00 ±0.00	98.81 ±0.39	0.19 ±0.39	-0.48 ±0.50	0.00 ±0.00	0.00 ±0.00
<b>Jun.</b>	92.00 ±0.00	7.25 ±0.43	1.00 ±0.00	98.81 ±0.39	0.56 ±0.50	-0.30 ±0.46	0.00 ±0.00	0.00 ±0.00

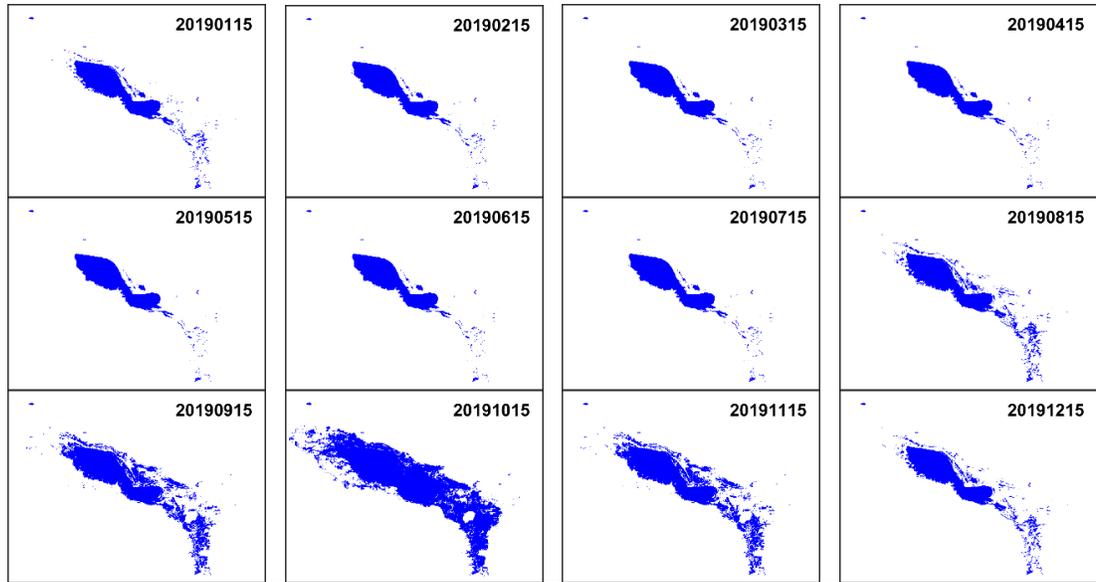
Table 4.7 (Continued)								
<b>Jul.</b>	91.77	6.04	2.66	99.00	0.86	-1.38	0.00	0.00
	$\pm 0.42$	$\pm 0.71$	$\pm 0.83$	$\pm 0.00$	$\pm 0.84$	$\pm 1.11$	$\pm 0.00$	$\pm 0.00$

For cross-comparison with Sentinel-1-derived inundation maps, climatological monthly averages of CSI, omission error and commission error and overall accuracy are listed in Table 4.8. Monthly average CSIs are from 84% to 98%, omission errors are from 0.5% to about 1%, commission errors are from 0.2% to 15%, and overall accuracies are from 98% to 100%. These statistics were compared with those obtained by cross-comparing with MODIS-derived inundation maps (For cross-comparison results with MODIS-derived inundation maps, please refer to Table 4.6). Results show that when cross-comparing with Sentinel-1-derived inundation maps, omission errors are much lower than with MODIS-derived ones, while commission errors are at the same level except for the higher value in January. These may lead to slightly higher CSIs and overall accuracies. The much lower omission errors may be due to the fact that Sentinel-1-derived inundation maps were used as the reference dataset for cross-comparison. Since the framework we proposed is based on Sentinel-1 SAR imagery, both can be influenced by the limited vegetation-penetrating capacity of the C-band signal which would lead to a similar extent of underestimation, and thus omission errors for the inundation extents estimated by our framework were reduced. Vegetation-penetrating capacity is also a main inherent difference between SAR imagery and optical imagery. By using Sentinel-1-derived inundation maps as the reference data for cross-comparison instead of MODIS-derived ones, the impact of such inherent difference on cross-comparison results was mitigated.

**Table 4.8** Climatological monthly averages and STDs of CSI, omission and commission error, and overall accuracy of our forecasting from January to July 2019, which were obtained by cross-comparing with inundation extents directly estimated from updated Sentinel-1A/-1B SAR imagery with K-means clustering algorithm.

Month	CSI (%)	Omission Error (%)	Commission Error (%)	Overall Accuracy (%)
<b>Jan.</b>	84.20	1.00	15.20	97.80
	±4.02	±0.00	±4.15	±0.84
<b>Feb.</b>	95.00	1.00	4.00	99.50
	±1.41	±0.00	±1.41	±0.58
<b>Mar.</b>	97.50	1.17	1.33	100.00
	±0.84	±0.41	±1.03	±0.00
<b>Apr.</b>	97.80	1.20	1.00	100.00
	±0.45	±0.45	±0.00	±0.00
<b>May</b>	98.00	1.60	0.20	100.00
	±0.71	±0.89	±0.45	±0.00
<b>Jun.</b>	97.80	1.60	0.20	100.00
	±0.84	±1.14	±0.45	±0.00
<b>Jul.</b>	96.25	0.50	3.00	99.75
	±0.96	±0.58	±1.41	±0.50

By using MEI-forecasted altimetry-derived TSL levels (See Figure 4.31(b)) and the REOF-based inundation extent estimation framework, forecasted inundation extents from January 1<sup>st</sup> to December 31<sup>st</sup>, 2019 were estimated. Figure 4.32 shows the forecasted inundation extents on the 15<sup>th</sup> day of each month in 2019, displaying the evolution of inundation extent from long-term forecasting perspective with months of lead time. Considering the huge impact of inundation extent over the TSLF on local fishery, livelihoods and economy the long-term forecasted inundation extent with months of lead time has the potential to help stakeholders make plans for effective water resource management in advance.



**Figure 4.32** Forecasted inundation extents on the 15<sup>th</sup> day of each month in 2019.

#### 4.5 Conclusions

In this study, we proposed an innovative REOF-based daily inundation extent estimation framework by exploiting multi-temporal stacks of Sentinel-1A SAR intensity images and Jason altimetry-derived water levels. The framework takes advantage of the physical interpretability of results for REOF analysis, the capacity of cloud penetration, weather and sunlight independence of SAR imagery, and the short revisit time and consistent data acquisition for Sentinel-1 and Jason altimetry satellites and has features including: (1) daily synthesis of SAR intensity images and estimation of areal inundation extents at any time as long as altimetry-derived water levels are available; (2) Fully remote sensing-based in which a computationally expensive model is not required; (3) Cloud-free daily inundation extents estimation. The framework has potential to be applied to the floodplains of other major river basins such as the Amazon and Congo River Basin. In this study, the framework was implemented for the TSLF. A

method capable of performing long-term TSL level forecasting with months of lead time was also proposed to fulfill the forecasting capacity of the proposed framework.

We first hindcasted historical inundation extents from 2003 to 2015 with the use of historical altimetry-derived TSL levels. The skills of our framework were evaluated by cross-comparing hindcasted historical inundation extents with 8-day composite MODIS-derived inundation maps. The connection between framework skills and various possible factors, including input altimetry-derived water level and its RMSE were then analyzed. Based on the evaluation of long-term hindcasted historical inundation extents, climatological monthly average CSIs for our estimated inundation extents are from 70% to 91% and have a significant negative correlation with omission errors. Both CSIs and omission errors have no significant connection with the level of RMSEs of altimetry-derived TSL levels in this study. Interestingly, they have a significant connection with altimetry-derived TSL levels. Since Sentinel-1 is a C-band SAR satellite with limited vegetation-penetrating ability, its radar backscattering characteristics may change between surface backscattering or volume scattering and specular scattering depending on TSL levels. Therefore, the band of SAR electromagnetic radiation is important in the proposed framework and should be selected carefully by taking local vegetation type into account. The CSIs increase from 75% to 91% when excluding high-elevation inundation. This is because our framework is based on connecting TSL levels with temporal variations of intensities in Sentinel-1A imagery, High-elevation inundation, which may be caused by regional rainfall or upstream river overbank flooding but not necessarily caused by TSL level change, may not be well captured by our framework.

In the forecasting case, the forecasted TSL levels from January to July 2019 were used as inputs to the proposed framework to estimate forecasted inundation extents. The skills of forecasted inundation extents were evaluated by MODIS-derived inundation maps as well as Sentinel-1-derived inundation maps. When using MODIS-derived inundation maps as the reference dataset for cross-comparison, CSIs are from 80% to 91% and increase to 81% to 92% when excluding high-elevation inundated pixels. On the other hand, when using Sentinel-1-derived inundation maps as the reference dataset for cross-comparison, CSIs are from 84% to 98%. The improvement of CSIs when cross-comparing with Sentinel-1-derived inundation maps is probably due to the reduction of omission error. Since our framework is based on Sentinel-1 SAR imagery, there is omission error caused by inherent differences between MODIS optical imagery and Sentinel-1 SAR imagery, as the latter can be influenced by limited vegetation penetration. When using Sentinel-1-derived inundation maps as the reference dataset, the influence of such inherent differences between data sources was mitigated. However, since only data from January to July of 2019 were used for forecasting and cross-comparison, more data is needed to have a comprehensive understanding of framework skill in the future.

Considering potential future anthropogenic and climatic impacts on the hydrology of MRB and TSL, our daily forecasted inundation maps can make a great contribution to water resource management, socioeconomic impact evaluation, and decision-making purposes without dependence on measurements from upstream countries. Since our forecast inundation extents are based on forecasts of TSL levels, which currently only considers the influence of ENSO, further investigation is needed

to address the impact of anthropogenic factors, such as the construction of upstream dams, to have a more comprehensive understanding about how future TSL levels and TSLF inundation extents will evolve.

## **5. FORECASTING INUNDATION EXTENTS USING ROTATED EMPIRICAL ORTHOGONAL FUNCTION ANALYSIS IN MEKONG RIVER BASIN FLOODPLAIN**

### **5.1 Introduction**

In the Mekong River Basin (MRB) floodplain, including the Tonle Sap Lake Floodplain (TSLF), the Cambodian Floodplain (CF) and Mekong Delta (MD), agriculture, especially rice cultivation, has been a major supporter of livelihoods for local inhabitants and is the foundation of the national economy and stabilizes not only regional but also global food supply (Maitah et al., 2020; Matsubara et al., 2020; Okazumi et al., 2014; Triet et al., 2018). However, persistent flooding can be detrimental to agriculture, economy and human lives. In fact, rice paddies die after being submerged in water for days (Mekong River Commission, MRC, 2005). Extreme flood events can also be highly disastrous, leading to casualties and tremendous damages in rice paddies, infrastructures and economic loss (Center for Excellence in Disaster Management and Humanitarian Assistance, 2017). Therefore, it is vital to predict future inundation extents as it may help the authorities do a risk assessment and proactively mitigate damages.

For flood inundation forecasting, hydrodynamic models based on solving one-dimensional Saint-Venant equations or two-dimensional shallow water equations have been widely used to transform discharge outputs from rainfall-runoff models to distributed inundation extents. However, these models suffer from several sources of error (Bates et al., 2014), including model structural errors and uncertainties in: (1) the

model input data including the rainfall-runoff data to set the boundary and initial conditions, (2) input Digital Elevation Model (DEM) and channel bathymetry data, (3) friction coefficients to represent energy loss mechanisms, and (4) information about hydraulic structures in the reach. All these uncertainties in hydrodynamic model calibration, boundary conditions, and topographic data significantly influence flood inundation predictions (Teng et al., 2017; Bates et al., 2014). Furthermore, the required spatial parametric inputs may not always be available (Chen et al., 2019; Leandro et al., 2014; Teng et al., 2017). In addition, they carry a heavy computational burden, especially for a high-resolution large-scale forecasting framework, that could affect forecast lead-time and accuracy.

A non-modeling, terrain-based approach, such as Height Above Nearest Drainage (HAND), that employs a planar approximation has also been used (Nobre et al., 2016; Zheng et al., 2018a). The HAND approach normalizes topography according to local relative heights found along the drainage network. Then, a rating curve is used to transform streamflow forecasts to depths for a given river cell. Finally, flood inundation extents are determined by selecting surrounding land cells whose HAND values are less than the given water depth in the stream (Nobre et al., 2011; Teng et al., 2017). The HAND approach requires significantly less computational power than hydrodynamic models and may work well on confined floodplains with steep valleys and straight river reaches (Bates and De Roo, 2000; Wing et al., 2019). However, a recent study by Johnson et al. (2019) demonstrated that the method overpredicts in regions of low relief. In addition, the HAND approach does not account for backwater

effects caused by infrastructure and coastal flooding. With flat terrain, especially over the CF and MD (Balica et al., 2014), the HAND approach may not be applicable.

Recently, space-borne remote sensing, especially Synthetic Aperture Radar (SAR), has emerged as a powerful tool to depict areal inundation and its variation with repeated views over extensive spatial areas (e.g., Liu et al., 2017; Smith, 1997). However, the sole use of these observations cannot directly forecast response to future or changed conditions (Teng et al., 2017). Accordingly, the role of SAR-derived inundation extents in most areal inundation forecasting studies is mainly limited to the calibration of hydrodynamic models (Teng et al., 2017).

Currently, there is no operational inundation forecasting system in the MRB, to the best of our knowledge. Most studies applied hydrodynamic models for scenario analysis, such as analyzing long-term impacts of future climate change or streamflow alteration, but do not provide a frequent and near real-time inundation extent forecasting (Shin et al., 2018; Try et al., 2020a, 2020b). Although the Region Flood Management and Mitigation Center (RFMMC) of MRC performs forecasts for Mekong River (MR) levels, there is a lack of publicly available information about forecasted inundation extents.

Chang et al. (2020), using the TSLF as a test bed, proposed a satellite imagery-based inundation forecasting framework that addresses the need for computationally efficient areal inundation estimation and forecasts with high temporal resolution. This framework is named Forecasting Inundation Extents using Rotated empirical orthogonal function analysis (FIER). Chang et al. (2020) successfully implemented this in the TSLF for daily hindcast and forecast of areal inundations using multi-temporal Sentinel-1A

imagery and daily interpolated Jason-series altimetry-derived TSL levels. This study further expanded the study area to the flood-prone MRB floodplain, including the TSLF and CF (Hereafter, TSLF and CF will be abbreviated as TCF when they appear together for conciseness.), and MD. The FIER framework over the MRB floodplain was constructed using a multi-temporal stack of Sentinel-1A intensity images and MR levels. With the water level forecasting system for MR developed by Chang et al. (2019) (Details in Section 5.3.3), this study serves as a pilot to test the feasibility of its integration with FIER over the MRB floodplain. An application of FIER in flood risk forecasting is also demonstrated.

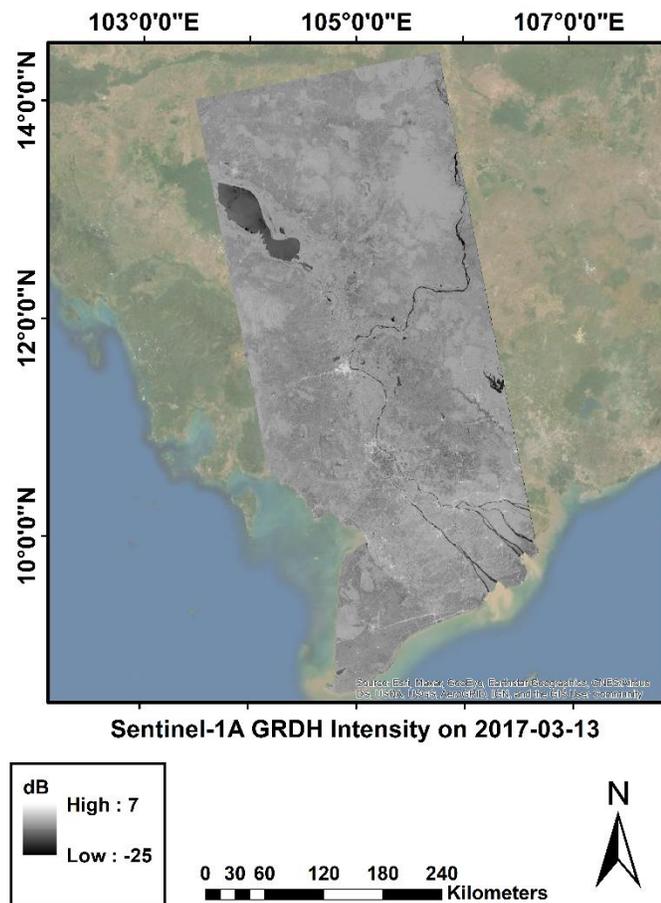
## 5.2 Data

### 5.2.1 Sentinel-1 SAR Data

Sentinel-1 is a two-satellite-constellation mission (Sentinel-1A/-1B) under the Copernicus Earth observation program of the European Space Agency (ESA). Sentinel-1A was launched on April 3<sup>rd</sup>, 2014, while Sentinel-1B was launched on April 25<sup>th</sup>, 2016. Both satellites are equipped with C-band (5.405 GHz) SAR sensors with 12 days of revisit time and free product accessibility. In this study, Sentinel-1A VV-polarized images of Ground Range Detection High-resolution (GRDH) product were obtained from the Alaska Satellite Facility (ASF). Only Sentinel-1A data were used to avoid the influence of potential systematic difference between backscattering intensities acquired by two satellites (Sentinel-1A/-1B) on the results of the Rotated Empirical Orthogonal

Function (REOF) analysis. The VV-polarization was chosen because of its superior surface water mapping capability (Markert et al., 2020; Twele et al., 2016).

To cover the TCF and MD in one scene, four Sentinel-1A frames including Frames 23, 29, 34 and 39 of Path 26 were mosaicked (see Figure 5.1). After mosaicking, 85 images acquired from March 13<sup>th</sup>, 2017 to December 28<sup>th</sup>, 2019, were used for REOF analysis, while 30 images from January 9<sup>th</sup> to December 22<sup>nd</sup>, 2020 were used for cross-comparison of hindcasted and pseudo-forecasted inundation extents. In this study, image pre-processing, including mosaicking, multi-looking, radiometric terrain correction, geocoding and co-registration (with respect to the image acquired on March 13<sup>th</sup>, 2017, which is the first acquired image) were done using GAMMA software (Werner et al., 2000) with the aid of the Multi-Error-Removed Improved-Terrain DEM (MERIT DEM) (Yamazaki et al., 2017). Note that we multi-looked the images to 500 m of spatial resolution for the purpose of this feasibility study as the pre-processing of fine-resolution (i.e., 10 m) multi-temporal SAR image stacks may take a prohibitively long time on a personal computer. However, it is worth mentioning that the pre-processing is one-time-only if there is no severer flooding that occurs during the time span of images where REOF analysis extracts spatiotemporal patterns (See Section 5.3.1). Moreover, the pre-processing time can also be alleviated by using cloud-based platforms, such as Google Earth Engine and Google Colab.



**Figure 5.1** Example of Sentinel-1A VV-polarized GRDH intensity image used in this study acquired on March 13<sup>th</sup>, 2017, covering the TCF and MD. The basemap is World Imagery provided by ArcMap 10.7.

### 5.2.2 MODIS Surface Reflectance Data and Yearly Water Mask

Terra and Aqua satellites with MODerate resolution Imaging Spectroradiometers (MODIS) onboard were launched in 1999 and 2002, respectively. They continuously provide Earth surface radiance with high temporal and spectral resolution (<https://modis.gsfc.nasa.gov/>). MODIS products have been widely used for monitoring and long-term analysis of inundation extents in many studies due to their high temporal

resolution despite cloud cover (Frappart et al., 2018; Gumma et al., 2014; Huang et al., 2013, 2014; Islam et al., 2010; Normandin et al., 2018; Sakamoto et al., 2007, 2009). In this study, tiles H28V07 and H28V08 of MOD09A1 and MOD44W were obtained from the USGS Earth Explorer website (<https://earthexplorer.usgs.gov/>). MOD09A1 is an 8-day composite surface reflectance data derived from Terra satellite raw radiance measurements from bands 1 to 7 at 500 m spatial resolution with the atmospheric effect corrected. For each pixel, surface reflectance data with the best quality within the 8-day period was retained based on cloud cover and solar zenith angle criteria. For cross-comparison of FIER-generated hindcast of inundation maps in the case of an extreme flooding event, we generated inundation extents using MOD09A1 images acquired in October of 2011 since the National Aeronautics Space Agency (NASA) Goddard Space Flight Center (GSFC) Near Real-Time (NRT) flood maps (details in Section 5.2.3) cover the MD only partially.

MOD44W is a MODIS-based yearly water mask product with 250 m resolution available from 2000 to 2015. In this study, MOD44W Version 6 was adopted. It is derived using a decision tree-based classifier trained with MODIS data and validated with the previous version of the MOD44W product (<http://lpdaac.usgs.gov/products/mod44wv006>). The MOD44W water mask from 2015 was used as a permanent water body mask since the worst drought in decades occurred in 2015, impacting Thailand, Laos, Cambodia, and Vietnam (Guo et al., 2017). By masking out the permanent water bodies, the influence of surface roughness change-induced SAR intensity variation on REOF analysis results can be mitigated. Such variation is often caused by wind-induced surface roughness change.

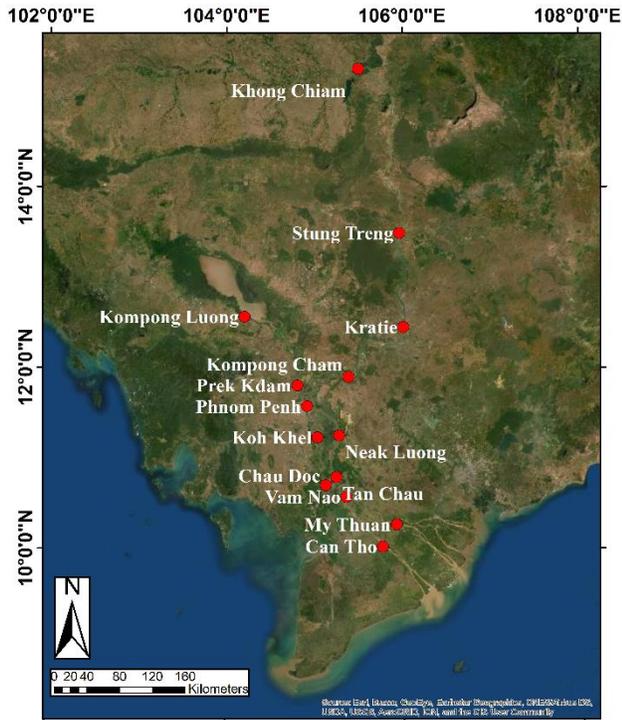
### 5.2.3 NASA GSFC NRT Global Flood Maps

NASA GSFC NRT global flood mapping project was developed in partnership with the Dartmouth Flood Observatory (now at the University of Colorado, Boulder) and NASA's GSFC funded by NASA's Applied Sciences program. Since late 2011, the project has produced daily global flood maps at 250 m spatial resolution using a threshold-based water detection algorithm by utilizing MODIS band-1 (red), band-2 (infra-red), and band-7 (shortwave infra-red) surface reflectance data (MOD09) acquired by both Terra and Aqua satellites (2 observations each day). Before temporally compositing, pre-processing, including terrain and cloud shadow masking, have been performed to reduce shadow-induced false-positive pixels (Slayback, 2021). The product has been adopted by several governmental agencies such as the Federal Emergency Management Agency (FEMA) and United Nations Office for the Coordination of Humanitarian Affairs (UN-OCHA), UN Operational Satellite Applications Program (UN-OSAT) and non-governmental organizations in the past decade (Slayback, 2021). In this study, the 2-day composite absolute surface water maps from March 13<sup>th</sup>, 2017 to December 28<sup>th</sup>, 2019, corresponding to the Sentinel-1A image acquisition dates, were used to select water classification thresholds (See Section 5.3.2 for details). The 2-day composite product is considered to have more representative flooded extents among other available products because of the strategy applied for cloud and terrain shadows masking (Nigro et al., 2014; Revilla-Romero et al., 2015; Slayback,

2021). The product is archived and available from <https://floodmap.modaps.eosdis.nasa.gov/>.

#### 5.2.4 MRC In-situ Water Levels

MRC is an inter-governmental organization established in 1995 with collaborating efforts of the Lower Mekong countries, including Thailand, Laos, Cambodia, and Vietnam, based on the 1995 Mekong Agreement. It serves as a diplomacy platform, knowledge exchange and data sharing hub between member countries, coordinating sustainable water resource management and regional development. Since 2008, MRC has installed hydrometeorological monitoring stations over MRB in collaboration with international organizations. The observed data are archived and distributed through MRC Data and Information Services (<https://portal.mrcmekong.org/home>). This study used in-situ river level observations at 14 stations in the MRB for analysis. Among these 14 stations, there are 4 stations located along the Mekong Mainstem (MM) (Khong Chiam to Kompong Cham), 5 stations located in the TCF (Kompong Luong to Koh Khel/Neak Luong), and another 5 stations located in the MD (Chau Doc/Tan Chau to My Thuan/Can Tho). Raw data are collected 1 – 6 times/day, and were averaged to provide daily data. Figure 5.2 shows the geographical location of in-situ stations used in this study.

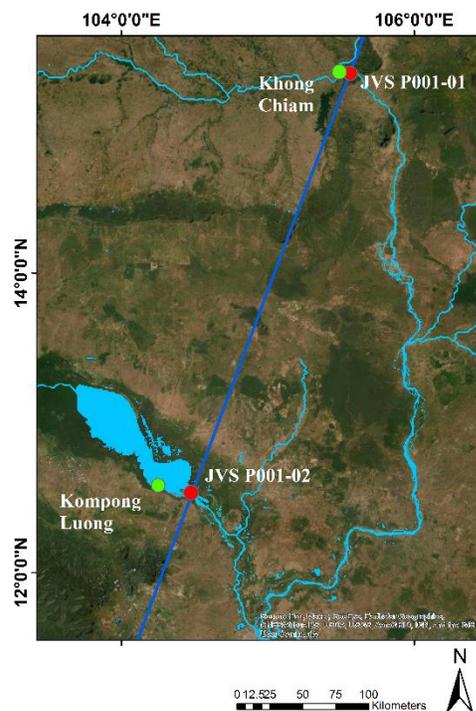


**Figure 5.2** Geographical locations of the MRC in-situ water level gauges used (Basemap is World Imagery provided by ArcMap 10.7).

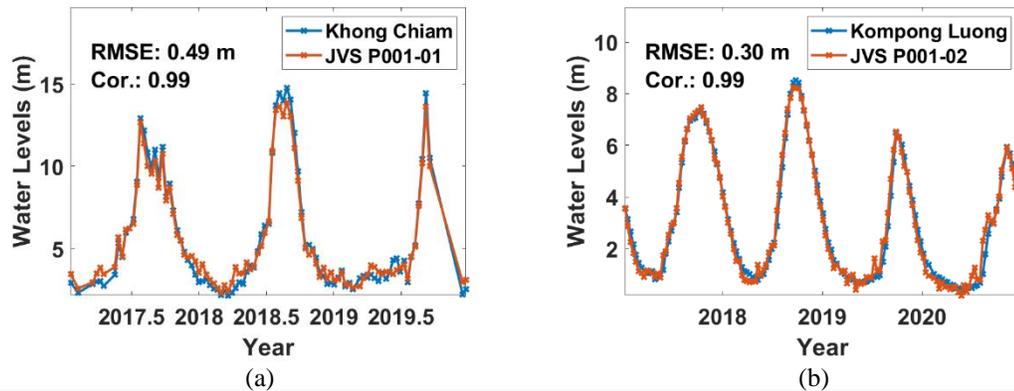
### 5.2.5 Jason-3 Altimetry-derived Water Levels

Jason-3 altimetry satellite, launched on January 17<sup>th</sup>, 2016, is the successor to the Topex/Poseidon (T/P) and Jason-1/-2 missions under the collaboration of NASA and Centre National d'Etudes Spatiales (CNES), National Oceanic and Atmospheric Administration (NOAA) and the European organization for the exploitation of METeorological SATellites (EUMETSAT). The Jason-3 satellite altimetry mission follows the same orbit configuration as its predecessors; thus, it continuously collects highly accurate altimetry data with the same spatial coverage with ~10 days revisit period. In this study, we used Jason-3 data from cycles 33 to 181, spanning from

December 30<sup>th</sup>, 2016 to January 6<sup>th</sup>, 2021. We used 20-Hz ICE-retracked ranges from the Geophysical Data Record (GDR)-D from cycles 33 to 177, while GDR-F was used starting with cycle 178 as it is the only standard data provided since then. After removing outliers in each cycle of altimetry measurements (Okeowo et al., 2017), water level time series (with respect to the WGS84 ellipsoid) at two Virtual Stations (VSs), JVS P001-01 and JVS P001-02, located in the MM near Khong Chiam and TSL, respectively, were derived (See Figure 5.3 for locations of VSs). The Jason-3 derived water levels were validated with MRC in-situ water levels. At JVS P001-01, Jason-3 derived water levels show 0.49 m of Root Mean Square Error (RMSE) and a 0.99 correlation coefficient (Cor.) with nearby Khong Chiam water levels (See Figure 5.4(a)). At JVS P001-02, Jason-3 derived water levels show 0.30 m of RMSE and a 0.99 Cor. with Kompong Luong water levels (See Figure 5.4(b)).



**Figure 5.3** Geographical location of VSs used and corresponding in-situ gauges for validation.



**Figure 5.4** Comparison of altimetry-derived water levels at (a) JVS P001-01 with Khong Chiam and (b) JVS P001-02 with Kompong Luong.

### 5.2.6 Modeled Discharges at Virtual Station

The river discharges at JVS P001-01 are obtained by the Variable Infiltration Capacity (VIC) hydrological model version 5 (Hamman et al., 2018; Liang et al., 1994) and a streamflow routing model (Lohmann et al., 1998). The VIC model is comprised of a three-layer soil column structure and considers land cover, soil type as well as meteorological forcings to characterize the hydrological mechanism in the soil column of given ground cells for runoff and baseflow estimation. Version 5 of the VIC model reconfigures the legacy VIC source code to enhance its reproducibility and integrability with other applications and makes it easier to maintain (Hamman et al., 2018). For this study, we used a 0.1-degree-resolution VIC model, which has been set up and calibrated for the MRB (Hossain et al., 2017). However, the model generates river discharges only down to Kompong Cham due to complicated hydraulic conditions, including bi-directional flow reversal into TSL and tide intrusion in the further downstream areas. Here, we used Integrated Multi-satellitE Retrievals for Global Precipitation

Measurement (IMERG) version 6B early run near real-time product (Tan et al., 2019) for precipitation forcing. For temperature and wind speed forcing, the National Centers for Environmental Prediction (NCEP) Final Operational Model Global Tropospheric Analysis data were used (NCEP, 2000). The streamflow routing model (Lohmann et al., 1998) was then used to route the surface runoff and baseflows estimated by the VIC model to the river channels based on the flow direction map derived from the SRTM DEM for river streamflow simulation at JVS P001-01 from 2017 to 2019.

### 5.2.7 **MERIT DEM**

MERIT DEM (Yamazaki et al., 2017) is a global DEM with respect to the Earth Gravitational Model 1996 (EGM96) with 3 arc-second spatial resolution (about 90 m at the equator). The baseline DEMs consist of 3 arc-second spatial resolution Shuttle Radar Topography Mission DEM (SRTM3 DEM) and the Advanced land observing satellite World 3D-30 m DEM (AW3D-30m DEM), in the regions 60° S to 60° N and 60° N to 90° N, respectively. The unobserved gaps in both SRTM3 and AW3D-30m DEMs are filled with the Viewfinder Panoramas DEM. The NASA Ice, Cloud, and land Elevation Satellite (ICESat) laser altimetry global land surface elevation data (GLAH14) is used as the reference ground elevations for DEM bias estimation. The University of Maryland Landsat forest cover dataset (Hansen et al., 2013) and NASA global forest height data (Simard et al., 2011) are used to estimate DEM errors due to forest canopy. For more details, readers are referred to Yamazaki et al. (2017). In this study, the DEM

was multi-looked to spatial resolutions of 500 m to match the preprocessed Sentinel-1A GRDH images.

## 5.3 Methods

### 5.3.1 FIER Framework

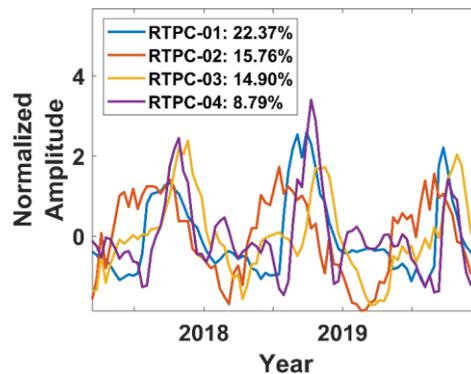
#### 5.3.1.1 REOF Analysis

Here, the first 4 dominant modes were retained for REOF analysis based on the “rule of thumb” of North et al. (1982), accounting for 60.4% of the total variance of the SAR imagery used. The temporally invariant spatial variations of SAR intensities are called Rotated Spatial Modes (RSMs). Their corresponding temporal variations are called Rotated Temporal Principal Components (RTPCs). Hereafter, the first mode of RSM (RTPC) will be called RSM-01 (RTPC-01), the second mode of RSM (RTPC) will be called RSM-02 (RTPC-02), and so forth. For technical details about REOF analysis, readers are referred to Section 4.3.1.

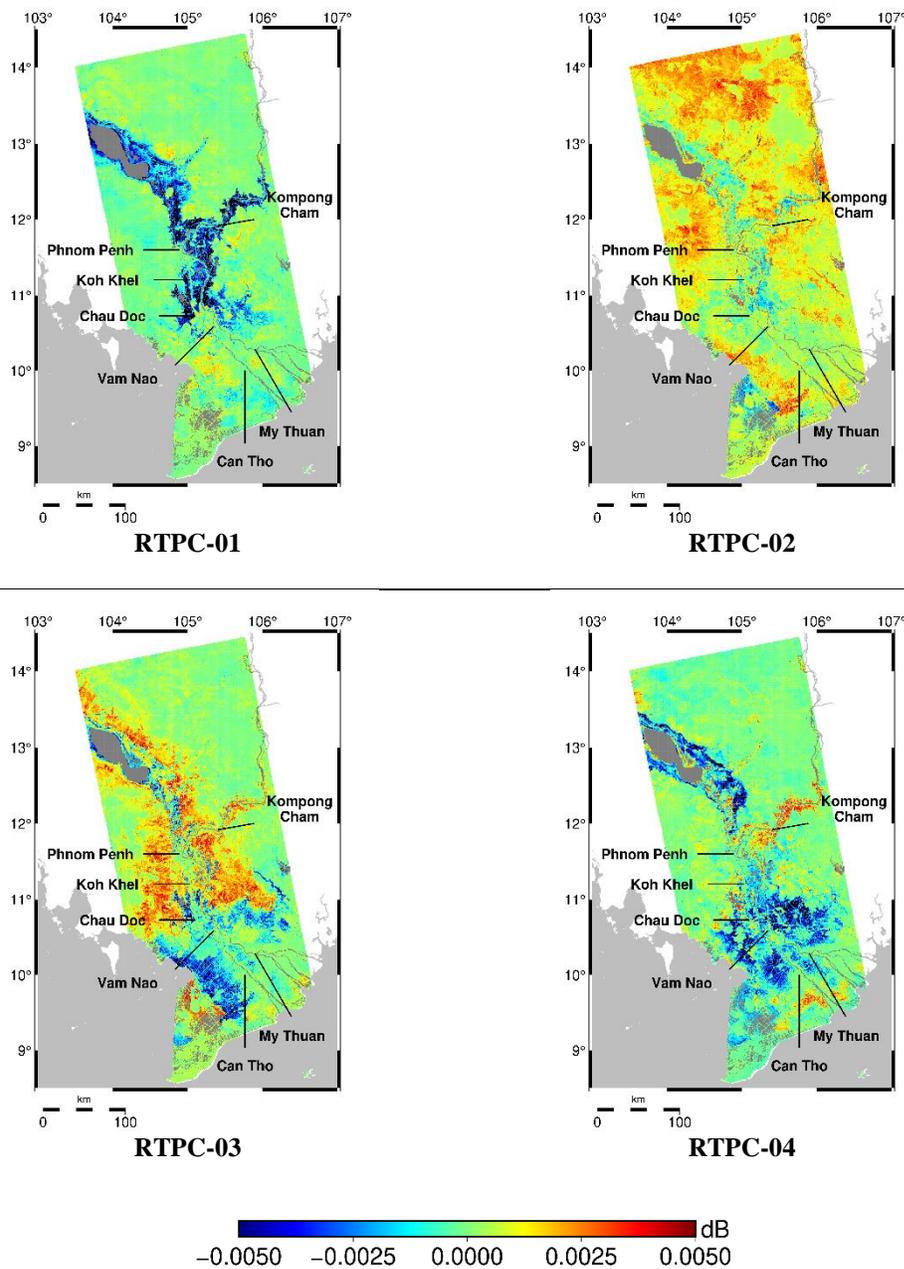
Figure 5.5 shows the normalized RTPCs, where annual seasonalities can be seen. In addition, there are also temporal phase differences between them. RTPC-01 and RTPC-02 seem to reach their peaks earlier than others, while the last to reach their peak is RTPC-03. These temporal phase differences may reflect different stages of floodwater distribution, which can be seen in the corresponding RSMs, as Figure 5.6 shows. In Figure 5.6, large negative-value pixels (deeper blue) represent flood signals as inundation leads to lower SAR intensities due to specular backscattering. In RSM-01, pixels with larger negative values are mainly distributed over the TCF. In RSM-02,

pixels with larger negative values are also mainly distributed over the CF (near Kompong Cham, Chau Doc). These areas are in a relatively upstream region of the MRB floodplain, where seasonal floods occur earlier than downstream. This echoes the patterns of RTPC-01 and RTPC-02, reaching peaks earlier than RTPC-03 and RTPC-04.

In RSM-03 and RSM-04, flood signals can be seen over the MD floodplain as well as the TSLF. In RTPC-03, flood signals over the MD floodplain reach the south of Can Tho, and are more distal from the MR than that in RTPC-04. On the other hand, in RTPC-03, flood signals over the TSLF are restricted to the surrounding areas of TSL, while those in RTPC-04 are distributed widely over the TSLF. This may indicate that mode-03 represents the stage when floodwater has mostly moved from the upstream floodplain to the wider downstream MD floodplain, which also echoes that RTPC-03 is the last to reach its peak (See Figure 5.5.). Mode-04, on the other hand, represents the stage when floodwater has partially moved to the downstream.



**Figure 5.5** Normalized RTPCs (with percentages of explained variance).



**Figure 5.6** REOF-extracted RSM-01 to RSM-04 for multi-temporal Sentinel-1A imagery used in this study. Grey areas are permanent water bodies.

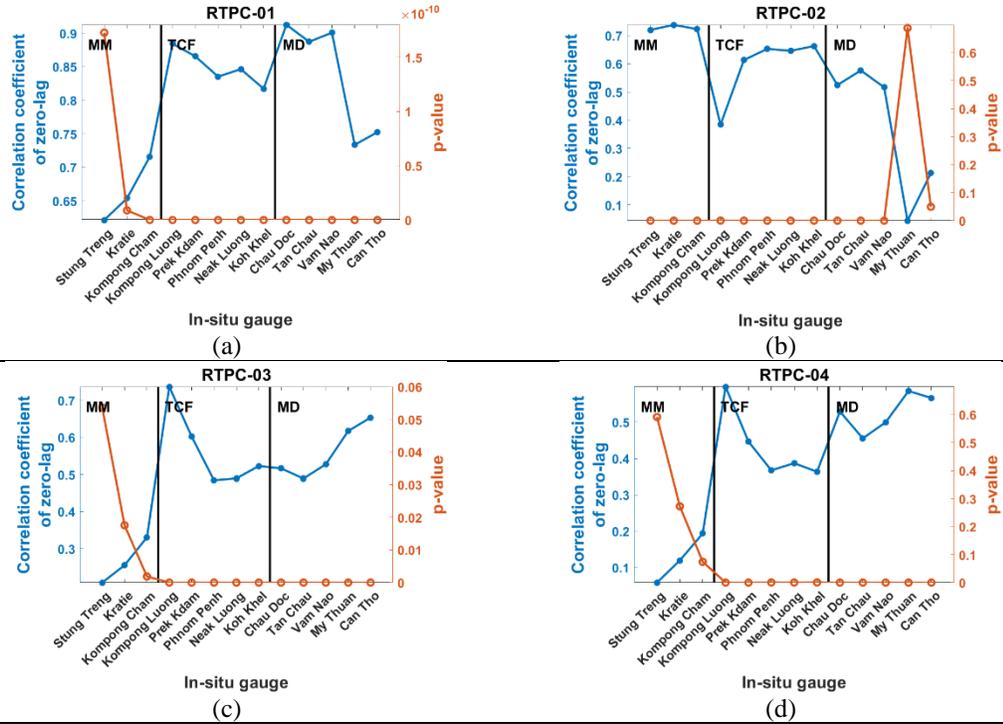
### 5.3.1.2 Synthesis of SAR-like Images

FIER synthesizes SAR-like images by summing up the products of each mode of temporally invariant RSMs and the corresponding temporally variant RTPCs

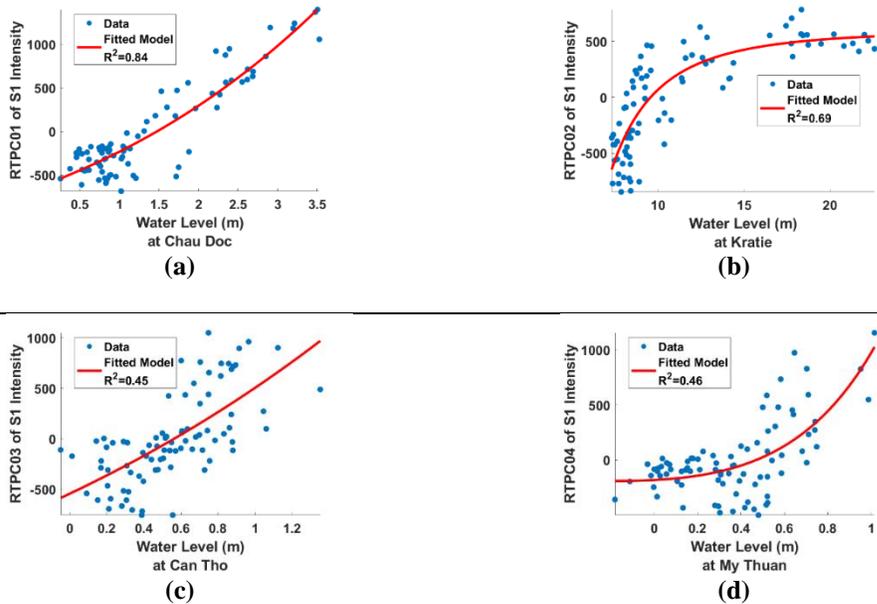
retained. Correlation analysis between the in-situ river levels and the RTPCs retained was conducted to decide which gauges to be used. As Figure 5.7 shows, RTPC-01 better correlates with river levels in the TCF and upper MD (Chau Doc – Vam Nao) with correlation coefficients of 0.8 – 0.9. RTPC-02 better correlates with river levels in the MM (Stung Treng – Kompong Cham) with correlation coefficients above 0.7. In the cases of RTPC-03 and RTPC-04, correlation increases with in-situ river levels closer to the lower MD (My Thuan and Can Tho). This indicates that mode-3 and mode-4 represent the flood signals in the lower MD. However, interestingly, both RTPC-03 and RTPC-04 have similar levels of correlation with water levels at Kompong Luong (located at the TSL) as with those in the lower MD. This may be because the water level at Kompong Luong reaches its peak later than those at downstream gauges in the CF and upper MD (Prek Kdam – Vam Nao), probably due to the role of TSL in the flow reversal mechanism (Kim et al., 2019). That is, TSL level reaches its peak later than downstream as water reversely flows from the MR to TSL through the floodplain.

Based on the correlation analysis between RTPCs and in-situ river levels, we then selected gauges whose river levels better correlate with the RTPCs and built Hydro-to-RTPC models through regression analysis. Consequently, Chau Doc, Kratie, Can Tho, and My Thuan were selected for RTPC-01 to RTPC-04, respectively. For RTPC-03 and RTPC-04, we chose Can Tho, and My Thuan rather than Kompong Luong to ensure the model inputs better reflect local hydrological variation, considering that these two modes are influenced by floods in the MD. Figure 5.8 shows the scatter plots of RTPCs and their corresponding river level data along with the Hydro-to-RTPC models built with regression analysis. With these Hydro-to-RTPC models, river level data at

any time can be fed into the models to estimate the corresponding RTPCs. By summing the products of each mode of RSMs and RTPCs, SAR-like images can be synthesized.



**Figure 5.7** Zero-lag correlation between MRC in-situ water levels at different gauges and RTPCs. Note that MM means Mekong Mainstem, TCF means Tonle Sap Lake and Cambodian Floodplain, while MD represents Mekong Delta

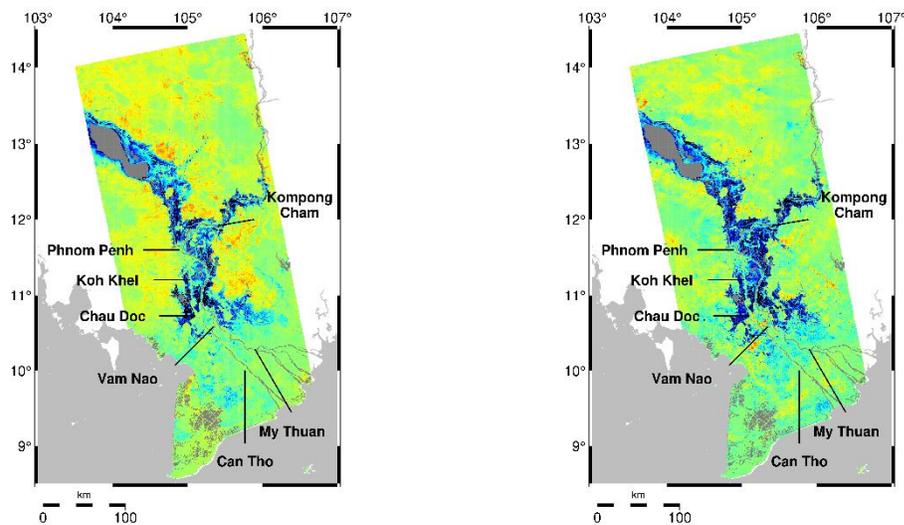


**Figure 5.8** Hydro-to-RTPC models between MRC in-situ river levels at selected gauges and RTPCs.

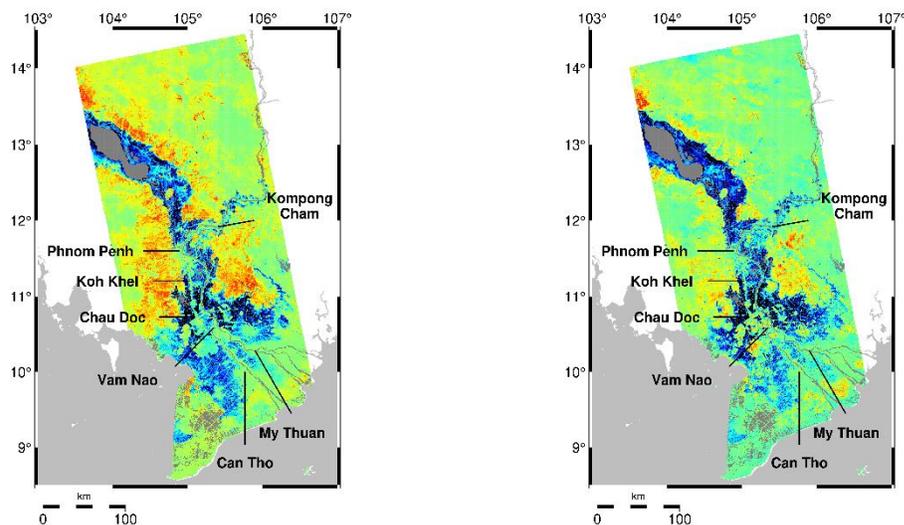
### 5.3.2 Iterative Threshold Calibration and Water Classification

After SAR-like images have been synthesized, the next step is to perform water classification. Here, a change detection approach was adopted, subtracting the driest condition image (hereafter called baseline image) from those of other dates to get difference images (Clement et al., 2018; Long et al., 2014; Singha et al., 2020). The historical difference images were then used for an iterative threshold calibration process. Once the threshold was determined, it was applied to the difference images on any other dates for water classification. The iterative threshold calibration was also performed on the Sentinel-1 SAR images acquired from March 13<sup>th</sup>, 2017 to December 28<sup>th</sup>, 2019, which will be later used for comparison with FIER-derived inundation extents.

As we adopted river levels at several gauges for FIER to synthesize SAR-like images, river levels of these gauges were taken into account to generate the baseline image. We identified the dates within the Sentinel-1 SAR image acquisition period from March 13<sup>th</sup>, 2017 to December 28<sup>th</sup>, 2019 when we have the lowest river levels at each of the selected gauges. Those were on July 1<sup>st</sup>, July 25<sup>th</sup> and December 16<sup>th</sup> of 2019. The baseline image was then generated as the median value composite of the images from these dry dates (Clement et al., 2018; Singha et al., 2020). Then, difference images were generated by subtracting the baseline image from the other FIER-synthesized SAR-like and real Sentinel-1 images. Due to specular backscattering, inundated areas in difference images would appear to be dark. Examples of difference images are shown in Figure 5.9.



August 23<sup>rd</sup>, 2018



October 10<sup>th</sup>, 2020



(a)

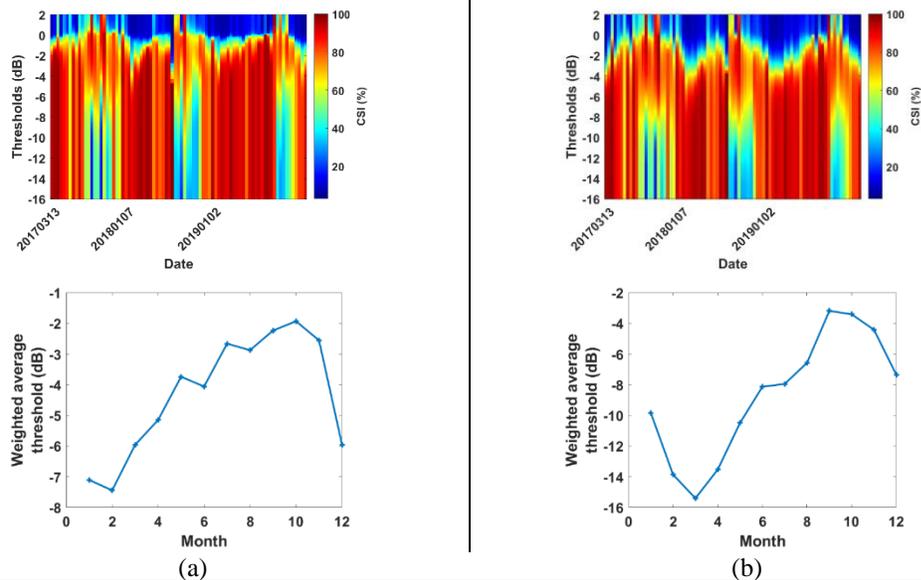
(b)

**Figure 5.9** Examples of difference images in the wet season generated from (a) FIER-synthesized SAR-like images and (b) Sentinel-1 SAR images where lower difference values over inundated areas can be clearly seen. Grey areas are permanent water bodies.

Next, the threshold candidates were iteratively applied to generate inundation extents for the period from 2017 – 2019, which were then compared with the NASA

GSFC NRT flood maps to obtain the CSI (top panel of Figure 5.10). The 2015 MOD44W water mask was added to the generated inundation extents as the permanent water body. For each calibration date, a threshold was selected which led to maximum CSI. Then, those thresholds from the same month were averaged using the numbers of valid pixels on their corresponding NASA GSFC NRT flood maps as weights to finally derive the monthly calibrated thresholds (bottom panel of Figure 5.10). The monthly calibrated thresholds for the Sentinel-1 images to derive the inundation maps for cross-comparison purpose were also determined in a similar manner.

For both the FIER-synthesized SAR-like images and Sentinel-1 SAR images, higher monthly calibrated thresholds are obtained during the wet season, as can be seen in the bottom panel of Figure 5.10. This may be because the inundated floodplains during the wet season have higher intensities than those from the permanent water bodies.



**Figure 5.10** (Top panel) CSIs from different threshold candidates for each calibration date, (bottom panel) calibrated monthly thresholds, for the case of (a) FIER-synthesized SAR-like images, and (b) Sentinel-1 SAR images.

### 5.3.3 Daily Water Level Forecasting

The forecasting of inundation extents using FIER requires forecasted river levels as inputs, and the model-aided satellite altimetry-based MR level forecasting system developed by Chang et al. (2019) was adopted for this study.

There are two upstream VSs used in this study; one located in the upstream of the MM near Khong Chiam and the other one located at the TSL: named JVS P001-01 and JVS P001-02, respectively. The reconstruction of JVS P001-01 altimetric river levels was based on a depth-discharge power-law rating curve (Leopold and Maddock, 1953)

$$d_w = e \cdot Q^f \quad (5.1)$$

where  $d_w$  is the water depth,  $Q$  is the discharge modeled with VIC and  $e$  and  $f$  are parameters to be estimated. Since satellite altimetry provides water levels with respect to a specific datum, (WGS84 reference ellipsoid in this case) not the river bed, we followed Kim et al. (2019a and 2019b) to estimate  $d_w$  in Eq. (5.1), which was reformatted as

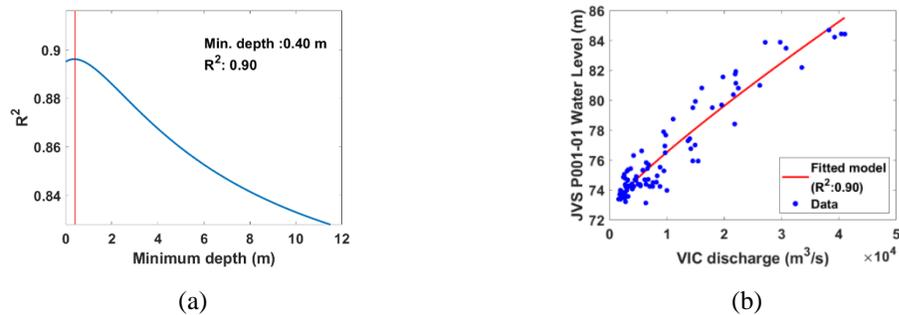
$$d_w = h_{Inst.}^{Alt.} - h_{min}^{Alt.} + d_{min} = e \cdot Q_{VIC}^f \quad (5.2)$$

where  $h_{Inst.}^{Alt.}$  is instantaneous altimetry-derived water level,  $h_{min}^{Alt.}$  is the minimum altimetry-derived water level within the observation period which has the corresponding minimum water depth of  $d_{min}$  and  $Q_{VIC}$  is the VIC-derived discharge. The minimum water depth ( $d_{min}$ ) was then determined by maximizing the coefficient of determination ( $R^2$ ) of the fitted depth-discharge rating curve (Kim et al., 2019a and 2019b) as in Figure 5.11(a). The final fitted depth-discharge rating curve is shown in Figure 5.11(b). Finally,

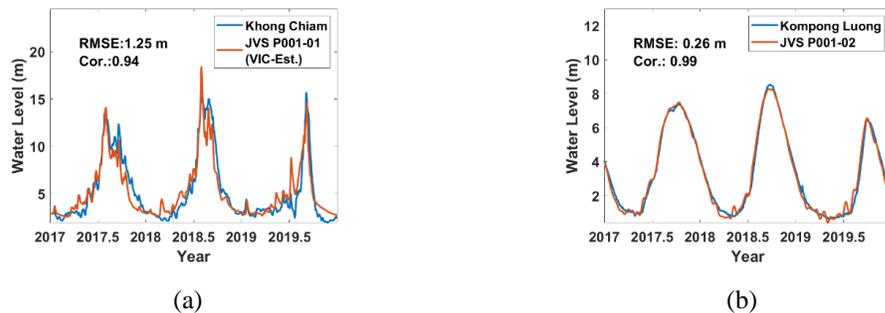
the daily reconstructed water levels at JVS P001-01 were obtained by reformatting Eq. (5.2) as

$$h_{Rec.}^{MK} = e \cdot Q_{VIC}^f + (h_{min}^{Alt.} - d_{min}). \quad (5.3)$$

Since the current version of the VIC model for the Mekong Basin (Hossain et al., 2017) does not simulate flooding of TSL, a different approach has to be considered for JVS P001-02 at TSL. Fortunately, since TSL water levels have smooth annual variation, daily TSL levels from Jason-3 altimeter data were simply reconstructed using a spline interpolation. Figure 5.12 shows the validation of daily reconstructed water levels from 2017 – 2019 at these two VSs with in-situ water levels from nearby gauges.



**Figure 5.11** (a) Iterative definition of the minimum depth at JVS P001-01 by maximizing  $R^2$  of fitted depth-discharge power-law relationship and (b) the final fitted rating used to reconstruct altimetry-derived water levels.



**Figure 5.12** Comparison of reconstructed daily water levels at (a) JVS P001-01 and (b) JVS P001-02 with in-situ water levels at Khong Chiam and Kompong Luong, respectively.

Based on the correlation analysis performed in section 5.3.1.2, water levels at the selected gauges of Kratie, Chau Doc, Can Tho, and My Thuan were forecasted to obtain RTPC-01 to RTPC-04, respectively. The water levels at Kratie were forecasted using a simple linear regression model with JVS P001-01 water levels as the only input since it is located in the middle reach of MM where its hydraulic condition is relatively simple (Chang et al., 2019)

$$h_{In-situ}(t+k) = E \cdot h_{Rec.}^{MK}(t) + F \quad (5.4)$$

where  $h_{Rec.}^{MK}(t)$  is the reconstructed water levels at JVS P001-01 at the time  $t$  that can be derived using Eq. (5.3),  $h_{In-situ}(t+k)$  is the historical water levels at Kratie at time  $(t+k)$  with  $k$  as lead time and  $E$  and  $F$  are parameters to be fitted.

To forecast the water levels at Chau Doc, Can Tho, and My Thuan, which are located in the MD, we need to take into account upstream water flows from the MM, TSL, and ocean tide intrusion from downstream (Pagano, 2014). Accordingly, the water levels at JVS P001-02 and tidal oscillations using a sinusoidal model are considered as (Chang et al., 2019)

$$h_{In-situ}(t+k) = G \cdot h_{Rec.}^{MK}(t) + H \cdot h_{Rec.}^{TSL}(t) + I + h_{Tide}(t+k), \quad (5.5)$$

$$h_{Tide}(t) = \sum_{i=1}^5 M_i \cdot \sin(2\pi \cdot \omega_i \cdot t + N_i)$$

where  $h_{Rec.}^{TSL}(t)$  is the reconstructed water level of time  $t$  at JVS P001-02 and  $h_{Tide}$  is the tidal fluctuation modeled by the 5-term sinusoidal function. Parameters to be fitted are  $G$ ,  $H$ , and  $F$  as well as magnitudes  $M_i$  and phase shift  $N_i$  of sinusoidal functions.

The known dominant tidal frequencies ( $\omega_i$ ) include the annual, semi-annual, monthly, fortnightly and synodic fortnightly tides (See Table 5.1).

**Table 5.1** Dominant ocean tide frequencies.

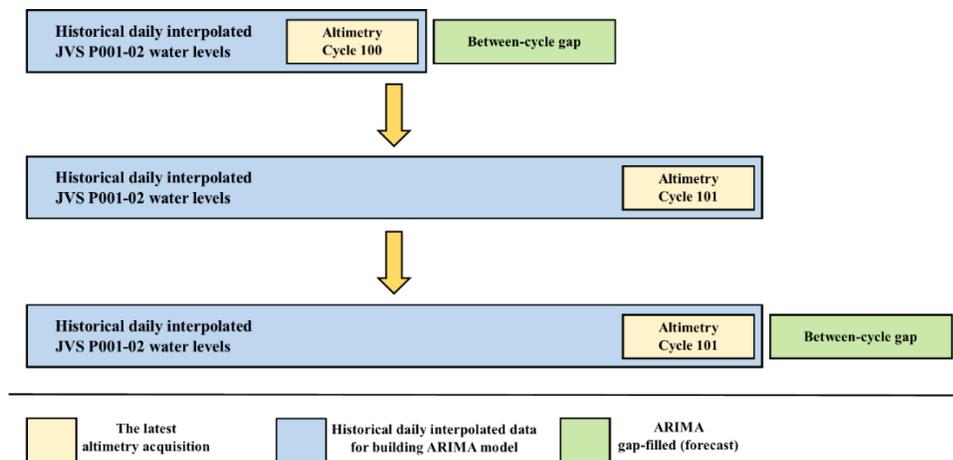
Name of tide	Period (Days)	Frequency (1/year)
<b>Annual</b>	365.26	1.00
<b>Semi-annual</b>	182.62	2.00
<b>Monthly</b>	27.55	13.26
<b>Fortnightly</b>	14.77	24.73
<b>Synodic fortnightly</b>	13.66	26.74

For pseudo-forecasting in 2020, the daily updated VIC-derived discharge was used to reconstruct the daily water levels at JVS P001-01 following Eq. (5.3). For JVS P001-02, since Jason-3 altimetry has a revisit period of ~10 days, the data gap between the latest available and the next Jason-3 data acquisition has to be filled in order to perform daily forecasting until the next visit of Jason-3. To fill the gap, a method called regression with Auto-Regressive Integrated Moving Average (ARIMA) errors was adopted following Hyndman (2013, 2014). The method adopts a Fourier term to deal with seasonality and models the remaining signals with the ARIMA process. Considering the fact that the annual signal (365.25 days of period) is dominant for the water level at JVS P001-02, the regression with ARIMA errors can be written as

$$h_{pre.}^{TSL} = c \cdot t + \left( \alpha \cdot \sin\left(\frac{2\pi \cdot t}{365.25}\right) + \beta \cdot \cos\left(\frac{2\pi \cdot t}{365.25}\right) \right) + \eta_t \quad (5.6)$$

where  $h_{pre.}^{TSL}$  is the daily spline-interpolated JVS P001-02 water levels up to the latest available Jason-3 data acquisition,  $t$  is the time,  $c$  is the temporal trend,  $\alpha$  and  $\beta$  are the parameters for the Fourier term with the annual seasonality and  $\eta_t$  is the remaining

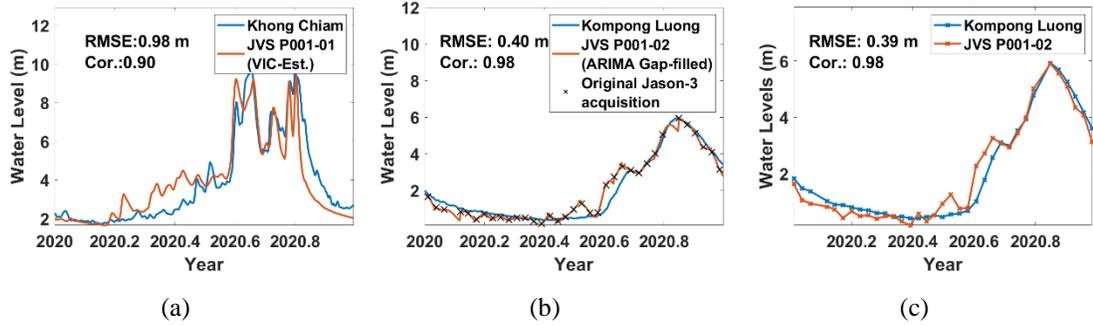
signal which is modeled by the ARIMA process. The gap-filling process was realized using the *auto.arima* function in the R language *forecast* package which automatically finds the best ARIMA structure by minimizing the Akaike Information Criterion (AIC) (Hyndman and Khandakar, 2008; Wang et al., 2006). Figure 5.13 illustrates this gap-filling process. First, the daily interpolated JVS P001-02 water levels were calculated up to the latest acquisition. The daily interpolated JVS P001-02 water levels were then used to build the ARIMA model and perform forecasting with lead times up to the between-cycle gap (i.e., 10 days). The process was repeated whenever new Jason-3 data becomes available.



**Figure 5.13** Illustration of the gap-filling process for daily JVS P001-02 water levels for pseudo forecasting.

The reconstructed daily water levels at the two VSs in 2020, for the purpose of pseudo-forecasting, are shown in Figure 5.14 with in-situ water levels at nearby gauges. As can be seen from Figure 5.14(b) and Figure 5.14(c), the RMSE between the original JVS P001-02 (with 10-day gaps) and the in-situ water levels is 0.39 m, which is close to that of 0.40 m between the ARIMA gap-filled JVS P001-02 and in-situ water levels.

This demonstrates that the method, regression with ARIMA errors, can successfully reconstruct daily water levels from the original JVS P001-02 water levels with 10-day gaps.



**Figure 5.14** Reconstructed and gap-filled daily water levels at (a) JVS P001-01 and (b) JVS P001-02 along with in-situ water levels at Khong Chiam, and Kompong Luong, respectively. The original JVS P001-02 water levels with the in-situ levels at Kompong Luong are shown separately in (c).

The skill of the forecasted water levels was evaluated by Mean Absolute Error (MAE) and Nash-Sutcliffe Efficiency (NSE). The MAE is a measure of error that can be calculated by

$$\text{MAE} = \frac{\sum_{i=1}^N |Y_i^{obs} - Y_i^{fct}|}{N} \quad (5.7)$$

where  $N$  is the total number of data, with  $i = 1 \sim N$ .  $Y_i^{obs}$  the in situ observations and  $Y_i^{fct}$  is the forecasting results. A MAE value of 0 indicates perfect forecasting.

The NSE is a normalized statistic which reflects the relative magnitude of the residual variance (noise) compared to the measured data variance (information) (Nash and Sutcliffe, 1970). It indicates how well the plot of observations versus forecasting results fits the 1:1 line and is commonly used to evaluate the performance of a forecasting model

$$\text{NSE} = 1 - \frac{\sum_{i=1}^N (Y_i^{obs} - Y_i^{fct})^2}{\sum_{i=1}^N (Y_i^{obs} - \overline{Y^{obs}})^2} . \quad (5.8)$$

NSE ranges from  $-\infty$  to 1.0. An NSE value of 1 indicates perfect forecasting skill, 0 means the forecasting skill is no better than adopting the average of the observations as forecasting results and a negative NSE means unacceptable skill (Moriassi et al., 2007).

### 5.3.4 Framework Skill Evaluation Statistics

Skill evaluation was conducted based on a  $2 \times 2$  confusion matrix (Kohavi and Provost, 1998) as Figure 5.15 illustrates, where  $a$  and  $d$  are the number of pixels which FIER correctly estimates as inundated (*hit*) and non-inundated (*correct negative*), respectively. Conversely,  $b$  and  $c$  are the number of pixels that are misestimated, representing *false alarm* and *miss*, respectively.

Confusion Matrix		Observation		Marginal Total
		Inundation	Non-inundation	
FIER estimation	Inundation	$a$ (Hits)	$b$ (false alarms)	$a + b$
	Non-inundation	$c$ (missed)	$d$ (correct negative)	$c + d$
Marginal Total		$a + c$	$b + d$	Total = $a + b + c + d$

**Figure 5.15** A  $2 \times 2$  confusion matrix, which displays the number of pixels that is *hit*, *false alarm*, *miss*, or *correct negative*.

We used pixel count-based skill evaluation indices including overall accuracy, CSI, omission error, and commission error. The overall accuracy is the percentage of pixels which were correctly estimated as either inundated or non-inundated (*hit* or *correct negative*) by the total number of pixels

$$\text{Overall accuracy} = \frac{a + d}{a + b + c + d} \times 100 (\%). \quad (5.9)$$

The range of overall accuracy is 0 – 100%, meaning zero skill to perfect skill.

CSI (Gilbert, 1884), also called the *threat score*, is the number of correctly estimated inundated pixels (*hit*) over the total number of pixels, which are either observed or estimated as inundated (*hit + false alarm + miss*)

$$\text{CSI} = \frac{a}{a + b + c} \times 100(\%). \quad (5.10)$$

CSI takes both *false alarm* and *miss* into account and is therefore considered to be more complete. It also avoids possible bias caused by *correct negative* (Wing et al., 2017). It ranges from 0 to 100%, meaning zero skill to perfect skill. CSI is often used as a standard validation measure (World Meteorological Organization, 2017).

Omission error represents the percentage of pixels which are actually inundated but are not captured over the total number of actually inundated pixels

$$\text{Omission error} = \frac{c}{a + c} \times 100 (\%). \quad (5.11)$$

It ranges from 0 to 100% with 0% meaning perfect skill. This indicates the extent of *missed* pixels.

Commission error is the percentage of estimated inundated pixels which are actually non-inundated over the total number of estimated inundated pixels

$$\text{Commission error} = \frac{b}{a + b} \times 100 (\%). \quad (5.12)$$

It ranges from 0 to 100% with 0% meaning perfect skill. This indicates the extent of *false alarm* pixels.

### 5.3.5 Inundation Depth Estimation

In this study, the inundation depths were determined by using the Floodwater Depth Estimation Tool (FwDET) version 2 (Cohen et al., 2019). The tool consists of

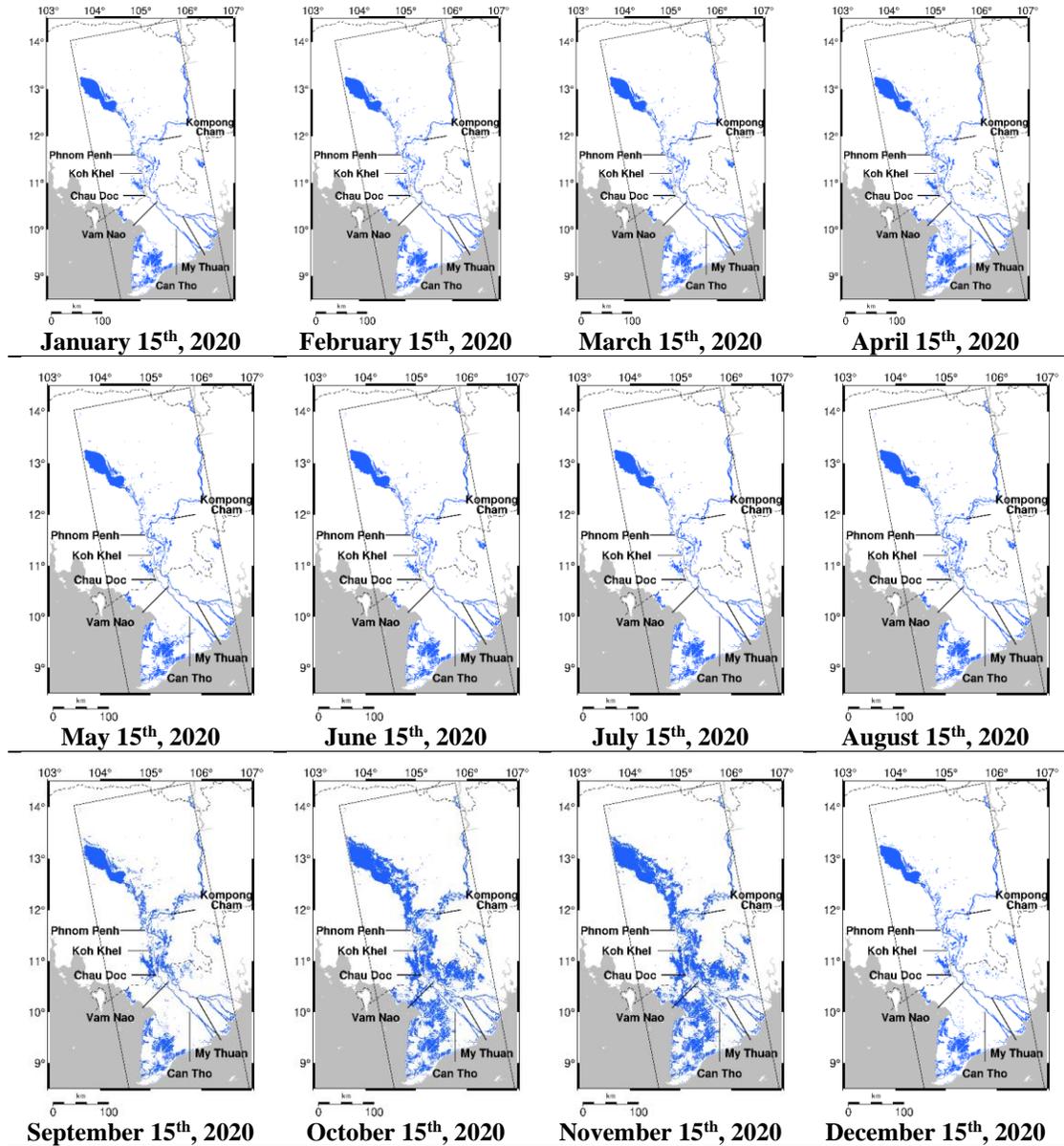
several steps including (1) identifying inundation boundary cells, (2) extracting the elevation of the inundation boundary cells, (3) assigning the elevations of the nearest inundation boundary cells to the non-boundary inundated cells, (4) floodwater depth estimation by subtracting the DEM value from the assigned boundary elevation at inundated cells, and (5) smoothing. The stand-alone ArcPy script is available on <https://github.com/csdms-contrib/fwdet>. The tool requires polygons of inundation extents as input, which can be generated by ArcGIS 10.7 using the “Raster to Polygon” function. We did not perform the smoothing process as it can also smooth the signal of real inundation depths. For a detailed technical description of FwDET, readers are referred to Cohen et al. (2018, 2019).

## 5.4 Results and Discussions

### 5.4.1 FIER-hindcasted Inundation Extents in 2020

FIER-hindcasted inundation extents over the MRB floodplain for 2020 were generated using in-situ water levels as inputs. Note that FIER is capable of generating daily inundation extents as long as daily water levels (or streamflows) are available. Figure 5.16 shows examples of FIER-hindcasted inundation extents on the 15<sup>th</sup> of each month in 2020, illustrating the dynamics of the areal inundation for a year in the MRB floodplain where widespread flooding can be seen in October and November. Skills of the FIER hindcast were then evaluated by cross-comparing with inundation extents derived from the original SAR images. Since the in-situ water levels are considered as ground truth and used as inputs in this case, the performance of FIER itself can be

assessed. For details about the water classification method, readers are referred to section 5.3.2.



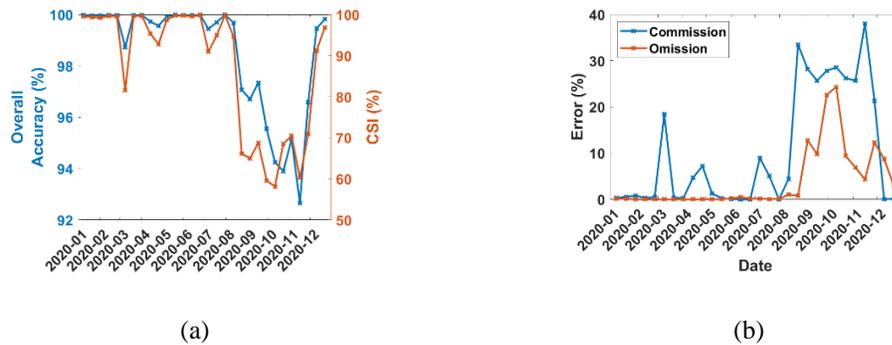
**Figure 5.16** Examples of FIER-hindcasted inundation extents on the 15<sup>th</sup> of each month for 2020 using MRC in-situ river levels as inputs. The black box in the background shows the boundary of Sentinel-1 frames.

The results of cross-comparison with concurrent Sentinel-1 images (See Section 5.2.1.) are shown in Figure 5.17. During the wet season, the overall accuracies and CSIs

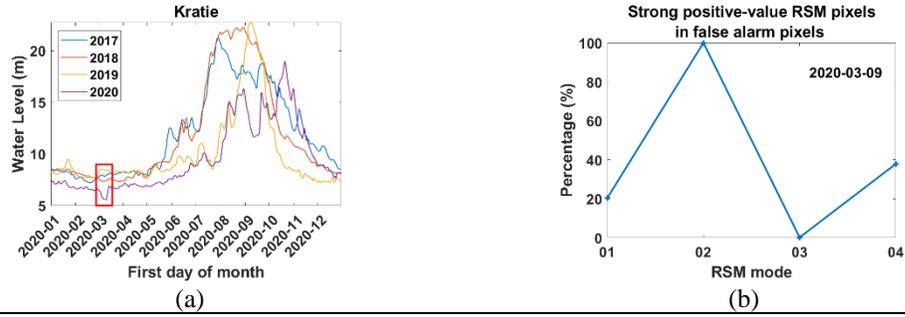
are around 93 – 97% and 60 – 70%, respectively (Figure 5.17(a)). The omission errors are about 5 – 25% during the wet season, while the commission errors are about 25 – 40% (Figure 5.17(b)). One of the potential error sources could be the less skillful Hydro-to-RTPC-03/-04 models which have relatively lower  $R^2$  than those of the first 2 modes. It should be noted that our FIER implementation in the MRB currently considers the dominant modes only for synthesizing the SAR-like images, which mostly retain the strong seasonal SAR intensity variations caused by the seasonal flood regime. Hence, the short-period and weaker variations in the in-situ MD water level may not be well coupled with RTPC-03/-04. Furthermore, MD is a region with intensive human water control with densely distributed canals and dykes as well as sluice gate operation and water pumping activity for enhancing agricultural productivity (Hung et al., 2012, 2014; Tran et al., 2019, 2020). Such human intervention weakens the natural hydraulic connectivity between the MD floodplain water levels and the MR water levels (Hung et al., 2012; 2014) which could also lead to less skillful couplings between the MD water levels with RTPC-03/-04 and the resulting errors when synthesizing SAR-like images.

On the other hand, as the accuracy of the synthesized SAR-like images relies on input water levels, short-period oscillations of in-situ water levels in the MD caused by tides can lead to errors in the SAR-like images. In addition, since water classification is based on monthly thresholds, strong intra-monthly variations in in-situ water levels in the MD could also result in overestimations or underestimations. Another factor that can cause errors, particularly during the dry season, is when the input water levels used for SAR-like image synthesis is lower than that used to generate the baseline image for the change detection-based water classification. Figure 5.18(a) shows an abrupt drop in

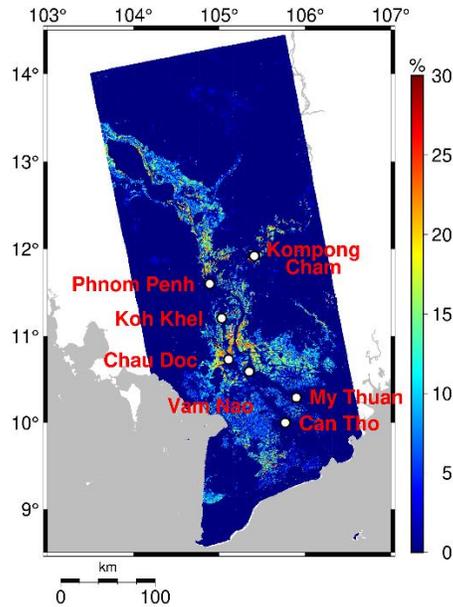
water levels at Kratie, the inputs to the Hydro-to-RTPC-02 model, in early March 2020 which makes water levels lower than 2017 – 2019. Figure 5.18(b) shows that all of the *false alarm* pixels on March 9<sup>th</sup>, 2020, have large positive values (> 90% percentile) in RSM-02. Since the change detection-based water classification we adopted was performed upon the difference images generated by subtracting the baseline image from the SAR-like images (See Section 5.3.2.), the “lower-than-baseline” input water levels can lead to larger negative difference values over positive RSM pixels, resulting in overestimated inundation extents. But this type of overestimation only occurs in the dry season and therefore would not be an issue if estimating the wet season inundation extents is the main interest. Figure 5.19 shows the occurrence (%) of either *miss* or *false alarm* pixels out of the 30 images used for cross-comparison indicating where the FIER hindcast in the MRB results should be used with more caution



**Figure 5.17** (a) Overall accuracies, and CSIs, (b) omission, and commission errors for the FIER-hindcasted inundation extents in 2020, compared with original SAR image-derived values.



**Figure 5.18** (a) In-situ river level at Kratie in 2017 – 2020 where an abrupt drop in early March 2020 is marked in the red box and (b) percentages of false alarm pixels on March 9<sup>th</sup>, 2020, that have high positive RSM values (> 90% percentile).



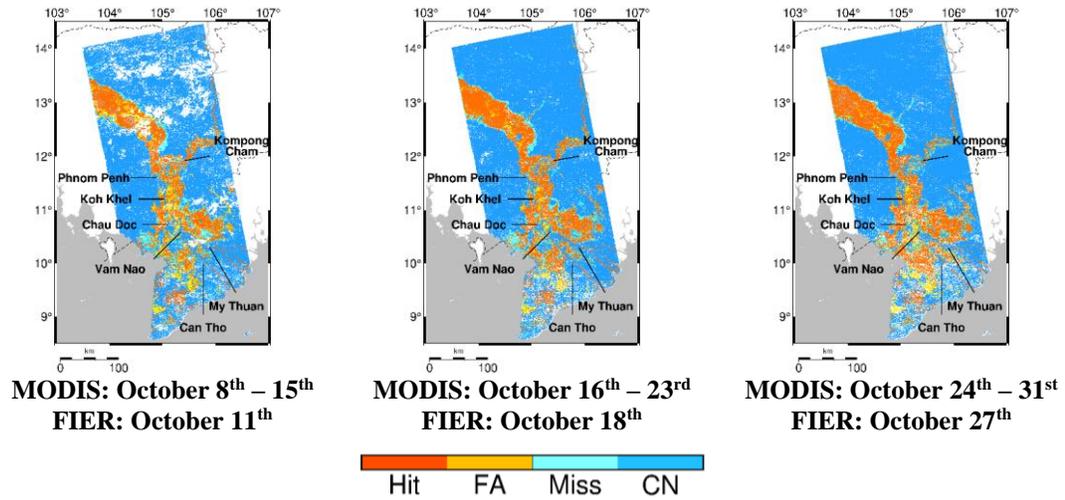
**Figure 5.19** Occurrence (%) of errors over 30 images used for cross-comparison.

We then assessed the skills of the FIER hindcast in the MRB under extreme flooding. As the 2011 flood was the most severe flooding event since 2000 and the water levels in October had the maximum exceedance for that year (MRC, 2014), the inundation extents in October 2011 were hindcasted. In this case, MODIS-based inundation maps were used for cross-comparison. Since the NASA GSFC NRT flood

maps available on the web portal do not cover the entire MRB floodplains back in 2011, we generated inundation maps from MOD09A1 8-day composite products using the same band math criteria as the NASA GSFC NRT flood maps (Slayback, 2021). Due to dynamic cloud cover over the MODIS images, the statistics were averaged by using the numbers of valid pixels as weights. Table 5.2 shows that the FIER hindcast provides about 92% of overall accuracies and 68% of CSIs with 21% of omission errors and 18% commission errors. The spatial distribution of each category of the cross-comparison results is shown in Figure 5.20. The CSIs of the FIER hindcast during the 2011 flood are of similar levels as (64% and 74%) in Triet et al. (2017, 2018) which simulated the inundation extents in 2011 over the MD using a quasi-two-dimensional hydraulic model.

**Table 5.2** Skills of the FIER hindcast in October 2011 cross-compared with MODIS-based inundations maps.

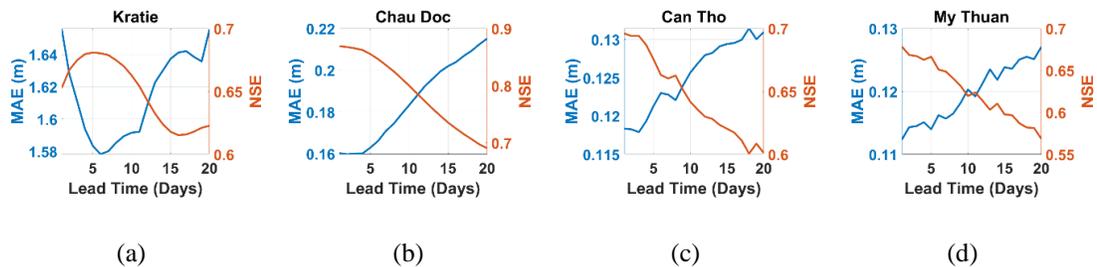
<b>Dates of MODIS MOD09A1 / FIER synthesis</b>	<b>Skills (%)</b>		<b>Errors (%)</b>		<b>Valid pixels</b>
	<b>Accuracy</b>	<b>CSI</b>	<b>Omission</b>	<b>Commission</b>	
<b>October 8<sup>th</sup> – 15<sup>th</sup> / October 11<sup>th</sup></b>	89.2	61.5	22.9	24.8	358261
<b>October 16<sup>th</sup> – 23<sup>rd</sup> / October 18<sup>th</sup></b>	92.1	69.5	22.2	13.2	428428
<b>October 24<sup>th</sup> – 31<sup>st</sup> / October 27<sup>th</sup></b>	93.0	71.5	17.8	15.5	406414
<b>Weighted average</b>	91.6	67.8	20.9	17.5	



**Figure 5.20** Cross-comparison between the FIER hindcast and MODIS-based inundation maps in October 2011 (FA: false alarm; CN: correct negative).

#### 5.4.2 FIER Pseudo-forecasted Inundation Extents in 2020

The daily FIER pseudo-forecasted inundation extents over the MRB floodplain for 2020 with up to 20-day lead time have been generated with forecasted water levels as inputs. Figure 5.21 shows MAEs and NSEs of forecasted water levels for different lead times at Kratie, Chau Doc, Can Tho, and My Thuan, that used as inputs to the FIER pseudo-forecast. For Kratie, MAEs are from 1.58 – 1.65 m with NSEs about 0.6 – 0.7. For Chau Doc, MAEs are from 0.16 to 0.22 with NSEs about 0.7 – 0.9. For Can Tho and My Thuan, MAEs are from about 0.11 – 0.13 m with NSEs about 0.6 – 0.7.



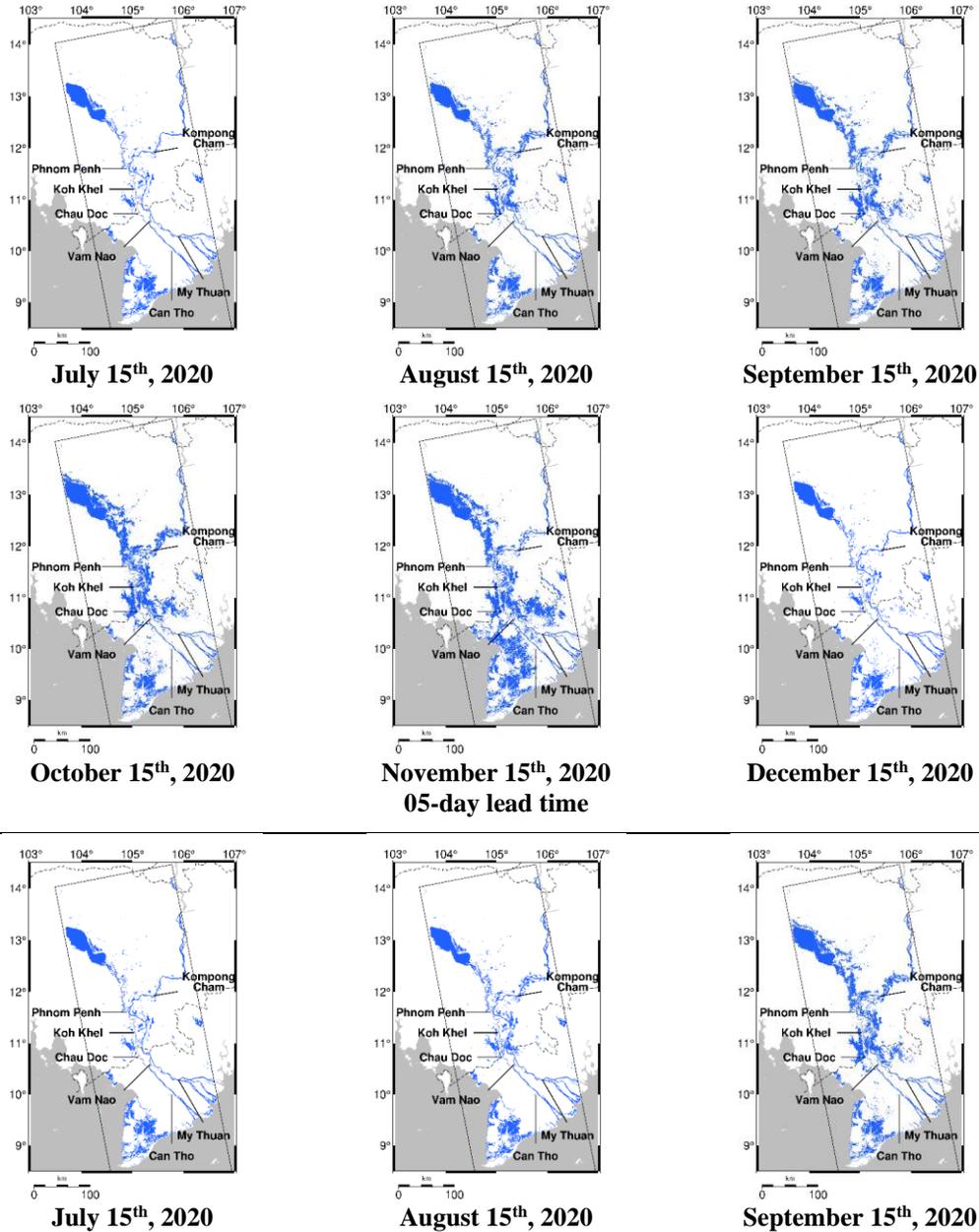
**Figure 5.21** Skills water level forecasting at selected gauges with up to 20-day lead time.

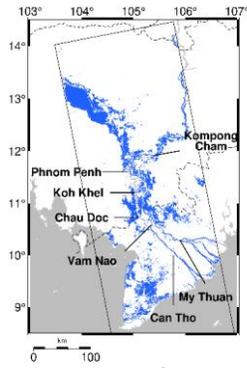
In Figure 5.22, FIER pseudo-forecasts with 5, 10, 15 and 20 days of lead time from the arbitrarily chosen 15<sup>th</sup> of July to December, the months during the critical flood

risk period (Wassmann et al., 2019), are shown as examples. The average skills of the FIER pseudo-forecast with 1-day to 20-day lead times are shown in Figure 5.23. In the wet season, the overall accuracies and CSIs are from 94 – 98%, and 50 – 70%, respectively (See Figure 5.23(a)). The omission and commission errors are up to 30% and 40%, respectively (See Figure 5.23(b)). Since the FIER forecast requires forecasted water levels as inputs, its skills can be influenced by not only the factors discussed in Section 5.4.1 but also the accuracy of forecasted water levels. Potential error sources of our forecasted water levels include the propagation of errors from the VIC-estimated discharges and altimetry-derived water levels at upstream VSs. For example, it can be seen in Figure 5.14(b) and Figure 5.14(c) that the altimetry water levels at JVS P001-02 (at the TSL) were about 1 – 2 m higher than the in-situ water levels at Kompong Luong in August and September 2020. An underestimation in reconstructed water levels at JVS P001-01 can be seen near early October 2020 in Figure 5.14(a). This could be an error source for the higher commission errors in August and September and higher omission errors in October 2020 as shown in Figure 5.23(b). For detailed discussion about the error sources of the forecasted water levels, readers are referred to Section 3.4.

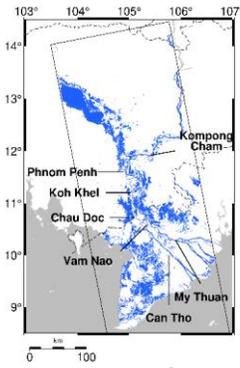
It is worth mentioning that the FIER forecast has the flexibility to take forecasted water levels generated by different approaches as inputs. In other words, FIER can take advantage of different water level (or streamflow) forecasting systems that may lead to improved inundation forecasting skill. For example, the water level forecasting system developed by Chang et al. (2019) is computationally lighter than the MRC's water level forecasting system and performs promisingly inside the MD, while the MRC's system has better skill at locations outside of the MD (such as at Kratie) but does not routinely

issue forecasting inside the MD due to a heavy computational burden (Chang et al., 2019; Pagano, 2014). With the flexibility of FIER, it is possible to assemble the forecasted water levels at different gauges generated by two different systems for more accurate forecasts of inundation extents. In the following section, applications of the daily FIER pseudo-forecasted inundation extents will be introduced.

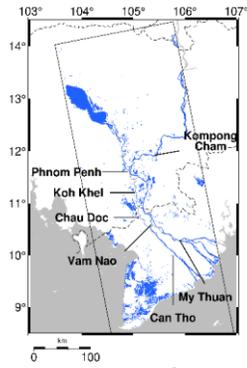




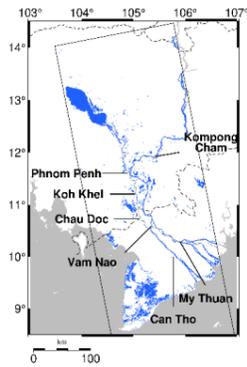
**October 15<sup>th</sup>, 2020**



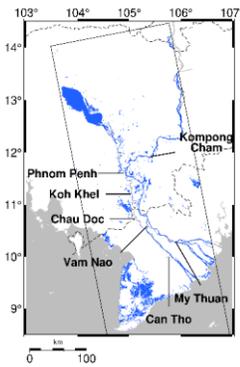
**November 15<sup>th</sup>, 2020  
10-day lead time**



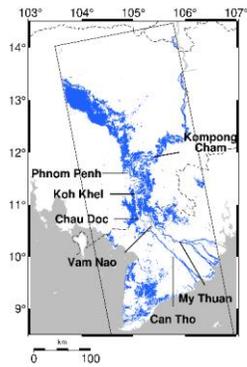
**December 15<sup>th</sup>, 2020**



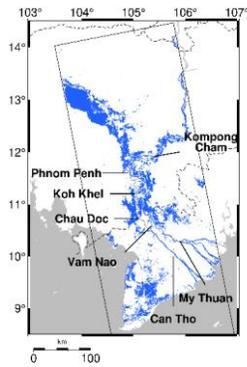
**July 15<sup>th</sup>, 2020**



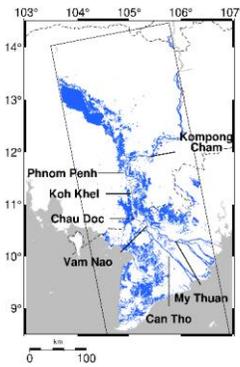
**August 15<sup>th</sup>, 2020**



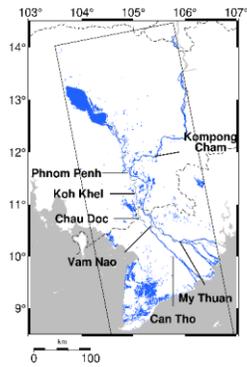
**September 15<sup>th</sup>, 2020**



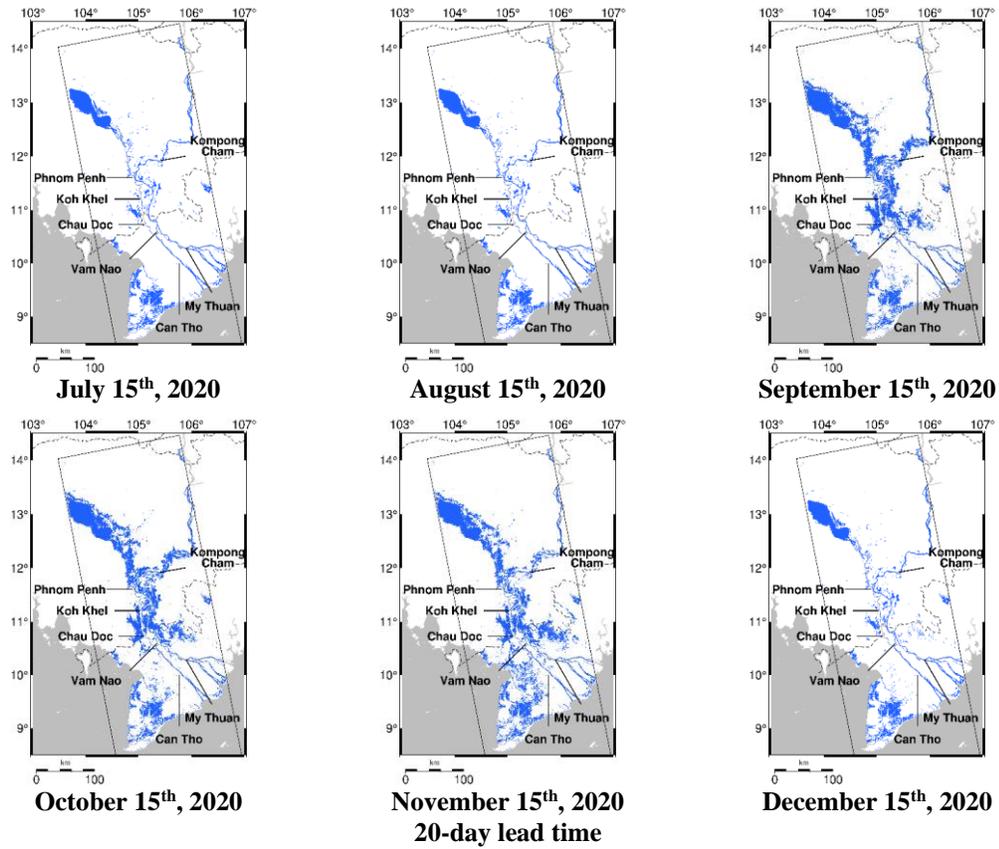
**October 15<sup>th</sup>, 2020**



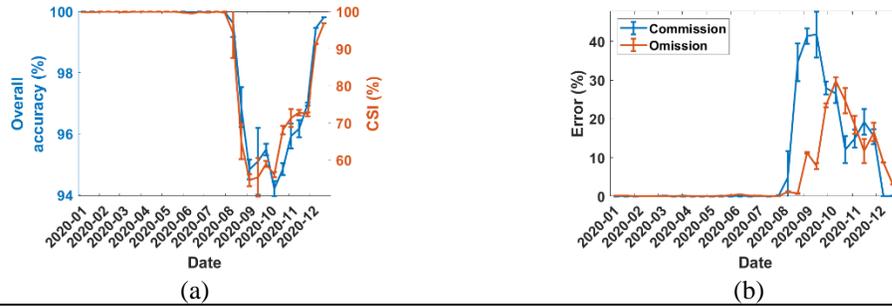
**November 15<sup>th</sup>, 2020  
15-day lead time**



**December 15<sup>th</sup>, 2020**



**Figure 5.22** FIER pseudo-forecasted inundation extents on the 15<sup>th</sup> of July to December, 2020.



**Figure 5.23** Averaged (a) overall accuracies, and CSIs, and (b) omission, and commission errors of FIER pseudo-forecasted inundation extents with 1-day to 20-day lead times.

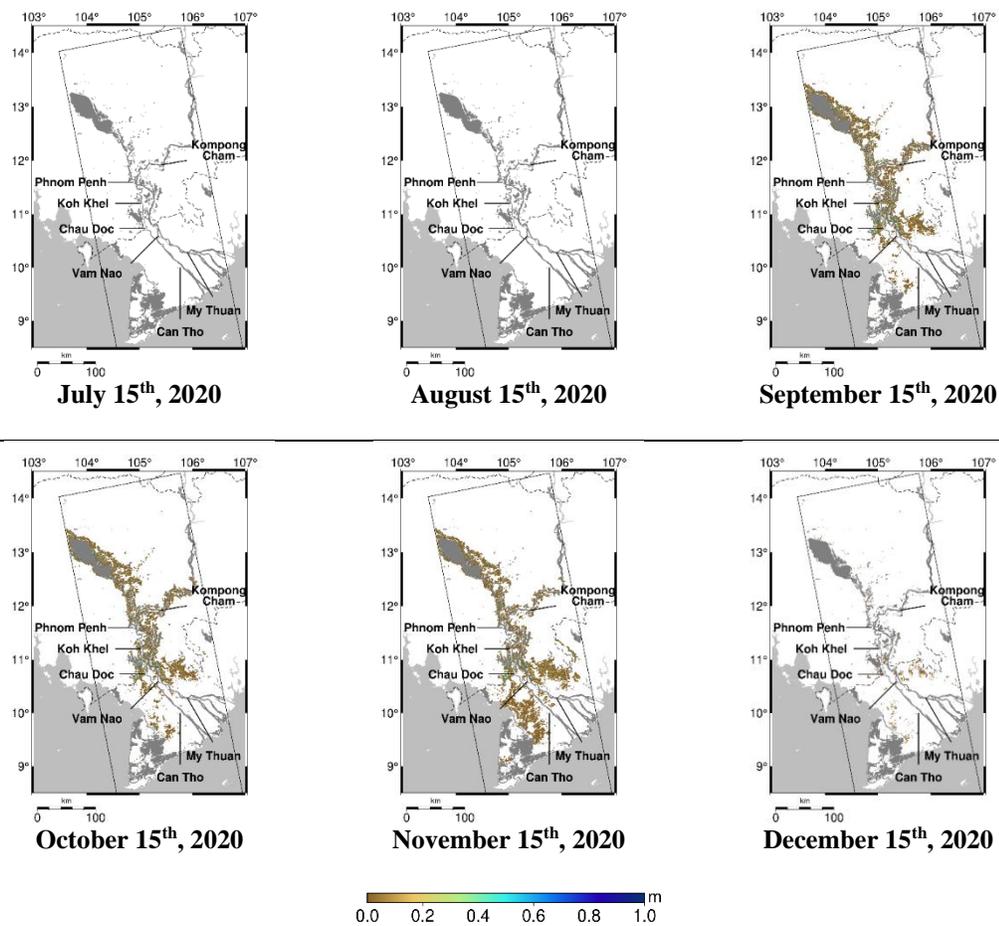
## 5.5 Application: Pseudo-forecasted Flood Risk in 2020

The capacity of FIER to forecast daily inundation extents can be applied to generate corresponding daily two-dimensional inundation depths that are commonly used for flood risk assessment. Since rice normally starts to die after being submerged

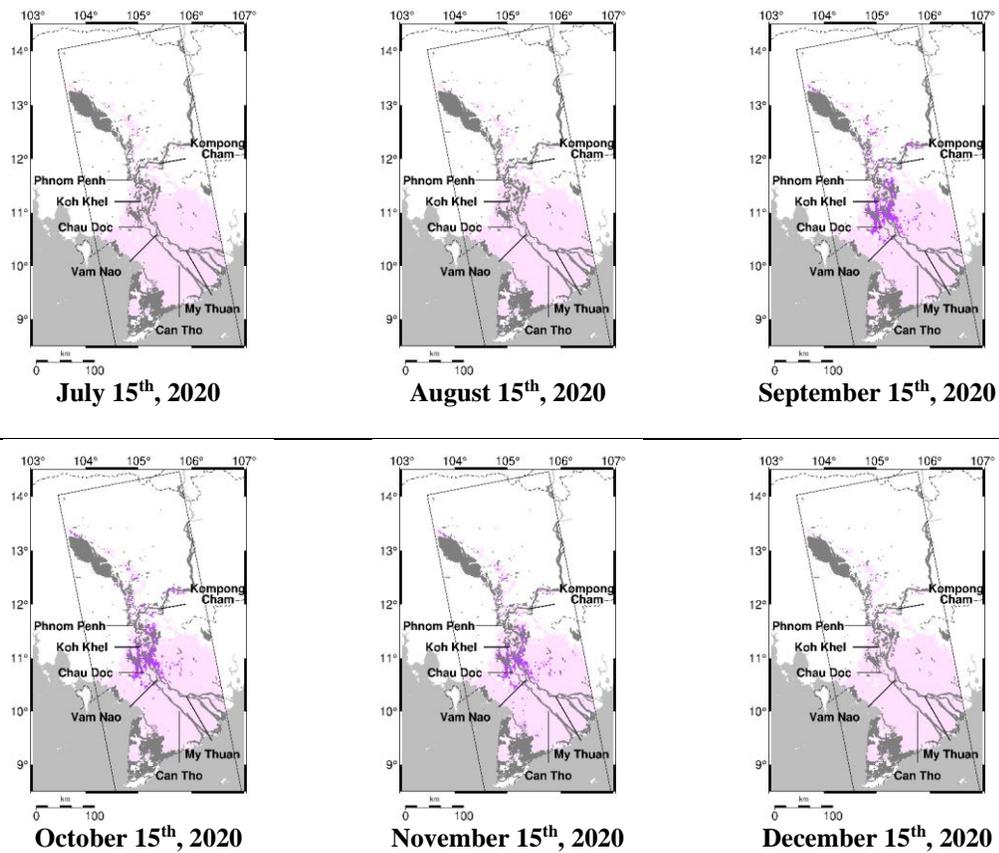
for 8 to 10 consecutive days (Pagano, 2014), the application of FIER forecasts for generating daily continuous inundation depth maps can help reveal where the rice paddies will be exposed to flood risk. Here, daily pseudo-forecasted inundation depths were generated using the FloodWater Depth Estimation Tool (FwDET) developed by Cohen et al. (2018, 2019) by integrating the daily FIER pseudo-forecasted inundation extents and the MERIT DEM. Figure 5.24 shows examples of 20-day lead time pseudo-forecasted inundation depths on the 15<sup>th</sup> of July to December 2020. We then followed Wassmann et al. (2019), defining rice paddies with flood risk as those that have been inundated with >0.4 m of water for 7 consecutive days. The pseudo-forecasted flood risk maps for rice cultivation were generated by combining the pseudo-forecasted inundation depth maps with 14 to 20-day lead time (7-day time span), showing the extents of rice paddies that could be under flood risk 20 days later.

Figure 5.25 shows examples of such pseudo-forecasted flood risk maps for rice paddies on the 15<sup>th</sup> of July to December 2020, the months in the critical risk period (Wassmann et al., 2019), in purple. Such maps can be generated on a daily basis so that the local government can frequently monitor when rice paddies will be threatened by floods. Figure 5.26 shows the start date (as Day-Of-Year, DOY) for pseudo-forecasted flood risk over rice paddies. It can be seen that based on the FIER pseudo-forecast, some rice paddies near the CF and upper MD could be exposed to flood risk at the end of August (DOY ~ 230). Since late August is when local farmers typically start to harvest (Triet et al., 2018), the spatial distribution of the rice paddies predicted to be under flood risk can serve as a reference to inform local stakeholders (1) where the inundation should be monitored with more caution, (2) where flood-preparedness resources (i.e.,

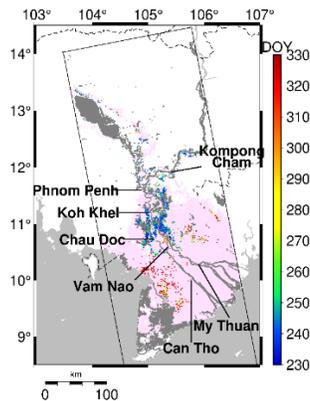
water pumps) should be allocated, and (3) where rice should be harvested with higher priority. Local stakeholders can then allocate flood-preparedness resources or issue a warning message to farmers for rice paddies at risk in advance to mitigate flood impact on rice productivity. In addition, Figure 5.26 also shows that some rice paddies in the lower MD could be under flood risk in November (DOY ~320). Since November is when local farmers typically start to sow, such information can help the local stakeholders decide where the sowing should be either postponed or switch to rice varieties with higher resiliency to floodwaters and then inform local farmers accordingly.



**Figure 5.24** 20-day lead time pseudo-forecasted inundation depths on the 15<sup>th</sup> of July to December. Grey areas are permanent water bodies.



**Figure 5.25** Rice paddies (light purple) in the Sentinel-1 frame with those pseudo-forecasted to be under flood risk on the 15<sup>th</sup> of July to December being marked in purple. Grey areas are permanent water bodies.



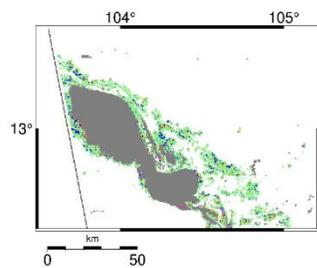
**Figure 5.26** Rice paddies (light purple) in the Sentinel-1 frame with the start DOY of pseudo-forecasted flood risk areas marked as colorbar shows. Grey areas are permanent water bodies.

Forecasted inundation depths derived from the FIER forecasts can also be used to predict the spatial distribution of flood hazard levels. We followed Balica et al. (2014), which classified flood hazards in the MD into five levels based on inundation depths (See Table 5.3 modified from Balica et al., 2014). Figure 5.27 is a detailed look at the 20-day lead time of pseudo-forecasted hazard levels over different areas in the MRB on the 15<sup>th</sup> of November 2020, the month that the FIER pseudo-forecast shows the most widespread inundation extents (See Figure 5.20.). While most areas have very low level of flood hazard (light green), some areas appeared to have low flood hazard (blue) where the movement of vehicles can be affected (See Table 5.3.). Areas with medium to very high hazard levels (yellow to purple) are sparsely distributed in the TSLF and MD, but widely distributed in the CF. Such flood hazard level maps can help decision-makers take necessary proactive actions, such as allocating essential resources (i.e., water pumps, and early reinforcement of first responders) or set up evacuation plans earlier if necessary so that flood damages can be mitigated. The applications of FIER forecasts are not restricted to the prevention of flood damage in rice cultivation but can also be more generally applied to multi-aspect damage prevention, and aiding decision-making for more comprehensive and efficient flood preparedness and early response.

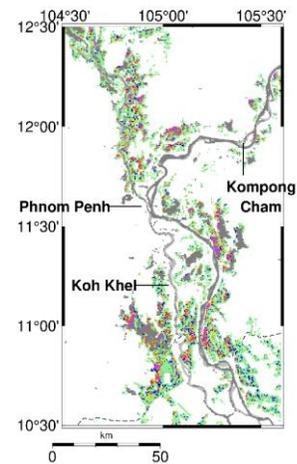
**Table 5.3** Definition of hazard levels categorized by inundation depths (modified from Balica et al., 2014).

<b>Inundation depth (m)</b>	<b>Hazard level</b>	<b>Definition</b>
0.0 – 0.2	Very low	<ul style="list-style-type: none"> <li>• Very low property damage</li> </ul>
0.2 – 0.5	Low	<ul style="list-style-type: none"> <li>• Insignificant flood-induced number of casualties</li> <li>• Low property damage</li> <li>• Affected vehicle movement but is safe walking through water</li> </ul>

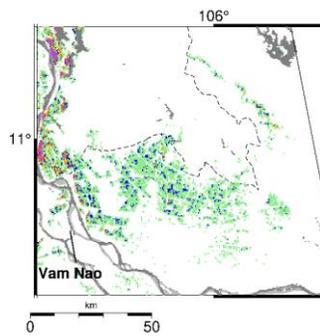
Table 5.3 (Continued)		
0.5 – 1.0	Medium	<ul style="list-style-type: none"> <li>• Considerable flood-induced casualties</li> <li>• High property damage</li> <li>• Not safe for vehicle movement and walking through water</li> </ul>
1.0 - 2.0	High	<ul style="list-style-type: none"> <li>• Extensive property damage</li> <li>• High probability of having dead and injured people</li> </ul>
> 2.0	Very high	<ul style="list-style-type: none"> <li>• Severe damages at all levels</li> <li>• Nothing is safe any longer</li> </ul>



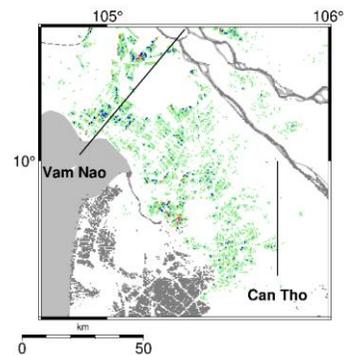
(a)



(b)



(c)



(d)



**Figure 5.27** Examples of 20-day pseudo-forecasted maps of flood hazard levels in November 2020 over the (a) TSLF, (b) CF, (c) northern MD, and (d) southern MD. Grey areas are permanent water bodies

## 5.6 Conclusions

MRB floodplain is a flat-terrain region with highly complex hydraulic conditions due to the impact of coexisting MR flows, flow reversal of TSL and tidal effects, that make the implementation of conventional inundation extent forecasting approaches, such as hydrodynamic modeling and GIS-based approaches, extremely challenging both in terms of computational burden and accuracy. In this study, we integrated FIER (Chang et al., 2020) and the daily water level forecasting system (Chang et al., 2019) that can more efficiently forecast daily inundation extent in the region.

The FIER-hindcasted inundation extents for 2020 were first generated with in-situ water levels at selected gauges as inputs. The skills were assessed with inundation extents derived from Sentinel-1A SAR images. To evaluate the skills of FIER hindcast under extreme flood events, FIER-hindcasted inundation extents for October 2011 were also generated and cross-compared with 8-day composite MODIS-derived inundation extents. Based on the cross-comparison, the FIER-hindcasted inundation extents can reach higher than 90% of accuracy and 60 to 70% of CSI during the 2020 wet season. The skills for the extreme 2011 flood event are compatible with those simulated by a quasi-two-dimensional hydraulic model (Triet et al., 2017, 2018). In the pseudo-forecasting case, up to 20-day lead time for FIER pseudo-forecasted inundation extents were generated and cross-compared with Sentinel-1 SAR image-derived inundation extents, showing higher than 90% accuracy and 50 – 70% CSI. One of the potential error sources for the FIER-estimated inundation extents can be the relatively lower

quality of Hydro-to-RTPC-03/-04 models which may be because: (1) the current FIER framework only considers dominant modes with strong seasonality that do not contain tidal influences with higher frequency; and (2) weaker hydraulic connection between the MD floodplain and the MR due to the intervention of human water controls. The dependence on synthesis of SAR-like images on input water levels as well as the adopted monthly water classification thresholds could also be another error source (See Section 5.4.1). In the case of FIER pseudo-forecasting, the skills can also be influenced by the accuracies of the pseudo-forecasted water levels.

It is expected that the quality of the Hydro-to-RTPC-03/-04 models and the skills of FIER can be further improved if later modes representing high-frequency tidal signals or flash floods can be identified and retained. On the other hand, if water levels in the MD floodplain are available as independent variables in the models, or more concrete information about how the MR levels can be linked to the MD floodplain water levels can be accessed, the skills of FIER could be improved as well. For the FIER forecast, note that FIER has the flexibility to take water levels forecasted by any system as inputs, meaning that it is possible to combine different water level forecasting systems that can be complementary (i.e., MRC's and Chang et al., 2019) in order to enhance the skill of inundation extent forecasting.

Lastly, the application of the FIER forecast for flood risk prediction was also demonstrated. By combining the FIER pseudo-forecasted inundation extents with a DEM, we were able to generate pseudo-forecasted inundation depths, which are commonly used for flood risk assessment. Note that an advantage of FIER is its capacity for generating daily forecasted inundation extents with daily forecasted water levels,

which allows forecasting of continuous daily inundation depths. Such continuous inundation depths are essential for assessing flood risk to rice cultivation. The forecasted inundation depths can also be used to generate forecasted flood hazard maps. With spatial information on the rice paddies exposed to future flood risks and future flood hazard levels, local stakeholders can make necessary decisions promptly and take proper proactive measures to have more efficient early response. Note that the accuracy of such applications can be influenced by not only the skill of FIER-estimated inundation extents but also the accuracy of the DEM. In addition, their spatial resolution can also be an influential factor, which can be addressed by the flexibility of FIER to be implemented on satellite images with different spatial resolutions. Although the pre-processing of satellite images (mosaic, co-registration, and RTC) with finer spatial resolution can take more time, such pre-processing is often a one-time-only event unless more severe floods, not observed by historical satellite images used for the REOF analysis, occur. The computational loading required for satellite imagery pre-processing can also be significantly reduced by leveraging recent cloud-based computation platforms such as Google Earth Engine and Google Colab.

The application of FIER, however, can be limited if historical floods were not observed by the satellite images due to their coarse temporal resolution. The implementation of FIER in this study used Sentinel-1A images with 12-day revisit times. Such a revisiting cycle may work fine in areas where the floods are more or less prolonged (i.e., wetlands that are seasonally flooded) but may not be temporally dense enough in areas where flash floods are more dominant. To address this limitation, a data fusion technique that merges optical imagery or multi-satellite SAR imagery should be

a topic of future study to generate a temporally longer and denser image stack with records of more flood signals.

## 6. CONCLUSION AND FUTURE WORK

### 6.1 Conclusion

Potential hydrology alteration that the downstream Mekong River Basin (MRB) may face, induced by climatic and anthropogenic changes, can influence the livelihoods of millions of local inhabitants. Currently, the Regional Flood Management and Mitigation Center (RFMMC) of the Mekong River Commission (MRC) does not publicly issue routine water level forecasting in the Mekong Delta (MD), except at Chau Doc and Tan Chau, located near the Cambodia-Vietnam national border. There is also no publicly and routinely available information about inundation extent forecasting. This could be due to: (1) the heavy computational burden and accuracy concerns of conventional approaches and (2) not very effective practical data exchange procedure between the MRC member countries (Anh et al., 2021). Hence, in this dissertation, one research question was introduced: How can we build skillful, computationally efficient, and sustainable flood forecasting systems of the Mekong River (MR) water levels and inundation extents for the MRB, with focus on the downstream areas?

To address this question, novel computationally efficient approaches for water level and inundation extent forecasting have been developed, that utilize remotely sensed observations to mitigate the geopolitical limitations on data sharing. In Chapter 3, a model-aided satellite altimetry-based daily water level forecasting system using simple regression analysis was proposed for the MRB. In the MD, where the RFMMC of MRC does not routinely issue comprehensive forecasting, ocean tides, which strongly impact water levels, were specifically addressed by a sum of sinusoidal functions. Forecasting skills for the system are promising in the MD. In contrast to the current

operating system, our system circumvents the need for frequent altimeter samplings in the upstream by using daily discharges estimated by the Variable Infiltration Capacity (VIC) model and is computationally efficient, without the need for complex hydrodynamic modeling.

In Chapter 4, a daily inundation extent estimation framework, named Forecasting Inundation Extents using Rotated empirical orthogonal function analysis (FIER), was proposed with the Tonle Sap Lake Floodplain (TSLF) as a test bed. By coupling spatiotemporal patterns of a multi-temporal Sentinel-1A Synthetic Aperture Radar (SAR) image stack extracted by Rotated Empirical Orthogonal Function (REOF) analysis with hydrological data, FIER is able to synthesize SAR-like images at any time when hydrological data is available. Inundation extents can then be generated through a water classification method. The estimated inundation extents are quite promising in the cases of hindcast and forecast. The advantage of FIER in generating inundation extents of any time allows retrospective and prospective studies on the continuous dynamics and impacts of inundation. FIER also has the potential to be implemented on imagery acquired by different satellites, over different regions, or at different spatial resolutions, as long as the connection between the temporal patterns of satellite imagery and hydrological data can be built.

In Chapter 5, the implementation of FIER was expanded to the whole MRB floodplain, encompassing the TSLF, Cambodian Floodplain (CF) and the MD where conventional inundation forecasting approaches are difficult to apply, by using a multi-temporal stack of Sentinel-1A images and multiple gauges for MR levels. Both hindcast and pseudo-forecasts were performed, that use historical in-situ and pseudo-forecasted

water levels (Chang et al., 2019) as inputs, respectively. Based on cross-comparison, FIER hindcast and pseudo-forecast both have fairly good skills. The hindcast skill for the extreme October 2011 flood is compatible with that simulated by the quasi-two-dimensional hydraulic model over the MD (Triet et al., 2017, 2018). FIER's capacity for quickly estimating inundation extents is advantageous when practically performing continuous areal inundation forecast, which allows instantaneous spatial prediction of flood hazard levels and flood risk for rice cultivation.

To conclude this dissertation,

- We have developed skillful, computationally efficient, and sustainable flood forecasting systems that address the challenges in the MRB Floodplain by using remotely sensed data (satellite altimetry, and satellite imagery), and low computational cost approaches (regression analysis, principal component analysis).
- By combining the water level forecasting system, FIER, and DEM allows spatial prediction of flood risk areas and flood hazard levels which can help local stakeholders and farmers more smartly manage rice cultivation activities, preventing and lifting the threat of flood to rice productivity. Such applications are extremely crucial for the MRB floodplain where is heavily populated with highly rice-dependent food intake and economy.
- FIER has potential to be applied to other flood-prone areas in the world. A freely accessible python package, named FIERpy, has been developed by our collaborator in NASA (Kel, 2021), which can also be implemented on cloud-computing-based platform, Google Colab. With this, FIER now can be easily scaled up to different areas of interest with enhanced accessibility.

Despite the advantages of FIER, its ability to reveal the spatiotemporal flood signals depends on whether the signals can be observed and recorded in the satellite imagery used. This means its skill is influenced by the electromagnetic characteristics of spectral bands, such as canopy penetrating capacity and the spatial resolution of the images used. Hence, proper satellite imagery and resolution should be chosen depending on the area of interest. In addition, whether the synthesized images are sufficient enough to reconstruct flood-related signals also influences the skill of FIER. This means the REOF modes retained, and the build of models, connecting hydrological data and temporal patterns of satellite imagery, should be carefully investigated. Water classification applied also affects prediction accuracy. Lastly, whether available input hydrological data represents what is really happening over the floodplains well can also be an influential factor for the skills of the models and the final estimated inundation extents. The MD, where floodplain water has weaker natural hydraulic connectivity with the MR due to intensive human water control, is an example of weak correlation. However, the FIER implementation in the TSLF and the whole MRB floodplain still shows promising skill and provides opportunities for valuable applications in fast and continuous spatial flood hazard and flood risk prediction. It is expected that FIER will greatly benefit local stakeholders with a more efficient decision-making process for better flood damage mitigation.

## 6.2 Future Works

Chapter 4 and Chapter 5 of this dissertation demonstrated the feasibility of FIER over the TSLF and the MRB floodplain using Sentinel-1 SAR imagery, respectively.

However, the implementation of FIER is not restricted to a specific geographical region, satellite imagery source and spatial resolution. Ideally, as long as skillful models between patterns in multi-temporal satellite images and hydrological data can be established, FIER should work. This gives FIER the potential to be applied globally in areas with different landcover types. Table 6.1 shows a list of implementations that have been tested so far.

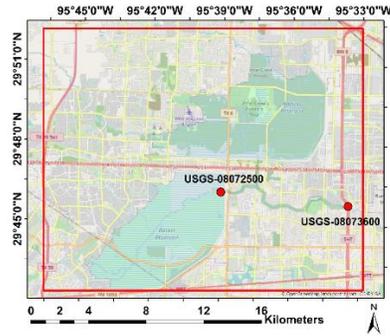
**Table 6.1** Implementation of FIER that have been tested in areas other than the MRB.

Satellite	Sensor type	Spatial resolution	Study area
Sentinel-1A/-1B	SAR	20 m	Houston
Visible Infrared Imaging Radiometer Suite (VIIRS)	Optical	10 arc-second	Mississippi River near New Madrid, Missouri

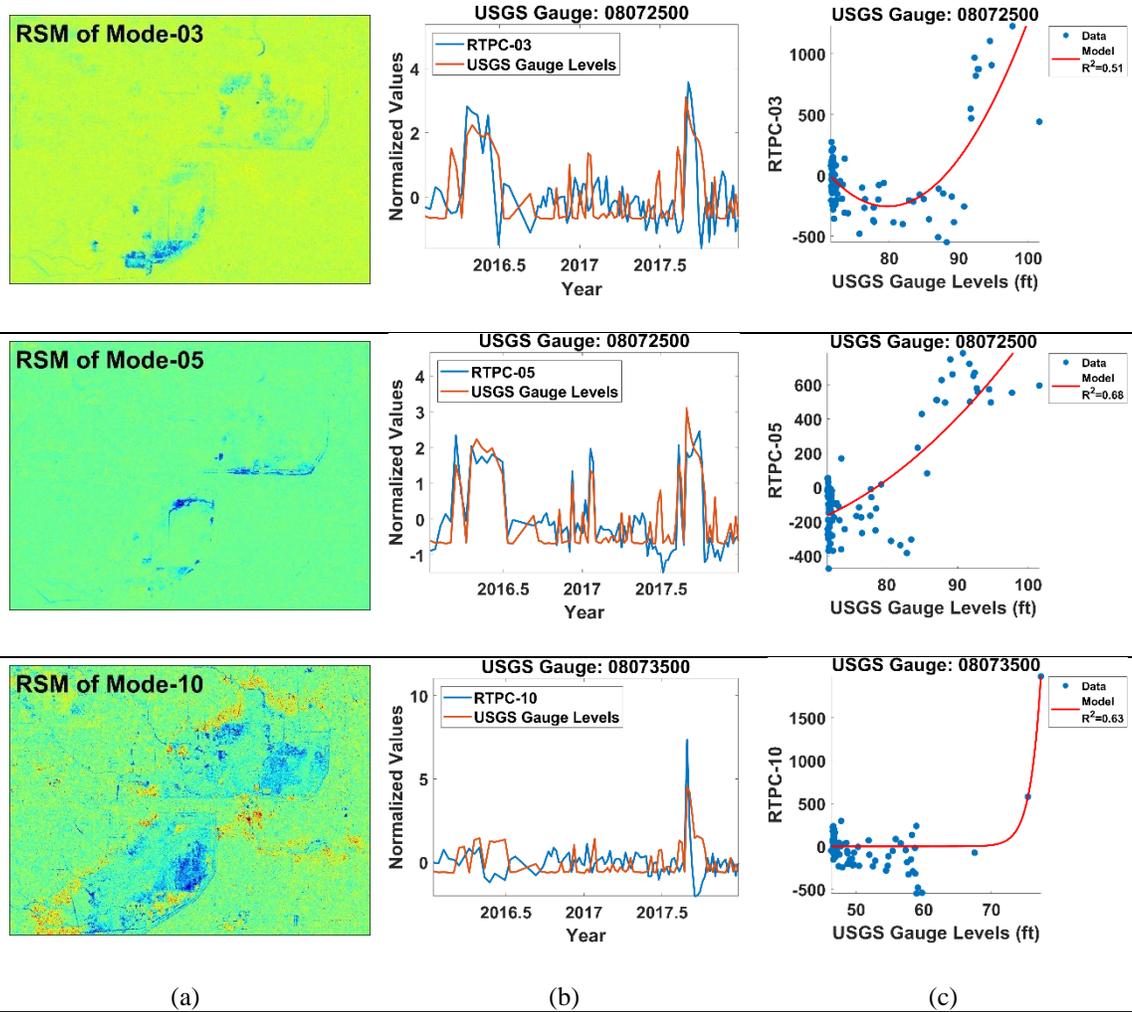
For testing the implementation over Houston, Sentinel-1A/-1B SAR imagery (Path 143, Frame 492) acquired within 2016 – 2017 were used. Figure 6.1 shows the test area in the red box where in-situ gauges are marked by red dots. Considering the distribution of floodwater in urban areas would require finer resolution for detection, therefore, images were multi-looked and pre-processed to 20-m resolution. Unlike the cases over the TSLF and the MRB floodplain, where seasonal flooding is the dominant signal, Houston has more flash floods caused by thunderstorms or tropical systems. The REOF analysis therefore decomposed flood-related signals to the later modes. Figure 6.2(a) shows the identified flood-related spatial patterns (Rotated Spatial Mode, RSMs) for SAR images. Figure 6.2(b) shows the normalized temporal patterns (Rotated

Temporal Principal Components, RTPCs) of SAR images and in-situ water levels. Figure 6.2(c) gives the corresponding scatter plots along with regression models. Preliminary hindcasted inundation extents along the Buffalo Bayou near the strike of Hurricane Harvey in late August 2017 are shown in Figure 6.3.

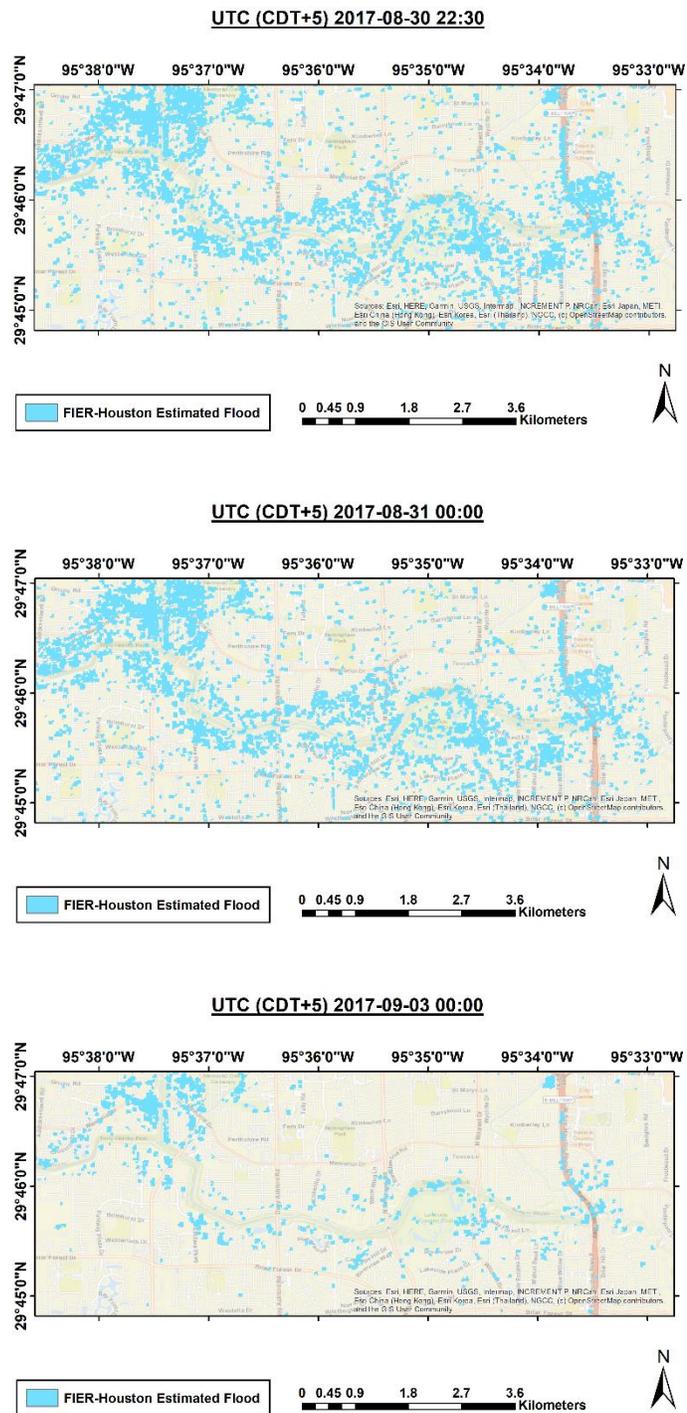
Note that since there are more flash floods in this area, the 6-day revisit time of Sentinel-1A/-1B constellation may not be temporally fine enough to record flash flood events. Consequently, only a limited number of data points in the scatter plot reflect the flooding, such as the case of mode-10 in Figure 6.2(c), which corresponds to the extreme floods caused by Hurricane Harvey. Recently, an open-source Python application for downloading, processing, and delivering surface water maps derived from remote sensing data, called the Hydrological Remote Sensing Analysis for Floods (HYDRAFloods), has been proposed. HYDRAFloods is built using Google Earth Engine and Google Cloud Platform and can be run on Google Colab, which leverages cloud computing for large-scale computations and handling high data volume outputs. It also provides machine learning modules for data fusion, which allows conversion of SAR indices into optical Modified Normalized Difference Water Index (MNDWI) (Haag et al., 2021). The combination of optical and SAR imagery can help generate a multi-temporal image stack with a denser temporal sampling rate and longer time span. With this, it is expected that more flood events can be observed and contained in the image stack where REOF analysis then extracts spatiotemporal patterns. This can also be beneficial for generating better linkages between hydrological data and RTPCs. Cloud-computing platforms can also alleviate the relatively longer computational time for pre-processing high-resolution SAR imagery.



**Figure 6.1** Area over the metropolitan area of Houston that FIER implementation has been tested.

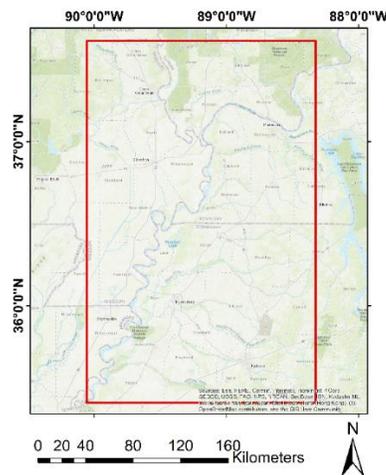


**Figure 6.2** Water-related (a) RSMs, (b) normalized RTPCs and in-situ water levels, and (c) scatter plots between RTPCs and in-situ water levels over the metropolitan areas of Houston.

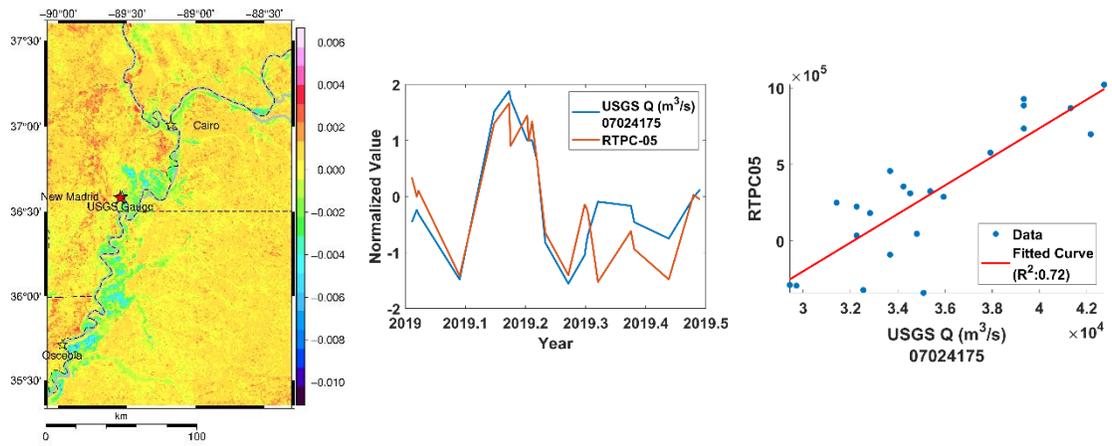
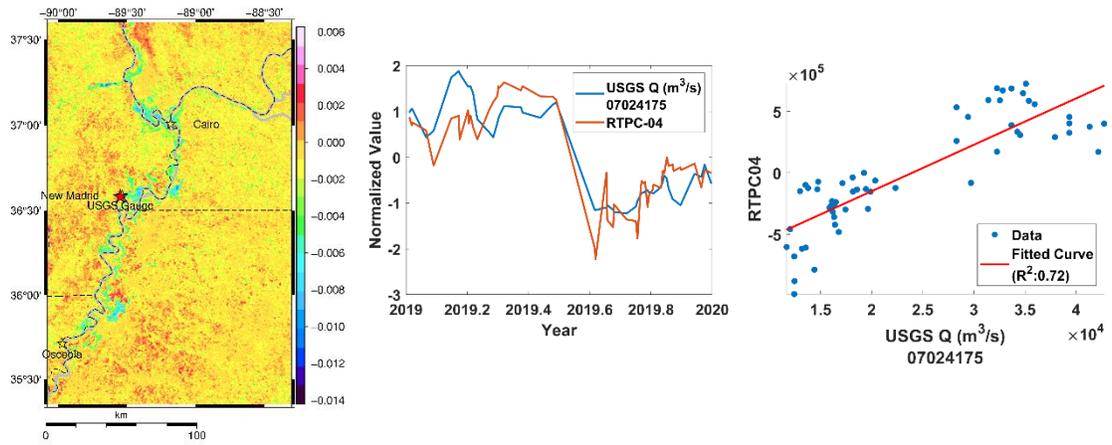


**Figure 6.3** Preliminary hindcast inundated extents along Buffalo Bayou in Houston in 2017 on the dates near the strike of Hurricane Harvey.

Implementation of FIER with the optical imagery has also been tested by using Visible Infrared Imaging Radiometer Suite (VIIRS) surface reflectance images. The test bed is the Mississippi River near New Madrid, Missouri (See Figure 6.4). The short-Wave InfraRed (SWIR) band is less affected by concentrations of sediments and other optical active constituents within the water, and therefore is a more stable indicator of inundated areas (Huang et al., 2018; Li et al., 2013), SWIR surface reflectance images with less cloud cover were selected to form a multi-temporal image stack for REOF analysis to extract spatiotemporal patterns. Figure 6.5 shows the identified water-related RSMs, normalized RTPCs and in-situ water levels as well as corresponding scatter plots and regression models. Preliminary hindcasted inundation extents are shown in Figure 6.6. It is worth mentioning that by implementing FIER on the selected less cloud-affected optical images, FIER can generate optical imagery-based inundation extents that are both temporally dense and with less cloud cover issue, which has previously been a major drawback of using optical imagery for continuous inundation mapping.



**Figure 6.4** Area over the Mississippi River near New Madrid, Missouri, that FIER implementation was tested.

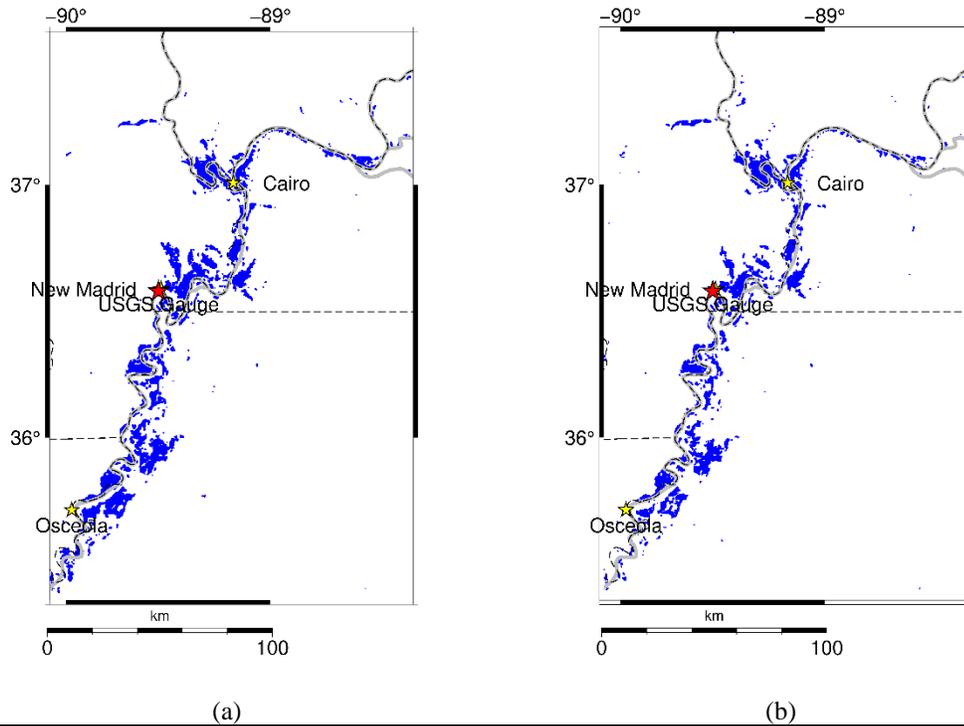


(a)

(b)

(c)

**Figure 6.5** Water-related (a) RSMs, (b) normalized RTPCs and in-situ water levels, and (c) scatter plots between RTPCs and in-situ water levels over the Mississippi River near New Madrid, Missouri.



**Figure 6.6** Preliminary hindcast inundation extents on (a) March 12<sup>th</sup>, and (b) March 25<sup>th</sup>, 2018.

To sum up, our future work will focus on directions of:

- Expanding the implementation of FIER over different geographical areas, including other wetlands or even urban areas that are flood-prone. For implementation in other wetlands, dominant local vegetation type has to be considered to properly select satellite imagery considering both canopy penetrating capacity and spatial resolution.
- Implementation of FIER for optical imagery, which has advantages of more frequent temporal sampling and more spectral information than SAR imagery. The more frequent temporal sampling enhances the chance that flood events can be captured. On the other hand, the ability to record more spectral information allows band-math to generate a water-sensitive index, such as MNDWI, which

might be beneficial when extracting flood-related spatiotemporal patterns using REOF analysis.

- Exploiting cloud-based database and computation platforms, such as Google Earth Engine, which is freely accessible and has archived remote sensing data, and Google Colab to alleviate processing time of satellite imagery pre-processing or data fusion.

## REFERENCES

- Ablain, M., Philipps, S., Picot, N., Bronner, E., 2010. “Jason-2 global statistical assessment and cross-calibration with Jason-1.” *Mar. Geod.* 33, 162–185.  
<https://doi.org/10.1080/01490419.2010.487805>
- Adamson, P., 2006. “Hydrological and water resources modelling in the Mekong region: A brief overview.” in: Explore Water Futures Together: Mekong Region Waters Dialogue, IUCN, TEI, IWMI, and M-POWER, pp. 69–74.
- Amitrano, D., Di Martino, G., Iodice, A., Riccio, D., Ruello, G., 2018. “Unsupervised rapid flood mapping using Sentinel-1 GRD SAR images.” *IEEE Trans. Geosci. Remote Sens.* 56, 3290–3299. <https://doi.org/10.1109/TGRS.2018.2797536>
- Anh, N.T.K., 2021. “Data exchange as key mechanism in the joint use of the Mekong River.” in: Modern Global Economic System: Evolutional Development vs. Revolutionary Leap, Popkova, E.G., Sergi, B.S. (Eds.), Springer International Publishing, Cham, pp. 990–996. [https://doi.org/10.1007/978-3-030-69415-9\\_110](https://doi.org/10.1007/978-3-030-69415-9_110)
- Arias, M.E., Cochrane, T.A., Piman, T., Kummu, M., Caruso, B.S., Killeen, T.J., 2012. “Quantifying changes in flooding and habitats in the Tonle Sap Lake (Cambodia) caused by water infrastructure development and climate change in the Mekong Basin.” *J. Environ. Manage.* 112, 53–66.  
<https://doi.org/https://doi.org/10.1016/j.jenvman.2012.07.003>
- Arthur, D., Vassilvitskii, S., 2007. “K-means++: The advantages of careful seeding.” in: Proceeding of the Annual ACM-SIAM Symposium on Discrete Algorithms.  
<https://doi.org/10.1145/1283383.1283494>
- AVISO+, <https://www.aviso.altimetry.fr/en/applications/hydrology-and-land.html>

(Accessed on June 22nd, 2021)

AVISO+, <https://www.aviso.altimetry.fr/en/techniques/altimetry/principle/pulses-and-waveforms.html> (Accessed on June 22nd, 2021)

Balica, S., Dinh, Q., Popescu, I., Vo, T.Q., Pham, D.Q., 2014. “Flood impact in the Mekong Delta, Vietnam.” *J. Maps* 10, 257–268. <https://doi.org/10.1080/17445647.2013.859636>

Bamber, J.L., 1994. “Ice sheet altimeter processing scheme.” *Int. J. Remote Sens.* 15, 925–938. <https://doi.org/10.1080/01431169408954125>

Banks, J.C., Camp, J. V., Abkowitz, M.D., 2014. “Adaptation planning for floods: A review of available tools.” *Nat. Hazards* 70, 1327–1337. <https://doi.org/10.1007/s11069-013-0876-7>

Bates, P.D., Pappenberger, F., Romanowicz, R.J., 2014. “Uncertainty in flood inundation modelling,” in: Beven, K.J., Hall, J. (Eds.), Applied Uncertainty Analysis for Flood Risk Management, Imperial College Press, pp. 232–269.

Bates, P.D., De Roo, A.P.J., 2000. “A simple raster-based model for flood inundation simulation.” *J. Hydrol.* 236, 54–77. [https://doi.org/10.1016/S0022-1694\(00\)00278-X](https://doi.org/10.1016/S0022-1694(00)00278-X)

Biancamaria, S., Frappart, F., Leleu, A.S., Marieu, V., Blumstein, D., Desjonquères, J.D., Boy, F., Sottolichio, A., Valle-Levinson, A., 2017. “Satellite radar altimetry water elevations performance over a 200 m wide river: Evaluation over the Garonne River.” *Adv. Sp. Res.* 59, 128–146. <https://doi.org/10.1016/j.asr.2016.10.008>

Biancamaria, S., Hossain, F., Lettenmaier, D.P., 2011. “Forecasting transboundary river

- water elevations from space.” *Geophys. Res. Lett.* 38, 1–5.  
<https://doi.org/10.1029/2011GL047290>
- Bioresita, F., Puissant, A., Stumpf, A., Malet, J.P., 2018. “A method for automatic and rapid mapping of water surfaces from Sentinel-1 imagery.” *Remote Sens.* 10.  
<https://doi.org/10.3390/rs10020217>
- Boergens, E., Dettmering, D., Seitz, F., 2019. “Observing water level extremes in the Mekong River Basin: The benefit of long-repeat orbit missions in a multi-mission satellite altimetry approach.” *J. Hydrol.* 570, 463–472.  
<https://doi.org/10.1016/j.jhydrol.2018.12.041>
- Bouvet, A., Le Toan, T., Lam-Dao, N., 2009. “Monitoring of the rice cropping system in the Mekong Delta using ENVISAT/ASAR dual polarization data.” *IEEE Trans. Geosci. Remote Sens.* 47, 517–526. <https://doi.org/10.1109/TGRS.2008.2007963>
- Bracher, A., Taylor, M.H., Taylor, B., Dinter, T., Röttgers, R., Steinmetz, F., 2015. “Using empirical orthogonal functions derived from remote-sensing reflectance for the prediction of phytoplankton pigment concentrations.” *Ocean Sci.* 11, 139–158.  
<https://doi.org/10.5194/os-11-139-2015>
- Brêda, J.P.L.F., Paiva, R.C.D., Bravo, J.M., Passaia, O.A., Moreira, D.M., 2019. “Assimilation of Satellite Altimetry Data for Effective River Bathymetry.” *Water Resour. Res.* <https://doi.org/10.1029/2018wr024010>
- Busker, T., De Roo, A., Gelati, E., Schwatke, C., Adamovic, M., Bisselink, B., Pekel, J.F., Cottam, A., 2019. “A global lake and reservoir volume analysis using a surface water dataset and satellite altimetry.” *Hydrol. Earth Syst. Sci.* 23, 669–690.  
<https://doi.org/10.5194/hess-23-669-2019>

- Calmant, S., Crétaux, J.F., Rémy, F., 2016. “Principles of radar satellite altimetry for application on inland waters.” in: Microwave Remote Sensing Land Surfaces, Baghdadi, N., and Zribi, M., (Eds.), pp. 175–218. <https://doi.org/10.1016/B978-1-78548-159-8.50004-9>
- Campbell, I.C., Say, S., Beardall, J., 2009. “Tonle Sap Lake, the heart of the Lower Mekong,” in: The Mekong: Biophysical Environment of an International River Basin, Campbell, I.C. (Ed.), pp. 251–272. <https://doi.org/10.1016/B978-0-12-374026-7.00010-3>
- Celik, T., 2009. “Unsupervised change detection in satellite images using principal component analysis and K-means clustering.” *IEEE Geosci. Remote Sens. Lett.* 6, 772–776. <https://doi.org/10.1109/LGRS.2009.2025059>
- Center for Excellence in Disaster Management and Humanitarian Assistance, 2017. Cambodia Disaster Management Reference Handbook. Center for Excellence in Disaster Management and Humanitarian Assistance, Joint Base Pearl Harbor – Hickam, Hawaii, U.S.
- Chang, C.-H., Lee, H., Hossain, F., Basnayake, S., Jayasinghe, S., Chishtie, F., Saah, D., Yu, H., Sothea, K., Du Bui, D., 2019. “A model-aided satellite-altimetry-based flood forecasting system for the Mekong River.” *Environ. Model. Softw.* 112, 112–127. <https://doi.org/https://doi.org/10.1016/j.envsoft.2018.11.017>
- Chang, C.-H., Lee, H., Kim, D., Hwang, E., Hossain, F., Chishtie, F., Jayasinghe, S., Basnayake, S., 2020. “Hindcast and forecast of daily inundation extents using satellite SAR and altimetry data with rotated empirical orthogonal function analysis: Case study in Tonle Sap Lake Floodplain.” *Remote Sens. Environ.* 241,

111732. <https://doi.org/https://doi.org/10.1016/j.rse.2020.111732>

Chang, M.J., Chang, H.K., Chen, Y.C., Lin, G.F., Chen, P.A., Lai, J.S., Tan, Y.C., 2018.

“A support vector machine forecasting model for typhoon flood inundation mapping and early flood warning systems.” *Water* 10.

<https://doi.org/10.3390/w10121734>

Chen, Z., Luo, J., Chen, N., Xu, R., Shen, G., 2019. “RFim: A real-time inundation

extent model for large floodplains based on remote sensing big data and water level observations.” *Remote Sens.* 11. <https://doi.org/10.3390/rs11131585>

Cheng, X., Nitsche, G., Wallace, J.M., 1995. “Robustness of low-frequency circulation

patterns derived from EOF and rotated EOF Analyses.” *J. Clim.* 8, 1709–1713.

[https://doi.org/10.1175/1520-0442\(1995\)008<1709:ROLFCP>2.0.CO;2](https://doi.org/10.1175/1520-0442(1995)008<1709:ROLFCP>2.0.CO;2)

Church, J.A., White, N.J., Coleman, R., Lambeck, K., Mitrovica, J.X., 2004. “Estimates

of the regional distribution of sea level rise over the 1950–2000 period.” *J. Clim.*

17, 2609c2625. [https://doi.org/10.1175/1520-](https://doi.org/10.1175/1520-0442(2004)017<2609:EOTRDO>2.0.CO;2)

[0442\(2004\)017<2609:EOTRDO>2.0.CO;2](https://doi.org/10.1175/1520-0442(2004)017<2609:EOTRDO>2.0.CO;2)

Clement, M.A., Kilsby, C.G., Moore, P., 2018. “Multi-temporal synthetic aperture radar

flood mapping using change detection.” *J. Flood Risk Manag.* 11, 152–168.

<https://doi.org/https://doi.org/10.1111/jfr3.12303>

Cohen, S., Brakenridge, G.R., Kettner, A., Bates, B., Nelson, J., McDonald, R., Huang,

Y.-F., Munasinghe, D., Zhang, J., 2018. “Estimating floodwater depths from flood inundation maps and topography.” *J. Am. Water Resour. Assoc.*, 54, 847–858.

<https://doi.org/10.1111/1752-1688.12609>

Cohen, S., Raney, A., Munasinghe, D., Loftis, J.D., Molthan, A., Bell, J., Rogers, L.,

- Galantowicz, J., Brakenridge, G.R., Kettner, A.J., Huang, Y.-F., Tsang, Y.-P., 2019. “The floodwater depth estimation tool (FwDET v2.0) for improved remote sensing analysis of coastal flooding.” *Nat. Hazards Earth Syst. Sci.*, 19, 2053–2065. <https://doi.org/10.5194/nhess-19-2053-2019>
- Cox, R., Bauer, B.L., Smith, T., 1998. “A mesoscale model intercomparison.” *Bull. Amer. Met. Soc.* 79, 265–283.
- Curlander, J. C., McDonough, R. N., 1991. Synthetic Aperture Radar: Systems and Signal Processing, John Wiley & Sons., New York, NY, U.S.
- Da Silva, J.S., Seyler, F., Calmant, S., Rotunno Filho, O.C., Roux, E., Araújo, A.A.M., Guyot, J.L., 2012. “Water level dynamics of Amazon wetlands at the watershed scale by satellite altimetry.” *Int. J. Remote Sens.* 33, 3323–3353. <https://doi.org/10.1080/01431161.2010.531914>
- Dang, T.D., Cochrane, T.A., Arias, M.E., Tri, V.P.D., 2018. “Future hydrological alterations in the Mekong Delta under the impact of water resources development, land subsidence and sea level rise.” *J. Hydrol. Reg. Stud.* 15, 119–133. <https://doi.org/10.1016/j.ejrh.2017.12.002>
- Dinh, D.A., Elmahrad, B., Leinenkugel, P., Newton, A., 2019. “Time series of flood mapping in the Mekong Delta using high resolution satellite images.” *IOP Conf. Ser. Earth Environ. Sci.* 266, 12011. <https://doi.org/10.1088/1755-1315/266/1/012011>
- Dommenget, D., Latif, M., 2002. “A cautionary note on the interpretation of EOFs.” *J. Clim.* 15, 216–225. [https://doi.org/10.1175/1520-0442\(2002\)015<0216:ACNOTI>2.0.CO;2](https://doi.org/10.1175/1520-0442(2002)015<0216:ACNOTI>2.0.CO;2)

- Du, T.L.T., Lee, H., Bui, D.D., Arheimer, B., Li, H.-Y., Olsson, J., Darby, S.E., Sheffield, J., Kim, D., Hwang, E., 2020. “Streamflow prediction in geopolitically ungauged basins using satellite observations and regionalization at subcontinental scale.” *J. Hydrol.* 588, 125016. <https://doi.org/https://doi.org/10.1016/j.jhydrol.2020.125016>
- Egbert, G.D., Ray, R.D., 2003. “Deviation of long-period tides from equilibrium: Kinematics and geostrophy.” *J. Phys. Oceanogr.* 33, 822–839. [https://doi.org/10.1175/1520-0485\(2003\)33<822:doltfe>2.0.co;2](https://doi.org/10.1175/1520-0485(2003)33<822:doltfe>2.0.co;2)
- Fisheries Administration of Cambodia, 2011. Status of the Fishery Sector in 2011 and Targets for 2012. Fisheries Administration, Phnom Penh, Cambodia.
- Fok, H.S., He, Q., Chun, K.P., Zhou, Z., Chu, T., 2018. “Application of ENSO and drought indices for water level reconstruction and prediction: A case study in the lower Mekong River estuary.” *Water* 10. <https://doi.org/10.3390/w10010058>
- Frappart, F., Biancamaria, S., Normandin, C., Blarel, F., Bourrel, L., Aumont, M., Azemar, P., Vu, P.L., Le Toan, T., Lubac, B., Darrozes, J., 2018. “Influence of recent climatic events on the surface water storage of the Tonle Sap Lake.” *Sci. Total Environ.* 636, 1520–1533. <https://doi.org/10.1016/j.scitotenv.2018.04.326>
- Frappart, F., Do Minh, K., L’Hermitte, J., Cazenave, A., Ramillien, G., Le Toan, T., Mognard-Campbell, N., 2006. “Water volume change in the lower Mekong from satellite altimetry and imagery data.” *Geophys. J. Int.* 167, 570–584. <https://doi.org/10.1111/j.1365-246X.2006.03184.x>
- Fredén, F., 2011. Impacts of dams on lowland agriculture in the Mekong River catchment. Lunds Universitets Naturgeografiska Institution-Seminarieuppsatser,

Lund, Sweden.

Gerlak, A.K., Lautze, J., Giordano, M., 2011. “Water resources data and information exchange in transboundary water treaties.” *Int. Environ. Agreements Polit. Law Econ.* 11, 179–199. <https://doi.org/10.1007/s10784-010-9144-4>

Gilbert, G. K., 1884. “Finley’s tornado predictions.” *Am. Meteorol. J.* 1, 166-172.

Gumma, M.K., Thenkabail, P.S., Maunahan, A., Islam, S., Nelson, A., 2014. “Mapping seasonal rice cropland extent and area in the high cropping intensity environment of Bangladesh using MODIS 500m data for the year 2010.” *ISPRS J. Photogramm. Remote Sens.* 91, 98–113. <https://doi.org/10.1016/j.isprsjprs.2014.02.007>

Haag, A., Markert, K., Markert, A., Mayer, T., Chisthie, F., Poortinga, A., Saah, D., Towashiraporn, P., Meechaiya, C., Thwal, N.S., Benito Lazaro, I., Nicolau, A.P., Bhandari, B., Wadhwa, A., Sundaram, J., Burja, K., 2021. “Current status and future developments of HYDRAFloods: Operational flood monitoring in Southeast Asia.” The 4<sup>th</sup> Hydrospace-GEOGloWS, Hosted as a virtual event from ESA-ESRIN, Frascati, Rome, Italy.

Hamman, J.J., Nijssen, B., Bohn, T.J., Gergel, D.R., Mao, Y., 2018. “The Variable Infiltration Capacity model version 5 (VIC-5): infrastructure improvements for new applications and reproducibility.” *Geosci. Model Dev.* 11, 3481–3496. <https://doi.org/10.5194/gmd-11-3481-2018>

Hannachi, A., Jolliffe, I.T., Stephenson, D.B., 2007. “Empirical orthogonal functions and related techniques in atmospheric science: A review.” *Int. J. Climatol.* 27, 1119–1152. <https://doi.org/10.1002/joc.1499>

Hannachi, A., Jolliffe, I.T., Stephenson, D.B., Trendafilov, N., 2006. “In search of

- simple structures in climate: simplifying EOFs.” *Int. J. Climatol.* 26, 7–28.  
<https://doi.org/10.1002/joc.1243>
- Hannachi, A., Jolliffe, I.T., Stephenson, D.B., Trendafilov, N., 2006. “In search of simple structures in climate: Simplifying EOFs.” *Int. J. Climatol.* 26, 7–28.  
<https://doi.org/10.1002/joc.1243>
- Hansen, M.C., Potapov, P. V, Moore, R., Hancher, M., Turubanova, S.A., Tyukavina, A., Thau, D., Stehman, S. V, Goetz, S.J., Loveland, T.R., Kommareddy, A., Egorov, A., Chini, L., Justice, C.O., Townshend, J.R.G., 2013. “High-resolution global maps of 21st-century forest cover change.” *Science* 342, 850 – 853.  
<https://doi.org/10.1126/science.1244693>
- Henry, J. -B., Chastanet, P., Fellah, K., Desnos, Y. -L., 2006. “Envisat multi-polarized ASAR data for flood mapping.” *Int. J. Remote Sens.* 27, 1921–1929.  
<https://doi.org/10.1080/01431160500486724>
- Hiep, N.H., Luong, N.D., Viet Nga, T.T., Hieu, B.T., Thuy Ha, U.T., Du Duong, B., Long, V.D., Hossain, F., Lee, H., 2018. “Hydrological model using ground- and satellite-based data for river flow simulation towards supporting water resource management in the Red River Basin, Vietnam.” *J. Environ. Manage.* 217, 346–355. <https://doi.org/https://doi.org/10.1016/j.jenvman.2018.03.100>
- Hoang, L.P., Lauri, H., Kummu, M., Koponen, J., van Vliet, M.T.H., Supit, I., Leemans, R., Kabat, P., Ludwig, F., 2016. “Mekong River flow and hydrological extremes under climate change.” *Hydrol. Earth Syst. Sci.* 20, 3027–3041.  
<https://doi.org/10.5194/hess-20-3027-2016>
- Hoanh, C.T., Jirayoot, K., Lacomme, G., Srunetr, V., 2010. “Impacts of climate change

and development on Mekong flow regimes. First assessment – 2009,” Mekong River Commission. Vientiane.

Hortle, K. G., Lieng, S., Valbo-Jorgensen, J., 2004. An Introduction to Cambodia’s Inland Fisheries, Mekong Development Series No. 4. Mekong River Commission, Phnom Penh, Cambodia.

Hossain, F., Maswood, M., Siddique-E-Akbor, A.H., Yigzaw, W., Mazumdar, L.C., Ahmed, T., Hossain, M., Shah-Newaz, S.M., Limaye, A., Lee, H., Pradhan, S., Shrestha, B., Bajracahrya, B., Biancamaria, S., Shum, C.K., Turk, F.J., 2014a. “A promising radar altimetry satellite system for operational flood forecasting in flood-prone bangladesh.” *IEEE Geosci. Remote Sens. Mag.* 2, 27–36. <https://doi.org/10.1109/MGRS.2014.2345414>

Hossain, F., Siddique-E-Akbor, A.H., Mazumder, L.C., Shahnewaz, S.M., Biancamaria, S., Lee, H., Shum, C.K., 2014b. “Proof of concept of an altimeter-based river forecasting system for transboundary flow inside Bangladesh.” *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* 7, 587–601. <https://doi.org/10.1109/JSTARS.2013.2283402>

Hossain, F., Sikder, S., Biswas, N., Bonnema, M., Lee, H., Luong, N.D., Hiep, N.H., Du Duong, B., Long, D., 2017. “Predicting water availability of the regulated Mekong River Basin using satellite observations and a physical model.” *Asian J. Water, Environ. Pollut.* 14, 39–48. <https://doi.org/10.3233/AJW-170024>

Hostache, R., Chini, M., Giustarini, L., Neal, J., Kavetski, D., Wood, M., Corato, G., Pelich, R.M., Matgen, P., 2018. “Near-real-time assimilation of SAR-derived flood maps for improving flood forecasts.” *Water Resour. Res.* 54, 5516–5535.

<https://doi.org/10.1029/2017WR022205>

Houghton, R.W., Tourre, Y.M., 1992. “Characteristics of low-frequency sea surface temperature fluctuations in the tropical Atlantic.” *J. Clim.* 5, 765–772.

[https://doi.org/10.1175/1520-0442\(1992\)005<0765:COLFSS>2.0.CO;2](https://doi.org/10.1175/1520-0442(1992)005<0765:COLFSS>2.0.CO;2)

Huang, C., Chen, Y., Wu, J., 2014. “Mapping spatio-temporal flood inundation dynamics at large riverbasin scale using time-series flow data and MODIS imagery.” *Int. J. Appl. Earth Obs. Geoinf.* 26, 350–362.

<https://doi.org/10.1016/j.jag.2013.09.002>

Huang, C., Chen, Y., Wu, J., 2013. “A dem-based modified pixel swapping algorithm for floodplain inundation mapping at subpixel scale.” *Int. Geosci. Remote Sens. Symp.* 3994–3997.

<https://doi.org/10.1109/IGARSS.2013.6723708>

Huete, A.R., Liu, H.Q., Batchily, K., van Leeuwen, W., 1997. “A comparison of vegetation indices over a global set of TM images for EOS-MODIS.” *Remote Sens. Environ.* 59, 440–451.

[https://doi.org/https://doi.org/10.1016/S0034-4257\(96\)00112-5](https://doi.org/https://doi.org/10.1016/S0034-4257(96)00112-5)

Hung, N.N., Delgado, J.M., Tri, V.K., Hung, L.M., Merz, B., Bárdossy, A., Apel, H., 2012. “Floodplain hydrology of the mekong delta, Vietnam.” *Hydrol. Process.* 26, 674–686.

<https://doi.org/10.1002/hyp.8183>

Hyndman, R.J., Khandakar, Y., 2008. “Automatic time series forecasting: The forecast Package for R.” *J. Stat. Software*, 1, 3.

<https://robjhyndman.com/hyndsight/dailydata/> (Accessed on June 22<sup>nd</sup>, 2021)

Hyndman, R. J., 2014. <https://robjhyndman.com/hyndsight/forecasting-weekly-data/>

(Accessed on June 22<sup>nd</sup>, 2021)

- Imani, M., Chen, Y., You, R., Lan, W., Kuo, C., Chang, J., Rateb, A., 2017. “Spatiotemporal prediction of satellite altimetry sea level anomalies in the tropical Pacific Ocean.” *IEEE Geosci. Remote Sens. Lett.* 14, 1126–1130. <https://doi.org/10.1109/LGRS.2017.2699668>
- Islam, A.S., Bala, S.K., Haque, M.A., 2010. “Flood inundation map of Bangladesh using MODIS time-series images.” *J. Flood Risk Manag.* 3, 210–222. <https://doi.org/10.1111/j.1753-318X.2010.01074.x>
- Jiang, L., Madsen, H., Bauer-Gottwein, P., 2019. “Simultaneous calibration of multiple hydrodynamic model parameters using satellite altimetry observations of water surface elevation in the Songhua River.” *Remote Sens. Environ.* 225, 229–247. <https://doi.org/10.1016/j.rse.2019.03.014>
- Johnson, J.M., Munasinghe, D., Eyelade, D., Cohen, S., 2019. “An integrated evaluation of the National Water Model (NWM)--Height Above Nearest Drainage (HAND) flood mapping methodology.” *Nat. Hazards Earth Syst. Sci.* 19, 2405–2420. <https://doi.org/10.5194/nhess-19-2405-2019>
- Johnston, R., Kumm, M., 2012. “Water resource models in the Mekong Basin: A review.” *Water Resour. Manag.* 26, 429–455. <https://doi.org/10.1007/s11269-011-9925-8>
- Kaiser, H.F., 1958. “The varimax criterion for analytic rotation in factor analysis.” *Psychometrika* 23, 187–200. <https://doi.org/10.1007/BF02289233>
- Kaula, K. M., 1970. The Terrestrial Environment: Solid Earth and Ocean Physics, NASA CR-1599, Massachusetts Institute of Technology, Cambridge, MA, U.S.

- Keskinen, M., 2006. “The lake with floating villages: Socio-economic analysis of the Tonle Sap Lake.” *Int. J. Water Resour. Dev.* 22, 463–480. <https://doi.org/10.1080/07900620500482568>
- Kim, D., Lee, H., Chang, C.-H., Bui, D.D., Jayasinghe, S., Basnayake, S., Chishtie, F., Hwang, E., 2019a. “Daily river discharge estimation using multi-mission radar altimetry data and ensemble learning regression in the Lower Mekong River Basin.” *Remote Sens.* <https://doi.org/10.3390/rs11222684>
- Kim, D., Yu, H., Lee, H., Beighley, E., Durand, M., Alsdorf, D.E., Hwang, E., 2019b. “Ensemble learning regression for estimating river discharges using satellite altimetry data: Central Congo River as a Test-bed.” *Remote Sens. Environ.* 221, 741–755. <https://doi.org/10.1016/j.rse.2018.12.010>
- Kobayashi, S., Ota, Y., Harada, Y., Ebita, A., Moriya, M., Onoda, H., Onogi, K., Kamahori, H., Kobayashi, C., Endo, H., Miyaoka, K., Takahashi, K., 2015. “The JRA-55 reanalysis: General specifications and basic characteristics.” *J. Meteorol. Soc. Japan. Ser. II* 93, 5–48. <https://doi.org/10.2151/jmsj.2015-001>
- Kohavi, R., Provost, F., 1998. “Glossary of terms.” *Mach. Learn.* 30, 271–274. <https://doi.org/10.1023/A:1017181826899>
- Kuenzer, C., Guo, H., Huth, J., Leinenkugel, P., Li, X., Dech, S., 2013. “Flood mapping and flood dynamics of the mekong delta: ENVISAT-ASAR-WSM based time series analyses.” *Remote Sens.* 5, 687–715. <https://doi.org/10.3390/rs5020687>
- Kummu, M., Tes, S., Yin, S., Adamson, P., Józsa, J., Koponen, J., Richey, J., Sarkkula, J., 2015. “Water balance analysis for the Tonle Sap Lake–floodplain system.” *Hydrol. Process.* 29, 5477. <https://doi.org/10.1002/hyp.10763>

- Kummu, M., Sarkkula, J., 2008. "Impact of the Mekong River flow alteration on the Tonle Sap flood pulse." *AMBIO A J. Hum. Environ.* 37, 185–192.
- Kummu, M., Sarkkula, J., Koponen, J., Nikula, J., 2006. "Ecosystem management of the Tonle Sap Lake: An integrated modelling approach." *Int. J. Water Resour. Dev.* 22, 497–519. <https://doi.org/10.1080/07900620500482915>
- Lamberts, D., 2006. "The Tonle Sap Lake as a productive ecosystem." *Int. J. Water Resour. Dev.* 22, 481–495. <https://doi.org/10.1080/07900620500482592>
- Lamberts, D., Koponen, J., 2008. "Flood pulse alterations and productivity of the Tonle Sap ecosystem: A model for impact assessment." *AMBIO A J. Hum. Environ.* 37, 178–184.
- Lauri, H., de Moel, H., Ward, P.J., Räsänen, T.A., Keskinen, M., Kummu, M., 2012. "Future changes in Mekong River hydrology: impact of climate change and reservoir operation on discharge." *Hydrol. Earth Syst. Sci.* 16, 4603–4619. <https://doi.org/10.5194/hess-16-4603-2012>
- Leandro, J., Chen, A.S., Schumann, A., 2014. "A 2D parallel diffusive wave model for floodplain inundation with variable time step (P-DWave)." *J. Hydrol.* 517, 250–259. <https://doi.org/https://doi.org/10.1016/j.jhydrol.2014.05.020>
- Lee, H., 2008. Radar altimetry methods for solid earth geodynamics studies. Ph.D. Diss. Ohio State Univ.
- Lee, H., Beighley, R.E., Alsdorf, D., Jung, H.C., Shum, C.K., Duan, J., Guo, J., Yamazaki, D., Andreadis, K., 2011. "Characterization of terrestrial water dynamics in the Congo Basin using GRACE and satellite radar altimetry." *Remote Sens. Environ.* 115, 3530–3538. <https://doi.org/10.1016/j.rse.2011.08.015>

- Li, Y., Martinis, S., Plank, S., Ludwig, R., 2018. “An automatic change detection approach for rapid flood mapping in Sentinel-1 SAR data.” *Int. J. Appl. Earth Obs. Geoinf.* 73, 123–135. <https://doi.org/10.1016/j.jag.2018.05.023>
- Lian, T., Chen, D., 2012. “An evaluation of rotated eof analysis and its application to tropical pacific sst variability.” *J. Clim.* 25, 5361–5373. <https://doi.org/10.1175/JCLI-D-11-00663.1>
- Liang, X., Lettenmaier, D.P., Wood, E.F., Burges, S.J., 1994. “A simple hydrologically based model of land surface water and energy fluxes for general circulation models.” *J. Geophys. Res.* 99. <https://doi.org/10.1029/94jd00483>
- Lin, G.F., Lin, H.Y., Chou, Y.C., 2013. “Development of a real-time regional-inundation forecasting model for the inundation warning system.” *J. Hydroinformatics* 15, 1391–1407. <https://doi.org/10.2166/hydro.2013.202>
- Liu, X., Sahli, H., Meng, Y., Huang, Q., Lin, L., 2017. “Flood inundation mapping from optical satellite images using spatiotemporal context learning and modest AdaBoost.” *Remote Sens.* . <https://doi.org/10.3390/rs9060617>
- Liu, Y.Y., Maidment, D.R., Tarboton, D.G., Zheng, X., Wang, S., 2018. “A cyberGIS integration and computation framework for high-resolution continental-scale flood inundation mapping.” *J. Am. Water Resour. Assoc.* 54, 770–784. <https://doi.org/https://doi.org/10.1111/1752-1688.12660>
- Lloyd, S., 1982. “Least squares quantization in PCM.” *IEEE Trans. Inf. Theory* 28, 129–137. <https://doi.org/10.1109/TIT.1982.1056489>
- Lohmann, D., Nolte-Holube, R., Raschke, E., 1996. “A large-scale horizontal routing model to be coupled to land surface parametrization schemes.” *Tellus, Ser. A Dyn.*

*Meteorol. Oceanogr.* <https://doi.org/10.3402/tellusa.v48i5.12200>

Lohmann, D., Raschke, E., Nijssen, B., Lettenmaier, D.P., 1998. “Regional scale hydrology: I. Formulation of the VIC-2L model coupled to a routing model.”

*Hydrol. Sci. J.* 43, 131–141. <https://doi.org/10.1080/02626669809492107>

Long, S., Fatoyinbo, T.E., Policelli, F., 2014. “Flood extent mapping for Namibia using change detection and thresholding with SAR.” *Environ. Res. Lett.* 9, 35002.

<https://doi.org/10.1088/1748-9326/9/3/035002>

Lorenz, E. N., 1956. Empirical Orthogonal Functions and Statistical Weather Prediction. Statistical Forecasting Project Scientific Report No. 1., Massachusetts Institute of Technology, Cambridge, MA, U.S.

Lutz, A.F., Immerzeel, W.W., Shrestha, A.B., Bierkens, M.F.P., 2014. “Consistent increase in high Asia’s runoff due to increasing glacier melt and precipitation.”

*Nat. Clim. Chang.* 4, 587–592. <https://doi.org/10.1038/nclimate2237>

Maçaira, P.M., Tavares Thomé, A.M., Cyrino Oliveira, F.L., Carvalho Ferrer, A.L., 2018. “Time series analysis with explanatory variables: A systematic literature review.”

*Environ. Model. Softw.* 107, 199–209.

<https://doi.org/https://doi.org/10.1016/j.envsoft.2018.06.004>

Markert, K.N., Chishtie, F., Anderson, E.R., Saah, D., Griffin, R.E., 2018. “On the merging of optical and SAR satellite imagery for surface water mapping applications.”

*Results Phys.* 9, 275–277.

<https://doi.org/10.1016/j.rinp.2018.02.054>

Martin-Puig, C., Leuliette, E., Lillibridge, J., Roca, M., 2016. “Evaluating the performance of Jason-2 open-loop and closed-loop tracker modes.” *J. Atmos.*

- Ocean. Technol.* 33, 2277–2288. <https://doi.org/10.1175/JTECH-D-16-0011.1>
- Martinis, S., Kersten, J., Twele, A., 2015. “A fully automated TerraSAR-X based flood service.” *ISPRS J. Photogramm. Remote Sens.* 104, 203–212. <https://doi.org/10.1016/j.isprsjprs.2014.07.014>
- Mason, D.C., Speck, R., Devereux, B., Schumann, G.J.P., Neal, J.C., Bates, P.D., 2010. “Flood detection in urban areas using TerraSAR-X.” *IEEE Trans. Geosci. Remote Sens.* 48, 882–894. <https://doi.org/10.1109/TGRS.2009.2029236>
- MRC, 2010. Mekong River Commission: State of the Basin Report 2010. MRC, Vientiane, Laos PDR.
- MRC, 2011. Lower Mekong River Basin. MRC, Vientiane, Laos PDR.
- MRC, 2020a. Situation Report on Hydrological Conditions in the Lower Mekong River Basin in January – July 2020. MRC, Vientiane, Laos PDR.
- MRC, 2020b. Weekly Wet Season Situation Report in the Lower Mekong River Basin for 25 – 31 August 2020. MRC, Vientiane, Laos PDR.
- Moreira, A., Prats-Iraola, P., Younis, M., Krieger, G., Hajnsek, I., Papathanassiou, K.P., 2013. “A tutorial on synthetic aperture radar.” *IEEE Geosci. Remote Sens. Mag.* 1, 6–43. <https://doi.org/10.1109/MGRS.2013.2248301>
- Moriasi, D.N., Arnold, J.G., Van Liew, M.W., Bingner, R.L., Harmel, R.D., Veith, T.L., 2007. “Model evaluation guidelines for systematic quantification of accuracy in watershed simulations.” *Trans. ASABE* 50, 885–900.
- MRC, 2005. Overview of the hydrology of the Mekong Basin, Mekong River Commission. Vientiane. <https://doi.org/10.1728/3248>
- Nash, J.E., Sutcliffe, J. V., 1970. “River flow forecasting through conceptual models part

- I — A discussion of principles.” *J. Hydrol.* 10, 282–290.  
[https://doi.org/https://doi.org/10.1016/0022-1694\(70\)90255-6](https://doi.org/https://doi.org/10.1016/0022-1694(70)90255-6)
- [Data] NCEP, 2000. FNL Operational Model Global Tropospheric Analyses, continuing from July 1999. <https://doi.org/10.5065/D6M043C6>
- [Software] NCL (Version 6.6.2), 2019. Boulder, Colorado. UCAR/NCAR/CISL/TDD.  
<http://dx.doi.org/10.5065/D6WD3XH5>
- Nguyen, V.K., Pittock, J., Connell, D., 2019. “Dikes, rice, and fish: how rapid changes in land use and hydrology have transformed agriculture and subsistence living in the Mekong Delta.” *Reg. Environ. Chang.* 19, 2069–2077.  
<https://doi.org/10.1007/s10113-019-01548-x>
- Nigro, J., Slayback, D., Policelli, F., Brakenridge, G. R., 2014. NASA/DFO MODIS near real-time (NRT) global flood mapping product evaluation of flood and permanent water detection. (Retrieved on June 22<sup>nd</sup>, 2021, from [https://floodmap.modaps.eosdis.nasa.gov/documents/NASAGlobalNRTEvaluationSummary\\_v4.pdf](https://floodmap.modaps.eosdis.nasa.gov/documents/NASAGlobalNRTEvaluationSummary_v4.pdf))
- Nobre, A.D., Cuartas, L.A., Hodnett, M., Rennó, C.D., Rodrigues, G., Silveira, A., Waterloo, M., Saleska, S., 2011. “Height Above the Nearest Drainage – a hydrologically relevant new terrain model.” *J. Hydrol.* 404, 13–29.  
<https://doi.org/https://doi.org/10.1016/j.jhydrol.2011.03.051>
- Nobre, A.D., Cuartas, L.A., Momo, M.R., Severo, D.L., Pinheiro, A., Nobre, C.A., 2016. “HAND contour: A new proxy predictor of inundation extent.” *Hydrol. Process.* 30, 320–333. <https://doi.org/10.1002/hyp.10581>
- Normandin, C., Frappart, F., Lubac, B., Bélanger, S., Marieu, V., Blarel, F., Robinet,

- A., Guiastrennec-Faugas, L., 2018. “Quantification of surface water volume changes in the Mackenzie Delta using satellite multi-mission data.” *Hydrol. Earth Syst. Sci.* 22, 1543–1561. <https://doi.org/10.5194/hess-22-1543-2018>
- North, G.R., Bell, T.L., Cahalan, R.F., Moeng, F.J., 1982. “Sampling errors in the estimation of empirical orthogonal functions.” *Mon. Weather Rev.* 110, 699–706. [https://doi.org/10.1175/1520-0493\(1982\)110<0699:SEITEO>2.0.CO;2](https://doi.org/10.1175/1520-0493(1982)110<0699:SEITEO>2.0.CO;2)
- Nuorteva, P., Keskinen, M., Varis, O., 2010. “Water, livelihoods and climate change adaptation in the Tonle Sap Lake area, Cambodia: learning from the past to understand the future.” *J. Water Clim. Chang.* 1, 87–101. <https://doi.org/10.2166/wcc.2010.010>
- Okeowo, M.A., Lee, H., Hossain, F., Getirana, A., 2017. “Automated generation of lakes and reservoirs water elevation changes from satellite radar altimetry.” *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* 10, 3465–3481. <https://doi.org/10.1109/JSTARS.2017.2684081>
- Pagano, T.C., 2014. “Evaluation of Mekong River commission operational flood forecasts, 2000-2012.” *Hydrol. Earth Syst. Sci.* 18, 2645–2656. <https://doi.org/10.5194/hess-18-2645-2014>
- Paris, A., Dias de Paiva, R., Santos da Silva, J., Medeiros Moreira, D., Calmant, S., Garambois, P.-A., Collischonn, W., Bonnet, M.-P., Seyler, F., 2016. “Stage-discharge rating curves based on satellite altimetry and modeled discharge in the Amazon basin.” *Water Resour. Res.* 52, 3787–3814. <https://doi.org/10.1002/2014WR016618>
- Pham-Duc, B., Prigent, C., Aires, F., 2017. “Surface water monitoring within cambodia

- and the Vietnamese Mekong Delta over a year, with Sentinel-1 SAR observations.” *Water* 9, 1–21. <https://doi.org/10.3390/w9060366>
- Pham, H.T., Marshall, L., Johnson, F., Sharma, A., 2018. “Deriving daily water levels from satellite altimetry and land surface temperature for sparsely gauged catchments: A case study for the Mekong River.” *Remote Sens. Environ.* 212, 31–46. <https://doi.org/10.1016/j.rse.2018.04.034>
- Pierdicca, N., Pulvirenti, L., Chini, M., Guerriero, L., Candela, L., 2013. “Observing floods from space: Experience gained from COSMO-SkyMed observations.” *Acta Astronaut.* 84, 122–133. <https://doi.org/10.1016/j.actaastro.2012.10.034>
- Pokhrel, Y., Burbano, M., Roush, J., Kang, H., Sridhar, V., Hyndman, D.W., 2018. “A review of the integrated effects of changing climate, land use, and dams on Mekong River hydrology.” *Water* 10, 1–25. <https://doi.org/10.3390/w10030266>
- Revilla-Romero, B., Hirpa, F.A., Pozo, J.T., Salamon, P., Brakenridge, R., Pappenberger, F., and De Groeve, T., 2015. “On the use of global flood forecasts and satellite-derived inundation maps for flood monitoring in data-sparse regions.” *Remote Sens.* 7, 15702–15728. <https://doi.org/10.3390/rs71115702>
- Rantz, S.E., 1982. Measurement and computation of stream flow. Volume 1: Measurement of stage and discharge, US Geol. Surv. water-supply Pap. 2175, 631. <https://doi.org/10.1029/WR017i001p00131>
- Räsänen, T.A., Kummu, M., 2013. “Spatiotemporal influences of ENSO on precipitation and flood pulse in the Mekong River Basin.” *J. Hydrol.* 476, 154–168. <https://doi.org/10.1016/j.jhydrol.2012.10.028>
- RFMMC, 2011. Seasonal Flood Situation Report for the Lower Mekong River Basin,

Phnom Penh, Cambodia.

RFMMC, 2012. Seasonal Flood Situation Report for the Lower Mekong River Basin,

Phnom Penh, Cambodia.

RFMMC, 2013. Seasonal Flood Situation Report for the Lower Mekong River Basin,

Phnom Penh, Cambodia.

Richman, M.B., 1986. “Rotation of principal components.” *J. Climatol.* 6, 293–335.

<https://doi.org/10.1002/joc.3370060305>

Ruzza, G., Guerriero, L., Grelle, G., Guadagno, F.M., Revellino, P., 2019. “Multi-method tracking of monsoon floods using Sentinel-1 imagery.” *Water* 11.

<https://doi.org/10.3390/w11112289>

Sáenz, L., Farrell, T., Olsson, A., Turner, W., Mulligan, M., Acero, N., Neugarten, R.,

Wright, M., McKinnon, M., Ruiz, C., Guerrero, J., 2016. “Mapping potential freshwater services, and their representation within Protected Areas (PAs), under conditions of sparse data. Pilot implementation for Cambodia.” *Glob. Ecol. Conserv.* 7, 107–121. <https://doi.org/10.1016/j.gecco.2016.05.007>

<https://doi.org/10.1016/j.gecco.2016.05.007>

Sakamoto, T., Van Nguyen, N., Kotera, A., Ohno, H., Ishitsuka, N., Yokozawa, M.,

2007. “Detecting temporal changes in the extent of annual flooding within the Cambodia and the Vietnamese Mekong Delta from MODIS time-series imagery.”

*Remote Sens. Environ.* 109, 295–313. <https://doi.org/10.1016/j.rse.2007.01.011>

Sakamoto, T., Van Phung, C., Kotera, A., Nguyen, K.D., Yokozawa, M., 2009.

“Analysis of rapid expansion of inland aquaculture and triple rice-cropping areas in a coastal area of the Vietnamese Mekong Delta using MODIS time-series

- imagery.” *Landsc. Urban Plan.* 92, 34–46.  
<https://doi.org/https://doi.org/10.1016/j.landurbplan.2009.02.002>
- Schumann, G.J.P., Moller, D.K., 2015. “Microwave remote sensing of flood inundation.” *Phys. Chem. Earth* 83–84, 84–95.  
<https://doi.org/10.1016/j.pce.2015.05.002>
- Shin, S., Pokhrel, Y., Yamazaki, D., Huang, X., Torbick, N., Qi, J., Pattanakiat, S., Ngo-Duc, T., Nguyen, T.D., 2020. “High resolution modeling of river-floodplain-reservoir inundation dynamics in the Mekong River Basin.” *Water Resour. Res.* 56. <https://doi.org/https://doi.org/10.1029/2019WR026449>
- Siddique-E-Akbor, A.H.M., Hossain, F., Sikder, S., Shum, C.K., Tseng, S., Yi, Y., Turk, F.J., Limaye, A., 2014. “Satellite precipitation data-driven hydrological modeling for water resources management in the Ganges, Brahmaputra, and Meghna Basins.” *Earth Interact.* 18, 1–25. <https://doi.org/10.1175/EI-D-14-0017.1>
- Sikder, M.S., Hossain, F., 2018. “Improving operational flood forecasting in monsoon climates with bias-corrected quantitative forecasting of precipitation.” *Int. J. River Basin Manag.* 0, 1–11. <https://doi.org/10.1080/15715124.2018.1476368>
- Simard, M., Pinto, N., Fisher, J.B., Baccini, A., 2011. “Mapping forest canopy height globally with spaceborne lidar.” *J. Geophys. Res. Biogeosciences* 116. <https://doi.org/10.1029/2011JG001708>
- Singha, M., Dong, J., Sarmah, S., You, N., Zhou, Y., Zhang, G., Doughty, R., Xiao, X., 2020. “Identifying floods and flood-affected paddy rice fields in Bangladesh based on Sentinel-1 imagery and Google Earth Engine.” *ISPRS J. Photogramm. Remote*

- Sens.* 166, 278–293. <https://doi.org/https://doi.org/10.1016/j.isprsjprs.2020.06.011>
- Slayback, 2021. LANCE MODIS Flood Product User Guide. (Retrieved on June 22<sup>nd</sup>, 2021, from <https://earthdata.nasa.gov/earth-observation-data/near-real-time/mcdwd-nrt>)
- Smith, L.C., 1997. “Satellite remote sensing of river inundation area, stage, and discharge: a review.” *Hydrol. Process.* 11, 1427–1439. [https://doi.org/10.1002/\(SICI\)1099-1085\(199708\)11:10<1427::AID-HYP473>3.0.CO;2-S](https://doi.org/10.1002/(SICI)1099-1085(199708)11:10<1427::AID-HYP473>3.0.CO;2-S)
- Stewart, C., Larson, V., 1999. “Synthetic Aperture Radar Algorithm,” in: The Digital Signal Processing Handbook, Madisetti, V. K., Williams, D. B. (Ed.), CRC Press LLC. New York, NY, U.S.
- Sulistioadi, Y.B., Tseng, K.H., Shum, C.K., Hidayat, H., Sumaryono, M., Suhardiman, A., Setiawan, F., Sunarso, S., 2015. “Satellite radar altimetry for monitoring small rivers and lakes in Indonesia.” *Hydrol. Earth Syst. Sci.* 19, 341–359. <https://doi.org/10.5194/hess-19-341-2015>
- Takagi, H., Ty, T. V., Thao, N.D., Esteban, M., 2015. “Ocean tides and the influence of sea-level rise on floods in urban areas of the Mekong Delta.” *J. Flood Risk Manag.* 8, 292–300. <https://doi.org/10.1111/jfr3.12094>
- Tan, J., Huffman, G.J., Bolvin, D.T., Nelkin, E.J., 2019. “IMERG V06: Changes to the morphing algorithm,” *J. Atmos. Ocean. Technol.* 36, 2471-2482. Retrieved June 22<sup>nd</sup>, 2021, from <https://journals.ametsoc.org/view/journals/atot/36/12/jtech-d-19-0114.1.xml>
- Tarpanelli, A., Santi, E., Tourian, M.J., Filippucci, P., Amarnath, G., Brocca, L., 2019.

- “Daily river discharge estimates by merging satellite optical sensors and radar altimetry through artificial neural network.” *IEEE Trans. Geosci. Remote Sens.* 57, 329–341. <https://doi.org/10.1109/TGRS.2018.2854625>
- Taylor, M.H., Losch, M., Wenzel, M., Schröter, J., 2013. “On the sensitivity of field reconstruction and prediction using empirical orthogonal functions derived from gappy Data.” *J. Clim.* 26, 9194–9205. <https://doi.org/10.1175/JCLI-D-13-00089.1>
- Teng, J., Jakeman, A.J., Vaze, J., Croke, B.F.W., Dutta, D., Kim, S., 2017. “Flood inundation modelling: A review of methods, recent advances and uncertainty analysis.” *Environ. Model. Softw.* 90, 201–216. <https://doi.org/10.1016/j.envsoft.2017.01.006>
- Tospornsampan, J., Malone, T., Katry, P., Pengel, B., Pich An, H., 2009. Flood Management and Mitigation Programme Component 1: Short and medium-term flood forecasting at the Regional Flood Management and Mitigation Center. 7<sup>th</sup> Annual Mekong Flood Forum, Integr. flood risk Manag. Mekong River Basin 155–164.
- Tourian, M.J., Schwatke, C., Sneeuw, N., 2017. “River discharge estimation at daily resolution from satellite altimetry over an entire river basin.” *J. Hydrol.* 546, 230–247. <https://doi.org/10.1016/j.jhydrol.2017.01.009>
- Tourian, M.J., Tarpanelli, A., Elmi, O., Qin, T., Brocca, L., Moramarco, T., Sneeuw, N., 2016. “Spatiotemporal densification of river water level time series by multitemission satellite altimetry.” *Water Resour. Res.* 52, 1140–1159. <https://doi.org/10.1002/2015WR017654>
- Try, S., Tanaka, S., Tanaka, K., Sayama, T., Hu, M., Sok, T., Oeurng, C., 2020a.

- “Projection of extreme flood inundation in the Mekong River basin under 4K increasing scenario using large ensemble climate data.” *Hydrol. Process.* 34, 4350–4364. <https://doi.org/10.1002/hyp.13859>
- Try, S., Tanaka, S., Tanaka, K., Sayama, T., Lee, G., Oeurng, C., 2020b. “Assessing the effects of climate change on flood inundation in the lower Mekong Basin using high-resolution AGCM outputs.” *Prog. Earth Planet. Sci.* 7, 34. <https://doi.org/10.1186/s40645-020-00353-z>
- Tsyganskaya, V., Martinis, S., Marzahn, P., Ludwig, R., 2018a. “SAR-based detection of flooded vegetation – a review of characteristics and approaches.” *Int. J. Remote Sens.* 39, 2255–2293. <https://doi.org/10.1080/01431161.2017.1420938>
- Tsyganskaya, V., Martinis, S., Marzahn, P., Ludwig, R., 2018b. “Detection of temporary flooded vegetation using Sentinel-1 time series data.” *Remote Sens.* <https://doi.org/10.3390/rs10081286>
- Tuan, V.A., Quang, N.H., Hang, L.T.T., 2021. “Optimizing flood mapping using multi-synthetic aperture radar images for regions of the Lower Mekong Basin in Vietnam.” *Eur. J. Remote Sens.* 54, 13–28. <https://doi.org/10.1080/22797254.2020.1859340>
- Twele, A., Cao, W., Plank, S., Martinis, S., 2016. “Sentinel-1-based flood mapping: a fully automated processing chain.” *Int. J. Remote Sens.* 37, 2990–3004. <https://doi.org/10.1080/01431161.2016.1192304>
- United Nations Country Team, 2016. Vietnam Drought and Saltwater Intrusion Situation Update No. 1, Hanoi, Vietnam.
- Van Trung, N., Choi, J.H., Won, J.S., 2013. “A land cover variation model of water

- level for the floodplain of Tonle Sap, Cambodia, derived from ALOS PALSAR and MODIS data.” *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* 6, 2238–2253. <https://doi.org/10.1109/JSTARS.2012.2226437>
- van Zyl, J., Kim, Y., 2011. “Synthetic Aperture Radar (SAR) Imaging Basics.” *Synth. Aperture Radar Polarim.* 1–22. <https://doi.org/10.1002/9781118116104.ch1>
- Västilä, K., Kummu, M., Sangmanee, C., Chinvano, S., 2010. “Modelling climate change impacts on the flood pulse in the Lower Mekong floodplains.” *J. Water Clim. Chang.* 1, 67–86. <https://doi.org/10.2166/wcc.2010.008>
- Wang, X., Smith, K., Hyndman, R., 2006. “Characteristic-based clustering for time series data.” *Data Min. Knowl. Discov.* 13, 335–364. <https://doi.org/10.1007/s10618-005-0039-x>
- Werner, C., Wegmüller, U., Strozzi, T., Wiesmann, A., 2000. “GAMMA SAR and interferometric processing software.” ERS-ENVISAT Sym. Gothenburg, Sweden.
- Werner, M., Schellekens, J., Gijsbers, P., van Dijk, M., van den Akker, O., Heynert, K., 2013. “The Delft-FEWS flow forecasting system.” *Environ. Model. Softw.* 40, 65–77. <https://doi.org/https://doi.org/10.1016/j.envsoft.2012.07.010>
- White, L., Brisco, B., Pregitzer, M., Tedford, B., Boychuk, L., 2014. “RADARSAT-2 beam mode selection for surface water and flooded vegetation mapping.” *Can. J. Remote Sens.* 40, 135–151. <https://doi.org/10.1080/07038992.2014.943393>
- Wilks, D.S., 2011. “Principal Component (EOF) Analysis,” in: Statistical Methods in the Atmospheric Sciences, Wilks, D.S. (Ed.), Academic Press, pp. 519–562. <https://doi.org/https://doi.org/10.1016/B978-0-12-385022-5.00012-9>
- Wing, O.E.J., Bates, P.D., Sampson, C.C., Smith, A.M., Johnson, K.A., Erickson, T.A.,

2017. “Validation of a 30 m resolution flood hazard model of the conterminous United States.” *Water Resour. Res.* 53, 7968–7986. <https://doi.org/10.1002/2017WR020917>
- Wing, O.E.J., Sampson, C.C., Bates, P.D., Quinn, N., Smith, A.M., Neal, J.C., 2019. “A flood inundation forecast of Hurricane Harvey using a continental-scale 2D hydrodynamic model.” *J. Hydrol.* X 4, 100039. <https://doi.org/https://doi.org/10.1016/j.hydroa.2019.100039>
- Wingham, D.J., Rapley, C.G., Griffiths, H., 1986. “New techniques in satellite altimeter tracking systems.” *Int. Geosci. Remote Sens. Symp.* 1339–1344.
- Wolter, K., Timlin, M.S., 2011. “El Niño/Southern Oscillation behaviour since 1871 as diagnosed in an extended multivariate ENSO index (MEI.ext).” *Int. J. Climatol.* 31, 1074–1087. <https://doi.org/10.1002/joc.2336>
- Wolter, K., Timlin, M.S., 1998. “Measuring the strength of ENSO events: How does 1997/98 rank?” *Weather* 53, 315–324. <https://doi.org/10.1002/j.1477-8696.1998.tb06408.x>
- Wolter, K., Timlin, M.S., 1993. “Monitoring ENSO in COADS with a seasonally adjusted principal component index,” in: Proceeding of the 17<sup>th</sup> Climate Diagnostics Workshop, pp. 52–57. Norman, OK, U.S.
- Wood, M., Hostache, R., Neal, J., Wagener, T., Giustarini, L., Chini, M., Corato, G., Matgen, P., Bates, P., 2016. “Calibration of channel depth and friction parameters in the LISFLOOD-FP hydraulic model using medium-resolution SAR data and identifiability techniques.” *Hydrol. Earth Syst. Sci.* 20, 4983–4997. <https://doi.org/10.5194/hess-20-4983-2016>

- World Meteorological Organization, 2017. Verification of flash flood warnings. 1<sup>st</sup> Steering Committee Meeting of the Southeastern Asia-Oceania Flash Flood Guidance System. Jakarta, Indonesia.
- World Meteorological Organization, 1980. Manual on Stream Gauging: Volume II - computation of discharge.
- Xiao, X., Boles, S., Liu, J., Zhuang, D., Liu, M., 2002. “Characterization of forest types in Northeastern China, using multi-temporal SPOT-4 vegetation sensor data.” *Remote Sens. Environ.* 82, 335–348. [https://doi.org/https://doi.org/10.1016/S0034-4257\(02\)00051-2](https://doi.org/10.1016/S0034-4257(02)00051-2)
- Yamazaki, D., Ikeshima, D., Tawatari, R., Yamaguchi, T., O’Loughlin, F., Neal, J.C., Sampson, C.C., Kanae, S., Bates, P.D., 2017. “A high-accuracy map of global terrain elevations.” *Geophys. Res. Lett.* 44, 5844–5853. <https://doi.org/10.1002/2017GL072874>
- Yan, K., Di Baldassarre, G., Solomatine, D.P., Schumann, G.J.-P., 2015. “A review of low-cost space-borne data for flood modelling: topography, flood extent and water level.” *Hydrol. Process.* 29, 3368–3387. <https://doi.org/10.1002/hyp.10449>
- Yosef, G., Alpert, P., Price, C., Rotenberg, E., Yakir, D., 2017. “Using EOF analysis over a large area for assessing the climate impact of small-scale afforestation in a semi-arid region.” *J. Appl. Meteorol. Climatol.* 56, 2545–2559. <https://doi.org/10.1175/JAMC-D-16-0253.1>
- Zheng, Z., Zhao, X., 1984. “A study of long-period sea level changes in the China Sea areas.” *Mar. Geophys. Res.* 7, 299–306. <https://doi.org/10.1007/BF00305429>
- Zheng, X., Maidment, D.R., Tarboton, D.G., Liu, Y.Y., Passalacqua, P., 2018a.

“GeoFlood: Large-scale flood inundation mapping based on high-resolution terrain analysis.” *Water Resour. Res.* 54, 10013–10033

<https://doi.org/https://doi.org/10.1029/2018WR023457>

Zheng, X., Tarboton, D.G., Maidment, D.R., Liu, Y.Y., Passalacqua, P., 2018b. “River channel geometry and rating curve estimation using Height Above the Nearest Drainage.” *J. Am. Water Resour. Assoc.* 54, 785–806.

<https://doi.org/https://doi.org/10.1111/1752-1688.12661>

Zheng, Y., Zhang, X., Hou, B., Liu, G., 2014. “Using combined difference image and K-means clustering for SAR image change detection.” *IEEE Geosci. Remote Sens. Lett.* 11, 691–695.

<https://doi.org/10.1109/LGRS.2013.2275738>

Zhou, Y., Jin, S., Tenzer, R., Feng, J., 2016. “Water storage variations in the Poyang Lake Basin estimated from GRACE and satellite altimetry.” *Geod. Geodyn.* 7,

108–116. <https://doi.org/10.1016/j.geog.2016.04.003>