

University of Louisville

ThinkIR: The University of Louisville's Institutional Repository

Electronic Theses and Dissertations

12-2016

Study of spatiotemporal rainfall structure and optimized local radar rainfall application to urban watershed, Louisville, Kentucky, 2010-2014.

Jin-Young Hyun
University of Louisville

Follow this and additional works at: <https://ir.library.louisville.edu/etd>



Part of the [Engineering Commons](#)

Recommended Citation

Hyun, Jin-Young, "Study of spatiotemporal rainfall structure and optimized local radar rainfall application to urban watershed, Louisville, Kentucky, 2010-2014." (2016). *Electronic Theses and Dissertations*. Paper 2604.
<https://doi.org/10.18297/etd/2604>

This Doctoral Dissertation is brought to you for free and open access by ThinkIR: The University of Louisville's Institutional Repository. It has been accepted for inclusion in Electronic Theses and Dissertations by an authorized administrator of ThinkIR: The University of Louisville's Institutional Repository. This title appears here courtesy of the author, who has retained all other copyrights. For more information, please contact thinkir@louisville.edu.

STUDY OF
SPATIOTEMPORAL RAINFALL STRUCTURE AND OPTIMIZED LOCAL RADAR
RAINFALL APPLICATION TO URBAN WATERSHED, LOUISVILLE,
KENTUCKY, 2010-2014

By
Jin-Young Hyun
B.S., University of Kentucky, 2009
M.S., University of Iowa, 2012

A Dissertation
Submitted to the Faculty of the
University of Louisville
J. B. Speed School of Engineering
in Partial Fulfillment of the Requirements
for the Degree of

Doctor of Philosophy in Civil Engineering

Department of Civil and Environmental Engineering
University of Louisville
Louisville, Kentucky

December 2016

Copyright 2016 by Jin-Young Hyun

All rights reserved

STUDY OF SPATIOTEMPORAL RAINFALL STRUCTURE AND OPTIMIZED
LOCAL RADAR RAINFALL APPLICATION TO URBAN WATERSHED,
LOUISVILLE, KENTUCKY

By

Jin-Young Hyun

B.S., University of Kentucky, 2009

M.S., University of Iowa, 2012

A Dissertation approved on

September 29, 2016

By the following Dissertation Committee:

Mark N. French, Dissertation Director

Thomas D. Rockaway

Nageshwar R. Bhaskar

Mehmed Kantardzic

DEDICATION

This thesis or dissertation is dedicated to my parents

Mr. Insung Hyun

and

Mrs. Jungsook Lee

who have given me invaluable educational opportunities.

and

to my wife

Mrs. Juyeon Gang

who supports me to fulfill my whole academic career.

ACKNOWLEDGMENTS

The entire research investigated in this dissertation was a completion of the interdisciplinary work of the coupled interactions of the extreme rainfall, urban sewer-shed and for their preparations. This study was unleashed by the local sewer district, Metropolitan Sewer District (MSD) to mitigate combined sewer water overflow.

I would like to thank sincerely my advisor, Dr. Mark N. French, for his endless guidance and patience on this research. The relationship with him is more than ten years since beginning of my engineering journey and his enthusiasm in the undergraduate class gave me initial motivation to study my main field, water resource. His comments and discussions improved this work.

I extended my appreciation to Dr. Thomas D. Rockaway, project investigator of the Green Infrastructure to mitigate the storm-sewer water discharge. I was involved practically by his preparation to achieve the practical results from this study.

I am also very grateful to Dr. Nageshar R. Bhaskar and Dr. Mehmed Kantardzic for serving as my advisory committees. Their guidance in the meetings and classes were helpful and are sincerely appreciated.

I want to thank my wife, lovely kids and parents for their patience, support and sacrifice. I definitely gives credit to them for this accomplishment of this research.

ABSTRACT

STUDY OF SPATIOTEMPORAL RAINFALL STRUCTURE AND OPTIMIZED LOCAL RADAR RAINFALL APPLICATION TO URBAN WATERSHED, LOUISVILLE, KENTUCKY

Jin-Young Hyun

September 29, 2016

In urban areas, a prevalence of combined sewer systems (CSS) exist that carry both storm water runoff and sanitary sewer flows in a single pipe, these system are considered combined sewers. In the absence of rainfall-runoff most of these systems function adequately, however CSS capacity is typically inadequate to carry peak stormwater runoff volume. In order to minimize sewage flooding into streets and backups into homes and businesses, most CSSs (as well as separate sanitary sewer systems) are designed to overflow into surface waters such as streams and rivers, lakes and seas. This occurrence is considered a combined sewer overflow (CSO) event and has a critical impact on urban aquatic environment and degrades downstream water quality. This investigation provides a framework for the application of radar-rainfall data to estimate the characteristics of rainfall events that produce a CSO event. The process addresses an urban sewer-shed, denoted as CSO 130, located in Louisville, Kentucky (USA). The characteristics of each heavy rainfall event; total volume, intensity, duration, continuity, and storm types govern the overflow in the approximately 13-ha (30-ac) sewer-shed.

In urban hydrology, accurate fine resolution temporal and spatial rainfall observations is a key factor for managing urban hydrologic systems and forecasting storm water runoff, particularly in the current era of higher variability in recent rainfall events. To study this issue, rain gauge data from a ground based rainfall measurement network operated by the local stormwater management agency, Metropolitan Sewer District (MSD), in Jefferson County Kentucky is studied. Rainfall spatial characteristics are evaluated through correlation spectrum by distance and this reveals a spatial rainfall variation concave relationship. Besides, the event based rainfall classification has been performed to provide a context for identification and description of rain events that may be useful as guidance for urban stormwater management. Based on this study, the limitation of the one dimensional rainfall monitoring system has been revealed by the severe variation of the rainfall characteristics.

In order to overcome this issue, the reliable areal rainfall measurement with fine spatiotemporal resolutions is urgently required to investigate the urban hydrologic issues. The radar data utilized in this study are from the weather radar associated with the National Weather Service (NWS) Forecast Office Louisville, Kentucky (denoted by call letters KLVX) and rain gauge data are from a regional network. The study applies fine resolution radar rainfall in this urban hydrologic system to reveal insights for planning CSO control and prevention under a range of rainfall event regimes. Weather radar data from the local NWS site is optimized using support vector classification (SVC) and serves as rainfall input for the urban sewer-shed. The radar-rainfall data were optimized through a comparison with NWS radar rainfall and a gauge network, the local stormwater and sewer agency. The optimized radar rainfall estimation has the highest spatiotemporal

correlation in quarter hourly temporal resolution. The rainfall and flow events are defined using the criteria proposed by United States Environmental Protection Agency (USEPA) to define the physical continuity of natural rainfall processes and the corresponding hydrologic response.

The optimized rainfall product has applied to the small scale urban watershed, CSO130 to investigate the sewer water overflow. In this setting, the extremity of the rainfall governs the overflow mainly with volumetric rainfall in the event based rainfall and its corresponding overflow with other decisive factors; rainfall intensity, duration, rain type as well as rainfall continuity. Discriminant analysis is introduced to classify these precipitation factors.

The objective of this study is that downscaled hydrologic application to the places where the sub-hourly rainfall data is required such as a complex urban watershed in order to investigate the fast inundated floods, overflows in the artificial watersheds or any hydrologic preparation.

TABLE OF CONTENTS

	PAGE
ACKNOWLEDGEMENTS.....	iv
ABSTRACT.....	v
LIST OF TABLES.....	x
LIST OF FIGURES.....	xi
CHAPTER 1 INTRODUCTION.....	1
CHAPTER 2 SPATIOTEMPORAL STRUCTURE OF RAINFALL.....	7
(Ground-level Rainfall Variation Over Jefferson County, Kentucky: published as a case study at Journal of Hydrologic Engineering, American Society of Civil Engineering, July 26 2016, (ASCE), DOI:10.1061/(ASCE)HE.1943-5584.0001438)	
2.1 Introduction.....	7
2.2 Data source and preprocessing.....	9
2.3 Methods.....	13
2.3.1 Spatiotemporal correlation coefficient.....	13
2.3.2 Quantile/threshold-range effect on correlation.....	19
2.4 Result and Conclusion.....	23
CHAPTER 3 RAINFALL CLASSIFICATION.....	26
(Rainfall event characterization with cluster and variogram analysis: in revision as a technical paper at Journal of Hydrometeorology, American Meteorological Society (AMS), October 19 2016)	
3.1 Introduction.....	26
3.2 Data and preprocessing.....	27
3.3 Methodology.....	29
3.4 Conclusion and Discussion.....	34
CHAPTER 4 OPTIMIZATION OF LOCAL WEATHER RADAR.....	36

(Urban sewershed overflow analysis using super-resolution weather radar rainfall-Part I: in revision as a technical paper at Journal of Hydrology and Earth System Sciences, November 16 2016, (DOI: 10.5194/hess-2016-362))

4.1	Introduction.....	36
4.2	Material and methods.....	39
4.2.1	Data sources.....	40
4.2.2	Data preprocessing.....	41
4.3	Theory and application.....	45
4.4	Results and iscussion.....	50
4.4.1	Results.....	50
4.4.2	Discussion.....	52
4.5	Conclusions.....	53
CHAPTER 5 APPLICATION OF RADAR RAINFALL TO URBAN AREA.....		55
(Urban sewershed overflow analysis using super-resolution weather radar rainfall-Part II: in review as a technical paper at Journal of Hydrology and Earth System Sciences, November 16 2016, (DOI: 10.5194/hess-2016-362))		
5.1	Introduction.....	55
5.2	Case Study.....	58
5.2.1	Urban sewershed setting and CSO location.....	58
5.2.2	Coupled radar rainfall and overflow event record.....	59
5.3	CSO130 overflow analysis.....	63
5.3.1	Quality control of coupled rainfall/overflow event.....	63
5.3.2	Analysis of coupled rainfall and overflow events.....	65
5.3.2.1	Overflow relation to rainfall depth, intensity and duration.....	65
5.3.2.2	Overflow relation to rainfall depth, storm type and continuity.....	67
5.3.2.3	Discriminant analysis in overflow inducing rainfall events.....	69
5.4	Conclusions and future works.....	74
CHAPTER 6 CONCLUSION AND FUTURE WORKS.....		77
REFERENCES.....		82
APPENDICES.....		96
Appendix A. Graphical Slides of Workflow		96

Appendix B. Theoretical equations.....	106
Appendix C. MATLAB codes.....	108
CURRICULUM VITA.....	144

LIST OF TABLES

Table 5.1	Group Mean Values of Rainfall Characteristics by Discriminant Analysis
Table 5.2	Classification Result and Predicted Group Membership by Discriminant Analysis.

LIST OF FIGURES

Figure 1.1 Areal map of the study area: locations of the MSD rain gauge (circle), NOAA rain gauge (triangle) and the boundary of Jefferson County, Kentucky

Figure 2.1 Years 2010-2014 gauge-averaged monthly rainfall accumulation (mm) from 2 co-regional gage networks: MSD (solid bar, 15 gauges) and NOAA (open bar, 12 gauges)

Figure 2.2 Daily rainfall accumulation for closest distance paired gauges using MSD and NOAA gauge networks (mm)

Figure 2.3 Comparison of rainfall data filtering on rainfall detection. Left side: Rainfall detection after elimination of isolated single tip signals and rainfall detection after discarding values 1mm or less (left), PPMCC sensitivity to heavy rainfall (right)

Figure 2.4 Spatiotemporal correlation structure: temporal resolution = 15-minute (upper-left), 30-minute (upper-right), 1-hour (lower-left), and 3-hour (lower-right). The critical distance (e-folding decay, $1/e$) and shape factor are shown in brackets

Figure 2.5 Quantile rainfall summary for time-averaged depth (mm) (left) and rainfall depth standard deviation (mm) (right) – for categories: entire domain, 50% or higher, 75% or higher, 90% or higher, and 95% or higher

Figure 2.6 Correlation spatial variation (distance and direction) for paired gauges according to thresholds on depth and duration (East-West is x axis and North-South is y

axis) across the study area. Correlations shown for time intervals: 15 min, 30 min, 1 hr, and 3 hr, and depths: all depths (entire), 0.50 inch or higher, 0.75 inch or higher, 1.0 inch or higher, and 2.0 inch or higher

Figure 2.7 Correlation structure variation for accumulation interval (15 min, 30 min, 1 hr, 3 hr) by rainfall threshold (0.50 inch or higher, 0.75 inch or higher, 1.0 inch or higher, 2.0 inch or higher) and distance range (0-40 km, 0-10 km, 10-20 km, 20-30km, 30-40 km)

Figure 3.1 Monthly accumulation of rainfall of TR17, TR18, average value of other MSD gauges, and NOAA rain gauge at Galena, IN for 2010-2014

Figure 3.2 Rainfall event clustering classification as light rainfall, high intensity, and high depth groups using the K-means clustering method: Initial status of scatter plot of averaged rain depth vs. maximum rainfall intensity (upper-left); Predefined centers for each group; light rainfall event at the origin, high intensity rainfall event at upper left corner, and high depth rainfall event at lower right corner in the 2-dimensional surface (upper-right); Final K-means clustered groups (lower-left), General statistics, mean and dispersion of duration for clustered groups (lower-right)

Figure 3.3 Spatial-variogram of the two rainfall groups, high intensity rainfall event group and high depth rainfall event group

Figure 4.1 NWS Z-R relationships by storm type, 0-60dBZ range (left) and 30-40dBZ range (right)

Figure 4.2 Histogram of Level 2, radar reflectivity at KLVX, Fort Knox, Kentucky, Jan. 2010 – Dec. 2014

Figure 4.3 Gauge and Radar rainfall depth in 2 dimensional space (gauge-radar volume) across temporal resolutions: monthly (upper- left), daily (upper-right), hourly (lower-left), quarter-hourly (lower-right)

Figure 4.4 Histogram of rain gauge values for cases where radar data is not observed (Not Available (NA)) or radar intensity is below the 5mm/15min threshold.

Figure 4.5 Scatter plots of rainfall volumes for each storm type after optimization (minimum RMSE error): convective type (upper-left), tropical type (upper-right), east-cool-stratiform type (lower-left), stratiform – Marshall/Palmer - type (lower-right)

Figure 4.6 Convective rainfall type: radar and gauge comparison with (a) standard NWS Z-R relation (light shade marker), and (b) optimized Z-R relation (minimum RMSE error)

Figure 4.7 Tropical type rainfall results: Optimized minimum RMSE error (upper-left); Decision schematic for SVC kernel within least RMSE error range (upper-right); SVC binary clustering hyperplane (lower-left), SVC-based optimization with two rainfall groups (lower-right)

Figure 4.8 Comparison of local radar rainfall estimations and corresponding Z-R relationships: Standard Z-R-based quarter-hourly rainfall accumulation (upper-left); four NWS Z-R relationships (upper-right); Optimal SVC-based quarter-hourly rainfall accumulation (lower-left); SVC-based optimal Z-R relationships (lower-right)

Figure 5.1 Rainfall data quality comparisons with the reference rainfall data (TR05): NWS radar data (upper-left), Quality-improved radar data by SVC (upper-right), MSD

rain gauge-TR12; 4.9km away from the study area (lower-left), and MSD rain gauge-TR03; 15.3km away from the study area (lower-right)

Figure 5.2 Selected CSO events and corresponding rainfall events; Event number denotes the rank of the overflow amount through the outfall structure to Beargrass Creek, Louisville, KY

Figure 5.3 Overflow ratio plots. Left side: Two-dimensional radar & gauge rainfall field; x-axis shows rain gauge ratio-MSD TR05 (nearest study area), y-axis shows radar ratio-NWS weather radar KLVX. Right side: Criteria threshold for valid event selection: 52 acceptable events (blue) below the 0.60 overflow/rainfall threshold, and 43 non-acceptable events (red) exceed the threshold

Figure 5.4 Event-based rain depth versus overflow depth (left), and rain event duration versus rainfall depth grouped by peak rain intensity (right). Intensity threshold peak is 4.87mm/15minute to identify weak (blue) and strong (red) peak event groups

Figure 5.5 Rainfall occupancy ratio (ratio of continuous rain duration to total event duration) and total rainfall event depth: convective event type in warm season (magenta), convective type in cold season (red) and stratiform (blue) (left). Event-based rainfall depth versus overflow depth: convective-warm season (magenta), convective-cold season (red), stratiform-warm season (cyan), and stratiform-cold season (blue) (right)

Figure 5.6 Overflow event and 1.5-mm depth threshold separating overflow events into two binary categories – significant (denoted as 1) and non-Significant (denoted as 0)

CHAPTER1 INTRODUCTION

Historically, the early human civilizations have prospered near sources of water as an essential substance to live. Almost all human activities require clean enough water for the purpose of drinking, livestock, sanitation, entertainment and irrigation. An accessibility to clean water for these various purposes is still essential a prerequisite for human's civil activities.

In recent years, damage from floods has been reported more frequently and severely around the world. Specially, people are concerned that water related natural disasters are related to an apocalyptic threat of the climate change. No one is undoubtedly sure of the strong correlation of these issues, but the current unstable tendency of the increasing severity of the floods is true and many strong scientific factors support it (Hlavčová et al. 2015, Riboust and Brissette 2015). The damages induced by climate related floods sweep densely populated areas. The imperviousness of the surface in this area is a general characteristic of crowded urban cities and the expected runoff is greater compared to rural areas. Therefore, this study focuses on the metro-city, Louisville, Kentucky to evaluate extreme storms and its impacts on the artificial urban watershed system. The study area, Louisville Metropolitan (Jefferson County) is a historical city in the United States. The city lies along the Ohio River which is the geographic border line between the states of Kentucky and Indiana. The city was built on the flat floodplain and swampland after drain out. In other words, the city is maintained by artificial structures

and draining systems to prevent intrusion of the water from the Ohio River which flows through the part of the northern borderline in the figure 1.1.

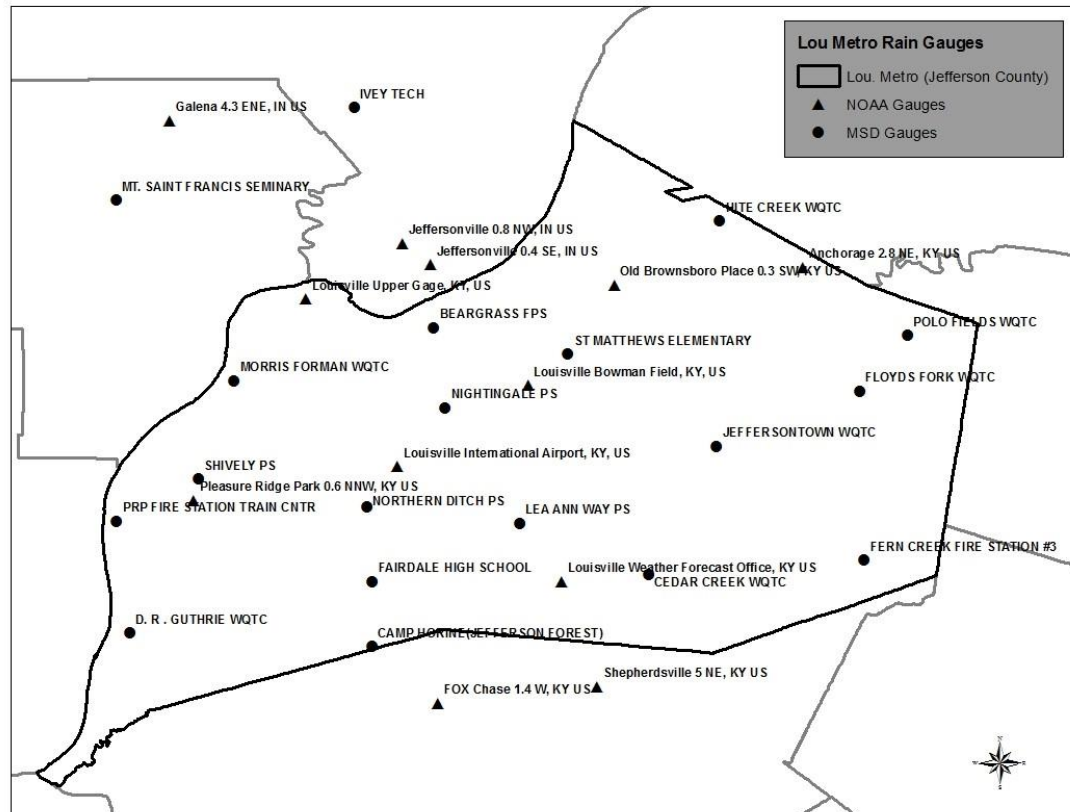


Figure1.1 Areal map of the study area: locations of the MSD rain gauge (circle), NOAA rain gauge (triangle) and the boundary of Jefferson County, Kentucky.

Fundamentally, this area is vulnerable to floods and the rational research and preparations are necessarily required to minimize damage from the floods. Recently, the type of sudden inundated flood referred to as the flash flood has emerged as a new storm-related disaster in the hydrologic field, and this study centered on fine resolution of spatiotemporal rainfalls over the study area for a downscaled hydrologic approach.

Overall, the interaction between heavy storms and small scaled urban watershed is investigated. The city watersheds are relatively small and responded faster than a natural watershed. Based on this condition, rainfall occurrence within a short period of time is a decisive factor for urban hydrology. The rainfall data accuracy is also an essential requirement to consider the rapid runoff response. Therefore, the quality of rainfall data is important for sub-hourly hydrologic research and evaluation of urban watershed response. The two types of the rainfall estimations incorporated in this work are from a ground-based rain gauge network of tipping bucket type devices and the indirect or remote-sensing of rainfall and ground-level rain estimation methods using reflectivity data from the local weather radar.

In chapter 2, the ground-based rainfall structure is investigated using the a common and historical data acquisition device, a tipping bucket rain gauge network at point locations shown in figure 1.1. The local municipal sewer district (MSD) operates this network of multiple rain gauges over the city and the spatiotemporal correlation variations were derived from these data as a representative of the rainfall structure. Moreover, the quantile analysis with different temporal resolutions are studied in order to understand the tendency of rain structure with rainfall severity. A gauge rainfall detection is a point measurement of the rainfall, and it has a serious limitation due to its narrow spatial coverage. This chapter reveals the limitation of the ground-based rainfall measurement system even in the densely deployed rain gauges throughout the city.

In chapter 3, the temporal resolution (minute-unit) of the gauge rainfall data were gathered by the continuous concept of rainfall, rainfall event, by use of public policy and regulatory definitions. The K-means clustering technique is used on the two-dimensional

organization of the axials of the total volume and the peak rainfall values to categorize the rainfall events into the three groups; a high intensity group, a high volume group and a light rainfall group. After that, the spatial variogram is applied to describe the spatial correlation and its spatial scope.

In chapter 4, a more sophisticated rainfall estimation using the weather radar is introduced. The radar rainfall estimation is an areal rainfall acquisition system which provides an improved definition of rainfall spatial variation measurement relative to the point gauge. Paradoxically, the gauge network in Louisville in figure1.1 is used as reference data to evaluate radar rainfall estimation. This is due to the nature of this application study whereby ground-level rainfall amounts (depths) are desired, and the ground-level gauge remains a useful record for point rainfall depth. The radar estimation is an indirect rainfall measurement system in which radar receives the scattering backed electromagnetic signal by the falling rain drop particle in the air. The NWS adapted the empirical conversion from the received reflectivity to the rainfall intensity (called Z-R relationship) which is meaningful value of rainfall volume. Still, the radar rainfall estimation may require consideration of error sources. This study focused on the calibration of an individual radar site where the use of the generalized Z-R relationship is not universally applicable. The local NWS radar, KLVX in Fort Knox, KY, which covers the study area is investigated by using supporting vector classification to derive an adjusted Z-R relationship for the purpose of local radar rainfall optimization.

From the previous study, the refined rainfall data has been studied to ensure the data quality. In this section, chapter 5, the rainfall data were applied to research the interaction of the extreme rainfall and the artificial urban watershed response using the

coupled radar rainfall and the sewer water overflow which flows into the urban aquatic environment. The small scaled Combined Sewer Overflow (CSO 130), located in the Louisville downtown area, is studied under extreme rainfall conditions to understand the rainfall conditions triggering the undesirable overflow events. The role of rain data is considered in terms of storm type (intensity, spatial scale) and storm characteristics (duration, inter-event time). This study suggests an approach for the evaluation and analysis for decreasing or mitigating the occurrence of CSO events. The two-dimensional ratio field of the rainfall and overflow was introduced to determine the reliable data set of coupled rainfall-runoff records prior to the analysis. The data driven rainfall-CSO overflow relationship demonstrates a dependency between the CSO overflow and the extremity of the rainfall. In this urbanized small scaled sewer-shed, the rainfall governs the overflow with interdisciplinary rainfall event factors; total rain volume, peak rainfall, duration, rainfall continuity as well as rain type. The discriminant analysis was used to separate the overflow into two severity groups by the rainfall characteristics.

One finding of this research is that urban hydrologic study requires use of accurate rainfall sources at high spatiotemporal time and space resolutions. The historic seasonal or event daily resolution of rainfall data with limited or no spatial variation does not represent rainfall variability at time and space scales required to evaluate runoff for urban areas in denser areas of cities. Therefore, this study investigated all the possible rainfall measurement equipment such as rain gauge network and the weather radar. The areal rainfall measurement instrument, the weather radar, was optimized in order to improve reliability at sub-hourly rainfall periods. The prepared rainfall product was

applied to demonstrate the relationship between rainfall extremes and overflow volume for the small scale sewer-shed. This study positively contributes to the urban hydrologic preparation and design.

The following four chapters consist of the academic journal publications developed from this dissertation research. Each chapter is a separate journal paper and is presented in a form identical to the published form or in the latest form available for publications submitted or under review. The final chapter provides a brief summary of this work.

CHAPTER 2 SPATIOTEMPORAL STRUCTURE OF RAINFALL

2.1 Introduction

In recent years, perceived flood severity has increased due to impacts on populations and economies through disruptions to transportation systems and displacement of residents. While flood risk has increased, hydrometeorologic measurement technology, rain gauges and weather radar, have also increased the capacity of municipal agencies to evaluate extreme rainfall events and develop strategies to mitigate adverse impacts (Karamouz 2013; Qin and Lu 2014; Salathé 2014; Wang 2014). More detailed observations provide improved definition of the spatial distribution and intensity variation of rainfall events and this information can be used to enhance estimates of flood impact at specific locations. Managing adverse impacts of severe heavy rainfall and resulting localized flooding may be reduced through a greater understanding of rainfall characteristics in terms of occurrence and magnitude.

In engineering design, use of design rainfall depth from a historical rainfall frequency duration curve, may not adequately represent observed rainfall-runoff characteristics (Einfalt 2009). Detailed rainfall spatiotemporal information can lead to insights about the runoff characteristics for specific watersheds. In urban areas, the prevalence of impervious land-cover and shorter runoff travel time, contribute to increased runoff volume and higher runoff discharge rates. Description of regional specific rainfall characteristics can lead to improved planning, management and design directed toward reducing sewer system overflows and reducing inundation of urban

properties. Historically, rainfall monitoring by ground-level rain gauges is considered the most reliable measurement system for hydrologic applications because it physically captures pluvial water. In hydrologic engineering and research, rain gauge measurements frequently serve as reference data for evaluation of indirect or remote sensing rainfall systems such as weather radars and satellites (Krajewski and Smith 2002; Seo and Breidenbach 2002; Habib 2012; Price 2014; Chen 2015; Fencil 2015). However, limitations must be considered since a single rain gauge is a point measurement with limited direct information defining spatial variation of rainfall across an area. Over a large areal region, or for a highly variable localized rain storm, a single rain gauge measurement may not describe the rainfall amount. In other words, when rain gauge data are measured at some distance from the area of interest, or a storm is non-uniform across the region, a network of several rain gauges is necessary to represent the variability of a storm event. For instance, a convective storm in a hot and humid season may generate significant rain on small areas and is spatially erratic across the spatial scale typical of a city region. A dense rain gauge network is needed to adequately observe storm characteristics. In this case study, the quantitative variability of rainfall observations from multiple rain gauges is investigated using measurement records from an operational rain gauge network.

The main objectives of this case study include evaluation of rainfall variation using spatio-temporal indicator statistics, and a quantile analysis to assess the variation of rainfall intensity across the study region. The region of interest is the city of Louisville Metropolitan area in Kentucky (USA), and focuses on a time period where mesoscale rainfall investigation is applicable, January 2010 to December 2014. As a part of data quality control review and pre-processing, data were validated using independent records

from the NWS/NOAA rain gauge network. Both first and second order statistics and critical correlation distance served as indicators to confirm data parsimony from the two independent sources. The spatial correlogram across a spectrum of temporal resolutions is investigated to indicate the spatio-temporal dynamics of rainfall. The quantile analysis of correlation is introduced to interpret spatio-temporal rainfall structure with rainfall intensity.

2.2 Data sources and preprocessing

The city of Louisville, Kentucky (metropolitan region covers Jefferson County, Kentucky) is geographically located in 38°15'N, 85°46'W along the Ohio River on the border between the states of Kentucky and Indiana. The area of the city is about 1022 km² (399 mi²) and this falls into the mesoscale range. The Metropolitan Sewer District (MSD) agency is responsible for flood protection, stormwater runoff, and wastewater treatment in the region. As part of its data collection systems, the MSD maintains a rain gauge network which is evenly spread over the Louisville Metro area. The minimum and maximum distance between nearest gauges are 4.52 km and 11.06 km respectively. The NOAA gauges serve as a reference for a data quality control comparison of characteristics developed from data recorded by the MSD network. There are 17 rain gauges in the MSD network with a complete record for the study period years 2011 to 2014. The MSD gauges are tipping bucket type with 0.254 mm (0.01 in) resolution and report accumulated rainfall every 5 minutes. The tipping bucket type rain gauge is desirable since it can record rainfall increments at a depth precision of 0.254 mm (0.01 in), however it is considered a high maintenance mechanical device. Many factors affect rainfall recordings by this type of device such as inadequate calibration of individual gauges, mechanical or electrical

problems, clogging, freezing, and wind effect as well as vandalism (Nešpor and Sevruk 1998; Wang 2008). Furthermore, there are errors associated with each precipitation type. The measurement error is most pronounced during extreme rainfall and light drizzle (Humphrey 1997).

A gauge record may be quality controlled by inter-gauge calibration using collocated gauges of similar type (Ciach and Krajewski 1999; Habib and Krajewski 2002; Ciach 2003; Ciach and Krajewski 2006; Tokay and Öztürk 2012; Jung 2014). However, inter-gauge analysis was not part of this study due to, as in most practical or operational network situations, no collocated gauges being available. Since quality control by inter-gauge calibration was not possible, the alternative approach was implemented whereby the MSD network data were reviewed and verified using coherent data from the NOAA rain gauge network.

Twelve NOAA rain gauges are deployed in the study area and provide daily accumulations. The purpose of comparison between these two network records is to evaluate rain gauge coherence and identify missing or erroneous data periods. Comparison of data records from the different institutes can serve as an acceptable substitute to validate the rain gauge system in practical situations. The main concerns regarding the gauge data reliability are the maintenance and calibration of the mechanical device rather than issues regarding detection of rainfall for a specific event. This means the case study comparing the MSD network data with the NOAA network is a sufficient indicator to explain the MSD data as a valid quantitative rainfall record. The daily total depth rain data from 12 NOAA rain gauges in the study area were retrieved from the NOAA National Climate Data Center

(NCDC). The daily data are a component of the archive of Global Historical Climatology Network–Daily (GHCN daily), Version 3 (Menne 2012).

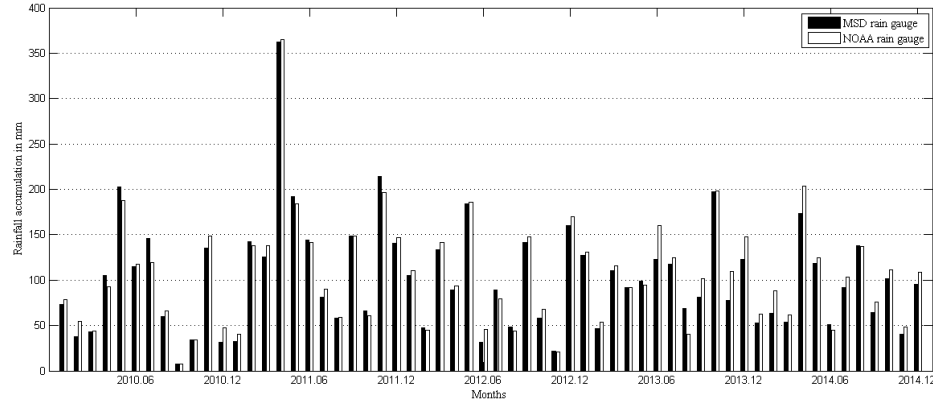


Figure 2.1 Years 2010-2014 gauge-averaged monthly rainfall accumulation (mm) from 2 co-regional gage networks: MSD (solid bar, 15 gauges) and NOAA (open bar, 12 gauges)

The monthly accumulated rainfall totals are shown in figure 2.1 indicating an overall match between the 2 rain gauge networks for the study period. The monthly rainfall totals by gauge were averaged for each rainfall gauge network, and the monthly accumulation comparison of the two networks yield a correlation coefficient, is 0.97. This indicates the rain gauge networks have a similar tendency in rainfall observation (with the implicit smoothing of spatial and temporal variation detection). Initial screening indicated two MSD gauges located the furthest from the study region consistently underreport rain depth and for this reason excluded from the case study. For daily accumulation, pairs of MSD and NOAA rain gauges within a minimum distance are selected to evaluate the MSD rain gauge data. Among twelve NOAA gauges, there are three NOAA referencing gauges. The reference locations were selected due to geographic location and data quality and the sites

are Louisville International Airport (Standiford Field), Louisville Bowman Field (Airport), and the National Weather Service Louisville Forecast Office.

These 3 gauges are centered in the study area in roughly equilateral triangle form which allows some detection of rainfall spatial variation. Moreover, these gauges are well maintained and have no missing data periods. The fifteen MSD rain gauges were paired with the closest NOAA gauge among these three. Gauge paring in this study means using the paired rain gauge records to investigate the spatial structure of rainfall. The distance between paired gauges was defined by the North Zone of Kentucky, State Plane Coordinate System (NOAA 1983). The averaged and maximum distances of the two closest gauges of the paired set of MSD and NOAA gauges are 8.9 km and 16.7 km respectively.

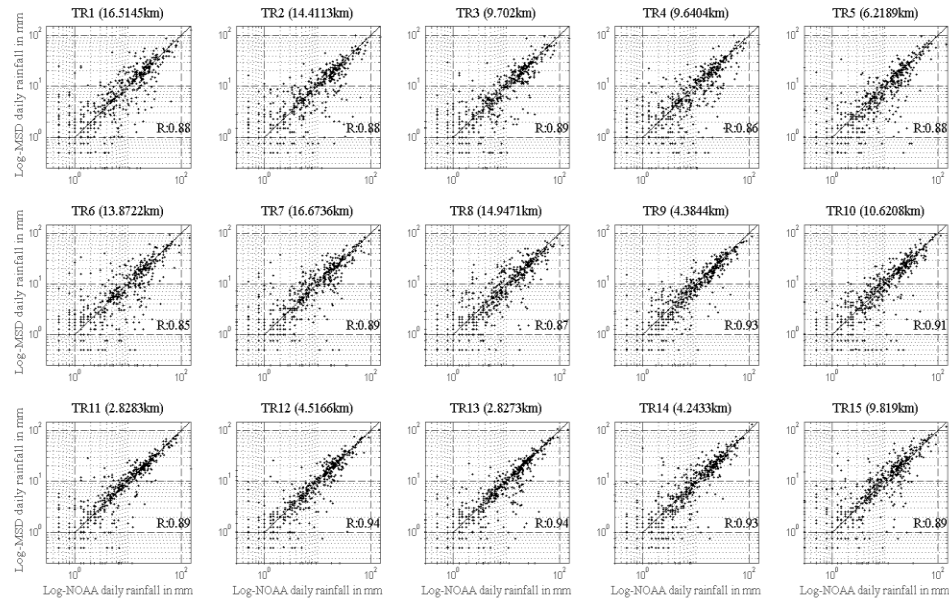


Figure 2.2 Daily rainfall accumulation for closest distance paired gauges using MSD and NOAA gauge networks (mm)

The MSD daily rainfall accumulations closely followed the depths recorded by the NOAA rain gauges as shown in Figure 2.2 All paired gauge data sets show a strong linearity with relatively high values of correlation coefficients, over the range of 0.85 to 0.94, irrespective of distance.

There is a decreasing tendency of daily rainfall correlation with distance. This result indicates that at a daily accumulation resolution, the quantitative comparison by depth magnitude and correlation coefficient indicates agreement between the NOAA and MSD rain gauge networks, and the MSD data are considered valid for use in this case study. Although this verification is limited to a daily temporal scale due to data resolution, it is a practical and meaningful validation for applications dealing with operational hydrologic systems since daily data incorporated any missing period in the MSD records.

2.3 Methodology and analysis

2.3.1 Spatiotemporal correlation coefficient

Understanding the spatial variation of rainfall is useful for understanding expected rainfall variations and for the management of stormwater. Observations recorded directly in the operational region and at a temporal scale useful for planning real-time actions are important issues for implementing management practices. Therefore, local rainfall spatial and temporal structure and variations must be understood or described at temporal scales within the sub-hour range. To describe the spatio-temporal structure of rainfall, a quantitative measure is possible through the Pearson's product moment correlation coefficient (PPMCC). The PPMCC, R , can describe the linear dependence of paired data

from the MSD network. The PPMCC is widely used for indicating spatial structure in hydrology (Mandapaka and Qin 2013; Jung 2014).

$$R_{i,j} = \frac{\sum_{i=1}^n (G_{1,i} - \overline{G_1})(G_{2,i} - \overline{G_2})}{\sqrt{\sum_{i=1}^n (G_{1,i} - \overline{G_1})^2} \sqrt{\sum_{i=1}^n (G_{2,i} - \overline{G_2})^2}} \quad (\text{Equation 2.1})$$

Equation (2.1) is the standard correlation equation for all possible paired rain gauge combinations in the network. The indices denoted as 1 and 2 indicate any 2 network gauges and n is the total number of gauges. Prior to determining the PPMCC values, it is useful to review the data characteristics necessary for the PPMCC method to be suited to a quantitative application (Habib 2001). First, there is a tendency of upward bias of PPMCC with a decrease in sample size. In this application, data records consist of a five year period with fine-scale temporal resolution, the finest resolution being a five minute time period.

A study of the data record determined the frequency of corresponding rain detection and non-detection between paired gauges (from 2 independent gauges in the MSD network) and the influence of this on the PPMCC values. The number of rain detections observed across the network of MSD rain gauges appears uniformly distributed throughout the data record. Meaning no single gauge indicates unusual frequency in the number of isolated positive records, such as a single data signal record within a multi-hour period. As summarized in figure 3.3 (left), rainfall detections at neighboring gauges becomes more uniform as threshold depth and temporal resolution increase.

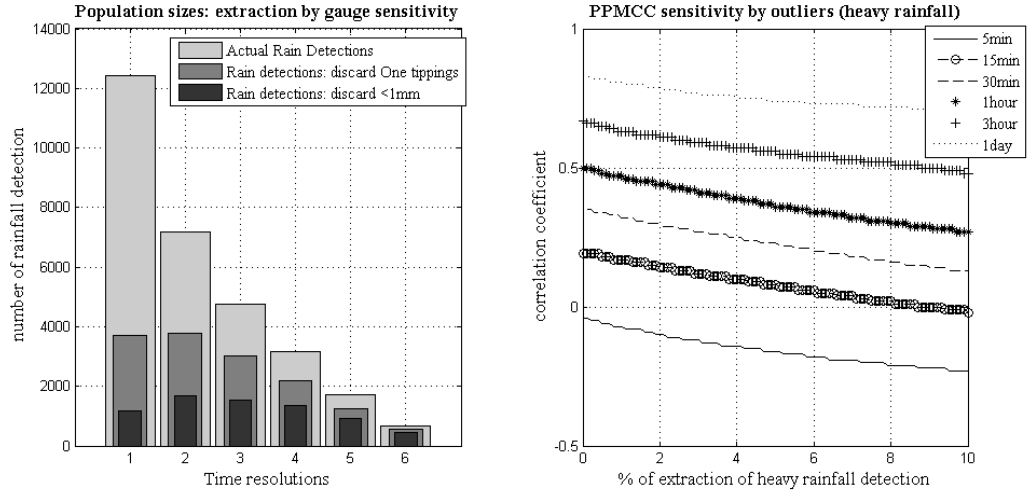


Figure 2.3 Comparison of rainfall data filtering on rainfall detection. Left side:

Rainfall detection after elimination of isolated single tip signals and rainfall detection after discarding values 1mm or less (left), PPMCC sensitivity to heavy rainfall (right)

This indicates a more consistent data record and smaller variation as possible outlier records are filtered. The number of isolated rain detections has the largest sample size regardless of temporal resolution. However, the sample size varies among the different temporal resolutions. Especially, at the daily resolution, there are relatively fewer numbers of paired rainfall detections. Therefore, the daily rainfall is not considered further in the analysis. In summary, this review indicates that more than two thirds of the rainfall detection records consist of an isolated one tipping in a 5 minute resolution, and the review of daily paired rain data indicate 20 percent of the records are “1-tip” rainfalls. These isolated 1-tip records may not be due to actual rainfall but are likely erroneous due to mechanical device calibration and sensitivity. A case for discarding rainfall less than 1-mm is evident since these isolated values have a small influence on sample size across temporal resolutions, and 1-mm rainfall depths would only influence applications at short

time scales. For these reasons, the 1-tip rainfall values were excluded. Another consideration is that the presence of even a few samples of extreme large rainfall depth may overestimate the representative PPMCC (Habib 2001). The nature of the correlation coefficient is sensitive to outliers, and so the influence of intense rainfall values is considered. On the other hand, the actual rainfall representative of the specific region must be incorporated, even when extreme values are present. Otherwise, the correlation coefficient cannot represent the entire domain of precipitation; the sensitivity and distortion of PPMCC by extreme rain values must be addressed carefully.

Figure 2.3 (right) shows the correlation moderately drops as larger portion of the high extreme rain values are extracted. The correlation coefficient values in this figure are the averaged correlation value for all pairs of gauges. The quality of the correlation at 5 minute temporal resolution is considered extremely poor and not used in the spatiotemporal correlation analysis.

As a result of this data quality control review, four temporal resolutions of gauge data remain for spatio-temporal correlation analysis exclusive of the daily and five minute resolution data. Although there remains uncertainty associated with the sample size necessary for this study, which addresses precipitation characteristics across a local region, the application of PPMCC has other obstacles that remain controversial in Statistics (Kowalski 1972; Hutchinson 1997).

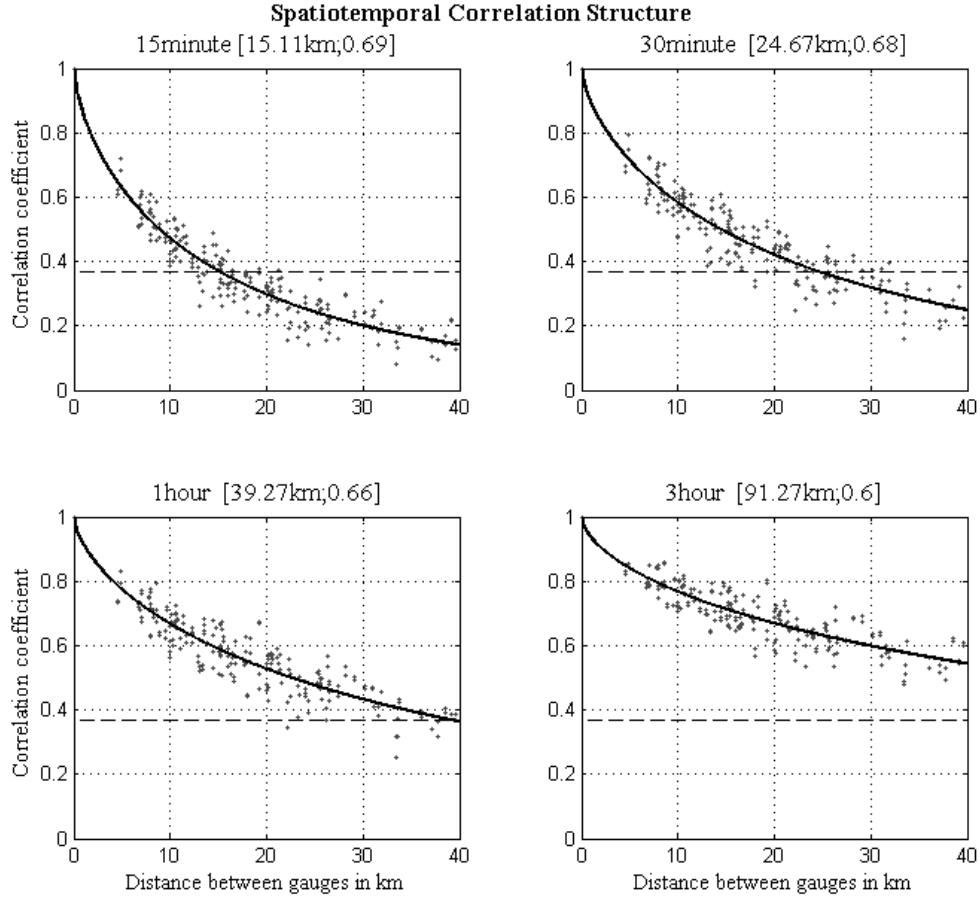


Figure 2.4 Spatiotemporal correlation structure: temporal resolution = 15-minute (upper-left), 30-minute (upper-right), 1-hour (lower-left), and 3-hour (lower-right). The critical distance (e-folding decay, $1/e$) and shape factor are shown in brackets

The nature of the distribution of rainfall intensity is skewed to the right even after trimming out the portion of 1-tip records. In other words, a Gaussian bell-shape distribution of precipitation and intensity is not physically possible in a hydro-meteorological field (Amburn 2015, Scheuerer and Hamill 2015). Nevertheless, this selection of data is meaningful and provides a way to define the spatiotemporal correlation while recognizing variation in sample size and dispersion. Finally, this

selection process identifies relevant data in a meaningful way to address concerns and allow objective application of the PPMCC to the precipitation phenomenon.

Results of the PPMCC study for the spatiotemporal rainfall variation in Louisville is shown in Figure 2.4. The concave tendencies of exponential decay are observed through the sub-hour temporal resolutions in the scattering plots. The best-fit line is derived from Equation (2) (Ciach 2006).

$$R_{i,j} = \exp \left[- \left(\frac{d_{i,j}}{d_c} \right)^s \right], 0 \leq s \leq 2 \quad (\text{Equation 2.2})$$

Due to occurrence of multiple gauge pairs in the same space within several meters, the nugget parameter is not considered, but the critical distance, d_c , and shape parameter, s , were found. The two parameters were found using the minimum value of Root-Mean-Square-Error (RMSE) and the exhaustive iteration of 0.01 km and 0.01 incremental resolutions of critical distance and shape parameter respectively. Using this approach, the critical distance corresponds to the e-folding decay (1/e) correlation. The critical distance is meaningful when comparing results of PPMCC applications from other regions. In the Louisville Metro region, the spatial variation is relatively smaller than other regions in other selected studies. The critical distance varies from 14.84 km to 89.32 km for the 15 minute to 3 hour temporal ranges respectively. The critical distance is primarily governed by the regional hydrologic climate regime and is slightly less than other study areas (Ha 2007; Villarini 2008).

The shape of the fit-line in figure 2.4 represents the decay rate and ranges from 0.72 km^{-1} to 0.61 km^{-1} for 15 minute to 3 hour temporal resolution respectively. Considering the smallest temporal scale (15 minute temporal resolution), the spatial variation is most

significant. The rate of correlation decrease within the 10 km range and 15 minute resolution is sharp and decay rate flattens as temporal resolution increases. At the 3 hour time resolution correlation remained above 0.8 to 7.5 km distance. The PPMCC variation with distance functions as a good descriptor of the linearity of rainfall across the space and time scales.

2.3.2 Quantile/threshold-range effect on correlation

The PPMCC depends on the distance between gauges as previously stated. Correlation decreases with distance in a concave form across all temporal resolutions. In consideration of hydrologic climate, a variety of seasonal and event-specific precipitation types and associated intensity regimes exist and become summarized in the rainfall record. In order to further understand the regional rainfall regime an analysis of correlation coefficient stratified by rainfall intensities is conducted. The variation in correlation coefficient with the rainfall intensity, as denoted by quantiles across a series of distance ranges is investigated.

The quantiles are established as shown in figure. 2.5. The left and right bar plots represent the average rainfall depth and corresponding standard deviation for which the rainfall data records fall into the quantile classes.

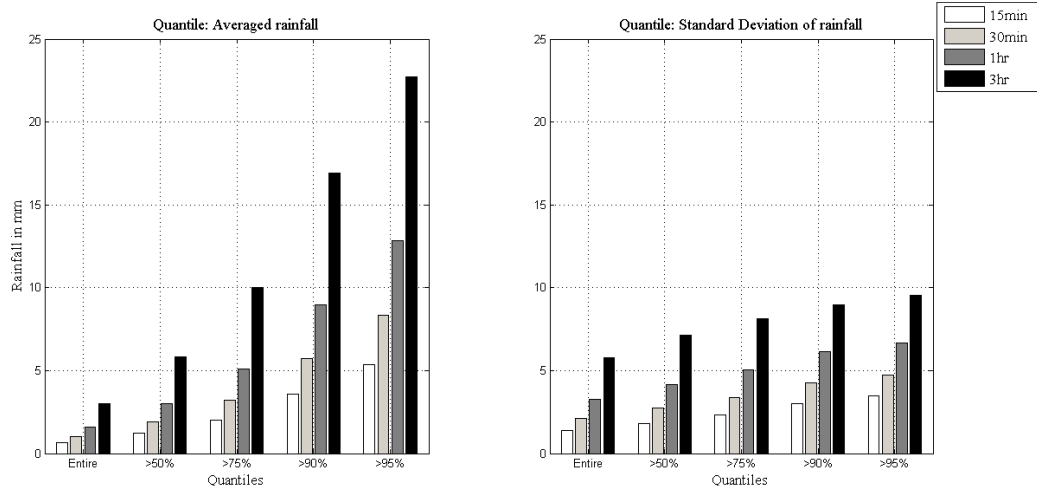


Figure 2.5 Quantile rainfall summary for time-averaged depth (mm) (left) and rainfall depth standard deviation (mm) (right) – for categories: entire domain, 50% or higher, 75% or higher, 90% or higher, and 95% or higher

Results of the study of correlation variation with distances and thresholds of rainfall are provided in figure 2.6. The horizontal and vertical axes represent the west-east and north-south directional distance. Due to the shape of the study area, the range in east-west direction is about 40 km, but the maximum north-south distance is less than 30 km. In figure 2.6, overall, the correlation increased proportionally with a decrease in temporal resolution. Correlation values with magnitude less than 0.5 dominate. Alternatively, Figure 2.6 provides a convenient and efficient format to visualization rainfall behavior and corresponding variation of correlation with range and specific direction. Correlation values greater than 0.7 remarkably stand out in each of the hourly temporal resolution results at the shorter distance range. This result indicates adjacent rainfall depths recorded within 20 km over a time period of 1-hour or longer, are expected to be associated with a correlation above 0.7. The correlation is shown to be strongly influenced by rainfall intensity and

distance range at finer temporal resolutions. The spatial correlations decrease as rainfall intensity increases.

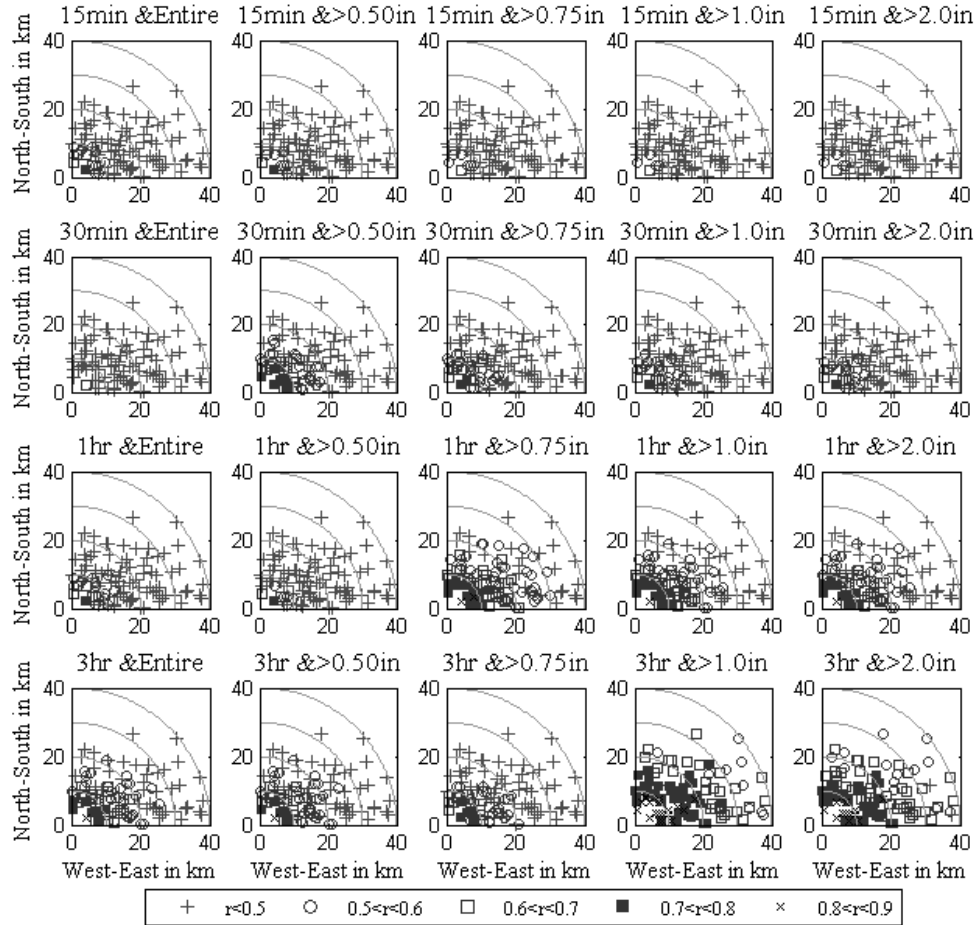


Figure 2.6 Correlation spatial variation (distance and direction) for paired gauges according to thresholds on depth and duration (East-West is x axis and North-South is y axis) across the study area. Correlations shown for time intervals: 15 min, 30 min, 1 hr, and 3 hr, and depths: all depths (entire), 0.50 inch or higher, 0.75 inch or higher, 1.0 inch or higher, and 2.0 inch or higher

This may indicate the higher spatial variation typical of severe and convective heavy rainfall at sub-hour scales. On the contrary, at the hourly temporal resolution, the tendency of increasing correlation was observed as long as rainfall intensity increased and a threshold of rainfall greater than 0.75inch. The similar tendency is observed at 3 hour temporal resolution and rainfall threshold 1inch or greater.

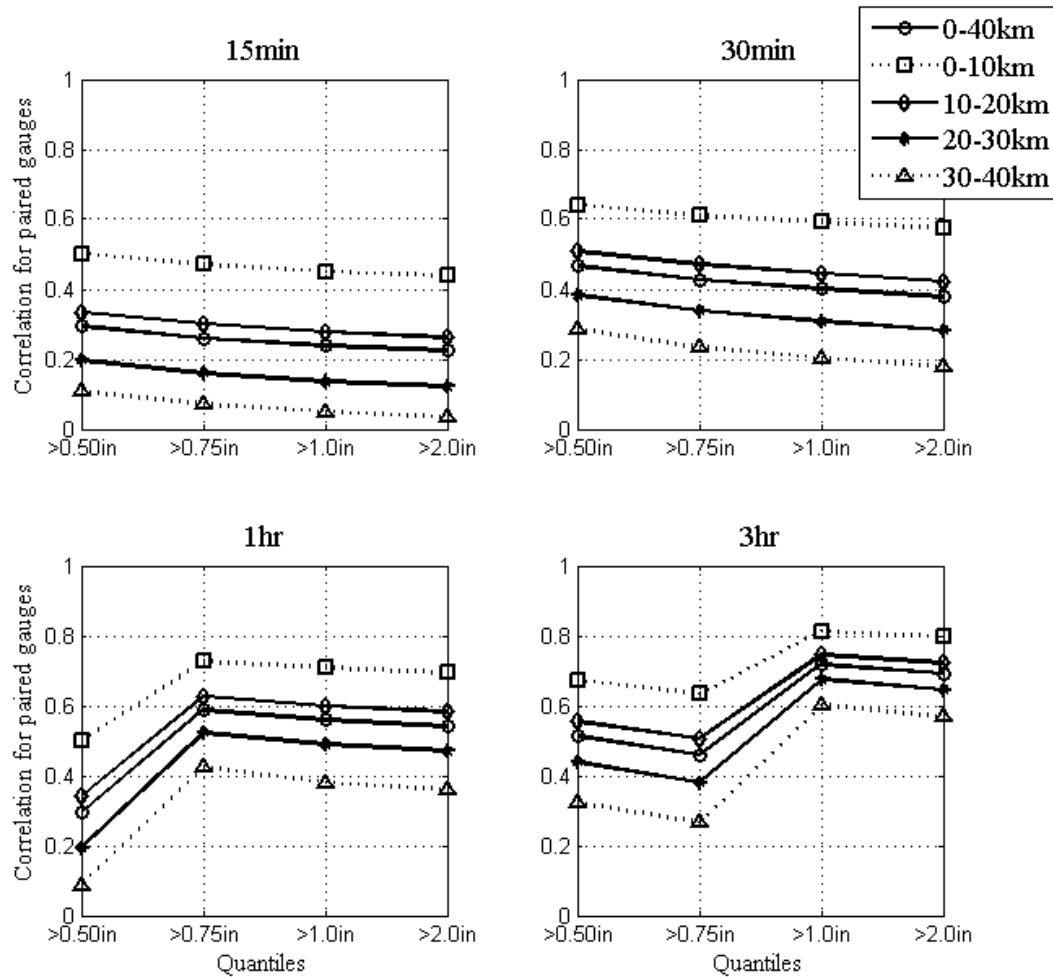


Figure 2.7 Correlation structure variation for accumulation interval (15 min, 30 min, 1 hr, 3 hr) by rainfall threshold (0.50 inch or higher, 0.75 inch or higher, 1.0 inch or higher, 2.0 inch or higher) and distance range (0-40 km, 0-10 km, 10-20 km, 20-30km, 30-40 km)

In figure 2.7, the plots show the variation in correlation coefficient averaged with distance range and for each rainfall intensity threshold. Overall, across all temporal resolution categories, the variation of correlation with distance interval is similar. The averaged correlation decreases with increase in distance, i.e., the nearest distance range has the highest correlations throughout all quantiles. The notable point in Figure 2.7 is that the response of the correlation due to the rainfall intensity differs for each temporal resolution. In the coarse temporal resolutions, 1-hourly and 3-hourly, there is a tendency toward increasing correlation magnitude as the rainfall intensity increases. The correlation magnitudes abruptly increase at the rainfall intensity of 0.75 inch and 1.0 inch for 1hourly and 3hourly the temporal resolutions respectively without reference to the distance ranges. It means that the quality of the rainfall data by rain gauge within sub-daily temporal resolutions is consistent, especially in heavy rainfall analysis. On the contrary, the correlations continuously fall in the sub-hourly the temporal resolutions (15 minute and 30 minute). In short, heavy rainfall data records for sub-hourly temporal resolution show a decrease in correlation with rain intensity. The hydrologic climate region, meteorological rainfall pattern, and areal extent of the study area contribute to these characteristics. In a mesoscale region, each gauge in the network influences and contributes to not only the distance data available but also to the definition of the temporal resolution and rainfall intensity details.

2.4 Results and conclusions

The PPMCC depends on the distance between gauges as previously stated. Correlation decreases with distance in a concave form across all temporal resolutions. In consideration of hydrologic climate, a variety of seasonal and event-specific precipitation types and

associated intensity regimes exist and become summarized in the rainfall record. In order to further understand the regional rainfall regime an analysis of correlation coefficient stratified by rainfall intensities is conducted. The variation in correlation coefficient with the rainfall intensity, as denoted by quantiles across a series of distance ranges is investigated.

The quantiles are established as shown in figure 2.5. The left and right bar plots represent the average rainfall depth and corresponding standard deviation for which the rainfall data records fall into the quantile classes. Results of the study of correlation variation with distances and thresholds of rainfall are provided in figure 2.6. The horizontal and vertical axes represent the west-east and north-south directional distance. Due to the shape of the study area, the range in east-west direction is about 40 km, but the maximum north-south distance is less than 30 km. In figure 2.6, overall, the correlation increased proportionally with a decrease in temporal resolution. Correlation values with magnitude less than 0.5 dominate. Alternatively, figure 2.6 provides a convenient and efficient format to visualization rainfall behavior and corresponding variation of correlation with range and specific direction. Correlation values greater than 0.7 remarkably stand out in each of the hourly temporal resolution results at the shorter distance range. This result indicates adjacent rainfall depths recorded within 20 km over a time period of 1-hour or longer, are expected to be associated with a correlation above 0.7. The correlation is shown to be strongly influenced by rainfall intensity and distance range at finer temporal resolutions. The spatial correlations decrease as rainfall intensity increases. This may indicate the higher spatial variation typical of severe and convective heavy rainfall at sub-hour scales. On the contrary, at the hourly temporal resolution, the tendency of increasing correlation

was observed as long as rainfall intensity increased and a threshold of rainfall greater than 0.75inch. The similar tendency is observed at 3 hour temporal resolution and rainfall threshold 1inch or greater. In figure 2.7, the plots show the variation in correlation coefficient averaged with distance range and for each rainfall intensity threshold. Overall, across all temporal resolution categories, the variation of correlation with distance interval is similar. The averaged correlation decreases with increase in distance, i.e., the nearest distance range has the highest correlations throughout all quantiles.

The notable point in figure 2.7 is that the response of the correlation due to the rainfall intensity differs for each temporal resolution. In the coarse temporal resolutions, 1-hourly and 3-hourly, there is a tendency toward increasing correlation magnitude as the rainfall intensity increases. The correlation magnitudes abruptly increase at the rainfall intensity of 0.75 inch and 1.0 inch for 1hourly and 3hourly the temporal resolutions respectively without reference to the distance ranges. It means that the quality of the rainfall data by rain gauge within sub-daily temporal resolutions is consistent, especially in heavy rainfall analysis. On the contrary, the correlations continuously fall in the sub-hourly the temporal resolutions (15 minute and 30 minute). In short, heavy rainfall data records for sub-hourly temporal resolution show a decrease in correlation with rain intensity. The hydrologic climate region, meteorological rainfall pattern, and areal extent of the study area contribute to these characteristics. In a mesoscale region, each gauge in the network influences and contributes to not only the distance data available but also to the definition of the temporal resolution and rainfall intensity details.

CHAPTER 3 RAINFALL CLASSIFICATION

3.1 Introduction

Throughout many regions of the world, densely populated urban communities face serious flooding issues. In recent years, the severity of damage has increased due to both higher variation in rainfall extremes and urban expansion (Qin and Lu 2014; Salathé et al. 2014; Wang et al. 2014). In urban areas, traditional hydrologic design methods, including selection of rainfall amount from historical rainfall frequency duration curves, often does not adequately describe local or more recent observed rainfall characteristics (Einfalt et al. 2009). An understanding of local and regional rainfall intensity and frequency characteristics for specific watersheds is necessary to implement plans and design procedures directed toward reducing sewer system overflows and reducing inundation of urban properties. This implies that traditional historic rainfall summaries may not characterize recent rainfall extremes, producing unexpected extreme floods. Improved understanding of ground-level rainfall intensity and spatial variability in urban areas can be thus be useful.

The US Environmental Protection Agency (EPA) regulates the amount and frequency of urban runoff from combined sewer systems that may overflow into natural streams, or combined sewer overflow (CSO) events. Municipal sewer agencies must attempt to meet EPA mandated constraints and conditions minimizing CSO occurrence or face penalties.

Typically, CSO occurrences are related to identifiable rainfall event characteristics such as depth and duration, and result in regulatory agencies such as the EPA having an interest in understanding rainfall variability. This work identifies rainfall events using the EPA inter-event time criteria for urban areas. The rain events are clustered using a K-means method based on three pre-assigned groups: low-intensity, high-intensity and high volume. The spatial variation of rainfall events which fluctuate in size, rainfall intensity, duration and total rainfall volume are evaluated using a variogram analysis. This information can provide guidance to plan and evaluate hydrologic measurement and design planning in the urban area.

3.2 Data and preprocessing

The study region is meso-scale city, the Louisville metropolitan area which is about 1022 km² (399 mi²) and a primarily urbanized area with a population of 1.3 million by 2013 census. This study focuses on a recent period with a complete record of rainfall available, January 2010 to December 2014. The investigation characterizes rainfall events in a practical and meaningful way in order to understand rainfall spatial and temporal variability in the context of rainfall events as defined by the United States Environmental Protection Agency (USEPA).

The data records from an operational rain gauge network, deployed and maintained by the local Metropolitan Sewer District (MSD) agency responsible for flood protection, stormwater runoff, and wastewater treatment in Jefferson County Kentucky, were utilized. The rain gauge locations are spread across the Louisville metro area as shown in Figure 3.1 (MSD gauges indicated by circle markers). There are 17 rain gauges in the MSD network as well as NOAA-NWS rain gauges (triangle markers). The minimum and

maximum distances between closest gauges are 4.52-km and 11.06-km respectively. The NOAA gauges serve as a reference for data quality control through comparison with data recorded by the MSD network. The MSD gauges are tipping bucket type with 0.254 mm (0.01 in) resolution and report accumulated rainfall every 5 minutes.

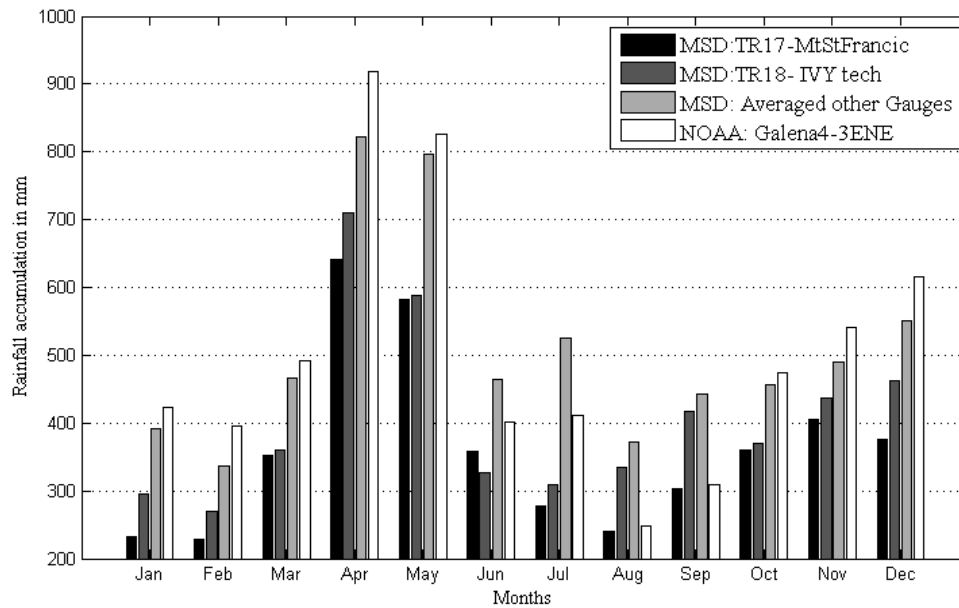


Figure 3.1 Monthly accumulation of rainfall of TR17, TR18, average value of other MSD gauges, and NOAA rain gauge at Galena, IN for 2010-2014

Two MSD gauges, Mt. Saint Francis Seminary (TR17) and Ivey Tech (TR18), were found to consistently underestimate rainfall depth relative to the network-averaged MSD monthly accumulation and NOAA gauge records as summarized in figure 2.2 (Hyun et al. 2016). Additionally, these two gauges are located in the northwestern region beyond the Jefferson County boundary. Data records from the single nearest neighboring NOAA gauge are compared to TR17 and TR18 as well as the average value from all MSD gauges (excluding TR17 and TR18). The NOAA comparison gauge, Galena 4_3ENE, was chosen due to its close proximity to the two MSD gauges, 5.8km and 9.4km from TR17 and TR18

respectively. Based on the information summarized in figure 3.1, both TR17 and TR18 data are considered biased for unknown reasons and these data are excluded from this work.

To address data quality validation further, twelve NOAA rain gauges in the study area (figure 2.2) were utilized for daily accumulation comparisons. Comparison of data records from different institutes can serve as an acceptable substitute to validate the rain gauge system in practical situations. The daily data are a component of the archive of Global Historical Climatology Network–Daily (GHCN daily) Version 3 (Menne et al. 2012). For daily accumulation comparison, each MSD gauge is paired with the nearest NOAA gauge. Figure 2.2 shows a Pearson’s coefficient range of 0.85 to 0.94 for all gauge pairs. The expected decrease in correlation with distance is observed and daily accumulation comparison indicates agreement between the NOAA and MSD rain gauge networks (Hyun et al. 2016).

3.3 Methodology

Initial work required identification of rainfall events from the data records. For hydrologic studies with regulatory implications, a rainfall event may be defined in terms of depth, duration, intensity, and runoff. For example, municipal agencies responsible for urban runoff management may be required to refer to the EPA regulatory definition of the inter-event time (IET). The concept of partitioning rainfall into events with specific durations or non-rain periods follows from the EPA rain event definition (Driscoll et al. 1989): “An underlying assumption necessary for the manipulation of probability density functions is that the event must be independent. One of the requirements associated with storm event analysis is selecting an appropriate inter-event time (IET). IET values of about 6-hours are found to be suitable for locations in the eastern part of the country”,

and this is done by grouping hourly values in accordance with the minimum number of consecutive dry hours. The EPA definition provides the inter-event time, IET, and further defines a rainfall event as (Driscoll et al. 1989): “A minimum storm volume of 2.54-mm (0.1-inch) was specified for the analysis performed, so that the analysis would produce statistics of ‘runoff producing’ events within 6-hours.”

This study adopted the EPA IET definition and considered a 6-hour time window as the initializing constraint defining the urban watershed state to be in a dry initial condition, and the constraint was found to provide an acceptable partitioning of rainfall-runoff events for the study area. Based on this IET, there are 558 rainfall events in the study period. When the minimum volume of rainfall was observed among 15 rain gauges, rainfall event duration continued until the last minimum rainfall was captured by the rain gauge network.

It is challenging to identify metrics to fully quantify the physical continuity of natural rainfall processes. In this section, the continuous spatial structure of rainfall events is investigated using the clustered-variogram. Prior to determining the variogram, it is critical to assume that the rainfall meets the stationary status. In general, climatic phenomena have non-stationary characteristics and a simple application of a mean value to the variogram may not be suitable for the entire study period. In order to lessen the influence of non-stationary characteristics, the K-means clustering method was applied to objectively partition rainfall events by similarity in intensity and similarity in average rain depth as shown in figure 4 (Ciach and Krajewski 2006; Dong et al. 2013; Jung et al. 2014; Khalid 2011; Tokay and Öztürk 2012; Zhang and Yan 2014). The rainfall events spread across the 2-dimensional rainfall intensity-volume field (upper-left) and three

cluster centers were selected based on the ideal position for each group. For example, the initial centroid at the upper-left corner represents the centroid of the high intensity group. The finalized cluster areas are: (1) high intensity events in the upper left cluster, (2) high depth events at the lower right cluster, and (3) lower depth, lower intensity events in the cluster area near the origin. The K-means process uses a weighting to define the centroid of each region, and once centroid locations are within a resolution of 0.001 in the normalized field the solution is complete; results are as shown in the lower right of figure 3.2. There are 358, 163, and 37 events in the light event, high intensity event, and high depth event clusters respectively. The light rainfall cluster data were not considered beyond this point since interest is in runoff and flood producing rainfall, and only the extreme events in intensity or volume were investigated.

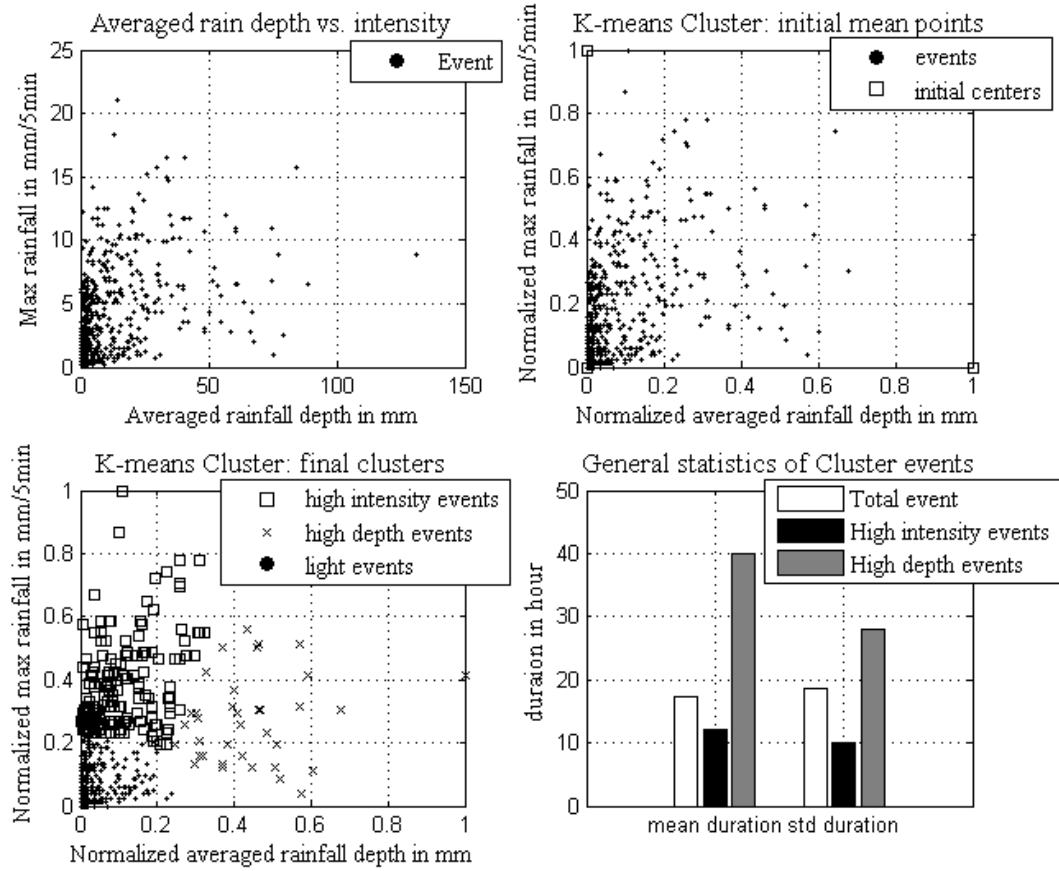


Figure 3.2 Rainfall event clustering classification as light rainfall, high intensity, and high depth groups using the K-means clustering method: Initial status of scatter plot of averaged rain depth vs. maximum rainfall intensity (upper-left); Predefined centers for each group; light rainfall event at the origin, high intensity rainfall event at upper left corner, and high depth rainfall event at lower right corner in the 2-dimensional surface (upper-right); Final K-means clustered groups (lower-left), General statistics, mean and dispersion of duration for clustered groups (lower-right)

In the lower right plot of figure 3.2, a notable difference in mean duration is indicated. Duration is an important factor to characterize rainfall and the high intensity rainfall cluster takes place well within a day and close to a half-day mean duration.

Meanwhile, the high depth rainfall events show duration closer to two days. These two mean values were applied using Equation 3.1 to establish the clustered-variogram (Cheng et al. 2008).

$$\gamma_{i,j} = \frac{1}{2}E \left[\left((G_i - G_c) - (G_j - G_c) \right)^2 \right] \text{ (Equation 3.1)}$$

As expected, the two rainfall clusters responded differently in spatial variation characteristics. The high intensity event cluster is more variable in terms of total rain depth compared to the high depth event cluster.

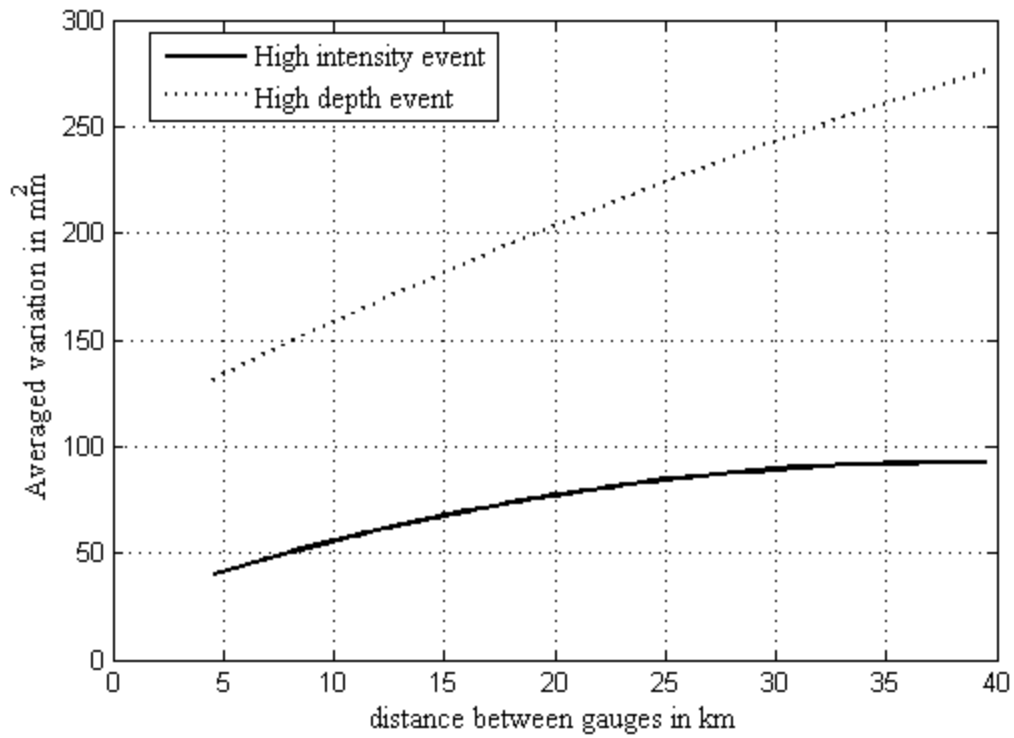


Figure 3.3 Spatial-variogram of the two rainfall groups, high intensity rainfall event group and high depth rainfall event group

Figure 3.3 shows the second order poly-fit lines for scatter plots of the variation in event accumulations. The high intensity cluster reaches a sill, with asymptotic value around 30-km, but the high depth event cluster sill is beyond the size of the study area. This indicates that storms with larger intensity exhibit less uniform rainfall intensity during an event (less than 1-day), or that a series of intense storms with strong convective cores pass through the study area. However, the larger volumetric rainfall events have relatively smaller variation spatially and this is likely associated with the longer duration of the event (2-day) and the storm characteristics are more stable and uniform. These common meteorological conditions may explain the lower spatial variation with longer range variogram in the high depth group.

3.4 Conclusion and Discussion

Rainfall event characteristics for the Jefferson County Kentucky Metro region were evaluated using data records from a local rain gauge network. This work is relevant in the context of understanding rainfall event characteristics in the context of regulatory requirements and identification of urban rainfall events leading to flooding and sewer overflows.

The K-mean clustering method identified the extreme rainfall event groups resulting in an intensity-based cluster and a volume-based cluster. The two cluster groups indicate a notable difference in mean event duration. The clustered-variogram reveals a change in magnitude and influential range (sill) of spatial variation for each cluster group. The high intensity cluster contains higher intensity, short duration events and indicates higher spatial variation relative to the volumetric extreme, longer duration cluster.

In conclusion, this study identifies rain event spatial structure using the clustered variogram of ground rainfall detected by a gauge network. This can be helpful in terms of understanding historical rain event patterns, spatial distribution, and managing deployment of rain gauges. The remarkable findings include that spatial variation in rain events depends not only on distance but also rainfall types: volume and intensity. Moreover, partitioning the rainfall record into rain events according to a defined IET may influence the number and characteristics of rain events identified in the historical record. This may further have an impact for small-scale watersheds in the urban area where runoff and sewer overflows are evaluated.

CHAPTER 4 RADAR RAINFALL OPTIMIZATION

4.1 Introduction

In this era of climate variation and weather extremes, heavy rain storms have emerged as a critical issue due to the associated increase in damage to community facilities and displacement of resident populations. Hydrological disasters, while anticipated in tropical and subtropical regions are becoming more common in other regions. Extreme rainfall is ubiquitous and remains difficult to predict with certainty. This is partially due to the rapid formation, high intensity and localized character of convective storms and results in a perceived increase in flood severity threat for entire communities and economies (Kundzewicz et al. 2014; Qin et al. 2014; Wernstedt and Carlet 2014). Scientists and engineers seek to improve infrastructure design and management of runoff control facilities to prevent or alleviate hydro-natural disasters (Wang et al. 2015; Woodward et al. 2014). A critical component in storm runoff management and mitigation is a real-time rainfall measurement system. Defining rainfall spatial and temporal quantities (depth, duration, intensity, areal extent) in real-time enables stormwater managers to plan and enact strategies to alleviate flood impacts.

To acquire real-time and historical records of rainfall, ground-based and remote-sensing measurement systems are commonly used. Instrumenting a region with a network of rain gauges is advantageous as the gauges directly measures pluvial water at the ground level. However, a rain gauge provides only a depth-time record with limited

spatial variation information. Thus, the quality of regional rainfall estimates from gauge networks is proportional to the density and distribution of gauges. In operational applications, the nearest rain gauge may be several kilometers from the catchment where rainfall amounts are required.

A means to define the spatial variation of rainfall and supplement the rain gauge network measurement is available through weather radar systems. The National Weather Services (NWS) operates the NEXRAD (Next Generation Radar) weather radar network which provides real-time monitoring of weather system occurrence, location and movement over the radar surveillance region. The challenge, however, is that observations are in form of radar reflectivity and this value must be transformed into a ground level rainfall estimate.

Many researchers have worked to develop algorithms that will transition radar reflectivity data to ground-level rainfall characteristics (Baeck and Smith 1998). As an example, rainfall intensity may be inferred from radar reflectivity (back-scattered radar power, dBZ) based on an exponential relationship (Z-R conversion) implicitly incorporating physics of Rayleigh scattering and assumptions regarding in-cloud drop size distribution (Schmidt et al. 2012). The exponential power relationship ($Z=aR^b$) between reflectivity (dBZ) and rainfall intensity (mm/hour) is illustrated in figure 4.1. The appropriate Z-R parameters (a, b) are typically identified empirically, and in the case of the NWS weather radar system there are four standard Z-R relations for the following rainfall types: convective, tropical, east cool stratiform and stratiform (Hogan 2007; Krajewski and Smith 2002). To generate rainfall products, one of the four Z-R relationships is assigned to each radar site (for a specific length of time or season); the default rainfall type

is convective where parameter “a” is 300 and “b” is 1.4. At a reflectivity level near 36 dBZ (vertical dashed line on the right-side Figure 4.1) the four Z-R relationships begin to diverge from one another. Thus, above this value, the importance of selecting the appropriate Z-R relationship of the Z-R relationship increases. Unfortunately, the Z-R relationship is not typically calibrated for a particular hydrologic climate or rainfall type and no real-time automated optimization is implemented (Chumchean et al. 2003; Ice 2014).

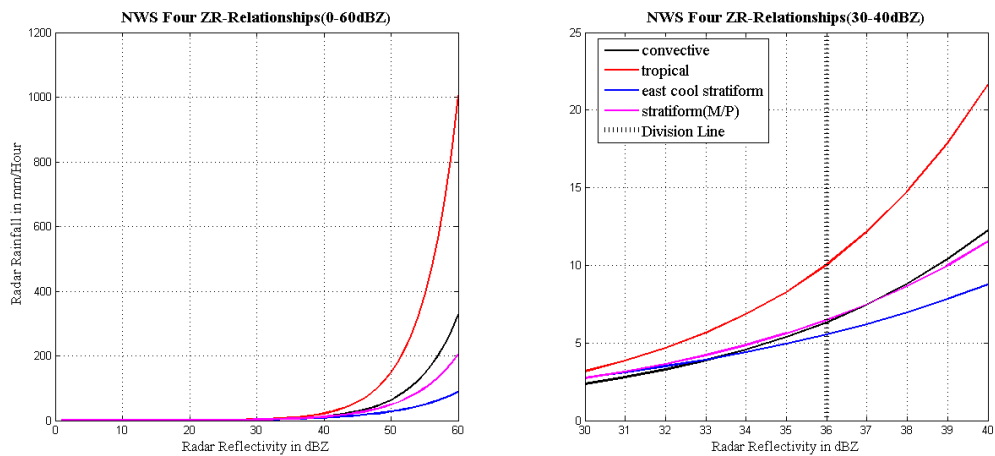


Figure 4.1 NWS Z-R relationships by storm type, 0-60dBZ range (left) and 30-40dBZ range (right)

While existing research has been directed to developing Z-R relationships for one-hour rainfall accumulations (Baek and Smith 1998), this study focused on linking radar reflectivity to rain gauge networks for short duration applications (less than one-hour). By synchronizing radar rainfall with rain gauge measurements the dependency (and associated uncertainty) of the Z-R conversion on storm type classification (convective, tropical, east cool stratiform and stratiform) is diminished.

Merging information from both the ground-level rain gauge and weather radar systems to assess real-time rainfall characteristics can improve stormwater management practices for small catchments in urban areas. Currently, NWS products available to support flash flood and storm warnings include the AFWS (Automated Flood Warning System) and IFLOWS (Integrated Flood Observing and Warning System) rain gauge network operates in remote and rural areas with inter-gauge distances from near 5 to over 15 km and reports 15-minute accumulations, to the Advanced Hydrologic Prediction Service product for daily precipitation at 16 km² for the entire USA (NWS 2016). While both NWS products serve a useful purpose, neither provides the type of rainfall spatial and temporal detail to meet the needs for urban catchment runoff estimation.

Synchronizing radar reflectivity data with the rain gauge network, a more precise estimate of rainfall (depth, spatial and temporal variations) indicate improved rainfall estimates at scales of 0.5 km² and 0.5 hours. Optimizing spatial and temporal estimates of rainfall variation can improve stormwater modeling and runoff estimation for small urban watersheds.

4.2 Material and methods

The study region for this work is the city of Louisville, Kentucky (USA) during the period January 2010 to December 2014. Rainfall data from an operational rain gauge network, managed by the local utility agency Metropolitan Sewer District (MSD), are the ground reference values (Hyun et al. 2016). The region's Next-Generation Radar (NEXRAD) (denoted by call letters KLVX) is located at Fort Knox, Kentucky about 40-km southwest of the city of Louisville.

In case studies of rainfall spatiotemporal structure, a correlation near 0.6 at 5-km distance for quarter-hour temporal resolution using ground based rain gauges is shown (Ciach and Krajewski 2006; Jung et al. 2014; Mandapaka and Qin 2013). The average inter-gauge distance in this application is slightly greater and the gauge network can therefore benefit from the complementary spatial detail provided by weather radar. Although the radar data are not explicitly filtered for error adjustment, the large quantity of data compiled for use, from gauge network and radar archives, is expected to minimize bias. Additionally, the proximity of the study area relative to the radar site, at about 40 km range, is expected to diminish the influence of common radar error influences such as range effects of signal attenuation, anomalous propagation, beam blockage, and beam spreading (Ciach et al. 2003; Gorgucci and Baldini 2015; Hunter 1996; Kalogiros et al. 2013; Krajewski and Vignal 2001; Morin et al. 2003; Seo et al. 2000; Vignal and Krajewski 2001).

The support vector classification (SVC) application not only serves to partition the storm events by underlying characteristics and identify the appropriate Z-R relationship, but includes an optimization process for Z-R parameter estimation.

4.2.1 Data sources

The fifteen rain gauges of the MSD network are mechanical tipping-bucket type with resolution of 0.01 inch and temporal interval of five minutes. The data records for radar base-scan reflectivity (Level II - NEXRAD dual polarization, 0.5 dBZ increment) were retrieved from the National Climatic Data Center (NCDC) at National Oceanic and Atmospheric Administration (NOAA). The data cover the entire five-year study period, 2010-2014, in the format of coded reflectivity volume scans. The raw reflectivity was

converted to a Cartesian coordinate (ESRI ASCII grid files) system using NCDC's Weather and Climate Tool Kit (WCT), version 3.7. The WCT provides visualization and export tools to manipulate radar data. The Constant Altitude Plan Position Indicator (CAPPI) data, 1-km above the radar elevation, forms the base-scan reflectivity array over the study area. The spatial pixels are approximately 220-m square in a Cartesian coordinate grid over the study area, thus pixel area is less than 5-hectare (15-acre).

4.2.2 Data preprocessing

The raw radar reflectivity data are instantaneous values and require conversion to rainfall intensity and accumulation to define volumetric rainfall. In order to geosynchronize reflectivity pixel locations with reference network gauge locations, radar reflectivity pixels with spatial locations corresponding to MSD gauge locations were identified. In the same vein, a temporal synchronization was performed to identify radar scan time stamps corresponding to the local time zone (accounting for daylight savings time as appropriate). For each rain gauge site, the collocated radar pixel and adjacent eight pixels were identified for use in data evaluation. Among these nine radar pixels, the single collocated radar pixel value was selected when reflectivity was within 50% of the average of surrounding pixel values. Where the difference between the center pixel value and surrounding values was more than 50%, the averaged reflectivity value was assigned. In the case where a majority of the 9 radar reflectivity data showed a status or condition as “not available” the pixel status was defined “not available”. The histogram in figure 4.2 shows the frequency distribution of reflectivity for the 15 gauge locations over the length of the study period. The distribution of reflectivity is understandable in the nature of the

precipitation, skewed to the right. The maximum frequency falls near 55 dBZ and the most frequent range of the reflectivity is near zero.

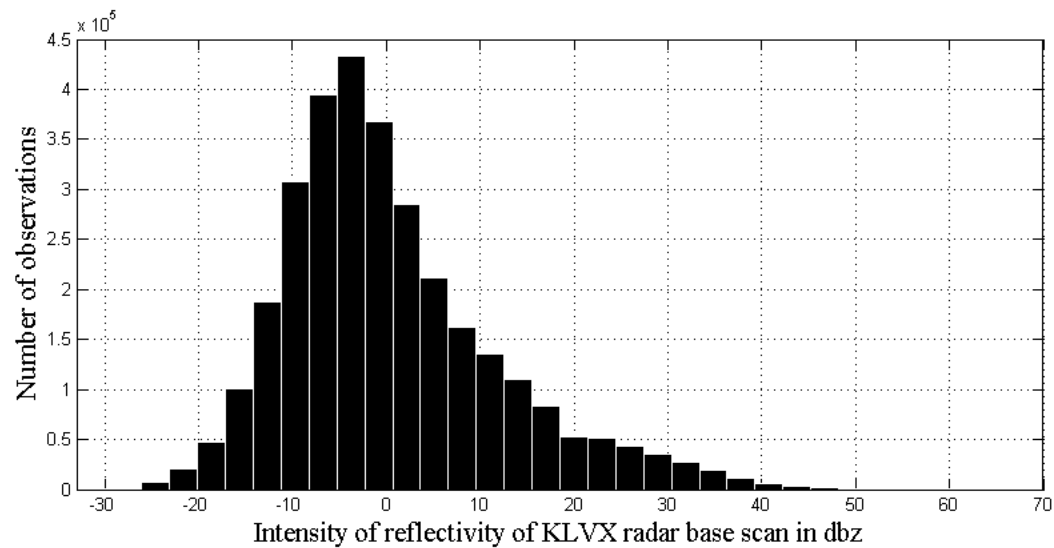


Figure 4.2 Histogram of Level 2, radar reflectivity at KLVX, Fort Knox, Kentucky,
Jan. 2010 – Dec. 2014

The KLVX radar data management system applies a Z-R relationship as described earlier according to four storm types: convective, tropical convective, stratiform, and east cool stratiform. Reflectivity transforms into rainfall intensity as an instantaneous value, whereas the gauge values are demonstrate accumulated rainfall over five minutes (Ulbrich and Lee 1999). Therefore, instantaneous intensity is further transformed into accumulated rainfall following application of the Z-R relationship. The first step considers all four reflectivity to rainfall intensity (mm/hr) conversions. Additionally, temporal synchronization was required since radar observations are not recorded at equal or uniform temporal intervals. Instead, reflectivity scan intervals cover a 4- to 10-minute range due to the operational mode of the weather radar. Generally, three or four instantaneous base scans cover the quarter-hourly period, and each volumetric scan is weighted according to

the inter-scan time interval within the quarter-hour interval. The product of rainfall intensity and time interval provides a part of the quarter-hourly volumetric rainfall accumulation.

$$R_Q(t, x) = \sum_{i=1}^n \left(\frac{R_i}{4} \right) \left(\frac{t_{i+1} - t_i}{2} \right) \quad (\text{Equation 4.1})$$

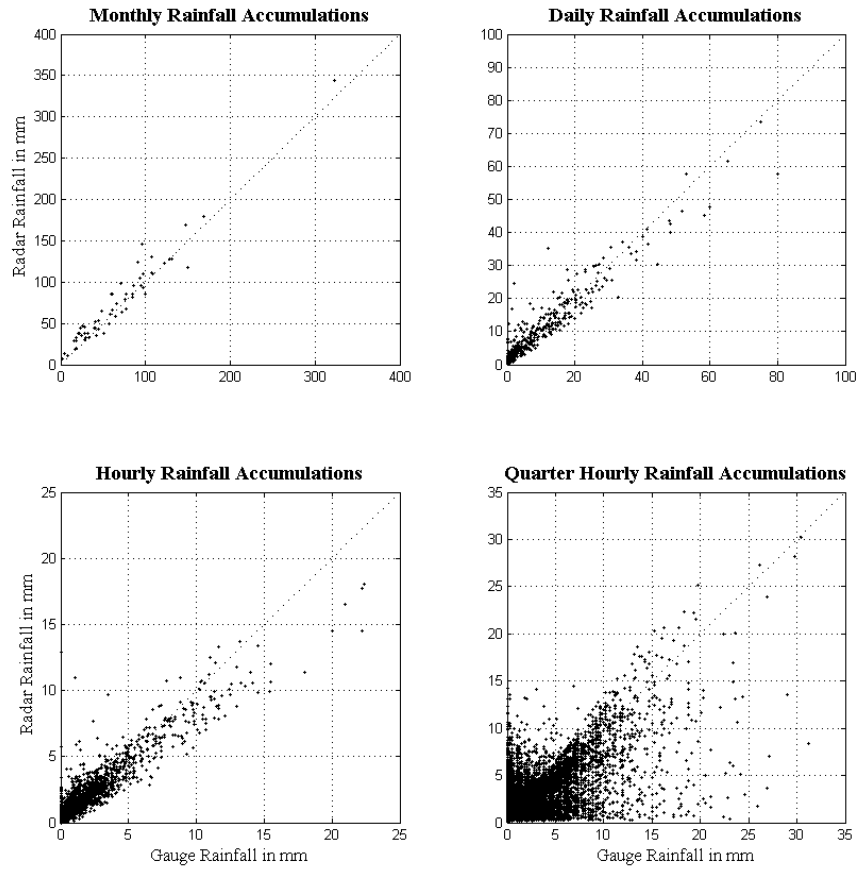


Figure 4.3 Gauge and Radar rainfall depth in 2 dimensional space (gauge-radar volume) across temporal resolutions: monthly (upper- left), daily (upper-right), hourly (lower-left), quarter-hourly (lower-right)

The process is defined in equation 4.1 based on a weighted time of occupation for each scan within the quarter-hourly interval. For each reflectivity volume scan, the time interval is defined as the duration from the observation until the next observation recorded (inter-scan interval) or the end of the fifteen-minute accumulation window.

Figure. 4.3 illustrates the volumetric radar rainfall products at monthly to quarter-hour temporal resolutions. The quality of quarter-hourly radar rainfall estimation is relatively low while hourly and longer accumulated rainfall products reveal better agreement with gauge rainfall. However, this study focuses on the shorter duration, quarter-hourly interval, in order to illustrate radar rainfall products for use in smaller urban catchment applications (Cunha et al. 2015; Krajewski et al. 2010; Smith et al. 2007; Wright et al. 2014).

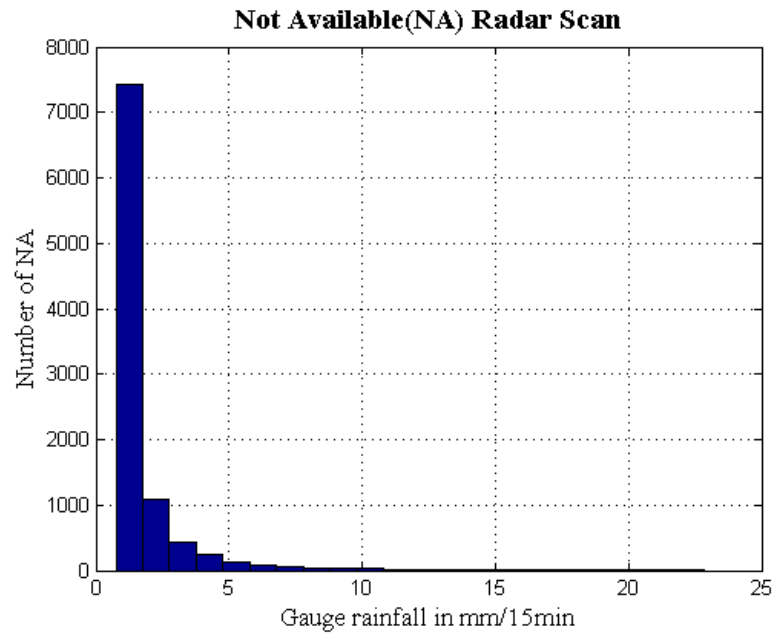


Figure 4.4 Histogram of rain gauge values for cases where radar data is not observed (Not Available (NA)) or radar intensity is below the 5mm/15min threshold

In figure 4.3, most data are found at the depth range of 5 mm or less for quarter-hourly rainfall and figure 4.4 demonstrates that a majority of this low- depth or low-intensity radar data is associated with low-intensity rainfall. Furthermore, rain gauge data reliability and detection is also sensitive during light rainfall (Ciach 2002; Humphrey et al. 1997). For these reasons, a quarter-hour rainfall accumulation threshold of 5 mm was implemented for evaluating the Z-R relationship in the remaining portion of this study. This is also in line with a focus on rainfall events relevant for potential urban inundation and flooding events.

4.3 Theory and applications

Application of the optimal Z-R relationship, selected as the one yielding lowest RMSE, for each 15-min rainfall accumulation and each rainfall type category, is summarized in figure 4.5 (origin at 5mm rainfall threshold). The rain type categories corresponding to tropical and east cool stratiform show a dispersed result. Conversely, the convective and stratiform types tend toward agreement with gauge values as indicated by the lower variance and narrower clustering along the one-to-one line. Further optimization processes are considered for the stratiform type rainfall values since the result shown is considered adequate for this work. In the case of east cool stratiform type, most rainfall depths are less than 10 mm, and this depth is less significant from a hydrologic runoff generation perspective. For this reason, the east cool stratiform rainfall type is not considered in the remaining part of this study. This leaves convective and tropical type rain categories for consideration, and the focus is on development of an optimization process to improve agreement of radar-based and gauge-measured rainfall accumulate.

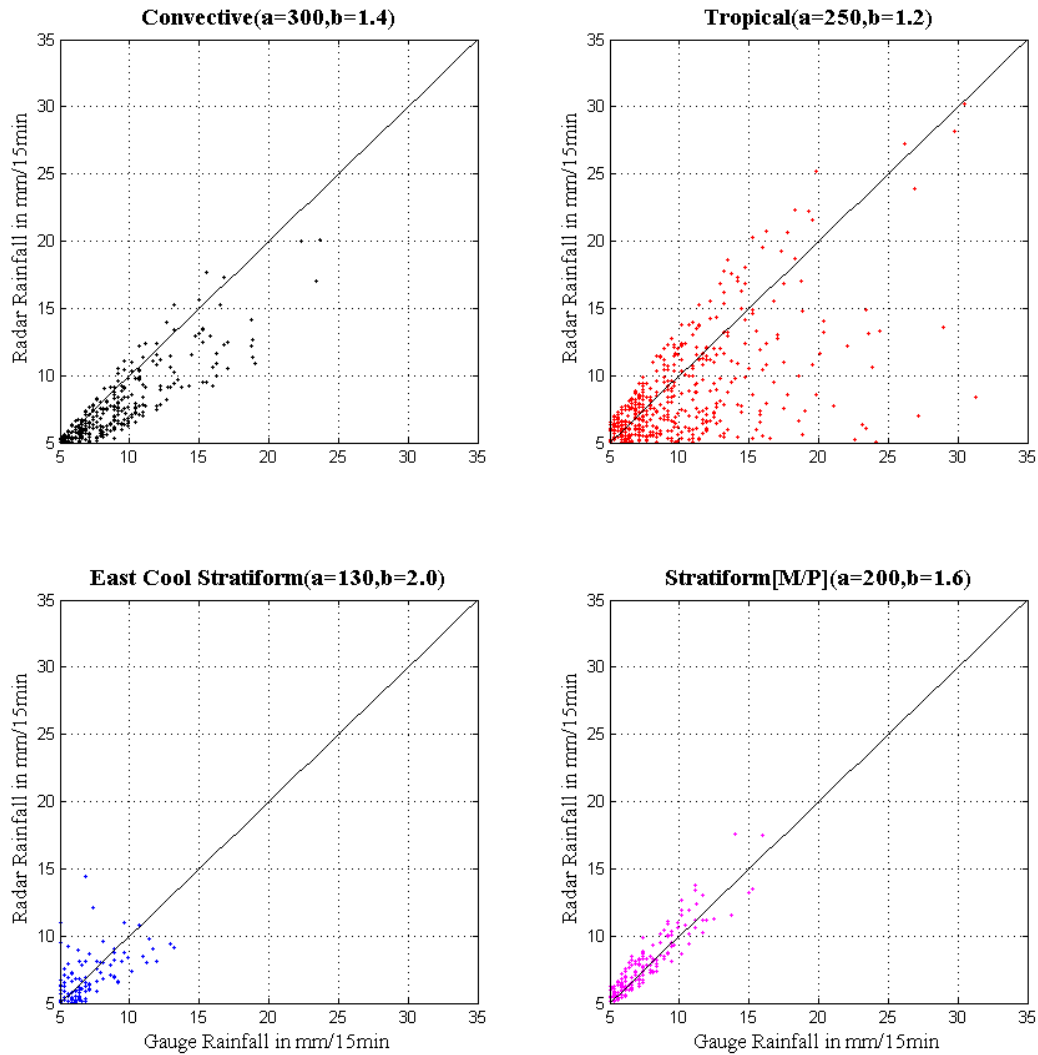


Figure 4.5 Scatter plots of rainfall volumes for each storm type after optimization (minimum RMSE error): convective type (upper-left), tropical type (upper-right), east-cool-stratiform type (lower-left), stratiform – Marshall/Palmer - type (lower-right)

Figure 4.6 presents a comparison of the convective storm type radar depths and the gauge rainfall depths. The lighter shaded markers indicate use of the standard Z-R transformation with parameters ($a:300$, $b:1.4$) and the darker marker dots indicate the optimized Z-R result. A simplex optimization procedure was applied to optimize the Z-R

parameters over value ranges of 10 to 500 and 0.5 to 3 for a and b respectively. Optimization decreased RMSE and Z-R parameters values of 300 and 1.4 were modified to 250 and 1.4 for a and b respectively. The optimized Z-R parameters eliminated the systematic underestimation but the dispersion is unchanged; the simplex optimization centered all values about the one-to-one line.

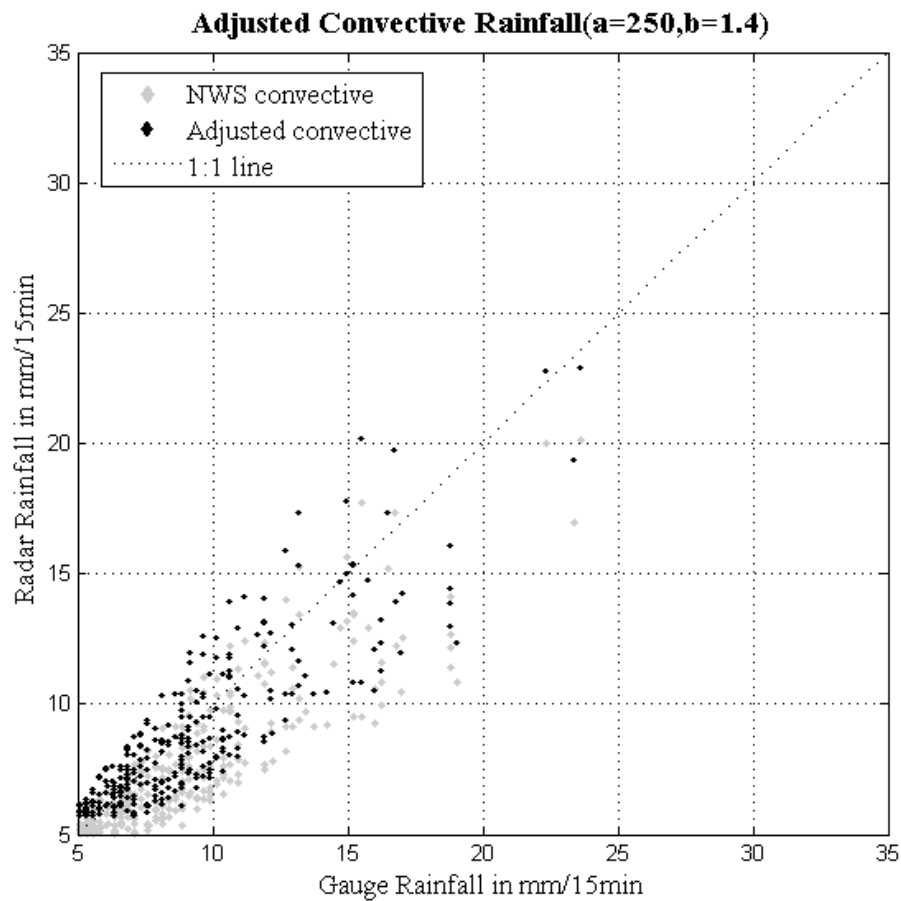


Figure 4.6 Convective rainfall type: radar and gauge comparison with (a) standard NWS Z-R relation (light shade marker), and (b) optimized Z-R relation (minimum RMSE error)

The tropical convective rainfall type contains a large number of high intensity values and has the widest spread of the comparison groups. An interesting and challenging issue is that bias cannot be eliminated by calibration of the Z-R parameters alone.

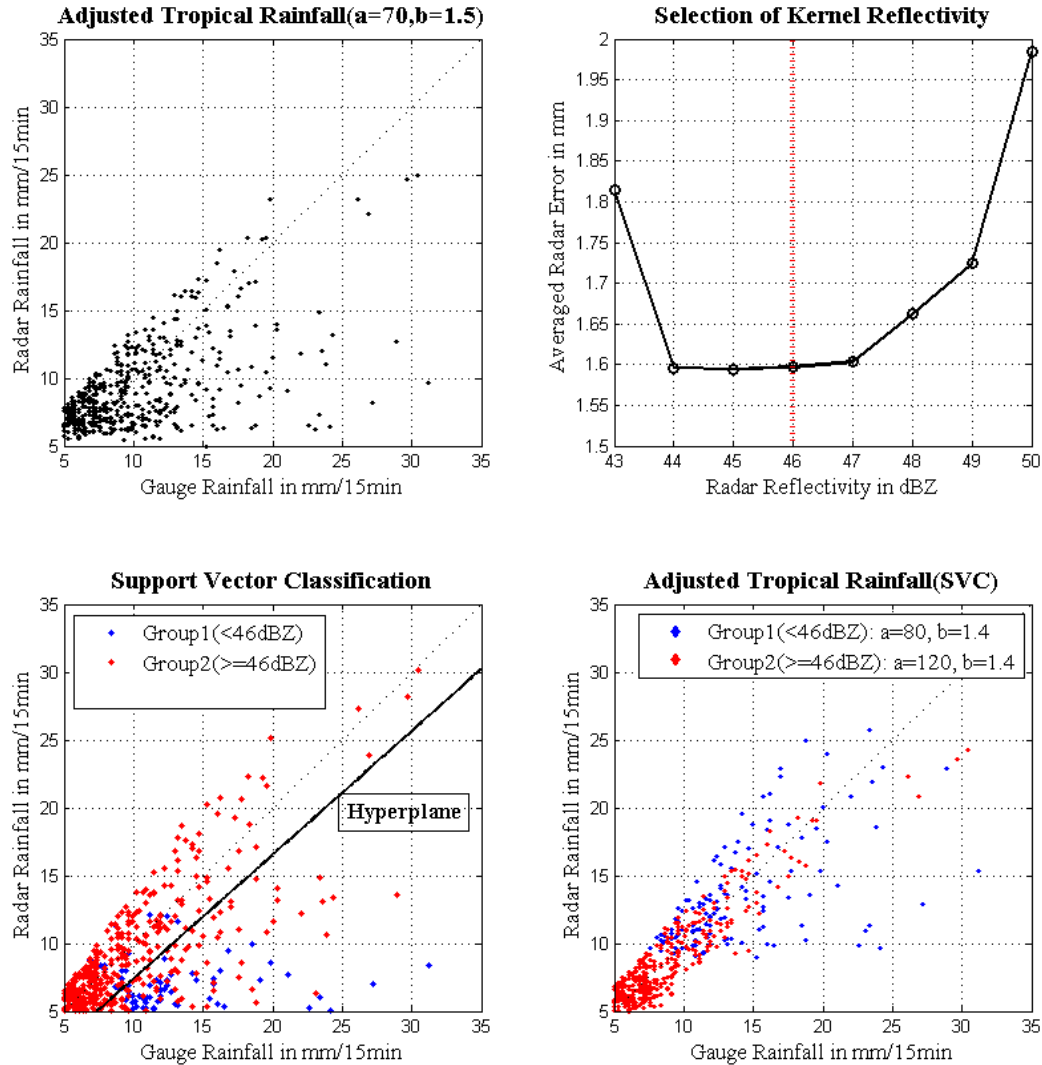


Figure 4.7 Tropical type rainfall results: Optimized minimum RMSE error (upper-left); Decision schematic for SVC kernel within least RMSE error range (upper-right); SVC binary clustering hyperplane (lower-left), SVC-based optimization with two rainfall groups (lower-right)

As shown in figure 4.7 (upper-left), underestimation of rainfall remains following optimization of Z-R parameters a and b. In figure 4.5 (upper-right, red), the result is shown for the best fit Z-R, yet the tropical type rainfall appears to form two distinct groups. The first group is slightly above the one-to-one line with limited dispersion, whereas a second group is under the one-to-one line with wider variability. This indicates that a single Z-R relationship for tropical type may not suffice to encompass the observed characteristics of tropical type rainfall for this region. In order to investigate a solution for this issue, a support vector classification (SVC) optimization procedure was developed. The SVC optimization incorporates an unsupervised learning algorithm applied in the context of a two dimensional surface (x-axis: gauge rainfall and y-axis: radar rainfall). The concept is a data-based learning process; the SVC creates a linear hyperplane separating two binary groups according to a separation margin criteria. The hyperplane forms a linear separation at the maximum margin and is highly efficient in differentiating the non-linear rainfall characteristics. The determinant in the SVC is a kernel method transformation into a feature vector (Cristianini and Shawe-Taylor 2003). In figure 4.7 (upper-right), the maximum instantaneous reflectivity among the group of influential reflectivity values for quarter-hourly rainfall accumulation defines the kernel. The averaged radar rainfall error (difference from gauge value) is at the range of 44 dBZ to 47 dBZ and reflectivity of 46 dBZ is selected in order to balance the number of data values in each group.

$$\hat{y} = \text{sgn} \sum_{i=1}^n w_i y_i k(R_i, \hat{R}) \quad (\text{Equation 4.2})$$

The kernel method is described in equation 4.2 in the Appendix and defines the two groups through the linear hyperplane. The similarity function of the kernel method follows

training data instead of a fixed set of parameters, and this feature involves the similarity function k , denoted as kernel in equation 4.2. More simply, the kernel is a weighted sum of similarities between the trained example input and the new unknown input. The kernel is used as a binary classifier in terms of $\hat{y}$, the binary classifier for clustering the two tropical groups. In figure 4.7 (lower-left) the linear hyperplane shows two data groups, one group fit with the tropical type Z-R relationship and the second group in the underestimated region. Following the SVC process, calibration of the Z-R parameters (RMSE minimum) was completed and the result is shown in figure 4.7 (lower-right). This result demonstrates gains in information for the tropical type rainfall when two Z-R parameter sets (a, b) are used. The original fan-shaped dispersion is greatly reduced, as well as the original underestimation issue. The SVC-based solution process provides a multi-category classification and overcomes limitations of binary classification (Xie et al. 2013).

4.4 Results and discussion

4.4.1 Results

Use of an uncalibrated Z-R relationship for conversion of reflectivity to rainfall intensity for short-duration accumulations may result in differences from ground-level rain gauge observations. An example of these differences is presented in Figure 4.8 (upper-left) where results show a fan-shaped spreading with correlation of 0.68 between gauge and radar rainfall. Coincidentally, underestimation of rainfall may be more problematic than overestimation in applications of hydrological management and design, and this result is relevant for the tropical type rainfall category. In figure 4.8 (upper-left), the comparison shows the best fit standard Z-R relationship (minimum RMSE), and the tropical Z-R relationship corresponding to the solid red line in figure 4.8 (upper-right). This indicates

the tropical Z-R may not capture a complete description of rainfall variability at higher intensity rainfall rates. This study introduced an alternate Z-R relationship formed using an SVC optimization process. Use of the alternate Z-R relationship produced the results in figure 4.8 (lower-left) and figure 4.8 (lower-right). A comparison of the alternate Z-R (red dotted line) and NWS tropical Z-R (red solid line) relationships are included. This alternate Z-R relationship, designated as “tropical-2”, is more influential in the higher reflectivity range (greater or equal to 46 dBZ). The Z-R relations shown in figure 4.8 (lower-right) illustrate the placement of the existing NWS tropical (tropical-1) Z-R relationship between the convective (black solid line) and tropical-2 (red dotted line). Graphically, the tropical-1 relationship fills the gap between the other two convective Z-R relationships. A notable point is that a large portion of the quarter-hourly rainfall accumulations occurring at lower radar reflectivity, less than 46 dBZ, are well represented by the tropical-1 Z-R relationship. The focus of this study on a relatively shorter, quarter-hourly, rainfall accumulation interval, and the focus on more intense rainfall values are factors influencing the need to partition this extreme- type rainfall (NWS tropical) into two sub-groups.

In general, the two tropical sub-groups are similar in the low reflectivity range and deviate more from one another for reflectivity above 45 dBZ to 50 dBZ. Based on the use of categorized Z-R relationships in the region of more extreme rainfall intensity, quarter-hourly rainfall estimation is improved with a correlation of 0.72 in figure 4.8 (lower-left). For applications in urban hydrologic designs and simulation of historical events, the shorter temporal resolution of rainfall is useful.

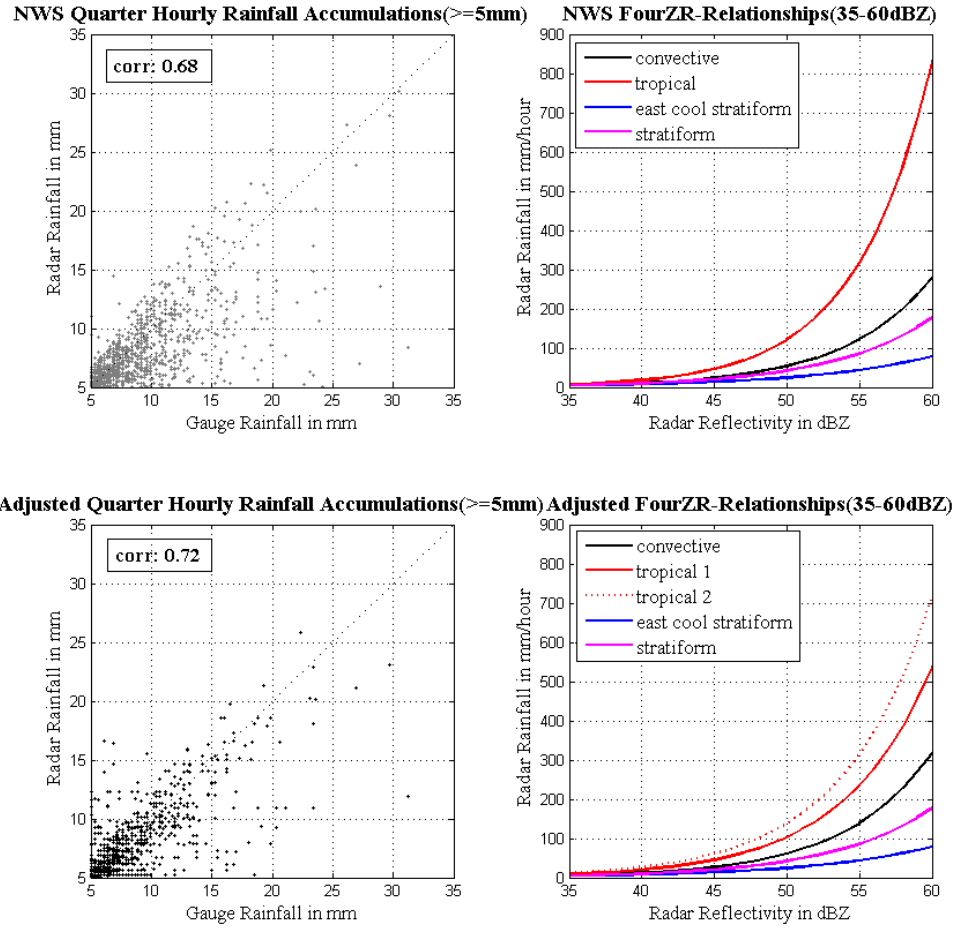


Figure 4.8 Comparison of local radar rainfall estimations and corresponding Z-R relationships: Standard Z-R-based quarter-hourly rainfall accumulation (upper-left); four NWS Z-R relationships (upper-right); Optimal SVC-based quarter-hourly rainfall accumulation (lower-left); SVC-based optimal Z-R relationships (lower-right)

4.4.2 Discussion

The inherent characteristics of rainfall are spatiotemporally dynamic and it remains difficult to define short-time interval volumetric rainfall for a local specific region such as a small urban catchment. Rain storm dynamics are challenging to predict due to

uncertainties in storm characteristics and the influential environmental factors (Coniglio et al. 2010; Panthou et al. 2014). A study of cloud microphysics can identify and record observations of key environmental variables yet the complex interactions between water vapour, in-cloud liquid droplets, and falling hydrometeors make rainfall estimation at the surface using parameterization methods challenging (Cui et al. 2014; Hu and Feng 2002; Morrison and Milbrandt 2015; Nogueira and Keim 2010). Radar rainfall estimation is a remote-sensing method measuring reflectivity rather than directly measuring rainfall rate. The Z-R relationship is the most common approach used to transform the radar signal into rainfall intensity. The value in recognizing rainfall type or category is one issue considered in this study. In particular, the influence of the reflectivity-rainfall transformation for higher spatiotemporal resolution is illustrated. This study illustrates an approach, the SVC classification procedure, as a means of improving the estimation of tropical type rainfall through transformation of radar reflectivity.

4.5 Conclusions

This study investigated optimization of the reflectivity-rainfall (Z-R) relationship to determine rainfall accumulations for quarter-hour temporal intervals in an operational setting. The study involved 3 primary components: 1 - optimization of Z-R relationship selection from a number of existing forms; 2 - calibration of Z-R relationship parameters; and 3 – implementation of an SVC optimization algorithm to classify radar reflectivity data in two groups and calibrate Z-R parameters for these groups. The selection or application of a Z-R relationship according to rainfall type at an individual radar site can provide adequate results for coarse temporal resolution rainfall products. However, differences between rainfall observed at ground level are increased at finer spatiotemporal scales

approaching sub-kilometer and sub-hour. Understandably, the operational approach to radar rainfall estimation does not address localized fine resolution rainfall estimation for urban hydrology applications. In this work, a binary clustering algorithm, SVC, partitioned tropical rainfall type data into two groups using a kernel optimization method. Results indicate application of the SVC classification algorithm, followed by application of Z-R relationships based on identified rainfall type groups, improved radar rainfall accumulation estimates for short duration time intervals.

CHAPTER 5 APPLICATION OF RADAR RAINFALL TO URBAN AREA

5.1 Introduction

In many urban areas, combined sewer systems (CSS) carry both storm water runoff and sanitary sewer flows in a single pipe. In the absence of rainfall-runoff most CSS adequately convey waste water flow, however system capacity may be overwhelmed when it must also transport significant stormwater runoff. In order to prevent sewage from flooding streets and backups into homes and businesses, most CSS (as well as separated sanitary sewer systems) are designed to overflow into surface waters such as streams and rivers, lakes and seas. This overflow occurrence is considered a combined sewer overflow (CSO) event and has a detrimental impact on aquatic environments and degrades downstream water quality.

In the United States (USA), regulations were established to eliminate CSO events in urban areas (EPA 1994). Although CSS are considered an outdated approach to waste water collection, these legacy water collection systems form a considerable portion of the sanitary sewer network in the United States. It is estimated that 860 communities across the USA are served by combined sewer systems with over 10,000 CSO outfalls directed into natural surface waters. These communities include approximately 40 million people in more than 30 states (EPA 2004). The direct solution to eliminate overflow occurrence through modification or replacement of CSS with separate sewer and storm drains is cost

prohibitive, disruptive to the community, and difficult or infeasible to accomplish in existing urban environments (Lyandres and Welch, 2012).

When a CSO event occurs, the effect on receiving waters can be significant. The flow transports microbial pathogens, oxygen depleting substances, suspended solids, toxics, nutrients, and debris including floatables and trash directly into the natural aquatic environment (EPA 2004). Furthermore, in most urban areas CSO occurrence is often a sudden phenomenon, due to both characteristics of the triggering rain storm and hydrologic conditions in the sewershed, resulting in a surge of runoff (Romnée et al. 2015).

Understanding characteristics of CSO triggering rain storm events is important to utilities as when specific event characteristics are identified, a plan to mitigate unsafe occurrences can be developed. The challenge, however, is that in most sewersheds, the common operationally measured rainfall characteristics (rain event duration, total rain volume, intensity, continuity (inter-event time, IET), seasonality and storm type (e.g., stratiform, convective, frontal, orographic, tropical storm remnants), are determined based on spatially distant point source rain gauges and lack catchment specific detail. Rain events identified in a precipitation record may be compared with sewershed CSO flow records to identify a record of coupled rainfall-CSO events.

Historically, rainfall monitoring by ground-level rain gauge networks is considered a reliable measurement system for many hydrologic applications because it physically captures pluvial water. In hydrologic engineering and research, rain gauge measurements frequently serve as a reference for evaluation of indirect or remote sensing rainfall estimation systems such as weather radars and satellites (Habib et al., 2012; Price

et al., 2014; Chen et al., 2015; Fencel et al., 2015). However, the spatial variation detection of rainfall measurement by gauge networks is limited, in particular at finer temporal resolution. Thus, characteristics of the localized rainfall events, which can vary significantly from spatial averages and are influential in triggering CSO events, is not captured.

Hydrometeorologic rainfall monitoring and measurement technology has advanced in recent decades (Karamouz and Nazif, 2013; Morita, 2011; Yang et al., 2013). The availability of quantitative two-dimensional measurements from weather radar has increased the availability of rain detection at higher spatial and temporal resolution. From radar rainfall archives, areal rainfall observations are available at the sub-hectare (radar polar coordinated pixel size) resolution at sub-hour temporal intervals. Identification of rainfall events, for example, using the EPA criteria for urban areas and a defined inter-event time (IET), provides the context for identifying rainfall spatial and temporal characteristics associated with CSO overflow events. Accordingly, preparation of accurate rainfall data, quality controlled weather radar data, identification of independent rainfall events and the corresponding CSO event hydrographs are essential to developing a quantitative understanding of this phenomenon.

To this end, an objective for this study includes application of locally optimized radar-rainfall to an urban sewershed (watershed) using fine-scale spatiotemporal resolution data, and evaluation of rain event characteristics resulting in CSO events. Categorization of the severe rainfall events inducing CSO can provide insights for hydrologic and hydraulic design guidelines to reduce sewer overflows from combined sewer systems in an urban area.

5.2 Case Study

5.2.1 Urban sewershed setting and CSO location

The sewershed CSO130 is part of the urban CSS and located in an older neighbourhood, called Buchertown, in Louisville, Kentucky. The specific location of CSO130 is along Webster Street and its overflow control structure type is a diversion dam. The sewershed is approximately 13-ha (30-ac) and land-use is a mixture of commercial and dense residential. The land-use is about 75% impervious with the portions distributed as residential (24%), commercial (25%), industrial (32%), vacant land (6%), and roads and other uses (13%). The CSO130 outfall discharges into the nearby stream, Beargrass Creek, which is a tributary of the Ohio River.

The regional sewer agency, Metropolitan Sewer District (MSD) operates a rain gauge network across the city region and one rain gauge is located near the study area.

However, data from this gauge serves only as a reference to evaluate radar rainfall quality rather than as rainfall for the sewershed. Hyun et al. (2016a, 2016b) characterize the spatial and temporal variation of rainfall using the MSD rain gauge network for the meso-scale region of the city of Louisville, Kentucky. The benefit and challenge of using weather radar data for operational applications is illustrated and the spatial variation of rainfall derived from weather radar products described (Hyun et al. 2016a, 2016b).

Additionally, radar-rainfall data from the National Climatic Data Center (NCDC) archive of National Weather Service (NWS) NEXRAD radar, for the Fort Knox, Kentucky location (denoted by call letters KLVX) are incorporated. The radar records are extracted and optimized, with spatiotemporal resolution of quarter-hour and less than five-hectare (15-acre), in order to be applicable for the urban hydrologic scale.

5.2.2 Coupled radar rainfall and overflow event record

Only rain events resulting in a CSO event are considered and a quality control threshold is applied to select rain-overflow events exceeding a ratio of 0.60 for overflow depth to rainfall depth. Discriminant analysis is used to categorize the coupled rainfall-overflow events according to overflow severity; a threshold of the overflow depth of 1.5-mm partitions the event categories. Results indicate that overflow depth has a strong linear relationship with rainfall depth and other environmental factors are influential.

Identification of rainfall events, using the EPA criteria for urban areas and a defined inter-event time (IET), provides the context for identifying overflow events in the CSO flow record. Accordingly, preparation of accurate rainfall data, quality controlled weather radar data, and a record of independent rainfall events, are essential. In radar rainfall estimation, areal rainfall observation is available at the sub-hectare (pixel size) radar resolution. The NCDC archive of NWS radar data provides super-resolution radar rainfall at a pixel scale of 250-m by 0.5° in polar coordinates, and this product is suitable to define rainfall variation over urban areas. However, the quality control of the data record provides an improved rainfall product (Hunter, 1996). To address this issue, Hyun et al. (2016c) focused on data quality control, calibrating the radar to local conditions through adjustment of the empirical Z-R relationship transforming observed radar reflectivity (Z dBZ) to rainfall intensity (R mm/hr). The radar rainfall optimization identified the Z-R relationship shown in figure 4.8. Figure 4.8 shows quarter-hourly radar derived accumulation compared to rain gauge values have a correlation of 0.68, and an underestimation tendency by the radar. To address this issue, rainfall data were clustered and partitioned using support vector classification (SVC), with the underestimated group

assigned a new Z-R relationship (lower-right figure 4.8). This resulted in improved data agreement for quarter-hour temporal resolution indicated by the correlation of 0.72.

Additionally, figure 5.1 shows radar rainfall comparisons with gauge rainfall data sources near the study area. An MSD rain gauge, gauge number TR05, is located about 600-m away from the study area, and this data record is used as a reference to evaluate spatial variation of rainfall. The improved radar data shown in figure 5.1 (upper-right) has improved the correlation to 0.79 compared with a 0.70 correlation in the original estimate (figure 5.1 upper-left). Two other rain gauges, TR12 and TR03 show correlation of 0.55 and 0.05 with the TR05 data respectively. Within a distance of 15-km, rainfall is spatially uncorrelated ($\rho=0.05$), and correlation decreased to 0.55 within 5km distance at TR12. This reveals the high spatial variation of the rainfall at quarter-hourly temporal resolution and the benefit of radar-rainfall over the limitation of reliance solely on ground-based rainfall measurement..

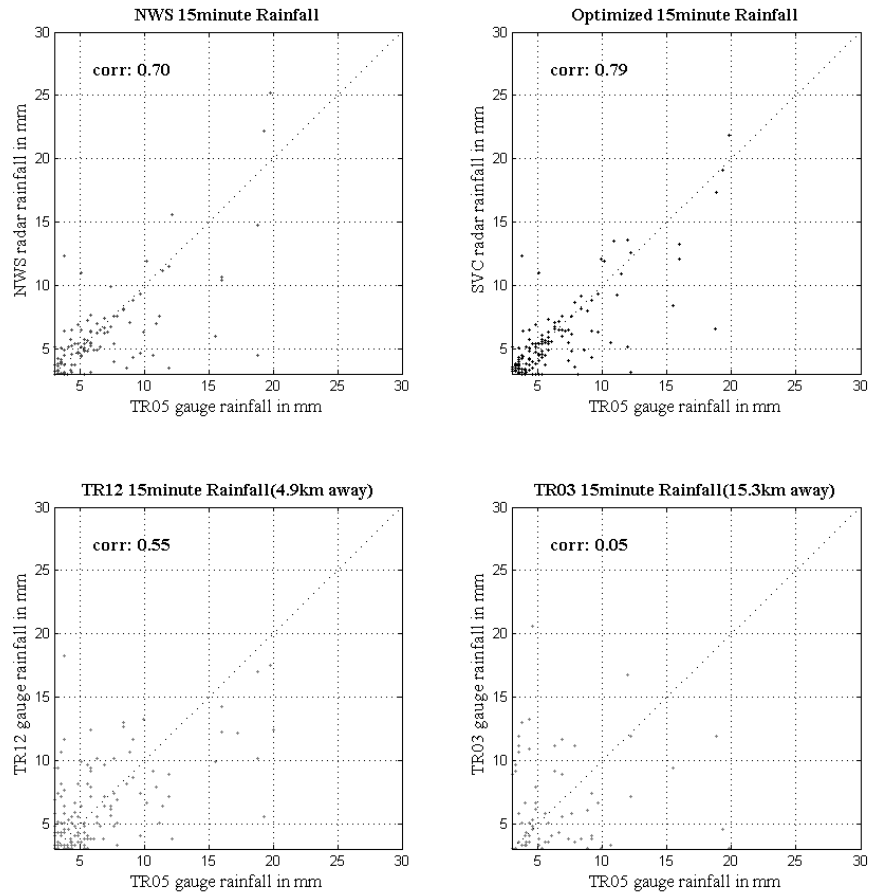


Figure 5.1 Rainfall data quality comparisons with the reference rainfall data (TR05): NWS radar data (upper-left), Quality-improved radar data by SVC (upper-right), MSD rain gauge-TR12; 4.9km away from the study area (lower-left), and MSD rain gauge-TR03; 15.3km away from the study area (lower-right)

The CSO mechanism in the sewershed is not only related to the rainfall characteristics; depth, duration and intensity, but also the continuity of the rainfall event.

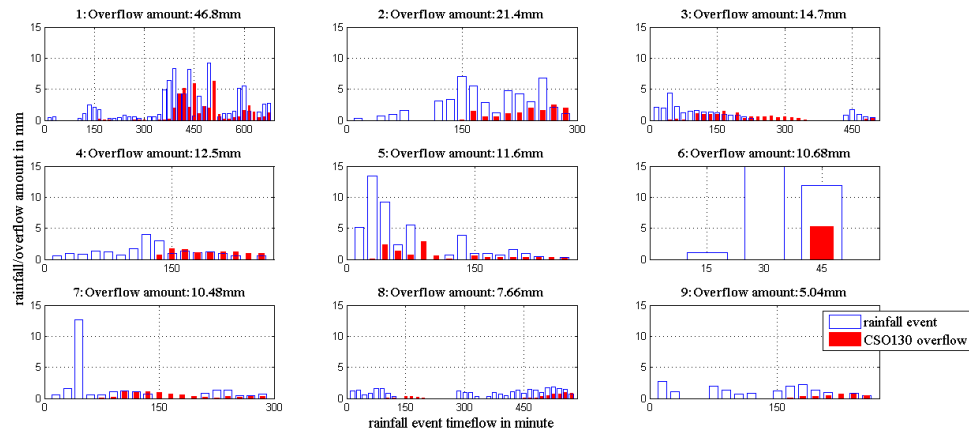


Figure 5.2 Selected CSO events and corresponding rainfall events; Event number denotes the rank of the overflow amount through the outfall structure to Beargrass Creek, Louisville, KY

Figure 5.2 illustrates the extreme overflow cases at CSO130 in time-flow manner and the hydrologic response is related to rainfall variations within the rain event. The nine overflow events shown indicate that most were triggered immediately following the heaviest rainfall interval. Naturally, the rainfall volume is the primary influence on overflow amount, but it is not the only factor. The more sizeable rainfall peaks affect the overflow amount and time distribution. For example, the sixth greatest overflow in figure 5.2, with overflow amount of 10.68-mm, has precipitation duration less than an hour but the overflow was significant because of high intensity rainfall. On the contrary, the overflow event which ranked in fourth has no clear intense rainfall observed; instead rainfall is steady and uninterrupted. These results indicate rainfall event continuity as another factor triggering overflow event occurrence.

One definition of a rainfall event is provided by the EPA for rainfall event in the context for urban regulatory settings (Driscoll et al. 1989). The EPA document defines a

rainfall event as “A minimum storm volume of 0.1-inch (2.54-mm) was specified for the analysis that were performed, so that the analysis would produce statistics of ‘runoff producing’ events within 6 hours.” In short, a single rainfall event is completely independent if no sizeable rainfall, greater than 2.54-mm (0.1-inch), occurs within six hours. The rainfall event defined by EPA regulation and corresponding time for the overflow event were determined from the time rainfall began until six hours from the end of the rainfall event. By EPA definition, a rainfall event is followed by at least a six-hour dry period and so the implied time available for overflow to occur is limited to six hours following the rain event. Based on this, there are 95 rain events with coupled CSO occurrence in the sewershed for CSO130 over the three-year study period, January 2011 to December 2013.

5.3 CSO130 overflow analysis

The CSS CSO130 control structure is a 0.61-m (24-inch) circular brick sewer pipe flowing with an average of 12 overflow incidents (events) per year (averaged 2.33 hours of duration and 90,000 gallons of combined sewer per incident) (MSD, 2014).

5.3.1 Quality control of coupled rainfall/overflow event

The number of CSO events identified directly from data records of rainfall and overflow analysis indicates the number of incidents is 95, and this is a greater number than the 35 to 40 otherwise expected according to the average number of reported incidents over the same three-year study period (MSD 2104). Potentially, the method of identification and event partitioning (IET) may influence the number, but proper quality confirmation is required for data reliability. To this end, a means of screening outliers and

poor quality data records is utilized. Application of a common rainfall-runoff index threshold screening, based on watershed characteristics, is not possible since the total runoff for each CSO event partitions flow into two directions, one part is the overflow and the remaining portion continues to the water treatment plant. During an overflow event it is not possible to separate the portion attributed only to stormwater runoff. To address this issue, CSO event data are partitioned into acceptable and non-acceptable clusters. Figure 5.3 (left) shows the normalized runoff-rainfall index ratios of overflow depth to radar-rainfall depth and overflow depth to gauge rain depth for each event. The plot spreads in a two-dimensional field; with x-axis: ratio for rain gauge MSD TR05 (600m from CSO130), and y-axis: ratio for radar rainfall. A ratio greater than 1 indicates runoff greater than rainfall, and this unlikely occurrence may indicate data error or other issues; for this reason, these data are excluded from the study. The use of two rainfall sources lessens the uncertainty concerning rainfall occurrence and incorporates both these practical hydrologic observations into this study.

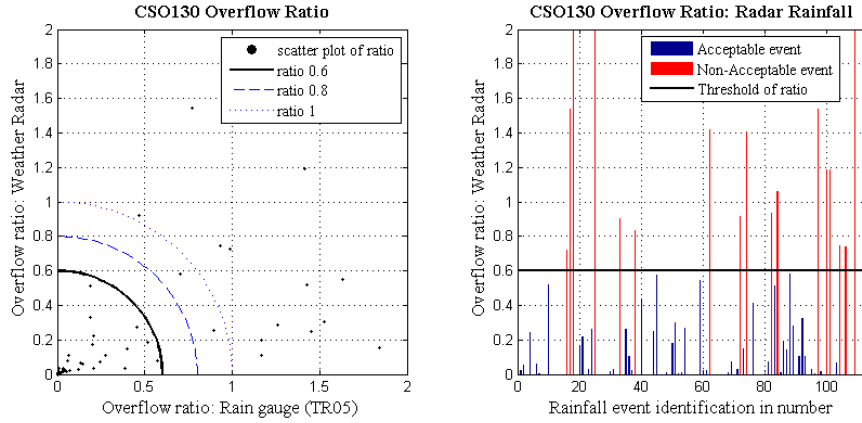


Figure 5.3 Overflow ratio plots. Left side: Two-dimensional radar & gauge rainfall field; x-axis shows rain gauge ratio-MSD TR05 (nearest study area), y-axis shows radar ratio-NWS weather radar KLVX. Right side: Criteria threshold for valid event selection: 52 acceptable events (blue) below the 0.60 overflow/rainfall threshold, and 43 non-acceptable events (red) exceed the threshold

Figure 5.3 shows a notable absence of overflow occurrence between ratio values of 0.60 and 0.80. Therefore, a threshold ratio of 0.60 was defined as the acceptable coupled rainfall and overflow event criteria; this is the boundary where data are densely populated and shown as the inner region defined by a bold solid line forming a quarter-circle in figure 5.3 (left). The right portion of figure 5.3 shows all CSO events and the bold solid line indicates a value of 0.6 for the radar overflow ratio. The result identifies two groups of CSO events: acceptable (blue) and non-acceptable (red). This process indicates 52 coupled rain-overflow events and this corresponds well with the expected number as suggested by the MSD report for the three-year study period (MSD, 2014).

5.3.2 Analysis of coupled rainfall and overflow events

5.3.2.1 Overflow relation to rainfall depth, intensity and duration

The coupled rain and overflow record for CSO130 shows the sewershed runoff response is prompt with a hydrograph form similar to a smoothed and time-lagged reflection of the hyetograph. That being so, understanding the rain event characteristics provides insights into the timing, intensity and amount of overflow. Fundamentally, the quantitative relation between rainfall and overflow has a visible linearity as in figure 5.4 (left). Rain volume is an important factor and shows a linear relationship with overflow. As shown in figure 5.4, when rainfall is less than about 8-mm a low overflow volume occurs and overflow amount increases linearly above this rainfall depth. For overflow values above a 0.40 ratio of rainfall volume the sewer overflow volume is more significant and likely to impact environmental quality in the Beargrass Creek. It is expected that total rainfall depth is a significant factor triggering an overflow, but this simplified conclusion cannot completely explain the behaviour and a search to understand the contributing factors causing overflow events is warranted.

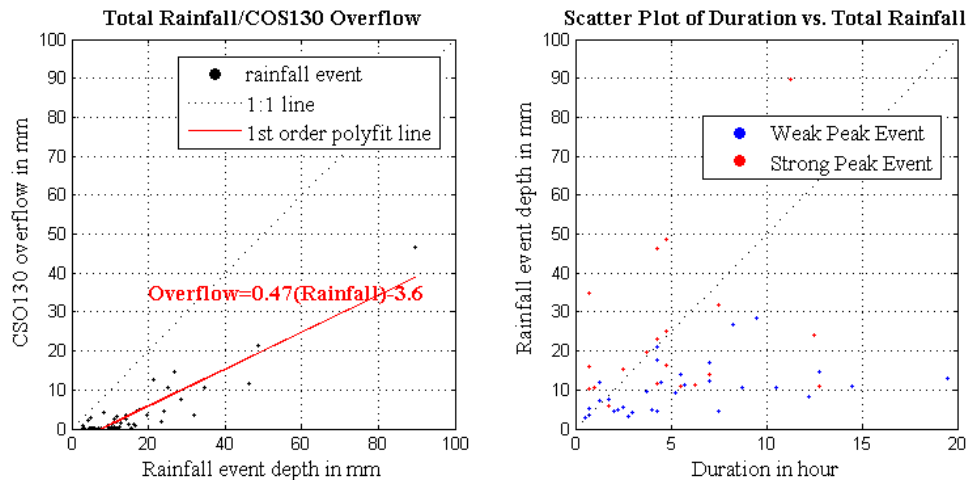


Figure 5.4 Event-based rain depth versus overflow depth (left), and rain event duration versus rainfall depth grouped by peak rain intensity (right). Intensity threshold peak is 4.87mm/15minute to identify weak (blue) and strong (red) peak event groups

The rain event duration and peak intensity (15-minute temporal resolution) are important in determining overflow volume in figure 5.4 (right). In the duration versus rainfall depth field (figure 5.4 right), the events clearly divide into two groups when clustered by peak rainfall intensity. The two groups have somewhat different tendencies in the two-dimensional space with the strong peak intensity group showing a relatively short duration but larger volume of rainfall. On the other hand, rainfall volume tends to be relatively stable and less relevant as an overflow trigger no matter the event duration. In the small-scale urban watershed setting, existence of high-intensity peak rainfall may produce significant volumetric rainfall, thus, rainfall intensity significantly impacts drainage system performance in urban areas (Arnbjerg-Nielsen et al., 2013; Mamo, 2015). Considering this result in a practical application, rainfall depth, intensity and duration are all factors indirectly incorporated into historic intensity-duration-frequency (IDF) curve used to define volumetric rainfall for urban hydrologic design. However, variations between rainfall observations and IDF design values illustrates the uncertainty for applications requiring fine spatiotemporal resolution such as urban sewersheds where runoff response occurs well under the sub-hour temporal scale.

5.3.2.2 Overflow relation to rainfall depth, storm type and continuity of rain

There is a thread of inter-connection between instantaneous heavy rainfall, storm type and resulting overflow in this small watershed. A concurrence of rainfall characteristics and watershed condition, for instance, existing antecedent moisture, wet or dry surfaces and soils, which effect rainfall retention and percolation, may influence overflow occurrence. Therefore, rainfall continuity in terms of single events is considered as an additional factor. The extreme overflow events in figure 5.2 reveal the

importance of the rainfall continuity since there are four high ranked overflows (rank number: 3, 4, 8, and 9) associated with rainfall events with relatively insignificant peak intensity (below 5-mm/15minute), but continuous and uninterrupted rainfall. In other words, the length of the duration of rainfall within a single event is an influencing factor related to the CSS capacity and resulting CSO for this sewershed.

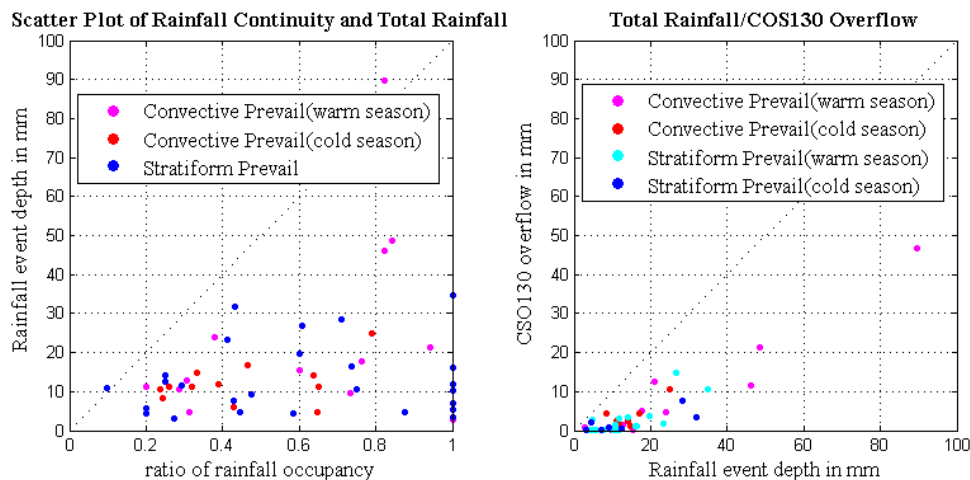


Figure 5.5 Rainfall occupancy ratio (ratio of continuous rain duration to total event duration) and total rainfall event depth: convective event type in warm season (magenta), convective type in cold season (red) and stratiform (blue) (left). Event-based rainfall depth versus overflow depth: convective-warm season (magenta), convective-cold season (red), stratiform-warm season (cyan), and stratiform-cold season (blue) (right)

The ratio of the time rain falls during a rain “event” to total event duration represents the continuity of the rainfall or rainfall occupancy ratio. Figure 5.5 (left) illustrates the relationship between rainfall depth and rainfall occupancy by rain type and season; warm season (April to September) and cold season (October to March). The radar rainfall product indicates rainfall type for each 15-minute rainfall accumulation. The characteristics of a single storm, in motion over CSO130 sewershed, are dynamic and a

series of storm cells may move over the area. The convective storm type may have a single or multiple cells within the developed storm structure associated with severe rainfall (Caine et al, 2013; Cetrone and Houze, 2006; Feng et al., 2014; Peter et al., 2015; Zawilski and Brzezińska, 2014).

Identification of rainfall type is based on the ratio of number of convective type radar pixels to total rain pixels in the storm. Applying a threshold ratio of 0.45 results in the two event groups; a convective prevalence group and a stratiform prevalence group. The stratiform group has no discernible spatial pattern features other than the continuity of rainfall coverage, while the convective group has a tendency toward increasing intensity beginning around a ratio of 0.60. The highest three rain depth events are in the convective group. The reappearance of the rainfall overflow plot (figure 5.5, right), with seasonal rainfall group details added, demonstrates the characteristics of the overflow inducing rainfall events. Prior to presentation of this figure, the nine ranked overflow inducing rainfall events show a 0.81 ratio of rainfall occupancy and no event with lower than a 0.60 ratio. This indicates the rainfall event group most likely to generate a CSO are the convective rain group in summer season. The mitigation of combined sewer overflow events can use this information in hydrologic design to improve future approaches to stormwater runoff reduction. The overflow of CSO130 is a response to the interaction of natural rainfall variability, the urban landscape, land-use and hydrologic environment. In addition to rainfall variability, other qualitative factors influence the likelihood of overflow occurrence. Therefore, understanding the temporal and spatial structure of overflow inducing rain events is useful to estimate overflows in CSOs.

5.3.2.3 Discriminant analysis in overflow inducing rainfall events

This study shows that an overflow event in a sewershed is induced through the integration of factors from two fields; natural rainfall variability and the constructed sewershed conditions. The fundamental assumption is that rainfall induces the overflow event in a small-scale sewershed because the runoff response is rapid and the hydrograph structure resembles the hyetograph. Thus, preventing overflow events inevitably requires understanding of rainfall characteristics. The volumetric rainfall depth was shown to relate linearly with overflow and other factors, such as rainfall intensity, duration, and continuity of rainfall (as a ratio of rainfall occupancy) as influential factors. Discriminant analysis is introduced to classify these precipitation factors. The discriminant analysis uses the combination of features from the continuous independent variable (rainfall characteristics) to define a separation of the discrete dependent variable (Martinez and Kak, 2001; Tahmasebi et al., 2010) and is applied broadly in water resource (Sangam et al., 2008, Boyacioglu, 2010). In order to apply discriminant analysis, the dependent variable (overflow) must be categorical unlike the continuous independent variables. In Figure 5.6, a threshold overflow depth set at 1.5-mm, for the CSO130 sewershed, and categorizes the coupled rainfall-overflow events into two groups; a significant overflow group (23 events referred to as group-1) and non-significant overflow group (29 events referred to as group- 0).

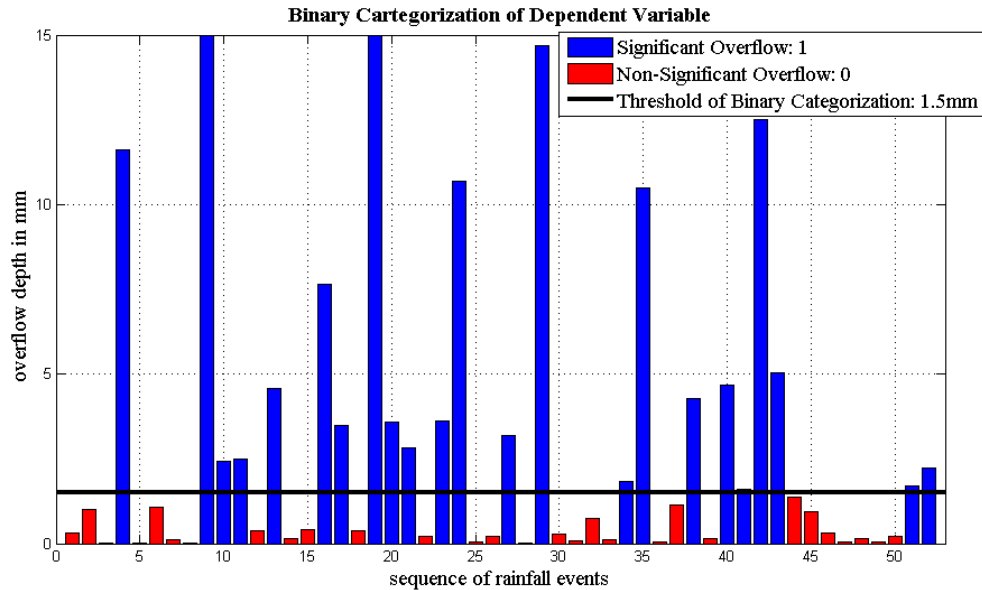


Figure 5.6 Overflow event and 1.5-mm depth threshold separating overflow events into two binary categories – significant (denoted as 1) and non-Significant (denoted as 0)

In table 5.1, the mean values of the rainfall related variables influencing the overflow events are presented in discriminant groups. As expected, this highlights differences between groups and provides a quantitative distinction of the decisive overflow factors. The mean rainfall depth is 8.85-mm in the non-significant group (group-0) and 23.70-mm in the significant group (group-1). Overall, volumetric rainfall governs the overflow in this small scale sewershed. The peak rain intensity has a similar tendency showing 3.76-mm per 15-minute and 6.20-mm per 15-minute for group-0 and group-1, respectively. Commonly, the rainfall depth and the rainfall intensity (peak) have positive correlation with overflow amount. However, the duration of rainfall indicates a contrast to this expected result. In figure 5.4 (right), the majority high rainfall depth events have shorter durations. These shorter rainfall duration events are expected to fall into group-1 considering the relationship between rainfall depth, intensity and overflow

occurrence. In this case, the continuous, uninterrupted and longer duration rainfall induces the overflow. This is due to the inverse correlation between rainfall duration and rainfall continuity, where a correlation of -0.64 is indicated between these independent variables. A longer rainfall duration is more likely to also contain a non-rain period resulting in proportionally lower continuity.

Groups	Variables	Mean
Non-Significant Overflow (group 0)	Duration (hour)	4.57
	Rain Total (mm)	8.85
	Rain Peak (mm/15min)	3.76
	Rain Type (convective ratio)	0.47
	Rain Continuity ratio	0.56
Significant Overflow (group 1)	Duration (hour)	6.14
	Rain Total (mm)	23.7
	Rain Peak (mm/15min)	6.20
	Rain Type (convective ratio)	0.50
	Rain Continuity ratio	0.60
Total Events	Duration (hour)	5.37
	Rain Total (mm)	15.4
	Rain Peak (mm/15min)	4.84
	Rain Type (convective ratio)	0.48
	Rain Continuity	0.58

Table 5.1 Group Mean Values of Rainfall Characteristics by Discriminant Analysis

Another matter that merits mention is the definition of rainfall event duration. The study incorporates the USEPA (EPA, 2004) definition for continuous rainfall and independent event identification. The emphasis is on whether event independence, using

the temporal separation of a six-hour dry period for a small-scale sewershed, is appropriate since both the sequence of rain depth and continuity of rainfall within the rain event are influencing factors. The differences are investigated here using discriminant grouping by rain type and rainfall continuity, nevertheless overflow occurrence is associated more with rain events in group-1. The definition of a rainfall event may be improved with incorporation of factors such as watersheds size, land-use characteristics, and hydrologic goals.

Overflow Severity			Predicted Group Membership		Total
			Non-Significant	Significant	
Original	Count	Non-Significant	29	0	29
		Significant	11	12	23
	Percentage	Non-Significant	100	0	100
		Significant	47.8	52.2	100

Table 5.2 Classification Result and Predicted Group Membership by Discriminant

Analysis

Table 5.2 shows 78.8% (41 of 52) rainfall events are correctly classified using the linear combinations identified by discriminant analysis. Under the predefined threshold overflow depth of 1.5-mm, the 29 and 23 coupled events fall into non-significant and significant groups respectively. This threshold considered a balance for the number of events in each group. The objective discriminant group clustering indicates 12 events in the significant group and 40 events in the non-significant group. The linear combination of rainfall factors, $w^T \cdot \mu^T$, include the mean and variance for the clustered factors. The cluster grouping decision includes the ratio of variances within and between the groups as defined by equation (5.1) (appendix). Based on this formulation, each group was

established by the lowest variance of rainfall factors. The corresponding overflow depth is found to be about 4.6-mm for a rain event with 24-mm depth (single rain event category).

This type of information, specifically identifying overflow volume associated with rainfall event characteristics, may serve as an indicator of overflow potential in a CSO sewershed. Information defining a gradation of overflow potential may be useful for operational optimization such as, real-time likelihood of an overflow occurrence, design of overflow dam height or pipe size, or design of retro-fit infrastructure to mitigate significant overflow events. In this study, CSO occurrence in a small-scale sewershed is investigated with a focus primarily on rainfall characteristics.

5.4 Conclusions and future works

Improving and preserving water quality and the aquatic environment in urban areas is a focus of the EPA and a component of regulatory guidelines limits the allowable occurrence of CSO (combined sewer overflow) events (EPA 1994, EPA 2004). The approaches for mitigating overflow events require information to define existing CSO conditions and event occurrence in terms of flow volume, seasonal variation, and pollutant type and concentration. In this study, the volumetric approach of CSS overflow event study in a small-scale sewershed was presented using the radar-rainfall characteristics. The study incorporated details of radar-rainfall data evaluation, rain event definition, and reveals the dependency of CSS overflow events on rainfall depth, duration, intensity, type and continuity. The radar-rainfall data are optimized by incorporating a Z-R relationship for the extreme rainfall group (tropical rainfall type). The radar data are validated with local rainfall sources; rain gauges from NWS and MSD

networks at locations closest to the study area. The radar derived rainfall is necessary to determine rain depth over the region of interest where the coverage of rain gauges is limited.

Fundamentally, a linear relation exists between rainfall and overflow depths governing the occurrence of CSO events in this small-scale sewershed. However, other factors influence the overflow hydrograph shape and flow volume, as surmised from the record of coupled rainfall and overflow events. The identification of corresponding rainfall and overflow events requires evaluation of coupled rain-overflow events and the study determined an overflow depth to rainfall depth ratio of 0.60 indicative of valid events. The discriminant analysis clustered overflow events into overflow severity classes. The objective classification categorized most events correctly and the discriminant analysis provided an indication of the volumetric relationship between the rainfall and overflow in this sewershed system.

In our era of climate change, the rainfall has a tendency to increase in intensity and spatial variation which is expected to induce localized flash floods, and in turn generate increased CSS overflows. In the same vein, quality two-dimensional rainfall observations at suitable spatiotemporal resolutions provide a means to evaluate existing hydrologic infrastructure and implement optimal designs in specific in urban area. The more significant sewer overflow events occur rapidly, typically within a few hours following rainfall and from rainfall event durations less than a half day. This means daily or longer rainfall records may not be suitable for overflow analysis for small-scale sewersheds. This is in part due to the lack of independence in the identification of overflow inducing rain events where rain intensity, rain continuity and variability

definition within a single rain event are necessary. An in-depth investigation of rainfall and overflow relationships across a range of hydrologic settings and sewershed characteristics may reveal an index for the practical design of a sewer overflow prevention structures. This type of study is essential for optimal development of objective and quantitative methods to mitigate CSS overflows in urban environmental systems.

CHAPTER 6 CONCLUSIONS AND FUTURE WORKS

A new technology has changed people's lives and make the impossible possible. The fine resolution of the radar rainfall with much more sophisticated error removing algorithms is one of the cases. The size of the radar pixel is just 15 acres with fine resolution of the temporal resolutions as in chapter 3. In chapter 2, the limitation of the one dimensional rainfall measurement came to the surface due to the spatial variation, particularly in the fine temporal resolution by the spatiotemporal structure of the ground-based rainfall. Although the study area, Louisville Metropolitan has a well maintained gauge network (15 rain MSD rain gauges are in operation), the rainfall spatial variation causes the limited use of the ground-based rainfall estimation. The urban hydrologic application requires the fine resolutions of rainfall data to investigate the local floods which take place suddenly.

In chapter 2, the spatiotemporal structure was studied by using PPMCC variation over the ground for four different temporal resolutions; quarter hourly, semi hourly, hourly and three hourly. All correlations of the rainfall dropped rapidly within the 5km range and it hit 0.6 around 5km range in the quarter hourly temporal resolution. The sub-hourly rainfall application is required for the urban hydrologic modelings and preparations and the expected correlation is only about 0.6 due to the spread of this gauge network (the distance range between gauges is 4.6km to 11km). Quantile analysis reveals another aspect of the undesirable output. The correlations have a decreasing tendency in

sub-hourly accumulations. Most of the hydrologic preparation and forecasting of the extreme storm focuses on the heavy rainfall, but less correlated tendencies are observed along with the magnitude of the rainfall. Chapter 2 informed the spatiotemporal limitations of the ground-based rainfall measurement system which engineers have believed to be the most reliable rainfall detection system.

Chapter 3 focuses on the continuity of the rainfall according to the hydrologic response of the watershed. This encapsulation is defined as the ‘rainfall event’ using EPA’s definition for urban applications for the rain gauge network. During the five years of the study period, 558 rainfall events were created by the definition. In other words, more than two rainfall events occurred weekly over the study area. The variogram is introduced to analyze the spatial correlation, and it requires the stationarity of the data characteristics. Rainfall is a natural phenomenon, and it is hard to expect the stationarity. The clustering of the rainfall events was performed to fit this requirement by using K-means clustering. In the normalized two dimensional field of rainfall volume and intensity, the three groups; high intensity rainfall, high volume rainfall, low rainfall, were prepared in order to maximize the rainfall stationarity within a group. Due to the variogram, the high intensity group reached the sill around the 20km and the correlated range is much shorter than the high volumetric group. The pouring rain may cause a sizeable amount of the runoff in the impervious urban situation and the spatial variation is huge and problematic. This chapter gives the meaningful unit of the rainfall for the hydrologic application, and the clustering of the rainfall events will give tips to access to the urban hydrologic modeling.

For the last few decades, the radar rainfall estimation has evolved and, currently, it is practical to get fine enough spatiotemporal data. However, the radar rainfall estimation is a fundamentally indirect method to measure the rainfall, and the user should consider the limitations. Chapter 4 addressed the local radar rainfall optimization by redefining the Z-R relationship based on the hardship of the calibration of the individual radar. The concept of this approach is like a mosaic of the radars over the nation. Currently, NWS used a single index of the reflectivity-rainfall intensity relationship due to the storm type. It causes serious errors because a storm has many dynamic factors; advection, developing, decaying, and spatial variation especially in the sub-hourly temporal resolution. Using the SVC, the extreme case of the storms were binarily divided into two different groups. As a result, an additional Z-R relationship improved the radar data quality, and it is applicable to the practical hydrologic situation with the highest correlation with the reference data.

In the last chapter, chapter 5, the application of the locally optimized radar rainfall to the artificial watershed is investigated to explain the causal relationship of the rainfall extremity and the overflow of the sewer water into the urban aquatic environment. More specifically, characteristics of the rainfall in a single event influence the overflow in this small scaled sewer-shed, CSO130. The study trimmed out the doubtful data before use using the two different rainfall sources, the radar and the nearest gauge. As the coupled rainfall-overflow events which created the maximum volume of overflows, the simplization of the relationship by the volumetric approach will fail in some cases. In short, the many factors such as intensity, duration, storm type, and continuity are

important with the volumetric governance. The discriminant analysis confirmed the result and suggested the numerical guideline for each factors.

In this climate changing era, this is an understandable effort to change the old-fashioned rainfall estimation system such as Intensity-Duration-Frequency (IDF) curve. It may still work in coarse resolution of the data, but far beyond the capacity of the current requirement in hydrology, peak flow analysis for sub-daily or sub-hourly. Furthermore, the rainfall variation is severe using the ground-based rainfall and rain gauge network for the small scaled watershed. Practically, the radar rainfall estimation could be a solution to fulfill all the requirements, particularly for the complex urban hydrologic application which required fine spatiotemporal resolutions. However, the users need to be concern with the quality issues on the radar rainfall estimation, even though continuous development is currently going on. The one of the main issues of the radar rainfall quality is the hardship of the radar calibration nationally. A single radar involves the local aspect of the climate and rainfall characteristics within its reliable range and an independent algorithms and relations should developed for the downscaling purpose in the hydrologic field. One more concern is that the the timely assigned storm type for the Z-R relationship application cannot represent the storms passing through the radar range. The statistical approach is required which considers the storm's characteristics; advection, developing, decaying, uniting and breaking off of the rainy cloud for the better weather forecast and hydrologic preparations.

This research focused on the production of reliable rainfall estimation and its application over the small scale urban watershed. In short, the urban hydrology required the sub-hourly analysis to suggest the new modeling to protect from the degrading water

quality phenomenon, overflow of sewer water. With the production of the fine enough resolutions of rainfall data as a input source for the hydrologic modeling, better solutions will be produced to save both life and property. Previous coarse resolution of the rainfall product or point measurement of the rainfall does not satisfy and solve the current hydrologic problems. The hydrologic researchers and engineers have a sense of duty to prepare efficient solutions to minimize the undesirable water related natural disasters.

REFERENCES

- Amburn, S. A., Lang, A. S. I. D., and Buonaiuto, M. A. (2015). "Precipitation forecasting with gamma distribution models for gridded precipitation events in eastern Oklahoma and northwestern Arkansas" *Weather and Forecasting*. Vol. 30, Issue 2, p349-367, DOI:10.1175/WAR-D-14-00054.1
- Arnbjerg-Nielsen, K., Willems, P., Olsson, J., Beecham, S., Pathirana, A., Bülow, G., I., Madsen, H.; Nguyen, V-T.-V. (2013). "Impacts of climate change on rainfall extremes and urban drainage systems: a review" *Water Science and Technology*, Vol. 68, Issue 1, p16-28, DOI: 10.2166/wst.2013.251
- Baeck, M. L., Smith, J. A. (1998). "Rainfall Estimation by the WSR-88D for Heavy Rainfall Events" *Weather and Forecasting*, AMS, Vol. 13
- Boyacioglu, H. (2010). "Detection of seasonal variations in surface water quality using discriminant analysis" *Environmental Monitoring and Assessment*, Vol. 162, p15-20
- Caine, S., Lane, T. P., May, P. T., Jakob, C., Siems, S. T., Manton, M. J., and Pinto, J. (2013). "Statistical assessment of tropical convection-permitting model simulations using a cell-tracking algorithm" *Monthly Weather Review*, Vol. 141, Issue 2, p557-581, DOI:10.1175/MWR-D-11-00274.1
- Cetrone, J. and Houze, R. A. (2006). "Characteristics of tropical convection over the ocean near Kwajalein" *Monthly Weather Review*, Vol. 134, Issue 3, p834-853

Chen, Y., Liu, H., An, J., Görsdorf, U. and Berger, F. H. (2015). “A field experiment on the small-scale variability of rainfall based on a network of micro rain radars and rain gauges” *Journal of Applied Meteorology & Climatology*, Vol. 54, Issue 1, p243-255, DOI: 10.1175/JAMC-D-13-0210.1

Cheng, K. S., Lin, Y. C. and Liou, J. J. (2008). “Rain-gauge network evaluation and augmentation using geostatistics” *Hydrological Processes*, Vol. 22, No. 15, p2554-2564

Chumchean, S., Sharma, A., and Seed, A. (2003). “Radar rainfall error variance and its impact on radar rainfall calibration” *Physics and Chemistry of the Earth, Elsevier*, Vol 28, p 27-39, DOI:10.1016/S1474-7065(03)00005-6

Ciach, G. J. (2003). “Local random errors in tipping-bucket rain gauge measurements” *Journal of Atmospheric & Oceanic Technology*, Vol. 20, Issue 5, p752-759

Ciach, G. J., Habib, E., and Krajewski, W. F. (2003). “Zero-covariance hypothesis in the error variance separation method of radar rainfall verification” *Advances in Water Resources*, DOI: 10.1016/S0309-1708(02)00163-X

Ciach, G. J., and Krajewski, W. F. (1999). “On the estimation of radar rainfall error variance” *Advances in water resources*, Vol. 22, p585-595

Ciach, G. J. and Krajewski, W. F. (2006). “Analysis and modeling of spatial correlation structure in small-scale rainfall in central Oklahoma” *Advances in water resources*, Vol. 29, Issue 10, p1450-1463, DOI: 10.1016/j.advwatres.2005.11.003

Coniglio, M. C., Hwang, J. Y., and Stensrud, D. J. (2010). “Environmental Factors in the Upscale Growth and Longevity of MCSs Derived from rapid update Cycle Analyses”

Monthly Weather Review, AMS, Vol 138, Issue 9, p3514-3539, DOI: 10.1175/2010MWR3233.1

Cristianini, N. and Shawe-Taylor, J. (2003). “An Introduction to Support Vector Machines and other kernel-based learning methods” Cambridge University Press, Cambridge, UK

Cui, Z., Gadian, A., Blyth, A., Crosier, J., and Crawford, I. (2014). “Observations of the Variation in Aerosol and Cloud Microphysics along the 20°S Transect on 13 November 2008 during VOCALS-Rex” Journal of the atmospheric sciences, AMS, Vol 71, DOI:10.1175/JAS-D-13-0245.1

Cunha, L. K., Smith, J. A., Baeck, M. L., and Krajewski, W. F. (2015). “An Early Performance Evaluation of the NEXRAD Dual-Polarization Radar Rainfall Estimates for Urban Flood Applications” Weather and Forecasting, AMS, DOI: 10.1175/WAF-D-13-00046.1

Dong, S., Zhou, D., Ding, W., Gong, J. (2013). “Flow cluster algorithm based on improved K-means method” IETE Journal of Research, Vol 59, Issue 4, p326-333 DOI: 10.4103/0377-2063.118021

Driscoll, E. D., Palhegyi, G. E., Strecker, E. W. and Shelley, P. E. (1989). “Analysis of storm event characteristics for selected rainfall gages throughout the United States” National Service Center for Environmental Publications (NSCEP), Environmental Protection Agency (EPA), Washington DC

Einfalt, T., Hatzfeld, F., Wagner, A., Seltsmann, J., Castro, D., and Frerichs, S. (2009). “URBAS: forecasting and management of flash floods in urban areas.” Urban Water Journal, Vol. 6, Issue 5, p369-374, DOI: 10.1080/15730620902934819

EPA: Environmental Protection Agency (1994). “Part VII, Combined Sewer Overflow (CSO) Control Policy” Notice. Federal Register Vol. 59, No. 75

EPA: Environmental Protection Agency (2004). “Environmental Protection Agency report to congress on impacts and control of combined sewer overflows and sanitary sewer overflows” factsheet files/ 2015-10/documents/csosso_rtc_ factsheet. Pdf

EPA (2004). “Report to Congress: Impacts and control of CSOs and SSOs”
Environmental Protection Agency, EPA 833-R-04-001

Fencel, M., Rieckermann, J., Sýkora, P., Stránský, D. and Bareš, V. (2015). “Commercial microwave links instead of rain gauges: fiction or reality?” Water Science and Technology, Vol. 71, DOI: 10.2166/wst.2014.466

Feng, Z., McFarlane, S. A., Schumacher, C., Ellis, S., Comstock, J., and Bharadwaj, N. (2014). “Constructing a merged cloud-precipitation radar dataset for tropical convective clouds during the DYNAMO/AMIE experiment at Addu Atoll” Journal of Atmospheric & Oceanic Technology, Vol. 31, Issue 5, p1021-1042, DOI: 10.1175/JTECH-D-13-00132.1

Gorgucci, E. and Baldini, L. (2015). “Influence of Beam Broadening on the Accuracy of Radar Polarimetric Rainfall Estimation” Journal of hydrometeorology, Vol 16, DOI: 10.1175/JHM-D-14-0084.1

Ha, K. J., Jeon, E. H. and Oh, H. M. (2007). “Spatial and temporal characteristics of precipitation using an extensive network of ground gauge in the Korean peninsula” Atmospheric Research, Vol. 86, Issue ¾, p330-339, DOI: 10.1016/j.atmosres.2007.07.002.

- Habib, E., Haile, A. T., Tian, Y. and Joyce, R. J. (2012). "Evaluation of the high-resolution CMORPH satellite rainfall product using dense rain gauge observations and radar-based estimates." *Journal of Hydrometeorology*, Vol. 13, Issue 6, p1784-1798, DOI: 10.1175/JHM-D-12-017.1
- Habib, E. and Krajewski, W. F. (2002). "Uncertainty analysis of the TRMM ground-validation radar rainfall products: application to the TEFLUN-B field campaign" *Journal of Applied Meteorology*, Vol. 41, Issue 5, p558-572
- Habib, E., Krajewski, W. F. and Ciach, G. J. (2001). "Estimation of rainfall interstation correlation" *Journal of Hydrometeorology*, Vol. 2, Issue 6, p621-629
- Hlavčová, K., Lapin, M., Valent, P., Szolgay, J., Kohnová, S., Rončák, P. (2015). "Estimation of the impact of climate change-induced extreme precipitation events on floods" *Contributions to Geophysics & Geodesy*, Vol. 45, Issue 3, p173-192, DOI: 10.1515/congeo-2015-0019
- Hogan, R. J. (2007). "A Variational Scheme for Retrieving Rainfall Rate and Hail Reflectivity Fraction from Polarization Radar" *Journal of applied meteorology and climatology*, Vol 46, DOI: 10.1175/JAM2550.1
- Hu, Q. and Feng, S. (2002). "Interannual Rainfall Variations in the North American Summer Monsoon Region: 1900-98" *Journal of climate*, AMS, Vol 15, Issue 10
- Humphrey, M. D., Istok, J. D., Lee, J. Y., Hevesi, J. A. and Flint, A. L. (1997). "A new method for automated dynamic calibration of tipping-bucket rain gauges." *Journal of Atmospheric & Oceanic Technology*, Vol. 14, Issue 6, p1513-1539

- Hunter, S. M. (1996). "WSR-88D Radar Rainfall Estimation: Capabilities, Limitations and Potential Improvements" *Natl. Wea. Dig* 20.4, p26-38
- Hutchinson, T. P. (1997). "A comment on correlation in skewed distributions" *Journal of General Psychology*, Vol. 124, Issue 2, p211-215
- Hyun, J., Rockaway, T. D., French, M. N. (2016). "Case Study: Ground-level Rainfall Variation Over Jefferson County, Kentucky" *ASCE Journal of Hydrologic Engineering*, in press, DOI: 10.1061/(ASCE)HE.1943-5584.0001438
- Hyun, J., Rockaway, T. D. and French, M. N. (2016) "Weather radar rainfall estimation using support vector classification" *Journal of Hydrology*, Elsevier, under review, 2016
- Ice, R. L., Heck, A. K., Cunningham, J. G., and Zittel, W. D. (2014). "Challenges of Polarimetric Weather Radar Calibration" *The eighth european converence on radar in meteorology and hydrology*
- Jung, Y., Kim, H., Baik, J. and Chio, M. (2014). "Rain-gauge network evaluation using spatiotemporal correlation structure for semi-mountainous regions" *Atmospheric & Oceanic Sciences*, Vol. 25, Issue 2, p267-278, DOI: 10.3319/TAO.2013.10.31.01(Hy)
- Kalogiros, J., Anagnostou, M. N., Anagonostou, E. N., Montopoli, M., Picciotti, E., and Marzano, F. S. (2013). "Correction of Polarimetric Radar Reflectivity measurements and Rainfall Estimates for Apparent Vertical Profile in Stratiform Rain" *Journal of applied meteorology and climatology*, Vol 52, DOI: <http://dx.doi.org/10.1175/JAMC-D-12-0140.1>
- Karamouz, M., Nazif, S. (2013). "Reliability-based flood management in urban watersheds considering climate change impacts" *Journal of Water Resources Planning &*

Management, Vol. 139, Issue 5, p520-533, DOI: 10.1061/(ASCE)WR.1943-5452.0000345

Khalid, M. N. (2011). "Cluster analysis – a standard setting technique in measurement and testing." *Journal of Applied Quantitative Methods*, Vol. 6, Issue 2, p46-58

Kowalski, C. J. (1972). "On the effect of non-normality on the distribution of the sample product-moment correlation coefficient" *Applied Statistics*, Vol. 21, Issue 1, p1-12

Krajewski, W. F. and Smith, J. A. (2002). "Radar hydrology: rainfall estimation." *Advances in Water Resources*, Vol. 25, Issue 8-12, p1387-1394

Krajewski, W. F. and Vignal, B. (2001). "Evaluation of anomalous propagation echo detection in WSR-88D data: A large sample case study" *Journal of Atmospheric and Oceanic Technology*, Vol. 18

Krajewski, W. F., Villarini, G., and Smith J. A. (2010). "Radar-rainfall uncertainties: Where are we after thirty years of effort?" *American Meteorological Society (AMS)*, DOI: 10.1175/2009BAMS2741.1

Kundzewicz, Z. W., Kanae, S., Seneviratne, S. I., Handmer, J., Nicholls, N., Peduzzi, P., Mechler, R., Bouwer, L. M., Arnell, N., Mach, K., Muir-Wood, R., Brakenridge, G. R., Kron, W., Benito, G., Honda, Y., Takahashi, K., and Sherstyukov, B. (2014). "Flood risk and climate change: global and regional perspectives" *Hydrological Sciences Journal/Journal des Sciences Hydrologiques*, Vol. 59, Issue 1, p1-28. 28p. DOI: 10.1080/02626667.2013.857411.

Louisville Metropolitan Sewer District (2014). “CSO characteristics report-Combined Sewer Overflow Fact Sheet -CSO130” Metropolitan Sewer District (MSD), Louisville, Kentucky, 2014

Lyandres, O. and Welch, L. C. (2012). “Reducing combined sewer overflows in the Great Lakes: Why investing in infrastructure is critical to improving water quality”, Alliance for the Great Lakes

Mamo, T. G. (2015). “Evaluation of the potential impact of rainfall intensity variation due to climate change on existing drainage infrastructure” Journal of Irrigation and Drainage Engineering, Vol. 141, Issue 10, p1-7, DOI: 10.1061/(ASCE)IR.1943-4774.0000887, 2015

Mandapaka, P. V. and Qin, X. (2013). “Analysis and characterization of probability distribution and small-scale spatial variability of rainfall in Singapore using a dense gauge network” Journal of Applied Meteorology & Climatology, Vol. 52, Issue 12, p2781-2796, DOI: 10.1175/JAMC-D-13-0115.1

Martinez, A. M. and Kak, A. C. (2001). “PCA versus LDA” IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 23, Issue 2, p228-233, DOI:10.1109/34.908974, 2001

Menne, M. J., Durre, I., Vose, R. S., Gleason, B. E. and Houston, T. G. (2012). “An overview of the global historical climatology network-daily database” Global historical climatology network - daily (GHCN-Daily), Journal of Atmospheric & Oceanic Technology, Vol. 29, Issue 7, p897-910, DOI: 10.1175/JTECH-D-11-00103.1

- Morin, E., Krajewski, W. F., Goodrich, D. C., Gao, X., and Sorooshian, S. (2003). “Estimating Rainfall Intensities from Weather Radar Data: The Scale-Dependency Problem” *Journal of hydrometeorology*, Vol 4
- Morrison, H. and Milbrandt, J. A. (2015). “Parameterization of Cloud Microphysics Based on the Prediction of Bulk Ice Particle Properties. Part I: Scheme Description and Idealized Tests” *Journal of the atmospheric sciences*, AMS, Vol 72, DOI: 10.1175/JAS-D-14-0065.1
- Morita, M. (2011). “Quantification of increased flood risk due to global climate change for urban river management planning” *Water Science & Technology*, Vol. 63, Issue 12, p2967-2974, DOI: 10.2166/wst.2011.172, 2011
- MSD (2014) “CSO characteristics report-Combined Sewer Overflow Fact Sheet (CSO130)” Metropolitan Sewer District (MSD), Louisville, Kentucky, 2014
- Nešpor, V. and Sevruk, B. (1998). “Estimation of wind-induced error of rainfall gauge measurements using a numerical simulation” *Journal of atmospheric and oceanic technology*, Vol. 16
- Nogueira, R. C. and Keim, B. D. (2010). “Annual Volume and Area Variations in Tropical Cyclone Rainfall over the Eastern United States” *Journal of climate*, Vol 23, Issue 16, p 4363-4374, DOI: 10.1175/2010JCLI3443.1
- National Weather Service (2016). “Internet Services Team, US Dept of Commerce, National Oceanic and Atmospheric Administration” National Weather Service, water.weather.gov/precip

- Panthou, G., Mailhot, A., Laurence, E., and Talbot, G. (2014). "Relationship between Surface Temperature and Extreme Rainfalls: A Multi-Time-Scale and Event-Based Analysis" *Journal of hydrometeorology*, Vol 15, DOI: 10.1175/JHM-D-14.0020.1
- Price, K., Purucker, S. T., Kraemer, S. R., Babendreier, J. E. and Knighters, C. D. (2014). "Comparison of radar and gauge precipitation data in watershed models across varying spatial and temporal scales" *Hydrological Processes*, Vol. 28, No. 9, p3505-3520
- Peter, J. R., Manton, M. J., Potts, R. J., May, P. T., Collis, S. M., and Wilson, L. (2015). "Radar-derived statistics of convective storms in Southeast Queensland" *Journal of applied meteorology & climatology*, Vol. 54, Issue 10, p1985-2008, DOI: 10.1175/JAMC-D-13-0347.1, 2015
- Qin, X. S. and Lu, Y. (2014). "Study of climate change impact on flood frequencies: a combined weather generator and hydrological modeling approach" *Journal of Hydrometeorology*, Vol. 15, Issue 3, p1205-1219, DOI: 10.1175/JHM-D-13-0126.1
- Riboust, P. and Brissette, F. (2015). "Climate Change Impacts and Uncertainties on Spring Flooding of Lake Champlain and the Richelieu River" *Journal of the American Water Resources Association*, Vol. 51 Issue 3, p776-793, DOI: 10.1111/jawr.12271
- Romnée, A., Evrard, A. and Trachte, S. (2015). "Methodology for a stormwater sensitive urban watershed design" *Journal of Hydrology*, Vol. 530, p87-102, DOI: 10.1016/j.jhydrol.2015.09.054
- Salathé, E. P., Jr., Hamlet, A. F., Mass, C. F., Lee, S., Stumbaugh, M. and Steed, R. (2014). "Estimates of twenty-first-century flood risk in the pacific northwest based on

regional climate model simulations” *Journal of Hydrometeorology*, Vol. 15, Issue 5, p1881-1899, DOI: 10.1175/JHM-D-13-0137.1.

Sangam, S., Futaba, K. and Takashi, N. (2008). “Use of principal component analysis, factor analysis and discriminant analysis to evaluate spatial and temporal variations in water quality of the Mekong River” *Journal of Hydroinformatics*, Vol. 10, Issue 1, p43-56

Scheuerer, M. and Hamill T. M. (2015). “Statistical postprocessing of ensemble precipitation forecasts by fitting censored, shifted gamma distribution” *Monthly Weather Review*. Vol. 143, Issue 11, p4578-4596, DOI:10.1175/MWR-D-15-0061.1.

Schmidt, J. M., Flatau, P. J., Harasti, P. R., Yates, R. D., Littleton, R., Pritchard, M. S., Fischer, J. M., Fischer, E. J., Kohri, W. J., Better, J. R., Richman, S., Baranowski, D. B., Anderson, M. J., Fletcher, E., and Lando, D. W. (2012). “Radar Observations of individual rain drops in the free atmosphere” *J. PNAS*, Vol 109, DOI: 10.1073/pnas.1117776109/-/DCSupplemental

Seo, D. J. and Breidenbach, J. P. (2002). “Real-time correction of spatially nonuniform bias in radar rainfall data using rain gauge measurements.” *Journal of Hydrometeorology*. Vol. 3, Issue 2, p93

Seo, D. J., Breidenbach, J. P, Fulton, R., and Miller, D. (2000). “Real-time adjustment of range-dependent biases in WSR-88D rainfall estimates due to nonuniform vertical profile of reflectivity” *Journal of Hydrometeorology*, 2000, Vol 1

- Smith, J. A., Baeck, M. L., Meierdiercks, K. L., Miller A. J., Krajewski, W. F. (2007). "Radar Rainfall Estimation for Flash Flood forecasting in Small Urban watersheds" *Advances in Water Resources*, Elsevier, doi:10.1016/j.advwatres.2006.09.00
- Tahmasebi, P., Hezarkhani, A., and Mortazavi, M. (2010). "Application of discriminant analysis for alteration separation; sungun copper deposit, East Azerbaijan, Iran" *Australian Journal of Basic and Applied Sciences*, Vol. 6, Issue 4, p564-576
- Tokay, A. and Öztürk, K. (2012). "An experimental study of the small-scale variability of rainfall" *Journal of Hydrometeorology*, Vol. 13, Issue 1, p351-365, DOI: 10.1175/JHM-D-11-014.1
- Ulbrich, C. W. and Lee, L. C. (1999). "Rainfall Measurement Error by WSR-88D Radars due to Variations in Z-R Law Parameters and the Radar Constant" *Journal of atmospheric and oceanic technology*, Vol 16
- Vignal, B. and Krajewski, W. F. (2001). "Large-sample evaluation of two methods to correct range-dependent error for WSR-88D rainfall estimates" *Journal of Hydrometeorology*, 2001, Vol 2
- Villarini, G., Mandapaka, P. V., Krajewski, W. F. and Moore, R. J. (2008). "Rainfall and sampling uncertainties: A rain gauge perspective" *Journal of Geophysical Research: Atmospheres*, Vol. 113, Issue 11, DOI: 10.1029/2007JD009214
- Wang, J., Fisher, B. L. and Wolff, D. B. (2008). "Estimating rain rates from tipping-bucket rain gauge measurements" *Journal of Atmospheric & Oceanic Technology*, Vol. 25, Issue 1, p43-56, DOI: 10.1175/2007JTECHA895.1.

Wang, X., Liao, J., Zhang, J., Shen, C., Chen, W., Xia, B. and Wang, T. (2014). “A numeric study of regional climate change induced by urban expansion in the Pearl River delta, China” *Journal of Applied Meteorology & Climatology*, Vol. 53, Issue 2, p346-362, DOI: 10.1175/JAMC-D-13-054.1

Wang, Y., Huang, Y., Rattanachot, W., Lau, K. K. (W.), and Suwansawas, S. (2015). “Improvement of Pavement Design and Management for More Frequent Flooding Caused by Climate Change” *Advances in Structural Engineering*, Vol. 18 Issue 4, p487-496. 10p. DOI: 10.1260/1369-4332.18.4.487

Wernstedt, K. and Carlet, F. (2014). “Climate Change, Urban Development and Storm Water: Perspective from the field” *Journal of Water Resource Planning and Management*, ASCE, Vol. 140, No. 4, ISSN 0733-9496/2014/4-543-552/\$25.00

Woodward, M., Kapelan, Z., and Gouldby, B. (2014). “Adaptive Flood Risk Management Under Climate Change Uncertainty Using Real Options and Optimization” *Risk Analysis*, An International Journal, Vol. 34 Issue 1, p75-92. 18p. DOI: 10.1111/risa.12088

Wright, D. B., Smith, J. A., Villarini, G., and Baeck, M. L. (2014). “Long-Term High-Resolution Radar rainfall Fields for Urban Hydrology” *Journal of the American water resources association*, Vol 50, No 3

Xie, J., Hone, K., Xie, W., Gao, X., Shi, Y., and Liu, X. (2013). “Extending twin support vector machine classifier for multi-category classification problems” *Intelligent Data Analysis*, Vol 17, Issue 4, p649-664, DOI: 10.3233/IDA-130598

Yang, L., Smith, J. A., Wright, D. B., Baeck, M. L., Villarini, G., Tian, F., and Hu, H. (2013). “Urbanization and climate change: An examination of nonstationarities in urban

flooding” Journal of Hydrometeorology, Vol. 14, Issue 6, p1791-1801, DOI:
10.1175/JHM-D-12-095.1

Zawilski, M and Brzeziska, A. (2014). Areal rainfall intensity distribution over an urban area and its effect on a combined sewerage system, Urban Water Journal, Vol. 11, Issue 7, p532-542, DOI:10.1080/1573062X.2013.831909

Zhang, X. and Yan, X. (2014). “Temporal change of climate zones in China in the context of climate warming” Theoretical and Applied Climatology, Vol. 115, p167-175, DOI 10.1007/s00704-013-0887-z

APPENDIX

Appendix A. Graphical Slides of Workflow



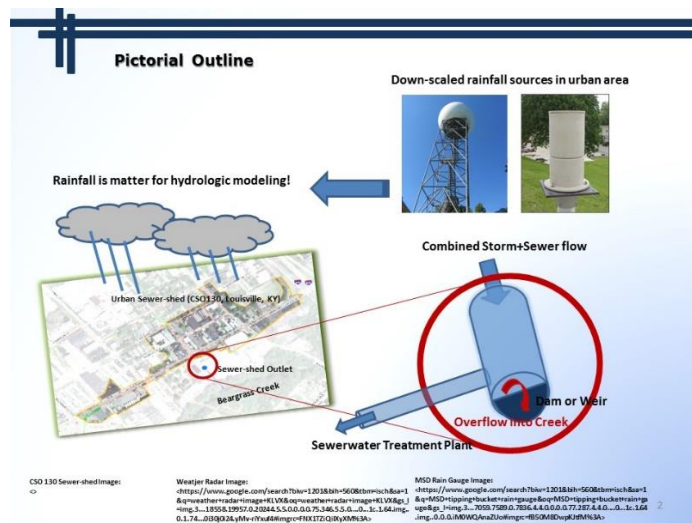
Final Oral Presentation:

SPATIOTEMPORAL RAINFALL STRUCTURE AND OPTIMIZED LOCAL RADAR RAINFALL APPLICATION TO URBAN WATERSHED, LOUISVILLE, KENTUCKY, 2010-2014

Jin-Young Hyun

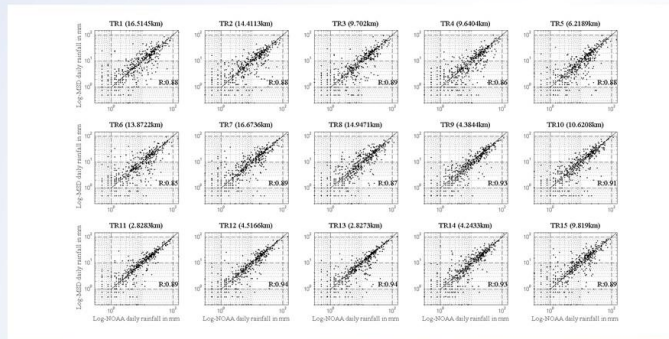
PhD Candidate
Department of Civil & Environmental Engineering

1



II. SPATIOTEMPORAL STRUCTURE OF RAINFALL

Rain Gauge Validation: daily

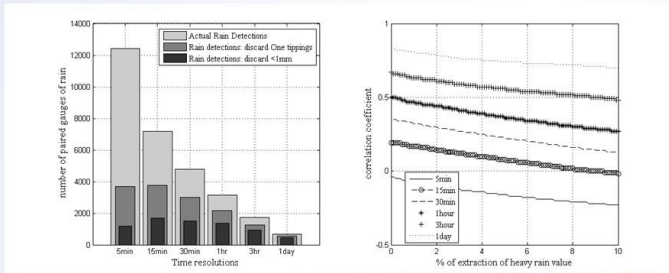


High correlations are observed, higher than 0.86 and moderately related to the spatial correlation in daily resolution.

6

II. SPATIOTEMPORAL STRUCTURE OF RAINFALL

Rain Gauge Validation: statistical consideration



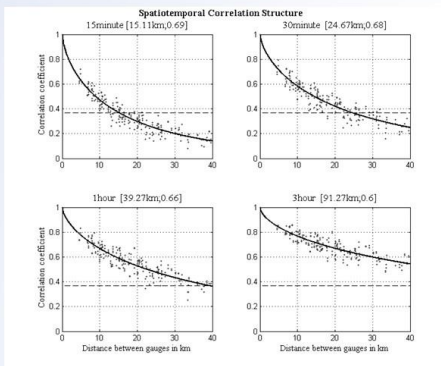
One tipping (0.254mm) data values are discarded.

No cascade drop observed by removing the extreme data.

7

II. SPATIOTEMPORAL STRUCTURE OF RAINFALL

Spatiotemporal Correlation Structure



Pearson Product-Moment
Correlation Coefficient
(PPMCC)

Concave shapes by distance.

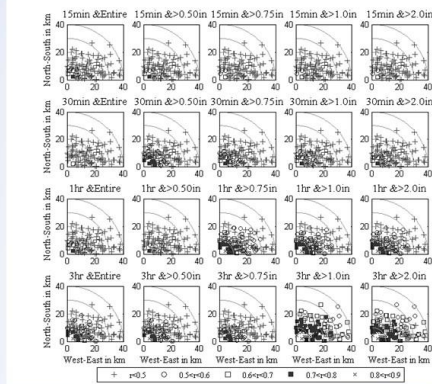
The correlation drops
significantly as long as the
temporal resolution decreases.

$$R_{ij} = \frac{\sum_{l=1}^n (G_{1l} - \bar{G}_1)(G_{2l} - \bar{G}_2)}{\sqrt{\sum_{l=1}^n (G_{1l} - \bar{G}_1)^2} \sqrt{\sum_{l=1}^n (G_{2l} - \bar{G}_2)^2}} \quad (\text{Equation 2.1})$$

8

II. SPATIOTEMPORAL STRUCTURE OF RAINFALL

Spatiotemporal Correlation Structure



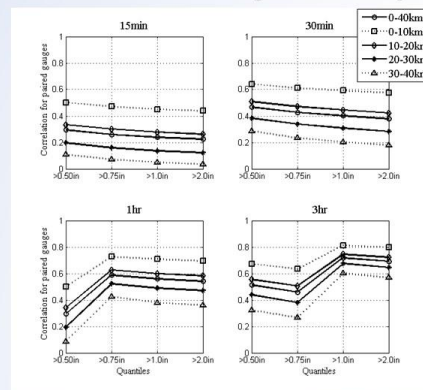
Directional spatiotemporal correlation

In sub-hourly temporal scale, the correlation dropped by increasing rainfall intensity.

9

II. SPATIOTEMPORAL STRUCTURE OF RAINFALL

Quantile analysis

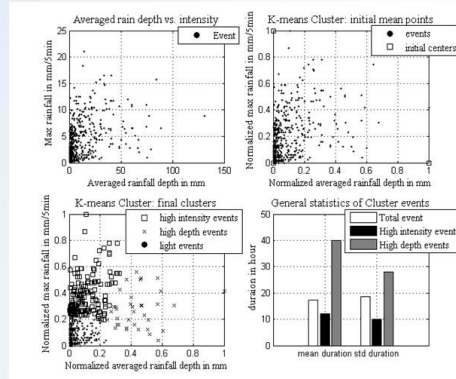


Decreasing tendencies observed with increasing rainfall intensity in sub-hourly temporal resolution.

10

III. RAINFALL CLASSIFICATION

Rainfall Event Classification: K-means



Consider rainfall's continuity

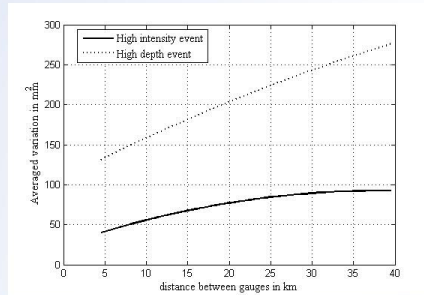
Using rainfall event definition/regulation by Environmental Protection Agency (EPA)

K-means clustering is used to cluster the rainfall event to three groups.

11

III. RAINFALL CLASSIFICATION

Variogram analysis



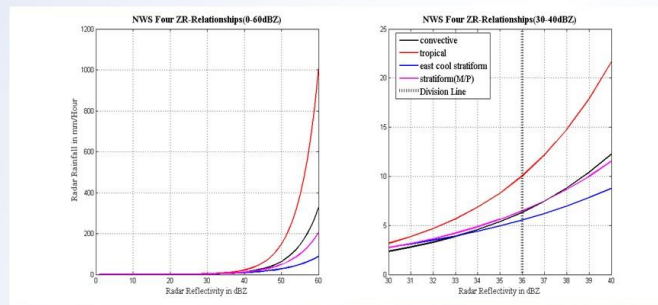
Variogram demonstrates the spatial boundary of the storm events – high intensity event and high depth event groups

$$\gamma_{i,j} = \frac{1}{2} E \left[\left((G_i - G_c) - (G_j - G_c) \right)^2 \right] \text{ (Equation 3.1)}$$

12

IV. RADAR RAINFALL OPTIMIZATION

NWS Z-R relationships

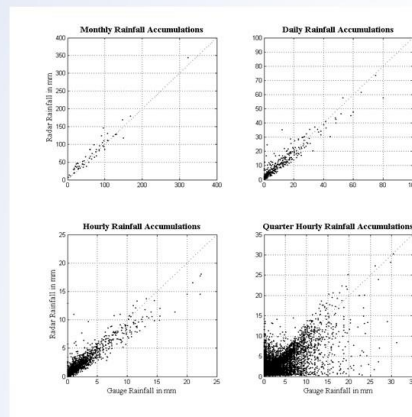


National Weather Service (NWS) uses four different reflectivity-rainfall intensity (Z-R) relationships for this region.

13

IV. RADAR RAINFALL OPTIMIZATION

NWS radar rainfall validation

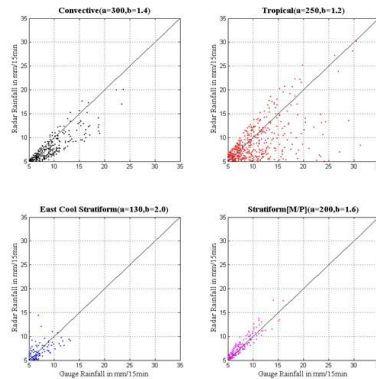


Quality of the NWS radar rainfall estimation dropped in sub-hourly temporal resolution.

14

IV. RADAR RAINFALL OPTIMIZATION

Best Matching by storm types

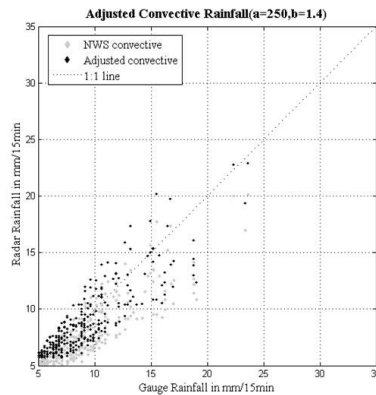


Convective and tropical type of the rainfalls need adjusted Z-R relationships for the local radar rainfall optimization.

15

IV. RADAR RAINFALL OPTIMIZATION

Convective storm adjustment

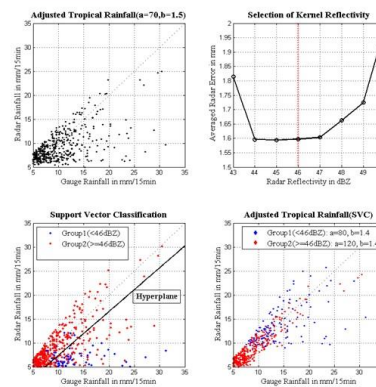


The underestimation problem is solved by the minimum RMSE error.

16

IV. RADAR RAINFALL OPTIMIZATION

Tropical storm adjustment: SVC



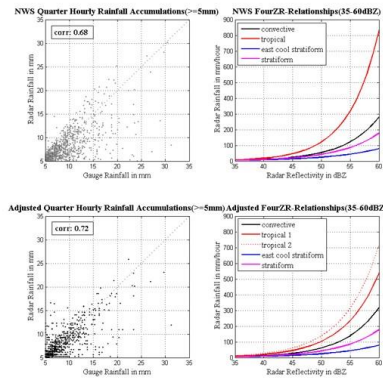
Support Vector Classification (SVC) is used to create a new Z-R relationship for the underestimated group.

$$\hat{y} = \text{sgn} \sum_{i=1}^n w_i y_i k(R_i, \hat{R}) \quad (\text{Equation 4.2})$$

17

IV. RADAR RAINFALL OPTIMIZATION

Optimized Z-R relationships

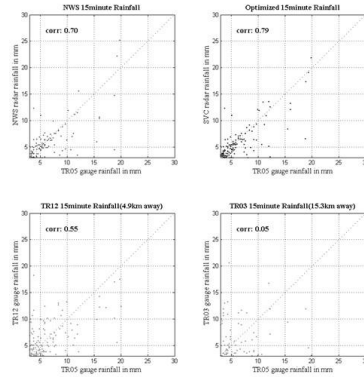


The quality of the radar rainfall has been improved by the newly assigned Z-R relationships.

18

V. APPLICATION OF RADAR RAINFALL TO URBAN AREA

Validation of radar rainfall

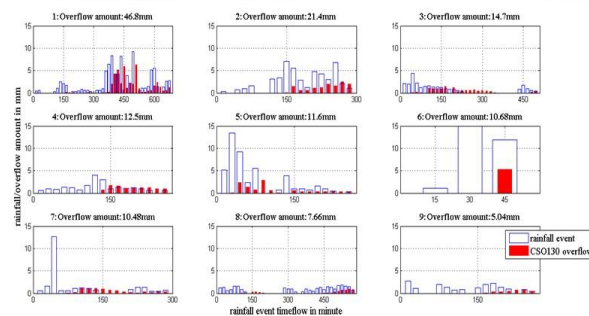


The optimized radar rainfall product (quarter hourly) is validated by rain gauge (TR05) and compared with other rainfall measuring sources – NWS radar rainfall observation, MSD gauges (TR12 and TR03)

19

V. APPLICATION OF RADAR RAINFALL TO URBAN AREA

Coupled rainfall-CSO130 overflow

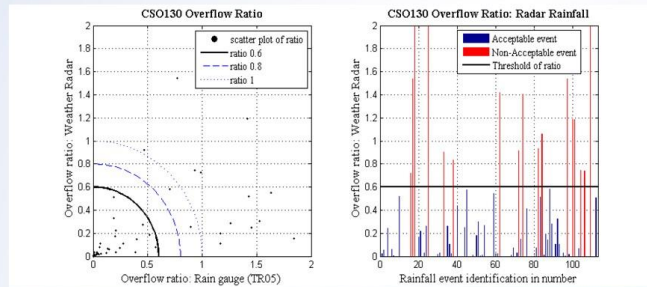


By EPA rainfall event definition rainfall events were predefined and the corresponding overflows are coupled. The overflows does not governed by the rainfall volume.

20

V. APPLICATION OF RADAR RAINFALL TO URBAN AREA

Filtering coupled events

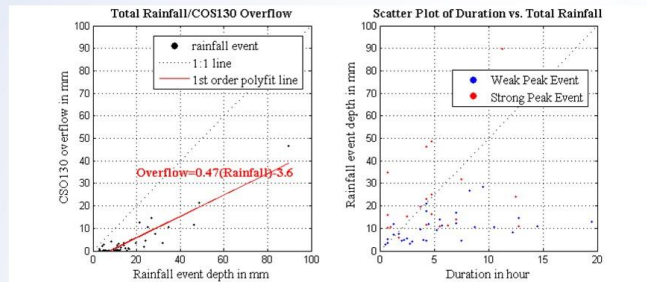


52 overflow events remains for 3 years of study period.

21

V. APPLICATION OF RADAR RAINFALL TO URBAN AREA

Rainfall characteristics and overflow

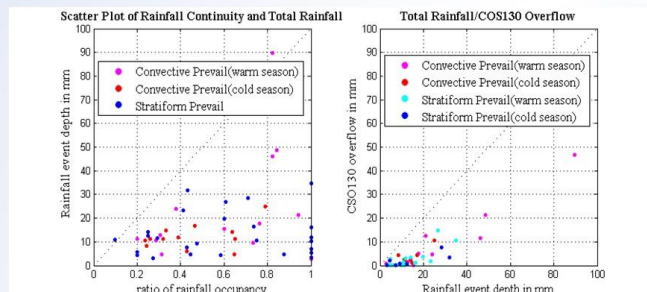


Mainly, the rainfall volume has strong linearity with overflow, but other factors need to be considered to analyze the overflows in small scale sewer-shed, CSO130.

22

V. APPLICATION OF RADAR RAINFALL TO URBAN AREA

Rainfall characteristics and overflow

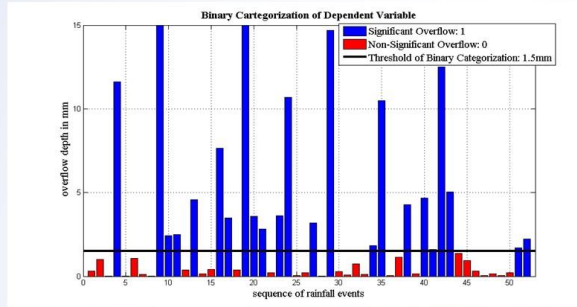


Mainly, the rainfall volume has strong linearity with overflow, but other factors need to be considered to analyze the overflows in small scale sewer-shed, CSO130.

23

V. APPLICATION OF RADAR RAINFALL TO URBAN AREA

Discriminant analysis



Discriminant analysis (DA) is used to classify the factors of the rainfall characteristics.

24

V. APPLICATION OF RADAR RAINFALL TO URBAN AREA

Discriminant analysis

Groups	Variables	Mean
Non-Significant Overflow (group 0)	Duration (hour)	4.57
	Rain Total (mm)	8.85
	Rain Peak (mm/15min)	3.76
	Rain Type (convective ratio)	0.47
	Rain Continuity ratio	0.56
Significant Overflow (group 1)	Duration (hour)	6.14
	Rain Total (mm)	23.7
	Rain Peak (mm/15min)	6.20
	Rain Type (convective ratio)	0.50
Total Events	Rain Continuity ratio	0.60
	Duration (hour)	5.37
	Rain Total (mm)	15.4
	Rain Peak (mm/15min)	4.84
	Rain Type (convective ratio)	0.48
	Rain Continuity	0.58

25

V. APPLICATION OF RADAR RAINFALL TO URBAN AREA

Discriminant analysis

Overflow Severity			Predicted Group Membership		Total
			Non-Significant	Significant	
Original	Count	Non-Significant	29	0	29
		Significant	11	12	23
	Percentage	Non-Significant	100	0	100
		Significant	47.8	52.2	100

26

Results

1. PPMCC reveals the limitation of the 1-D rainfall measurement system, in particular sub-hourly rainfall application.
2. Rainfall events are created by EPA and grouped by K-means clustering method. Variogram shows the boundaries of the correlation for grouped rainfall event.
3. Radar rainfall need to be optimized by SVC, Imbedding the local radar and regional rainfall characteristics.
4. Optimized radar rainfall has the highest correlation to apply to the small scale sewer-shed, CSO130. Rainfall volume is mainly governed the sewer overflow in combined sewer system, but the other factors; intensity, duration, continuity and storm types, should be considered.

27

Conclusion

Rainfall event types with CSO overflow pattern can be used for the hydrologic modeling to mitigate sewer water overflow.



Considering the combination of the stormwater-mitigating structures (Gray and Green) can applied due to the sewershed size(shape), land use, imperviousness, and sewer water distribution system.

The fulfillment of the overflow regulations (frequency and volume) can be achieved efficiently and economically.

28

Future works

This research focused on the production of reliable rainfall estimation and its application over the small scale urban watershed. In short, the urban hydrology required the sub-hourly analysis to suggest the new modeling to protect from the degrading water quality phenomenon, overflow of sewer water. With the production of the fine enough resolutions of rainfall data as a input source for the hydrologic modeling, better solutions will be produced to save both life and property. Previous coarse resolution of the rainfall product or point measurement of the rainfall does not satisfy and solve the current hydrologic problems. The hydrologic researchers and engineers have a sense of duty to prepare efficient solutions to minimize the undesirable water related natural disasters.

29

Appendix B. Theoretical Equations

Equation 2.1

$$R_{i,j} = \frac{\sum_{i=1}^n (G_{1,i} - \overline{G_1})(G_{2,i} - \overline{G_2})}{\sqrt{\sum_{i=1}^n (G_{1,i} - \overline{G_1})^2} \sqrt{\sum_{i=1}^n (G_{2,i} - \overline{G_2})^2}}$$

$R_{i,j}$ = Pearson's product moment correlation coefficient between paired gauges

$G_{1,i}, G_{2,i}$ = indices of any possible two network gauges

$\overline{G_1}, \overline{G_2}$ = mean values of each gauge in the pair

Equation 2.2

$$R_{i,j} = \exp \left[- \left(\frac{d_{i,j}}{d_c} \right)^s \right], 0 \leq s \leq 2$$

d_c = critical distance

$d_{i,j}$ = distance between two gauges in paired set

s = shape parameter

Equation 3.1

$$\gamma_{i,j} = \frac{1}{2} E \left[\left((G_i - G_c) - (G_j - G_c) \right)^2 \right]$$

$\gamma_{i,j}$ = variogram

G_c = mean value of gauge accumulation of rainfall event for the same cluster

G_i, G_j = gauge rainfall accumulation in an event

Equation 4.1

$$R_Q(t, x) = \sum_{i=1}^n \left(\frac{R_i}{4} \right) \left(\frac{t_{i+1} - t_i}{2} \right)$$

$R_Q(t, x)$ = weighted accumulation of quarter hourly rainfall capsule at the fixed location where reference data are corresponded in arbitrary time, t

n = influential number of reflectivity values for $R_Q(t, x)$ which fall into the quarter hourly capsule and front and the rear reflectivity value when it is influential to the capsule.

R_i = converted rainfall rate in millimeter per hour among four different NWS Z-R relationships

t_i = time of observation of instantaneous radar base scan

Equation 4.2

$$\hat{y} = \text{sgn} \sum_{i=1}^n w_i y_i k(R_i, \hat{R})$$

R_i and y_i are instance-based learners for the i -th training example

w_i = a corresponding weight

$\hat{R}$ = an unlabeled input

k = a function of the weighted sum of similarities which is called kernel

$\hat{y}$ = a prediction label of kernelized binary classifier

Equation 5.1

$$T = \frac{(\vec{W} \cdot \vec{\mu}_{\text{Significant}} - \vec{W} \cdot \vec{\mu}_{\text{Non-Significant}})^2}{\vec{W}^T \text{COV}(R_{i,j})_{\text{Significant}} \vec{W} + \vec{W}^T \text{COV}(R_{i,j})_{\text{Non-Significant}} \vec{W}}$$

T = a decision factor for the clustering of Significant and Non-Significant groups of the overflow event

$\vec{\mu}_i$ = mean value for the clustered group.

$R_{i,j}$ = rainfall variables among rainfall depth, duration, intensity, type, continuity.

Appendix C. MATLAB Codes

This research requires programming intensive procedures to conducting data extraction, assimilation and evaluation. MATLAB is mostly used as a programming tool to handling the analysis and creating figures.

The gauge and radar rainfall data are extracted and paired with spatiotemporal synchronization (local time which considers daylight saving) [Appendix C. Chapter2. b. Temporal matching MSD gauge network (15 gauges over Louisville area)]. The 1 tipping amount of the rainfall (0.01inch) in 5 minute time resolution for chapter 2 is discarded due to the data quality issue as well as in 15 minute resolution for chapter 3, 4, and 5. The gauge rainfall matrix consists of 5 minute rainfall values from 15 multiples gauges which is operated by MSD [Appendix C. Chapter2. a. gauge data extraction and preprocessing]. The smallest unit of the gauge rainfall data were accumulated by quarter hourly, half hourly, one hourly and three hourly for the spatial variation with different temporal scales [Appendix C. Chapter 2. c. Accumulation of raw data into hourly]. The general statistics and spatiotemporal structure of the ground-based rainfall is investigated by the code of [Appendix C. Chapter 2. d. General statistics and spatiotemporal correlation for MSD rain gauge network, Louisville, KY].

In chapter 3, the rainfalls are encapsulized by the EPA rainfall event definition and clustered by [Appendix C. Chapter3. a. Clustering of rainfall events]. After the grouping of the rainfall event by intense and volumetric considerations, the variogram analysis has been performed using [Appendix C. Chapter3. b. Clustered variogram].

The core optimization of the local radar rainfall estimation by using the [Appendix C. Chapter 4. a. Support Vector Classification of tropical type rainfall values].

The support vector classification created an additional Z-R relationship in the tropical storm group.

Chapter 5 illustrates the relationship between the encapsulated rainfall event and its corresponding watershed overflow. [Appendix C. Chapter 5. a. Generate matched Hydrograph and CSO130 overflow Hydrograph] performed the coupling of these two hydrologic phenomena.

Chapter 2

a. gauge data extraction and preprocessing

```

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%% Data Preprocessing
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%% data quality check

% 1. annual accumulation

% col 6-22: MSD rainfall 5minute Accumulation in inch &
mm

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

% 01/20/2015

% Developed by Jin-Young Hyun

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

clear all;

accum5minTR17n18='Z:/DATA/GroundBasedRainfall/MSD
rainGauge/MSDrainfall_inch_mm/RainGaugeMSDincludeT
R1718inMM.out';

open5min=dlmread(accum5minTR17n18);

% year accumulation

RaccumYr=zeros(5,17);

for i=2010:2014

    for j=6:22

        findYr=find(open5min(:,1)==i);

        yrSum=sum(open5min(findYr,j));

        RaccumYr((i-2009),(j-5))=yrSum;

        disp(yrSum);

    end

end

% year average

RyrAvg=zeros(5,1);

for i=1:5

    RyrAvg(i,1)=mean(RaccumYr(i,:));

    disp(RyrAvg(i,1));

end

totalAvg=mean(RyrAvg);

% plot the bar graph of yearly accumulation

figure;

YEARlist={'2010','2011','2012','2013','2014'};

```

```

%XTICKS={'01','02','03','04','05','06','07','08','09','10','11','12',
,'13','14','15','17','18'};

XTICKS=['01','02','03','04','05','06','07','08','09','10','11','12','
13','14','15','17','18'];

Ravg=zeros(1,17);

for i=1:5

    subplot(5,1,i);

    bar(1:17,RaccumYr(i,:),0.3,'black');

    avgYr=mean(RaccumYr(i,:));

    Ravg(1,1:17)=avgYr;

    hold on;

    plot(1:17,Ravg,'black:');

    hold off;

    titleName=strcat('Year: ',YEARlist(i));

    title(titleName,'fontweight','bold','fontsize',11);

    legend('Annual rainfall for each gauge','Averaged annual
rainfall');

    ylim([500 2000]);

    set(gca,'XLim',[0.5 17.5])

    set(gca,'XTick',[1:1:17])

    set(gca,'XTickLabel',XTICKS)

    if i==3

        ylabel('Yearly rainfall accumulated rainfall in mm');

    elseif i==5

        xlabel('MSD rain gauges ID');

    end

end

% open NOAA rain gauge Galena4_3ENE and TR17 and
TR18

TR17='Z:/DATA/GroundBasedRainfall/MSDrainGauge/MS
Drainfall_inch_mm/numericRG_MSD_TR17_MtStFrancis.o
ut';

TR18='Z:/DATA/GroundBasedRainfall/MSDrainGauge/MS
Drainfall_inch_mm/numericRG_MSD_TR18_IVY_Tech.out
';

openTR17=dlmread(TR17);

openTR18=dlmread(TR18);

NOAAgalena='Z:/DATA/GroundBasedRainfall/Accumulatio
nMSDgaugeDaily_discardedTR17_18MM.out';

openNOAA=dlmread(NOAAgalena);

Rmonth=zeros(12,5);

%findMissTR17=find(openTR17(:,7)<0);

```



```

%findMissTR18=find(openTR18(:,7)<0);

for i=1:12

    Rmonth(i,1)=i;

    findMonTR17=find(openTR17(:,2)==i);

    findMonTR18=find(openTR18(:,2)==i);

    findMonNOAA=find(openNOAA(:,2)==i);

    sumTR17=sum(openTR17(findMonTR17,7));

    sumTR18=sum(openTR18(findMonTR18,7));

    % other MSD gauges

    sumOtherTRs=sum(openNOAA(findMonNOAA,4:18));

    avgOtherTRs=mean(sumOtherTRs);

    % NOAA_Galena4_3ENE

    sumGalena=sum(openNOAA(findMonNOAA,22));

    % store

    Rmonth(i,2)=sumTR17;

    Rmonth(i,3)=sumTR18;

    Rmonth(i,4)=avgOtherTRs;

    Rmonth(i,5)=sumGalena;

end

% monthly bar graph for TR17 and TR18 validation

figure;

fH=gcf;

colormap('gray');

x=Rmonth(:,1);

y=Rmonth(:,2:5);

b=bar(x,y);

xlim([0 13]);

ylim([200 1000]);

xlabel('Months');

ylabel('Rainfall accumulation in mm (2010-2014)');

legend('MSD:TR17-MtStFrancic', 'MSD:TR18- IVY tech',
'MSD: Averaged other Gauges', 'NOAA: Galena4-3ENE');

% averaged monthly accumulation of TR01-15 and NOAA1-
12 for each year

RmonYr=zeros(12,3,5);

YEAR=2010;

% discard -999

MISS=0;

for i=1:length(openNOAA)

    for j=4:30

        value=openNOAA(i,j);

        if value>=0

            continue;

        else

            openNOAA(i,j)=0;

            MISS=MISS+1;

            XX=['Missing: ',num2str(MISS)];

            disp(XX);

        end

    end

end

for i=1:5

    Rm=zeros(12,2);

    for j=1:12

        findMonYr=find(openNOAA(:,1)==YEAR &
openNOAA(:,2)==j);

        sumTRs=sum(openNOAA(findMonYr,4:18));

        avgTRs=mean(sumTRs);

        sumNOAAs=sum(openNOAA(findMonYr,19:30));

        avgNOAAs=mean(sumNOAAs);

        RmonYr(j,1,i)=j;

        RmonYr(j,2,i)=avgTRs;

        RmonYr(j,3,i)=avgNOAAs;

    end

    YEAR=YEAR+1;

end

Rstreach=zeros(60,3);

for i=1:60

    Rstreach(i,1)=i;

end

Rstreach(1:12,2:3)=RmonYr(:,2:3,1);

Rstreach(13:24,2:3)=RmonYr(:,2:3,2);

Rstreach(25:36,2:3)=RmonYr(:,2:3,3);

Rstreach(37:48,2:3)=RmonYr(:,2:3,4);

Rstreach(49:60,2:3)=RmonYr(:,2:3,5);

figure;

```

```

colormap('gray');
x=Rstreach(:,1);
y=Rstreach(:,2:3);
b=bar(x,y);
xlim([0 61]);
ylim([0 400]);
legend('MSD rain gauges','NOAA rain gauges');
xlabel('Months');
ylabel('Rainfall accumulation in mm');
figure;
scatter(Rstreach(:,3),Rstreach(:,2),50,'filled','black');
% monthly comparison with 15 MSD and 12 gauges
RcorMon=corrcoef(Rstreach(:,2),Rstreach(:,3));
Rdiff=zeros(60,2);

for i=1:60
    diff=Rstreach(i,2)-Rstreach(i,3);
    Rdiff(i,1)=diff;
    Rdiff(i,2)=abs(diff);
    disp(Rdiff(i,1));
end
maxDiff=max(Rdiff(:,2));
avgDiff=mean(Rdiff(:,2));
figure;
for i=1:5
    subplot(5,1,i);
    %fH=gcf;
    colormap('gray');
    x=RmonYr(:,1,i);
    y=RmonYr(:,2:3,i);
    b=bar(x,y);
    xlim([0 13]);
    ylim([0 400]);
    titleName=strcat('Year: ',YEARlist(i));
    title(titleName,'fontweight','bold','fontsize',11);
end
legend('MSD rain gauges','NOAA rain gauges');
xlabel('Months');

```

```

ylabel('Rainfall accumulation in mm');
%
figure;
fH=gcf;
colormap('gray');
x=Rmonth(:,1);
y=Rmonth(:,2:5);
b=bar(x,y);
xlim([0 13]);
ylim([200 1000]);
xlabel('Months');
ylabel('Rainfall accumulation in mm (2010-2014)');
legend('MSD:TR17-MtStFrancis', 'MSD:TR18- IVY tech',
'MSD: Averaged other Gauges', 'NOAA: Galena4-3ENE');

% daily comparison with 15 MSD and 12 gauges
openDaily=dlmread(NOAAgalena);
% check
for i=4:30
    sumG=sum(openDaily(:,i));
    disp(sumG);
end
% pair the closest rain gauges from MSD and NOAA
locInfo='Y:/projects/CSOsMSD/LouMetroGauge/locationGN
etwork4KyStatePlaneNorthZone_KYNZ83.txt';
openLoc=dlmread(locInfo);
lengthLoc=length(openLoc);
Rpaired=zeros(15,5);
for i=1:15
    xIni=openLoc(i,3);
    yIni=openLoc(i,2);
    RdistList=zeros(12,2);
    for j=16:27
        xNOAA=openLoc(j,3);
        yNOAA=openLoc(j,2);
        xDist=abs(xIni-xNOAA);
        yDist=abs(yIni-yNOAA);
        dist=sqrt(xDist*xDist+yDist*yDist)*1200/3937*0.001;
        RdistList((j-15),1)=openLoc(j,1);
    end
end

```

```

        RdistList((j-15),2)=dist;
    end
    RdistList(1:2,:)=[];
    RdistList(2:4,:)=[];
    RdistList(4,:)=[];
    RdistList(5:6,:)=[];
    RdistList(1,:)=[];
    minD=min(RdistList(:,2));
    rowMin=find(RdistList(:,2)==minD);
    Rpaired(i,1)=i;
    Rpaired(i,2)=RdistList(rowMin,1);
    Rpaired(i,3)=RdistList(rowMin,2);
    Rpaired(i,4)=i+3;
    Rpaired(i,5)=Rpaired(i,2)/100+18;
end
avgDist=mean(Rpaired(:,3));
maxDist=max(Rpaired(:,3));
minDist=min(Rpaired(:,3));

% daily data comparison
figure;
Rcc=zeros(15,2);
for i=1:15
    RpairDay=zeros(length(openDaily),2);
    RpairDay(:,1)=openDaily(:,(i+3));
    pairNOAA=Rpaired(i,5);
    RpairDay(:,2)=openDaily(:,pairNOAA);
    subplot(3,5,i);
    findMissing=find(RpairDay(:,2)<0);
    missing=isempty(findMissing);
    %if findMissing==1
    %    continue;
    %else
    %    RpairDay(findMissing,:)=[];
    %end
    findNoRain=find(RpairDay(:,1)==0 & RpairDay(:,1)==0);
    RpairDay(findNoRain,:)=[];
    lenPair=length(RpairDay);

    %disp(lenPair);
    % calc fault rain detection
    findFRD=find(RpairDay(:,1)>5 & RpairDay(:,2)==0);
    lenFRD=length(findFRD);
    %disp(lenFRD);
    %scatter(RpairDay(:,2),RpairDay(:,1),2.5,'o','filled','black')
    ;
    scatter(RpairDay(:,2),RpairDay(:,1),3,'filled','black','o');
    grid on;
    box on;
    set(gca,'gridlinestyle','--');
    set(gca,'xcolor',[0.3 0.3 0.3]);
    set(gca,'ycolor',[0.3 0.3 0.3]);
    set(gca,'xscale','log');
    set(gca,'yscale','log');
    xlim([0 150]);
    ylim([0 150]);
    hold on;
    plot(1:150,1:150,'black');
    hold off;
    gaugeName=strcat('TR',num2str(i),'
    ',num2str(Rpaired(i,3)),'km');
    t=title(gaugeName);
    set(t,'FontSize',14);
    set(t,'FontName','Times')
    if i==11
        xt=xlabel('Log-NOAA daily rainfall in mm');
        yt=ylabel('Log-MSD daily rainfall in mm');
        set(xt,'FontSize',12);
        set(yt,'FontSize',12);
        set(xt,'FontName','Times')
        set(yt,'FontName','Times')
    elseif i>=12
        xt=xlabel('Log-NOAA daily rainfall in mm');
        set(xt,'FontSize',12);
        set(xt,'FontName','Times')
    elseif i==1 || i==6 || i==11
        yt=ylabel('Log-MSD daily rainfall in mm');
        set(yt,'FontSize',12);

```

```

        set(yt,'FontName','Times')
    end

    rCorr=corrcoef(RpairDay(:,1),RpairDay(:,2));

    Rcc(i,2)=rCorr(1,2);

    cc=strcat('R: ',num2str(rCorr(1,2),2));

    tex=text(35,1.5,cc);

    set(tex, 'FontSize',14);

    set(tex,'FontName','Times')

    disp(cc);

end

Rcc(:,1)=Rpaired(:,3);

figure;

scatter(Rcc(:,1),Rcc(:,2));

b. Temporal matching MSD gauge network (15 gauges over
Louisville area)

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%% Matching closest
Gauges: MSD and NOAA

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%% 11/04/2014

% Developed by Jin-Young Hyun

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%clear all

% calculate distace between two different gauges and choose
the closest one

MSD='Y:/projects/CSOsMSD/LouMetroGauge/lists/MSDloc
ations.txt';

NOAA='Y:/projects/CSOsMSD/LouMetroGauge/lists/NOA
Alocations.txt';

openMSD=dlmread(MSD);

openNOAA=dlmread(NOAA);

lenMSD=length(openMSD);

lenNOAA=length(openNOAA);

RcloseDist=zeros(lenMSD,4);

for i=1:lenMSD

    Rdist=zeros(lenNOAA,2);

    for j=1:lenNOAA

        LonDist=abs(openMSD(i,1)-openNOAA(j,1));

        LatDist=abs(openMSD(i,2)-openNOAA(j,2));

        dist=sqrt(LonDist*LonDist+LatDist*LatDist);

```

```

        Rdist(j,1)=j+3; % data stored column PRCP

        Rdist(j,2)=dist;

    end

    RcloseDist(i,1)=i; % TR number

    minDist=min(Rdist(:,2));

    findMinDist=find(Rdist(:,2)==minDist);

    RcloseDist(i,2)=Rdist(findMinDist,1); % corresponding
NOAA

    RcloseDist(i,3)=minDist; % dist in deg

    RcloseDist(i,4)=minDist*111.18; % dist in km

end

dlmwrite('GaugeOutput/matchedGaugesMSD_NOAA.out',R
closeDist,'delimiter',';', 'precision','%6.2f');

% create the matching rainfall

TRmm='Z:/DATA/GroundBasedRainfall/MSDrainGauge/M
SDrainfall_inch_mm/RainGaugeMSDinMM.out';

NOAAmm='Z:/DATA/GroundBasedRainfall/MSDrainGauge
/MSDrainfall_inch_mm/RainGaugeMSDinInch.out';

openTRmm=dlmread(TRmm,');

openTRinch=dlmread(TRinch,');

numTR=length(openTRmm);

% calculation JulianDate

jdate=zeros(numTR,1);

for i=1:numTR

    YYYY=openTRinch(i,1);

    MO=openTRinch(i,2);

    DD=openTRinch(i,3);

    HH=openTRinch(i,4);

    MM=openTRinch(i,5);

    JD=datenum(YYYY,MO,DD,HH,MM,0);

    jdate(i,1)=JD;

end

% Daily and Hourly Accumulations

DaysIn5yr=365+365+366+365+273;

StartTime=datenum(2010,1,1,0,0,0);

EndTime=datenum(2014,10,1,0,0,0);

DayItv=datenum(2010,1,2,0,0,0)-StartTime;

StartDay=StartTime;

Rdays=zeros(DaysIn5yr,21);

```

```

for i=1:DaysIn5yr
    EndDay=StartDay+DayItv;
    dayList=find(jdate(:,1)>StartDay & jdate(:,1)<=EndDay);
    numDayList=length(dayList);

    % days in Gregorian Days
    GDday=strsplit(datestr(StartDay,2),'/');
    Rdays(i,1)=2000+str2num(char(GDday(3)));
    Rdays(i,2)=str2num(char(GDday(1)));
    Rdays(i,3)=str2num(char(GDday(2)));

    if numDayList==0
        AccumDay=0;
        Rdays(i,4)=0;
    else
        for j=6:23
            AccumDay=sum(openTRmm(dayList,j));
            Rdays(i,(j-2))=AccumDay;
            disp(GDday);
        end
    end

    StartDay=StartDay+DayItv;
end

dlmwrite('Z:/DATA/GroundBasedRainfall/MSDrainGauge/
MSDrainfall_inch_mm/AccumulationMSDgaugeDailyMM.out',Rdays,'delimiter',';', 'precision','%6.2f');

c. Accumulation of raw data into hourly
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%% 1 Hourly
Accumulation MSD rain gauges
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%% end hourly
accumulated values
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%% 01/14/2015

% Developed by Jin-Young Hyun

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%clear all;

MSD5min='Z:/DATA/GroundBasedRainfall/MSDrainGauge/
MSDrainfall_inch_mm/RainGaugeMSDinMM.out';

openMSD5min=dlmread(MSD5min,');
numMSD5min=length(openMSD5min);

% trimming the data to discard TR17 and TR18

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%openMSD5min(:,21:22)=[];

% calculation JulianDate
R_jd=zeros(numMSD5min,1);

for i=1:numMSD5min;
    YYYY=openMSD5min(i,1);
    MO=openMSD5min(i,2);
    DD=openMSD5min(i,3);
    HH=openMSD5min(i,4);
    MM=openMSD5min(i,5);
    JD=datenum(YYYY,MO,DD,HH,MM,0);
    R_jd(i,1)=JD;
end

% hourly and 3 hourly Accumulations
HoursIn5yr=(365+365+366+365+365)*24;
StartTime=datenum(2010,1,1,0,0,0);
EndTime=datenum(2015,1,1,0,0,0);
HourItv=datenum(2010,1,1,1,0,0)-StartTime;
StartHour=StartTime;
Rhours=zeros(HoursIn4yr,19);

for i=1:HoursIn4yr
    EndHour=StartHour+HourItv;
    hoursList=find(R_jd(:,1)>StartHour &
R_jd(:,1)<=EndHour);
    numHourList=length(hoursList);

    % days in Gregorian Days
    GDday=datestr(EndHour,30);
    Rhours(i,1)=str2num(GDday(1:4));
    Rhours(i,2)=str2num(GDday(5:6));
    Rhours(i,3)=str2num(GDday(7:8));
    Rhours(i,4)=str2num(GDday(10:11));

    for j=6:20
        if numHourList==0
            Rhours(i,j)=0;
        else
            AccumHour=sum(openMSD5min(hoursList,j));
            Rhours(i,(j-1))=AccumHour;
            disp(GDday);

```

```

end
end

StartHour=StartHour+HourItr;

end

dlmwrite('Y:/projects/CSOsMSD/LouMetroGauge/AccumulationMSDgauge1hourlyMM.out',Rhours,'delimiter',' ','precision', '%6.2f');

dlmwrite('Z:/DATA/GroundBasedRainfall/MSDrainGauge/MSDrainfall_inch_mm/AccumulationMSDgauge1hourlyMM.out',Rhours,'delimiter',' ','precision', '%6.2f');

% check

% bar(1:HoursIn5yr,Rhours(:,19));

d. General statistics and spatiotemporal correlation for MSD
rain gauge network, Louisville, KY

%%%%%%%%%%%%%%
%%%%%%%%%%%%%%

% General Statistics and Correllogram of Combined
Rain Gauge Network (MSD & NOAA) for Lou. Metro

%%%%%%%%%%%%%%
%%%%%%%%%%%%%%

% 01/12/2015

% Developed by Jin-Young Hyun

%%%%%%%%%%%%%%
%%%%%%%%%%%%%%

clear all;

% open daily rainfall from network

gaugeNet='Z:/DATA/GroundBasedRainfall/AccumulationMSDgaugeDaily_discardedTR17_18MM.out';

openGNet=dlmread(gaugeNet,');

lenGNet=length(openGNet);

% open location data

% distance between gauges (27 gauges)

%R_4KyStateNorthZoneNAD83=zeros(lenLocGNet,4);

%R_4KyStateNorthZoneNAD83(:,1)=openLocGNet(:,1);

%R_4KyStateNorthZoneNAD83(:,2)=openLocGNet(:,3);

%R_4KyStateNorthZoneNAD83(:,3)=openLocGNet(:,2);

%R_4KyStateNorthZoneNAD83(:,4)=1:27;

%dlmwrite('locationGNetwork4KyStatePlaneNorthZone.txt',
R_4KyStateNorthZoneNAD83,'delimiter',' ','precision', '%6.2f
');

locGNET='Y:/projects/CSOsMSD/LouMetroGauge/location
GNetwork.txt';

openLocGNet=dlmread(locGNET,');

```

```

locKyNorthZoneStatePlane='Y:/projects/CSOsMSD/LouMetroGauge/locationGNetwork4KyStatePlaneNorthZone_KYN
Z83.txt';

openLocGNetKY=dlmread(locKyNorthZoneStatePlane,');

lenLocGNet=length(openLocGNet);

distNum=0;

R_dist=zeros(lenLocGNet*lenLocGNet,3);

for i=1:27

    idBaseG=openLocGNetKY(i,1);

    verBaseG=openLocGNetKY(i,2);

    horBaseG=openLocGNetKY(i,3);

    for j=1:27

        idG=openLocGNetKY(j,1);

        verG=openLocGNetKY(j,2);

        horG=openLocGNetKY(j,3);

        % distance in km

        yDist=abs(verBaseG-verG);

        xDist=abs(horBaseG-horG);

        dist=sqrt(xDist*xDist+yDist*yDist)*1200/3937*0.001;

        distNum=distNum+1;

        R_dist(distNum,1)=idBaseG;

        R_dist(distNum,2)=idG;

        R_dist(distNum,3)=dist;

    X=strcat('Gauge_A_',num2str(idBaseG), '::Gauge_B_',num2s
tr(idG),'__distance: ',num2str(dist),'km');

        disp(X);

    end

end

% NOAAs missing data control

R_MissAnal=zeros(100,9);

findMissing=0;

countMiss=0;

for i=19:30

    idG=openLocGNet((i-3),1);

    searchG=find(openGNet(:,i)<0);

    checkG=isempty(searchG);

    if checkG ==1

        numMissG=0;

        disp(numMissG);

```

```

        continue;
    else
        numMissG=length(searchG);
        for j=1:numMissG
            R_list=zeros(lenLocGNet,2);
            countMiss=countMiss+1;
            R_MissAnal(countMiss,9)=countMiss;
            missingDay=searchG(j);
            % averaged rainfall for other gauges
            a=sort(openGNet(missingDay,4:30));
            averagedOG=mean(a(2:27));
            disp(averagedOG);
            R_MissAnal(countMiss,1)=idG;

R_MissAnal(countMiss,2)=openGNet(missingDay,1);

R_MissAnal(countMiss,3)=openGNet(missingDay,2);

R_MissAnal(countMiss,4)=openGNet(missingDay,3);

            % find closest gauge and its number
            closestGlist=find(R_dist(:,1)==idG);
            R_list=R_dist(closestGlist,2:3);
            minDist=10000;
            minG=1000;
            for k=1:lenLocGNet
                distG2G=R_list(k,2);
                if distG2G==0
                    continue;
                elseif distG2G>0
                    if distG2G<minDist
                        minDist=distG2G;
                        minG=R_list(k,1);
                    else
                        continue;
                    end
                end
            end
            R_MissAnal(countMiss,5)=minG;
            R_MissAnal(countMiss,6)=minDist;

            R_MissAnal(countMiss,7)=averagedOG;
            if minG>=100
                nearG=(minG/100)+18;
            else
                nearG=minG+3;
            end
            R_MissAnal(countMiss,8)=openGNet(missingDay,nearG);
        end
    end
end
% General Statistics
% monthly averaged rainfall amount of MSD and NOAA
% correlogram
R_corr=zeros(lenLocGNet*lenLocGNet,4);
R_corr(:,1:3)=R_dist;
numCorr=0;
for i=1:27
    for j=1:27
        pair=zeros(lenGNet,2);
        pair(:,1)=openGNet(:,i+3);
        pair(:,2)=openGNet(:,j+3);
        % find missing data
        missingsGbase=find(pair(:,1)<0);
        missingsGcomp=find(pair(:,2)<0);
        missOXGbase=isempty(missingsGbase);
        missOXGcomp=isempty(missingsGcomp);
        if missOXGbase==1 && missOXGcomp
            numCorr=numCorr+1;
            R=corrcoef(pair(:,1),pair(:,2));
            R_corr(numCorr,4)=R(1,2);
            disp(R_corr(numCorr,4));
        else
            numCorr=numCorr+1;
            if missOXGbase~=1 && missOXGcomp~=1
                pair(missOXGbase,:)=[];
                pair(missOXGcomp,:)=[];
            elseif missOXGbase~=1

```

```

        pair(missOXGbase,:)=[];
    elseif missOXGcomp~=1
        pair(missOXGcomp,:)=[];
    end

    R=corrcoef(pair(:,1),pair(:,2));
    R_corr(numCorr,4)=R(1,2);
    disp(R_corr(numCorr,4));
end
end

figure;
scatter(R_corr(:,3),R_corr(:,4),40,'filled','black');

% correlogram MSD only
distNum=0;
R_dist=zeros(15*15,3);
for i=1:15
    idBaseG=openLocGNetKY(i,1);
    verBaseG=openLocGNetKY(i,2);
    horBaseG=openLocGNetKY(i,3);
    for j=1:15
        idG=openLocGNetKY(j,1);
        verG=openLocGNetKY(j,2);
        horG=openLocGNetKY(j,3);
        % distance in km
        yDist=abs(verBaseG-verG);
        xDist=abs(horBaseG-horG);
        dist=sqrt(xDist*xDist+yDist*yDist)*1200/3937*0.001;
        distNum=distNum+1;
        R_dist(distNum,1)=idBaseG;
        R_dist(distNum,2)=idG;
        R_dist(distNum,3)=dist;

X=strcat('Gauge_A_',num2str(idBaseG), '::Gauge_B_',num2str(idG), '_distance: ',num2str(dist),'km');
        disp(X);
    end
end

% call 5minute rainfall data

MSD5min='Z:/DATA/GroundBasedRainfall/MSDrainGauge/MSDrainfall_inch_mm/RainGaugeMSDinMM.out';
open5min=dlmread(MSD5min);
len5min=length(open5min);
open5min(:,21:22)=[];
R_corr=zeros(15*15,4);
R_corr(:,1:3)=R_dist;
numCorr=0;
for i=1:15
    for j=1:15
        pair=zeros(len5min,2);
        pair(:,1)=open5min(:,i+5);
        pair(:,2)=open5min(:,j+5);
        numCorr=numCorr+1;
        R=corrcoef(pair(:,1),pair(:,2));
        R_corr(numCorr,4)=R(1,2);
        disp(R_corr(numCorr,4));

        % find missing data
        %missingsGbase=find(pair(:,1)<0);
        %missingsGcomp=find(pair(:,2)<0);
        %missOXGbase=isempty(missingsGbase);
        %missOXGcomp=isempty(missingsGcomp);
        %if missOXGbase==1 && missOXGcomp
            %numCorr=numCorr+1;
            %R=corrcoef(pair(:,1),pair(:,2));
            %R_corr(numCorr,4)=R(1,2);
            %disp(R_corr(numCorr,4));
        %else
            %numCorr=numCorr+1;
            %if missOXGbase~=1 && missOXGcomp~=1
                %pair(missOXGbase,:)=[];
                %pair(missOXGcomp,:)=[];
            %elseif missOXGbase~=1
                %pair(missOXGbase,:)=[];
            %elseif missOXGcomp~=1
                %pair(missOXGcomp,:)=[];
            %end
        end
    end
end

```



```

        %R=corrcoef(pair(:,1),pair(:,2));

        %R_corr(numCorr,4)=R(1,2);

        %disp(R_corr(numCorr,4));

    %end

end

figure;

scatter(R_corr(:,3),R_corr(:,4),40,'filled','black');

R_MissAnal((countMiss+1):100,:)=[];

days = 0:5:35;

conc = [515,420,370,250,135,120,60,20];

temp = [29,23,27,25,20,23,23,27];

[ax,hBar,hLine] = plotyy(days,temp,days,conc,'bar','plot');

[ax,hBar,hLine] =
plotyy(R_MissAnal(:,9),R_MissAnal(:,8),R_MissAnal(:,9),R
_MissAnal(:,6),'bar','plot');

%NOAAmm='Z:/DATA/GroundBasedRainfall/NWSrainGau
ge/gaugeNOAA_PRCP_inMM.out';

NOAAmm='Z:/DATA/GroundBasedRainfall/NWSrainGaug
e/gaugeNOAA_PRCP_inMM.out';

openNOAAmm=dlmread(NOAAmm,');

% dimensions

lenDays=length(openMSDmm);

numCols=3+17+13;

% combining MSD and NOAA daily rainfall accumulations

Rcombine=zeros(lenDays,numCols);

Rcombine(:,1:20)=openMSDmm(:,1:20);

Rcombine(:,21:numCols)=openNOAAmm(:,4:16);

% find missing of NOAA

rainyDays=0;

for i=1:lenDays

    dayData=Rcombine(i,4:33);

    rainObs=length(find(Rcombine(i,4:33)>0));

    if rainObs>0

        rainyDays=rainyDays+1;

    end

    missings=find(dayData<0);

    numMissing=length(missings);

    if numMissing == 0

        continue;

    else

        %disp(numMissing);

        otherGauges=find(dayData>=0);

        avgOG=mean(dayData(otherGauges));

        date=strcat(num2str(Rcombine(i,1)),',',num2str(Rcombine(i,2
)),',',num2str(Rcombine(i,3)));

        X=strcat('Missing on_',date,':
',num2str(numMissing),'with other gauges avg rainfall
of_',num2str(avgOG));

        disp(X);

    end

end

disp(rainyDays);

% bar plots for daily data from MSD and NOAA

R_figure=zeros(lenDays,33);

%R_missing=zeros(lenDays,33);

R_figure(:,1:3)=Rcombine(:,1:3);

%R_missing(:,1:3)=Rcombine(:,1:3);

for i=4:33

    findRain=find(Rcombine(:,i)>0);

    %findMiss=find(Rcombine(:,i)<0);

    R_figure(findRain,i)=Rcombine(findRain,i);

    %R_missing(findMiss,i)=0;

    disp(i);

end

TR_MSD={'01','02','03','04','05','06','07','08','09','10','11','12','
13','14','15','17','18',...

'Anchorage2_8NE','Charlestown2_6N','FoxChase1_4W','Gal
ena4_3ENE','Jeffersonville0_4SE','Jeffersonville0_8NW',...

'LouBowmanField','LouIntAirport','LouUpperGage','LouWF
Office',...

'OldBrownsboroP0_3SW','PRP0_6NNW','Shepherdsville5N
E'};

h=figure;

figures=0;

for i=4:33

    if i<21

        PREFIX='MSD Daily Rainfall Accumulation_TR';

    else

```

```

    PREFIX='NOAA Daily Rainfall Accumulation';

end

% total sum of the gauge for the study period

totalSum=sum(R_figure(:,i));

numMissing=length(find(R_figure(:,i)<0));

%X=['Total Accumulation in mm of ',TRs(i-
5),char(num2str(totalSum)), ' with
',char(num2str(numMissing)), ' missings'];

X=strcat('Total Accumulation in mm of ',TR_MSD(i-
3),': ',num2str(totalSum), ' with ',num2str(numMissing), '
missings');

disp(char(X))

subplot(3,1,numFig);

bar(1:lenDays,R_figure(:,i),'black');

%hold on;

%bar(1:lenDays,R_missing(:,i),'black');

%hold off;

titleName=char(strcat(PREFIX,TR_MSD(i-3)));

title(titleName);

ylim([0 200]);

if numFig<3

    if numFig==2

        ylabel('Daily Rainfall Accum. in mm');

    end

    numFig=numFig+1;

elseif numFig==3

    xlabel('Daily Interval');

    figures=figures+1;

saveName=char(strcat('Z:/DATA/GroundBasedRainfall/figur
es/MSDdaily',num2str(figures),'.jpg'));

saveas(h,saveName);

numFig=1;

end

end

% discard TR17 and TR18 due to the too low intensities...

Rcombine(:,19:20)=[];

dlmwrite('Z:/DATA/GroundBasedRainfall/AccumulationMS
DgaugeDaily_discardedTR17_18MM.out',Rcombine,'delimit
er',' ','precision','%6.2f');

```

Chapter 3

a. Clustering of rainfall events

```

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

```

```

%    Clustering of Rainfall Event of MSD rain gauges by
distance between them for Lou. Metro

```

```

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

```

```

% rainfall definition by EPA 0.1inch(0.254mm for 6hours)

```

```

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

```

```

% 04/28/2015

```

```

% Developed by Jin-Young Hyun

```

```

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

```

```

clear all;

```

```

% startTime, EndTime

```

```

openEvt=dlmread('Z:/DATA/GroundBasedRainfall/events/E
ventsMSDgaugeRainfall_mm.out','');

```

```

numEvt=max(openEvt(:,21)); % 558

```

```

rainfallEvt=zeros(5,16,numEvt);

```

```

Rcategory=zeros(numEvt,1);

```

```

RmaxAccum=zeros(numEvt,1);

```

```

for i=1:numEvt

```

```

    EvtList=find(openEvt(:,21)==i);

```

```

    STARTE=EvtList(1);

```

```

    month=openEvt(STARTE,2);

```

```

    % categories by seasons

```

```

    if month>2 && month<10

```

```

        category=1;

```

```

    elseif month==3 || month==10

```

```

        category=2;

```

```

    else

```

```

        category=3;

```

```

    end

```

```

    Rcategory(i,1)=category;

```

```

    ENDE=EvtList(length(EvtList));

```

```

    rainfallEvt(2,1,i)=(ENDE-STARTE+1)*5/60; % duration
of event in hour;

```

```

    rainSum=sum(openEvt(EvtList,6:20));

```

```

    RmaxAccum(i,1)=max(rainSum);

```

```

rainfallEvt(3,2:16,i)=rainSum; % rainfall depth for each
gauge during the event

for j=6:20
    ObsList=find(openEvt(EvtList,j)>0);
    lenObs=length(ObsList);
    if lenObs==0
        disp('no rain for this gauge. ');
    else
        rainfallEvt(1,(j-4),i)=1;
        STARTG=EvtList(ObsList(1));
        ENDG=EvtList(ObsList(lenObs));
        rainfallEvt(2,(j-4),i)=(ENDG-STARTG+1)*5/60; %
duration of gauge site rainfall event in hour;

        rainfallEvt(4,(j-
4),i)=var(openEvt(STARTG:ENDG,j)); % variance of gauge
rainfall;

        rainfallEvt(5,(j-
4),i)=max(openEvt(STARTG:ENDG,j)); % max 5min
rainfall of gauge rainfall;

    end
end

rainfallEvt(1,1,i)=sum(rainfallEvt(1,2:16,i)); % number of
gauge which observed rainfall

%rainfallEvt(3,1,i)=sum(rainSum)/rainfallEvt(1,1,i); %
avg depth in the event

rainfallEvt(3,1,i)=sum(rainSum)/15; % avg depth in the
event

rainfallEvt(4,1,i)=var(rainfallEvt(3,2:16,i)); % var of var
for the event;

rainfallEvt(5,1,i)=max(rainfallEvt(5,2:16,i)); % max of
5min for the event;

disp(i);
end

% check the collection
Rplots=zeros(558,7);
for i=1:558
    Rplots(i,1)=rainfallEvt(1,1,i);
    Rplots(i,2)=rainfallEvt(2,1,i);
    Rplots(i,3)=rainfallEvt(3,1,i);
    Rplots(i,4)=rainfallEvt(4,1,i);
    Rplots(i,5)=rainfallEvt(5,1,i);
    Rplots(i,6)=Rcategory(i,1);

    Rplots(i,7)=RmaxAccum(i,1);
end

% extreme case only for the rain depth and max rainfall
figure;
subplot(2,2,1);
scatter(Rplots(:,2),Rplots(:,3),5,'filled','black');
%scatter(RplotsFS(:,2),RplotsFS(:,3),5,'filled','black');
grid on;
box on;
xlabel('Duration in hour');
ylabel('Averaged rainfall depth in mm');
subplot(2,2,2);
scatter(Rplots(:,2),Rplots(:,5),5,'filled','black');
%scatter(RplotsFS(:,2),RplotsFS(:,5),5,'filled','black');
grid on;
box on;
xlabel('Duration in hour');
ylabel('Max rainfall intensity in mm/5min');
subplot(2,2,3);
scatter(Rplots(:,3),Rplots(:,5),5,'filled','black');
%scatter(RplotsFS(:,3),RplotsFS(:,5),5,'filled','black');
grid on;
box on;
xlabel('Averaged rainfall depth in mm');
ylabel('Max rainfall intensity in mm/5min');
subplot(2,2,4);
scatter3(Rplots(:,2),Rplots(:,3),Rplots(:,5),5,'filled','black');
%scatter3(RplotsFS(:,2),RplotsFS(:,3),RplotsFS(:,5),5,'filled',
'black');
grid on;
box on;
xlabel('Duration in hour');
ylabel('Averaged rainfall depth in mm');
zlabel('Max rainfall intensity in mm/5min');

% feature scaling
RplotsFS=zeros(length(Rplots),8);
RplotsFS(:,6)=1:558;
RplotsFS(:,7)=Rplots(:,6);

```

```

RplotsFS(:,8)=Rplots(:,7);
for i=2:5
    maxVal=max(Rplots(:,i));
    minVal=min(Rplots(:,i));
    for j=1:length(Rplots)
        val=Rplots(j,i);
        RplotsFS(j,i)=(val-minVal)/(maxVal-minVal);
        disp(val);
    end
end
maxVal=max(Rplots(:,7));
minVal=min(Rplots(:,7));
for j=1:length(Rplots)
    val=Rplots(j,7);
    RplotsFS(j,8)=(val-minVal)/(maxVal-minVal);
    disp(val);
end
% Kmeans clustering: 3 initial centers.
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% three initial mean points: at the corners
r1x=0; % high intensity/ low vol.
r1y=1; % high intensity/ low vol.
r2x=1; % low intensity/ high vol.
r2y=0; % low intensity/ low vol.
r3x=0; % low intensity/ low vol.
r3y=0; % low intensity/ low vol.
%r4x=1; % high intensity/ high vol.
%r4y=1; % high intensity/ high vol.
BreakPoint=0.00001;
numIteration=0;
distMatrix=zeros(1,3);
newThresh=0;
while 1
    %storage
    storage1=zeros(length(RplotsFS),5);
    storage2=zeros(length(RplotsFS),5);
    storage3=zeros(length(RplotsFS),5);

    %storage4=zeros(length(RplotsFS),4);
    moveCenter=zeros(1,3);
    num1=0;
    num2=0;
    num3=0;
    %num4=0;
    % calculate possessions
    for i=1:length(RplotsFS)
        xPt=RplotsFS(i,3);
        yPt=RplotsFS(i,5);
        evtNum=RplotsFS(i,6);
        evtCategory=RplotsFS(i,7);
        dist2mean1=sqrt((xPt-r1x)*(xPt-r1x)+(yPt-r1y)*(yPt-r1y)); % atv1
        distMatrix(1,1)=dist2mean1;
        dist2mean2=sqrt((xPt-r2x)*(xPt-r2x)+(yPt-r2y)*(yPt-r2y)); % atv2
        distMatrix(1,2)=dist2mean2;
        dist2mean3=sqrt((xPt-r3x)*(xPt-r3x)+(yPt-r3y)*(yPt-r3y)); % iatv1
        distMatrix(1,3)=dist2mean3;
        %dist2mean4=sqrt((xPt-r4x)*(xPt-r4x)+(yPt-r4y)*(yPt-r4y)); % iatv2
        %distMatrix(1,4)=dist2mean4;
        minDist=min(distMatrix);
        posession=find(distMatrix(1,:)==minDist);
        %store values to the closest sub-cluster
        if posession==1
            num1=num1+1;
            storage1(num1,1)=num1;
            storage1(num1,2)=xPt;
            storage1(num1,3)=yPt;
            storage1(num1,4)=evtNum;
            storage1(num1,5)=evtCategory;
        elseif posession==2
            num2=num2+1;
            storage2(num2,1)=num2;
            storage2(num2,2)=xPt;
            storage2(num2,3)=yPt;
            storage2(num2,4)=evtNum;

```

```

        storage2(num2,5)=evtCategory;
elseif posession==3
    num3=num3+1;
    storage3(num3,1)=num3;
    storage3(num3,2)=xPt;
    storage3(num3,3)=yPt;
    storage3(num3,4)=evtNum;
    storage3(num3,5)=evtCategory;
%else
%   num4=num4+1;
%   storage4(num4,1)=num4;
%   storage4(num4,2)=xPt;
%   storage4(num4,3)=yPt;
%   storage4(num4,4)=evtNum;
end
end
storage1((num1+1):length(RplotsFS),:)=[];
storage2((num2+1):length(RplotsFS),:)=[];
storage3((num3+1):length(RplotsFS),:)=[];
%storage4((num4+1):length(RplotsFS),:)=[];
x1new=mean(storage1(:,2));
y1new=mean(storage1(:,3));
x2new=mean(storage2(:,2));
y2new=mean(storage2(:,3));
x3new=mean(storage3(:,2));
y3new=mean(storage3(:,3));
%x4new=mean(storage4(:,2));
%y4new=mean(storage4(:,3));
dist1=sqrt((r1x-x1new)*(r1x-x1new)+(r1y-y1new)*(r1y-
y1new));
dist2=sqrt((r2x-x2new)*(r2x-x2new)+(r2y-y2new)*(r2y-
y2new));
dist3=sqrt((r3x-x3new)*(r3x-x3new)+(r3y-y3new)*(r3y-
y3new));
%dist4=sqrt((r4x-x4new)*(r4x-x4new)+(r4y-y4new)*(r4y-
y4new));
moveCenter(1,1)=dist1;
moveCenter(1,2)=dist2;
moveCenter(1,3)=dist3;
%moveCenter(1,4)=dist4;

newThresh=max(moveCenter);
numIteration=numIteration+1;
disp(numIteration);
%disp(newThresh);
if newThresh<=BreakPoint
    break;
else
    r1x=x1new;
    r1y=y1new;
    r2x=x2new;
    r2y=y2new;
    r3x=x3new;
    r3y=y3new;
    %r4x=x4new;
    %r4y=y4new;
end
end
% export clustered rainfall event
% labeling: 1-high intensity, 2-high depth, 3-light
Rcluster=zeros(numEvt,2);
for i=1:numEvt
    Rcluster(i,1)=i;
    highInt=length(find(storage1(:,4)==i));
    highDepth=length(find(storage2(:,4)==i));
    light=length(find(storage3(:,4)==i));
    if highInt==1
        Rcluster(i,2)=1;
        disp('High intensity of rainfall event');
    elseif highDepth==1
        Rcluster(i,2)=2;
        disp('High depth of rainfall event');
    elseif light==1
        Rcluster(i,2)=3;
        disp('Light event');
    end
end
end
dlmwrite('Z:/DATA/GroundBasedRainfall/events/eventClus-
teredRainfall.out',Rcluster,'delimiter',' ','precision','%6.2f');

```

```

% num event in a cluster
LE=length(find(Rcluster(:,2)==3));
Hlevts=length(find(Rcluster(:,2)==1));
HDevts=length(find(Rcluster(:,2)==2));

figure;
subplot(2,2,1);
scatter(Rplots(:,3),Rplots(:,5),5,'filled','black');
%scatter(RplotsFS(:,3),RplotsFS(:,5),5,'filled','black');
grid on;
box on;

t=title('Averaged rain depth vs. intensity');
set(t,'FontSize',12);
set(t,'FontName','Times')
xt=xlabel('Averaged rainfall depth in mm');
set(xt,'FontSize',11);
set(xt,'FontName','Times')
yt=ylabel('Max rainfall in mm/5min');
set(yt,'FontSize',11);
set(yt,'FontName','Times')
lt=legend('Event');
set(lt,'FontSize',11);
set(lt,'FontName','Times')

% initial centroid
iniCnt=zeros(4,2);
iniCnt(1,1)=0;
iniCnt(1,2)=1;
iniCnt(2,1)=1;
iniCnt(2,2)=0;
iniCnt(3,1)=0;
iniCnt(3,2)=0;
%iniCnt(4,1)=1;
%iniCnt(4,2)=1;
subplot(2,2,2);
scatter(RplotsFS(:,3),RplotsFS(:,5),5,'filled','black');
%scatter(RplotsFS(:,3),RplotsFS(:,5),5,'filled','black');
grid on;
box on;

hold on;
scatter(iniCnt(:,1),iniCnt(:,2),30,'black','s');
hold off;

t=title('K-means Cluster: initial mean points');
set(t,'FontSize',12);
set(t,'FontName','Times');
xt=xlabel('Normalized averaged rainfall depth in mm');
set(xt,'FontSize',11);
set(xt,'FontName','Times')
yt=ylabel('Normalized max rainfall in mm/5min');
set(yt,'FontSize',11);
set(yt,'FontName','Times');
lt=legend('events','initial centers');
set(lt,'FontSize',11);
set(lt,'FontName','Times');

% final centroid and groups
fnCnt=zeros(3,2);
fnCnt(1,1)=r1x;
fnCnt(1,2)=r1y;
fnCnt(2,1)=r2x;
fnCnt(2,2)=r2y;
fnCnt(3,1)=r3x;
fnCnt(3,2)=r3y;
%fnCnt(4,1)=r4x;
%fnCnt(4,2)=r4y;
subplot(2,2,3);
scatter(storage1(:,2),storage1(:,3),25,'black','s');
grid on;
box on;
hold on;
scatter(storage2(:,2),storage2(:,3),35,'black','X');
hold on;
scatter(storage3(:,2),storage3(:,3),3,'black','filled','o');
%hold on;
%scatter(storage4(:,2),storage4(:,3),20,'black','d');
hold off;

t=title('K-means Cluster: final clusters');
set(t,'FontSize',12);

```

```

set(t,'FontName','Times');
xt=xlabel('Normalized averaged rainfall depth in mm');
set(xt,'FontSize',11);
set(xt,'FontName','Times')
yt=ylabel('Normalized max rainfall in mm/5min');
set(yt,'FontSize',11);
set(yt,'FontName','Times');
lt=legend('high intensity events','high depth events','light
events');
set(lt,'FontSize',11);
set(lt,'FontName','Times');
% eventClustering;
subplot(2,2,4);
RevtStat=zeros(2,3);
% mean duration over the area
RevtStat(1,1)=mean(Rplots(highR,2));
RevtStat(1,2)=mean(Rplots(storage1(:,4),2));
RevtStat(1,3)=mean(Rplots(storage2(:,4),2));
% std duration over the area
RevtStat(2,1)=std(Rplots(highR,2));
RevtStat(2,2)=std(Rplots(storage1(:,4),2));
RevtStat(2,3)=std(Rplots(storage2(:,4),2));
% mean max intensity
%RevtStat(3,1)=mean(Rplots(highR,5));
%RevtStat(3,2)=mean(Rplots(storage1(:,4),5));
%RevtStat(3,3)=mean(Rplots(storage2(:,4),5));
% std max intensity
%RevtStat(4,1)=std(Rplots(highR,5));
%RevtStat(4,2)=std(Rplots(storage1(:,4),5));
%RevtStat(4,3)=std(Rplots(storage2(:,4),5));
bar(RevtStat);
grid on;
box on;
xlim([0 3]);
set(gca,'XTickLabel',{'mean duration','std duration','mean
max intensity','std max intensity'})
t=title('General statistics of Cluster events');
set(t,'FontSize',12);
set(t,'FontName','Times');

```

```

%xt=xlabel('statistics');
%set(xt,'FontSize',11);
%set(xt,'FontName','Times')
yt=ylabel('duraion in hour');
set(yt,'FontSize',11);
set(yt,'FontName','Times');
lt=legend('Total event','High intensity events','High depth
events');
set(lt,'FontSize',11);
set(lt,'FontName','Times');
% additional
RANGEname={'0-5km','5-10km','10-15km','15-20km','20-
25km','25-30km','30-35km','35-40km'};
RANGEmp={'15minute','30minute','1hour','3hour'};
QUANTILES=[0.00000001,0.25,0.5,0.75,0.9,0.95];
xlabel('Normalized Averaged rainfall depth in mm');
ylabel('Normalized Max rainfall intensity in mm/5min');
legend();
% categorical by season
% hot season: 4-9
% transitional: 3,10
% cold season: 11-2
highIhot=find(storage1(:,5)==1);
highItran=find(storage1(:,5)==2);
highIcold=find(storage1(:,5)==3);
highDhot=find(storage2(:,5)==1);
highDtran=find(storage2(:,5)==2);
highDcold=find(storage2(:,5)==3);
%highLhot=find(storage3(:,5)==1);
%highLtran=find(storage3(:,5)==2);
%highLcold=find(storage3(:,5)==3);
% event statistics
% highI
HIhot=storage1(highIhot,4);
highIhot=Rplots(HIhot,:);
HIhotDuration=mean(highIhot(:,2));
HIhotDepth=mean(highIhot(:,3));
HIhotVar=mean(highIhot(:,4));
HIhotIntensity=mean(highIhot(:,5));

```

```

Hltran=storage1(highltran,4);
highltran=Rplots(Hltran,:);
HltranDuration=mean(highltran(:,2));
HltranDepth=mean(highltran(:,3));
HltranVar=mean(highltran(:,4));
HltranIntensity=mean(highltran(:,5));
Hlcold=storage1(highlcold,4);
highlcold=Rplots(Hlcold,:);
HlcoldDuration=mean(highlcold(:,2));
HlcoldDepth=mean(highlcold(:,3));
HlcoldVar=mean(highlcold(:,4));
HlcoldIntensity=mean(highlcold(:,5));
% highD
HDhot=storage2(highdhot,4);
highdhot=Rplots(HDhot,:);
HDhotDuration=mean(highdhot(:,2));
HDhotDepth=mean(highdhot(:,3));
HDhotVar=mean(highdhot(:,4));
HDhotIntensity=mean(highdhot(:,5));
HDtran=storage2(highdtran,4);
highdtran=Rplots(HDtran,:);
HDtranDuration=mean(highdtran(:,2));
HDtranDepth=mean(highdtran(:,3));
HDtranVar=mean(highdtran(:,4));
HDtranIntensity=mean(highdtran(:,5));
HDcold=storage3(highdcold,4);
highdcold=Rplots(HDcold,:);
HDcoldDuration=mean(highdcold(:,2));
HDcoldDepth=mean(highdcold(:,3));
HDcoldVar=mean(highdcold(:,4));
HDcoldIntensity=mean(highdcold(:,5));
% low
low=storage3(:,4);
lowEvt=Rplots(low,:);
LEduration=mean(lowEvt(:,2));
LEdepth=mean(lowEvt(:,3));
LEvar=mean(lowEvt(:,4));
LEintensity=mean(lowEvt(:,5));

```

```

% significant rainfall only
highR=zeros((length(storage1)+length(storage2)),1);
highR(1:length(storage1))=storage1(:,4);
highR((length(storage1)+1):length(highR))=storage2(:,4);

b. Clustered variogram
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%% Clustering of
Rainfall Event of MSD rain gauges by distance between them
for Lou. Metro
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%% rainfall definition by
EPA 0.1inch(0.254mm for 6hours)
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%% 04/28/2015
% Developed by Jin-Young Hyun
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

lenObs=length(ObsList);
if lenObs==0
    disp('no rain for this gauge. ');
else
    rainfallEvt(1,(j-4),i)=1;
    STARTG=EvtList(ObsList(1));
    ENDG=EvtList(ObsList(lenObs));
    rainfallEvt(2,(j-4),i)=(ENDG-STARTG+1)*5/60; %
duration of gauge site rainfall event in hour;
    rainfallEvt(4,(j-
4),i)=var(openEvt(STARTG:ENDG,i)); % variance of gauge
rainfall;
    rainfallEvt(5,(j-
4),i)=max(openEvt(STARTG:ENDG,i)); % max 5min
rainfall of gauge rainfall;
    end
end
    rainfallEvt(1,1,i)=sum(rainfallEvt(1,2:16,i)); % number of
gauge which observed rainfall
    rainfallEvt(3,1,i)=sum(rainSum)/rainfallEvt(1,1,i); % avg
depth in the event
    rainfallEvt(4,1,i)=var(rainfallEvt(3,2:16,i)); % var of var
for the event;
    rainfallEvt(5,1,i)=max(rainfallEvt(5,2:16,i)); % max of
5min for the event;
    disp(i);
end
% check the collection

```



```

Rplots=zeros(558,7);
for i=1:558
    Rplots(i,1)=rainfallEvt(1,1,i);
    Rplots(i,2)=rainfallEvt(2,1,i);
    Rplots(i,3)=rainfallEvt(3,1,i);
    Rplots(i,4)=rainfallEvt(4,1,i);
    Rplots(i,5)=rainfallEvt(5,1,i);
    Rplots(i,6)=Rcategory(i,1);
    Rplots(i,7)=RmaxAccum(i,1);
end
% check stability
figure;
subplot(3,1,1);
plot(1:numEvt,Rplots(:,3),'black');
subplot(3,1,2);
plot(1:numEvt,Rplots(:,5),'blue');
subplot(3,1,3);
plot(1:numEvt,Rplots(:,7),'red');
% moving average of the rainfall event
for i=1:(numEvt-4)
    windowList=i:(i+4)

% extreme case only for the rain depth and max rainfall
figure;
subplot(2,2,1);
scatter(Rplots(:,2),Rplots(:,3),5,'filled','black');
%scatter(RplotsFS(:,2),RplotsFS(:,3),5,'filled','black');
grid on;
box on;
xlabel('Duration in hour');
ylabel('Averaged rainfall depth in mm');
subplot(2,2,2);
scatter(Rplots(:,2),Rplots(:,5),5,'filled','black');
%scatter(RplotsFS(:,2),RplotsFS(:,5),5,'filled','black');
grid on;
box on;
xlabel('Duration in hour');
ylabel('Max rainfall intensity in mm/5min');
subplot(2,2,3);
scatter(Rplots(:,3),Rplots(:,5),5,'filled','black');
%scatter(RplotsFS(:,3),RplotsFS(:,5),5,'filled','black');
grid on;
box on;
xlabel('Averaged rainfall depth in mm');
ylabel('Max rainfall intensity in mm/5min');
subplot(2,2,4);
scatter3(Rplots(:,2),Rplots(:,3),Rplots(:,5),5,'filled','black');
%scatter3(RplotsFS(:,2),RplotsFS(:,3),RplotsFS(:,5),5,'filled','black');
grid on;
box on;
xlabel('Duration in hour');
ylabel('Averaged rainfall depth in mm');
zlabel('Max rainfall intensity in mm/5min');
% feature scaling
RplotsFS=zeros(length(Rplots),8);
RplotsFS(:,6)=1:558;
RplotsFS(:,7)=Rplots(:,6);
RplotsFS(:,8)=Rplots(:,7);
for i=2:5
    maxVal=max(Rplots(:,i));
    minVal=min(Rplots(:,i));
    for j=1:length(Rplots)
        val=Rplots(j,i);
        RplotsFS(j,i)=((val-minVal)/(maxVal-minVal));
        disp(val);
    end
end
% Kmeans clustering: 3 initial centers.
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% three initial mean points: at the corners
r1x=0; % high intensity/ low vol.
r1y=1; % high intensity/ low vol.

```

```

r2x=1; % low intensity/ high vol.
r2y=0; % low intensity/ low vol.
r3x=0; % low intensity/ low vol.
r3y=0; % low intensity/ low vol.
%r4x=1; % high intensity/ high vol.
%r4y=1; % high intensity/ high vol.
BreakPoint=0.00001;
numIteration=0;
distMatrix=zeros(1,3);
newThresh=0;
while 1
    %storage
    storage1=zeros(length(RplotsFS),5);
    storage2=zeros(length(RplotsFS),5);
    storage3=zeros(length(RplotsFS),5);
    %storage4=zeros(length(RplotsFS),4);
    moveCenter=zeros(1,3);
    num1=0;
    num2=0;
    num3=0;
    %num4=0;
    % calculate possessions
    for i=1:length(RplotsFS)
        xPt=RplotsFS(i,3);
        yPt=RplotsFS(i,5);
        evtNum=RplotsFS(i,6);
        evtCategory=RplotsFS(i,7);
        dist2mean1=sqrt((xPt-r1x)*(xPt-r1x)+(yPt-r1y)*(yPt-r1y)); % atv1
        distMatrix(1,1)=dist2mean1;
        dist2mean2=sqrt((xPt-r2x)*(xPt-r2x)+(yPt-r2y)*(yPt-r2y)); % atv2
        distMatrix(1,2)=dist2mean2;
        dist2mean3=sqrt((xPt-r3x)*(xPt-r3x)+(yPt-r3y)*(yPt-r3y)); % iatv1
        distMatrix(1,3)=dist2mean3;
        %dist2mean4=sqrt((xPt-r4x)*(xPt-r4x)+(yPt-r4y)*(yPt-r4y)); % iatv2
        %distMatrix(1,4)=dist2mean4;
        minDist=min(distMatrix);
        posession=find(distMatrix(1,:)==minDist);
        %store values to the closest sub-cluster
        if posession==1
            num1=num1+1;
            storage1(num1,1)=num1;
            storage1(num1,2)=xPt;
            storage1(num1,3)=yPt;
            storage1(num1,4)=evtNum;
            storage1(num1,5)=evtCategory;
        elseif posession==2
            num2=num2+1;
            storage2(num2,1)=num2;
            storage2(num2,2)=xPt;
            storage2(num2,3)=yPt;
            storage2(num2,4)=evtNum;
            storage2(num2,5)=evtCategory;
        elseif posession==3
            num3=num3+1;
            storage3(num3,1)=num3;
            storage3(num3,2)=xPt;
            storage3(num3,3)=yPt;
            storage3(num3,4)=evtNum;
            storage3(num3,5)=evtCategory;
        %else
        %   num4=num4+1;
        %   storage4(num4,1)=num4;
        %   storage4(num4,2)=xPt;
        %   storage4(num4,3)=yPt;
        %   storage4(num4,4)=evtNum;
        end
    end
    storage1((num1+1):length(RplotsFS),:)=[];
    storage2((num2+1):length(RplotsFS),:)=[];
    storage3((num3+1):length(RplotsFS),:)=[];
    %storage4((num4+1):length(RplotsFS),:)=[];
    x1new=mean(storage1(:,2));
    y1new=mean(storage1(:,3));
    x2new=mean(storage2(:,2));

```

```

y2new=mean(storage2(:,3));
x3new=mean(storage3(:,2));
y3new=mean(storage3(:,3));
%x4new=mean(storage4(:,3));
%y4new=mean(storage4(:,3));

dist1=sqrt((r1x-x1new)*(r1x-x1new)+(r1y-y1new)*(r1y-
y1new));
dist2=sqrt((r2x-x2new)*(r2x-x2new)+(r2y-y2new)*(r2y-
y2new));
dist3=sqrt((r3x-x3new)*(r3x-x3new)+(r3y-y3new)*(r3y-
y3new));
%dist4=sqrt((r4x-x4new)*(r4x-x4new)+(r4y-y4new)*(r4y-
y4new));

moveCenter(1,1)=dist1;
moveCenter(1,2)=dist2;
moveCenter(1,3)=dist3;
%moveCenter(1,4)=dist4;

newThresh=max(moveCenter);
numIteration=numIteration+1;
disp(numIteration);
%disp(newThresh);

if newThresh<=BreakPoint
    break;
else
    r1x=x1new;
    r1y=y1new;
    r2x=x2new;
    r2y=y2new;
    r3x=x3new;
    r3y=y3new;
    %r4x=x4new;
    %r4y=y4new;
end
end
% group analysis
figure;
subplot(2,2,1);
scatter(Rplots(:,3),Rplots(:,5),5,'filled','black');
%scatter(RplotsFS(:,3),RplotsFS(:,5),5,'filled','black');
grid on;

box on;
title('Scattering of rain depth and intensity');
xlabel('Averaged rainfall depth in mm');
ylabel('Max rainfall intensity in mm/5min');

% initial centroid
iniCnt=zeros(4,2);
iniCnt(1,1)=0;
iniCnt(1,2)=1;
iniCnt(2,1)=1;
iniCnt(2,2)=0;
iniCnt(3,1)=0;
iniCnt(3,2)=0;
%iniCnt(4,1)=1;
%iniCnt(4,2)=1;
subplot(2,2,2);
scatter(RplotsFS(:,3),RplotsFS(:,5),5,'filled','black');
%scatter(RplotsFS(:,3),RplotsFS(:,5),5,'filled','black');
grid on;
box on;
hold on;
scatter(iniCnt(:,1),iniCnt(:,2),30,'black','s');
hold off;
title('Kmeans: initial mean points');
xlabel('Normalized Averaged rainfall depth in mm');
ylabel('Normalized Max rainfall intensity in mm/5min');
legend('events','initial centers');

% final centroid and groups
fnCnt=zeros(3,2);
fnCnt(1,1)=r1x;
fnCnt(1,2)=r1y;
fnCnt(2,1)=r2x;
fnCnt(2,2)=r2y;
fnCnt(3,1)=r3x;
fnCnt(3,2)=r3y;
%fnCnt(4,1)=r4x;
%fnCnt(4,2)=r4y;
subplot(2,2,3);
scatter(storage1(:,2),storage1(:,3),25,'black','s');

```

```

grid on;
box on;
hold on;
scatter(storage2(:,2),storage2(:,3),35,'black','X');
hold on;
scatter(storage3(:,2),storage3(:,3),3,'black','filled','o');
%hold on;
%scatter(storage4(:,2),storage4(:,3),20,'black','d');
hold off;
title('Kmeans: final groups');
xlabel('Normalized Averaged rainfall depth in mm');
ylabel('Normalized Max rainfall intensity in mm/5min');
legend('high intensity events','high depth events','light
events');
% categorical by season
% hot season: 4-9
% transitional: 3,10
% cold season: 11-2
highIhot=find(storage1(:,5)==1);
highItran=find(storage1(:,5)==2);
highIcold=find(storage1(:,5)==3);
highDhot=find(storage2(:,5)==1);
highDtran=find(storage2(:,5)==2);
highDcold=find(storage2(:,5)==3);
%highLhot=find(storage3(:,5)==1);
%highLtran=find(storage3(:,5)==2);
%highLcold=find(storage3(:,5)==3);
% event statistics
% highI
HIhot=storage1(highIhot,4);
highIhot=Rplots(HIhot,:);
HIhotDuration=mean(highIhot(:,2));
HIhotDepth=mean(highIhot(:,3));
HIhotVar=mean(highIhot(:,4));
HIhotIntensity=mean(highIhot(:,5));
HItran=storage1(highItran,4);
highItran=Rplots(HItran,:);
HItranDuration=mean(highItran(:,2));

```

```

HItranDepth=mean(highItran(:,3));
HItranVar=mean(highItran(:,4));
HItranIntensity=mean(highItran(:,5));
HIcold=storage1(highIcold,4);
highIcold=Rplots(HIcold,:);
HIcoldDuration=mean(highIcold(:,2));
HIcoldDepth=mean(highIcold(:,3));
HIcoldVar=mean(highIcold(:,4));
HIcoldIntensity=mean(highIcold(:,5));
% highD
HDhot=storage2(highDhot,4);
highDhot=Rplots(HDhot,:);
HDhotDuration=mean(highDhot(:,2));
HDhotDepth=mean(highDhot(:,3));
HDhotVar=mean(highDhot(:,4));
HDhotIntensity=mean(highDhot(:,5));
HDtran=storage2(highDtran,4);
highDtran=Rplots(HDtran,:);
HDtranDuration=mean(highDtran(:,2));
HDtranDepth=mean(highDtran(:,3));
HDtranVar=mean(highDtran(:,4));
HDtranIntensity=mean(highDtran(:,5));
HDcold=storage3(highDcold,4);
highDcold=Rplots(HDcold,:);
HDcoldDuration=mean(highDcold(:,2));
HDcoldDepth=mean(highDcold(:,3));
HDcoldVar=mean(highDcold(:,4));
HDcoldIntensity=mean(highDcold(:,5));
% low
low=storage3(:,4);
lowEvt=Rplots(low,:);
LEduration=mean(lowEvt(:,2));
LEdepth=mean(lowEvt(:,3));
LEvar=mean(lowEvt(:,4));
LEintensity=mean(lowEvt(:,5));
% significant rainfall only
highR=zeros((length(storage1)+length(storage2)),1);
highR(1:length(storage1))=storage1(:,4);

```

```

highR((length(storage1)+1):length(highR))=storage2(:,4);
% eventClustering;
subplot(2,2,4);
%figure;
RevtStat=zeros(7,3);
RevtStat(1,1)=mean(Rplots(highR,3));
RevtStat(1,2)=mean(Rplots(HIhot,3));
RevtStat(1,3)=mean(Rplots(HItran,3));
RevtStat(1,4)=mean(Rplots(HIcold,3));
RevtStat(1,5)=mean(Rplots(HDhot,3));
RevtStat(1,6)=mean(Rplots(HDtran,3));
RevtStat(1,7)=mean(Rplots(HDcold,3));
RevtStat(2,1)=std(Rplots(highR,3));
RevtStat(2,2)=std(Rplots(HIhot,3));
RevtStat(2,3)=std(Rplots(HItran,3));
RevtStat(2,4)=std(Rplots(HIcold,3));
RevtStat(2,5)=std(Rplots(HDhot,3));
RevtStat(2,6)=std(Rplots(HDtran,3));
RevtStat(2,7)=std(Rplots(HDcold,3));
RevtStat(3,1)=std(Rplots(highR,7));
RevtStat(3,2)=std(Rplots(HIhot,7));
RevtStat(3,3)=std(Rplots(HItran,7));
RevtStat(3,4)=std(Rplots(HIcold,7));
RevtStat(3,5)=std(Rplots(HDhot,7));
RevtStat(3,6)=std(Rplots(HDtran,7));
RevtStat(3,7)=std(Rplots(HDcold,7));
bar(RevtStat);
legend('total event','high intensity events-hot','high intensity
events-tran','high intensity events-cold',...
'high depth events-hot','high depth events-tran','high depth
events-cold');
% additional
subplot(2,2,4);
figure;
scatter(storage1(highIhot,2),storage1(highIhot,3),25,'red','s');
grid on;
box on;
hold on;

```

```

scatter(storage1(highItran,2),storage1(highItran,3),25,'filled','
black','s');
hold on;
scatter(storage1(highIcold,2),storage1(highIcold,3),25,'filled',
'blue','s');
hold on;
scatter(storage2(highDhot,2),storage2(highDhot,3),25,'red','X
');
hold on;
scatter(storage2(highDtran,2),storage2(highDtran,3),25,'black',
'X');
hold on;
scatter(storage2(highDcold,2),storage2(highDcold,3),25,'blue',
'X');
hold on;
%scatter(storage3(highLhot,2),storage3(highLhot,3),25,'red','
o');
%hold on;
%scatter(storage3(highLtran,2),storage3(highLtran,3),25,'black',
'o');
%hold on;
%scatter(storage3(highLtran,2),storage3(highLtran,3),25,'o');
scatter(storage3(:,2),storage3(:,3),3,'black','filled','o');
scatter(storage2(:,2),storage2(:,3),35,'black','X');
hold on;
scatter(storage3(:,2),storage3(:,3),3,'black','filled','o');

```

Chapter 4

a. Support Vector Classification of tropical type rainfall values

%% Rainfall >=35dBZ onlyh

%% outputs

% 1-5: time information - year,month,day,hour,minute

% 6: TR ID #

% 7: total available inner scans of KLVX

% 8: # of NA (-999) in 15minute Scan

% 9: max interval within 15min capsule in min

% 10: type of rainfall - 1(convective), 2(tropical), 3(ECS), 4(stratiform)

% 11: gauge accumulation mm/15min (quarter hour)

% 12: KLVX radar accumulation mm/15min (quarter hour)

```

% 13: KLVX reflectivity in dBZ

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%% 10/03/2015

% Developed by Jin-Young Hyun

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%clear all;

openData=dlmread('Z:/DATA/KLVXoutput/tropicalHighRef
Only.out','');

lenData=length(openData);

winterT2=find(openData(:,2)>9 | openData(:,2)<4);

lenT2winter=length(winterT2);

summerT2=find(openData(:,2)>3 & openData(:,2)<10);

lenT2summer=length(summerT2);

% check high REFLECTIVITY in the first half water year

winterOnly=openData(winterT2,:);

winterKernel=zeros(length(winterOnly),1);

DIVw=32;

lowREFw=find(winterOnly(:,13)<DIVw);

highREFw=find(winterOnly(:,13)>=DIVw);

winterKernel(highREFw,1)=1;

winterKernel(lowREFw,1)=-1;

svmStruct =
svmtrain(winterOnly(:,11:12),winterKernel(:,1));

figure;

svmStructV =
svmtrain(winterOnly(:,11:12),winterKernel(:,1),'ShowPlot',tr
ue);

group=svmclassify(svmStruct,winterOnly(:,11:12));

G1=find(group(:,1)==1);

G2=find(group(:,1)==-1);

SVMhighREFw=winterOnly(G1,:);

ErrorW=zeros(length(SVMhighREFw),3);

numIteration=1;

for i=32.1:0.1:34.4

    DIVw=i;

    lowREFwH=find(SVMhighREFw(:,13)<DIVw);

    highREFwH=find(SVMhighREFw(:,13)>=DIVw);

    winterKernelH=zeros(length(SVMhighREFw),1);

    winterKernelH(highREFwH,1)=1;

    winterKernelH(highREFwH,1)=-1;

```

```

svmStructH=svmtrain(SVMhighREFw(:,11:12),winterKernel
H);

groupH=svmclassify(svmStructH,SVMhighREFw(:,11:12));

G1H=find(groupH(:,1)==-1);

G2H=find(groupH(:,1)==0);

prevE1=1000000000000;

for j=0.1:0.1:20

    for k=10:10:1000

rainRate=((((10.^(SVMhighREFw(G1H,13)./10))/k).^(1/j))./4;

error=sum(abs(rainRate-SVMhighREFw(G1H,11)));

if error<prevE1

    a1=k;

    b1=j;

    %XX=[a1,b1];

    %disp(XX);

    prevE1=error;

end

end

end

% find the best G2 for each group

prevE2=1000000000000;

for j=0.1:0.1:20

    for k=10:10:1000

rainRate=((((10.^(SVMhighREFw(G2H,13)./10))/k).^(1/j))./4;

error=sum(abs(rainRate-SVMhighREFw(G2H,11)));

if error<prevE2

    a2=k;

    b2=j;

    %XX=[a2,b2];

    %disp(XX);

    prevE2=error;

end

end

end

totalE=prevE1+prevE2;

ErrorW(numIteration,1)=i;

```

```

        RerrorW(numIteration,2)=totalE/length(SVMhighREFw);

RerrorW(numIteration,3)=length(G1H)/(length(G1H)+length
(G2H));

    disp(numIteration);

    numIteration=numIteration+1;
end
figure;
subplot(2,2,1)

svmStruct =
svmtrain(winterOnly(:,11:12),winterKernel(:,1),'ShowPlot',true);

xlim([0 30]);
ylim([0 30]);
hold on;
plot(0:30,0:30,'black');

box on;

grid on;

hold on;

subplot(2,2,2)

svmStructH=svmtrain(SVMhighREFw(:,11:12),winterKernel
H,'ShowPlot',true);

xlim([0 30]);
ylim([0 30]);
hold on;
plot(0:30,0:30,'black');

box on;

grid on;

hold on;

T2=scatter(winterOnly(:,11),winterOnly(:,12),25,'d','filled','black');

figure;

hist(SVMhighREFw(:,9),20);

figure;

DIV=33.5;

subplot(1,3,1);

T2=scatter(SVMhighREFw(:,11),SVMhighREFw(:,12),25,'d',
'filled','black');

xlim([0 30]);
ylim([0 30]);
hold on;

plot(0:30,0:30,'black');

box on;

grid on;

subplot(1,3,2);

extremeInt=find(SVMhighREFw(:,13)>=DIV);

T2extreme=scatter(SVMhighREFw(extremeInt,11),SVMhighREFw(extremeInt,12),25,'d','filled','red');

xlim([0 30]);
ylim([0 30]);
hold on;

plot(0:30,0:30,'black');

box on;

grid on;

subplot(1,3,3);

normInt=find(SVMhighREFw(:,13)<DIV);

T2extreme=scatter(SVMhighREFw(normInt,11),SVMhighREFw(normInt,12),25,'d','filled','blue');

xlim([0 30]);
ylim([0 30]);
hold on;

plot(0:30,0:30,'black');

box on;

grid on;

RcaliW(:,3)=(((10.^(openData(winterT2,13)./10))/a).^(1/b))./4;

svmStruct = svmtrain(dataW,RcaliW(:,5),'ShowPlot',true);

highREFw=find(RcaliW(:,6)>=DIVw);

lowREFw=find(RcaliW(:,6)<DIVw);

RcaliW(highREFw,5)=1;

RcaliW(lowREFw,5)=-1;

dataW=RcaliW(:,1:2);

group=svmclassify(svmStruct,RcaliW(:,1:2));

G1=find(group(:,1)==1);

G2=find(group(:,1)==-1);

highCheck=zeros((openData(:,2)>9 | openData(:,2)<4) &
openData(:,13)<32);

% T2 winter adjustment for entire domain

prevE=10000000000000;

for i=0.1:0.1:20

    for j=10:10:1000

```

```

rainRate=((10.^(openData(winterT2,13)./10))/j).^(1/i))./4;
error=sum(abs(rainRate-openData(winterT2,11)));
if error<prevE
    a=j;
    b=i;
    %XX=[a,b];
    %disp(XX);
    prevE=error;
end
end
end
RcaliW=zeros(lenT2winter,7);
RcaliW(:,1)=openData(winterT2,11);
RcaliW(:,2)=openData(winterT2,12);
RcaliW(:,3)=(((10.^(openData(winterT2,13)./10))/a).^(1/b))./4;
RcaliW(:,4)=openData(winterT2,9);
RcaliW(:,6)=openData(winterT2,13);
%support vector classifier winter
minWinterREF=min(RcaliW(:,6));
maxWinterREF=max(RcaliW(:,6));
prevTotalE=1000000000;
RerrorW=zeros(36,3);
DIVw=32;
highREFw=find(RcaliW(:,6)>=DIVw);
lowREFw=find(RcaliW(:,6)<DIVw);
RcaliW(highREFw,5)=1;
RcaliW(lowREFw,5)=-1;
dataW=RcaliW(:,1:2);
svmStruct = svmtrain(dataW,RcaliW(:,5));
group=svmclassify(svmStruct,RcaliW(:,1:2));
G1=find(group(:,1)==1);
G2=find(group(:,1)==-1);
% scatter plot of first half water year
figure;
subplot(2,2,1);
svmStruct = svmtrain(dataW,RcaliW(:,5),'ShowPlot',true);
box on;

```

```

grid on;
xlim([0 30]);
ylim([0 30]);
high1stREF=zeros(length(G1),3);
low1stREF=zeros(length(G2),3);
subplot(2,2,2);
high1stREF(:,1)=RcaliW(G1,1);
high1stREF(:,2)=RcaliW(G1,6);
high1stREF(:,3)=(((10.^(high1stREF(:,2)./10))/50).^(1/1.1))./4;
low1stREF(:,1)=RcaliW(G2,1);
low1stREF(:,2)=RcaliW(G2,6);
low1stREF(:,3)=(((10.^(low1stREF(:,2)./10))/230).^(1/0.5))./4;
high1st=scatter(high1stREF(:,1),high1stREF(:,3),30,'d','filled','red');
box on;
grid on;
xlim([0 30]);
ylim([0 30]);
hold on;
low1st=scatter(low1stREF(:,1),low1stREF(:,3),30,'d','filled','blue');
hold on;
high1stLL=lsline;
set(high1stLL,'Color','black');
set(high1stLL,'LineWidth',2);
for i=31.1:0.1:34.6
    DIV=i;
    highREFw=find(RcaliW(:,6)>=DIV);
    lowREFw=find(RcaliW(:,6)<DIV);
    RcaliW(highREFw,5)=1;
    RcaliW(lowREFw,5)=-1;
    dataW=RcaliW(:,1:2);
    svmStruct = svmtrain(dataW,RcaliW(:,5));
    group=svmclassify(svmStruct,RcaliW(:,1:2));
    G1=find(group(:,1)==1);
    G2=find(group(:,1)==-1);
    % find the best G1 for each group
    prevE1=1000000000000;

```



```

for j=0.1:0.1:20
    for k=10:10:1000
        rainRate=((10.^(RcaliW(G1,6)/10))/k).^(1/j))./4;
        error=sum(abs(rainRate-RcaliW(G1,1)));
        if error<prevE1
            a1=k;
            b1=j;
            %XX=[a1,b1];
            %disp(XX);
            prevE1=error;
        end
    end
end
% find the best G2 for each group
prevE2=1000000000000;
for j=0.1:0.1:20
    for k=10:10:1000
        rainRate=((10.^(RcaliW(G2,6)/10))/k).^(1/j))./4;
        error=sum(abs(rainRate-RcaliW(G2,1)));
        if error<prevE2
            a2=k;
            b2=j;
            %XX=[a2,b2];
            %disp(XX);
            prevE2=error;
        end
    end
end
totalE=prevE1+prevE2;
ErrorW(numIteration,1)=i;
ErrorW(numIteration,2)=totalE/lenT2winter;

ErrorW(numIteration,3)=length(G1)/(length(G1)+length(G2));
disp(numIteration);
numIteration=numIteration+1;
end
minErrorW=min(ErrorW(:,2));
rowMinW=find(ErrorW(:,2)==minErrorW);

optREFw=ErrorW(rowMinW,1);
% T2 winter adjustment for entire domain
prevE=1000000000000;
for i=0.1:0.1:20
    for j=10:10:1000
        rainRate=((10.^(openData(summerT2,13)/10))/j).^(1/i))./4;
        error=sum(abs(rainRate-openData(summerT2,11)));
        if error<prevE
            a=j;
            b=i;
            %XX=[a,b];
            %disp(XX);
            prevE=error;
        end
    end
end
RcaliS=zeros(lenT2summer,7);
RcaliS(:,1)=openData(summerT2,11);
RcaliS(:,2)=openData(summerT2,12);
RcaliS(:,3)=(((10.^(openData(summerT2,13)/10))/a).^(1/b))./4;
RcaliS(:,4)=openData(summerT2,9);
RcaliS(:,6)=openData(summerT2,13);
%support vector classifier summer
minSummerREF=min(RcaliS(:,6));
maxSummerREF=max(RcaliS(:,6));
prevTotalE=1000000000;
ErrorS=zeros(36,3);
numIteration=1;
for i=31.1:0.1:34.6
    DIV=i;
    highREFs=find(RcaliS(:,6)>=DIV);
    lowREFs=find(RcaliS(:,6)<DIV);
    RcaliS(highREFs,5)=1;
    RcaliS(lowREFs,5)=-1;
    dataS=RcaliS(:,1:2);
    svmStruct = svmtrain(dataS,RcaliS(:,5));
    group=svmclassify(svmStruct,RcaliS(:,1:2));

```

```

G1=find(group(:,1)==1);
G2=find(group(:,1)==-1);
% find the best G1 for each group
prevE1=1000000000000;
for j=0.1:0.1:20
    for k=10:10:1000
        rainRate=((10.^(RcaliS(G1,6)./10))/k).^(1/j))./4;
        error=sum(abs(rainRate-RcaliS(G1,1)));
        if error<prevE1
            a1=k;
            b1=j;
            %XX=[a1,b1];
            %disp(XX);
            prevE1=error;
        else
            continue;
        end
    end
end
% find the best G2 for each group
prevE2=1000000000000;
for j=0.1:0.1:20
    for k=10:10:1000
        rainRate=((10.^(RcaliS(G2,6)./10))/k).^(1/j))./4;
        error=sum(abs(rainRate-RcaliS(G2,1)));
        if error<prevE2
            a2=k;
            b2=j;
            %XX=[a2,b2];
            %disp(XX);
            prevE2=error;
        end
    end
end
totalE=prevE1+prevE2;
ErrorS(numIteration,1)=i;
ErrorS(numIteration,2)=totalE/lenT2summer;

ErrorS(numIteration,3)=length(G1)/(length(G1)+length(G2));
disp(numIteration);
numIteration=numIteration+1;
end
minErrorS=min(ErrorS(:,2));
rowMinS=find(ErrorS(:,2)==minErrorS);
optREFs=ErrorW(rowMinS,1);
% figure for critical reflectivity
maxErrorW=max(ErrorW(:,2));
normErrorW=ErrorW(:,2)./maxErrorW;
maxErrorS=max(ErrorS(:,2));
normErrorS=ErrorW(:,2)./maxErrorS;
lenS=length(ErrorS(:,3));
ErrorS(lenS,3)=ErrorS((lenS-1),3);
figure;
subplot(1,2,1);
[refW,RMSEw,ratioW]=plotyy(ErrorW(:,1),ErrorW(:,2),R
errorW(:,1),ErrorW(:,3),'plot');
set(RMSEw,'LineWidth',2);
set(ratioW,'LineWidth',2);
box on;
grid on;
hold on;
lineW=plot([32,32,32,32,32],1.8:0.2:2.6,'black');
set(lineW,'LineWidth',3);
hold off;
t=title('First Half Water Year');
set(t,'FontWeight','bold');
set(t,'FontSize',14);
set(t,'FontName','Times');
xt=xlabel('Radar Reflectivity in dBZ');
set(xt,'FontSize',12);
set(xt,'FontName','Times');
yt1=ylabel(refW(1),'Averaged Minimum RMSE in
mm/15min');
set(yt1,'FontSize',12);
set(yt1,'FontName','Times');

```

```

yt2=ylabel(refW(2),'Occupancy Ratio for Strong Reflectivity
Group');

set(yt2,'FontSize',12);

set(yt2,'FontName','Times');

subplot(1,2,2);

[refS,RMSEs,ratioS]=plotyy(ErrorS(:,1),ErrorS(:,2),Error
S(:,1),ErrorS(:,3),'plot');

set(RMSEs,'LineWidth',2);

set(ratioS,'LineWidth',2);

ylim([2.6 2.95]);

box on;

grid on;

hold on;

lineS=plot([31.4,31.4,31.4,31.4,31.4,31.4,31.4,31.4],2.6:0.05:
2.95,'black');

set(lineS,'LineWidth',3);

hold off;

t=title('Second Half Water Year');

set(t,'FontWeight','bold');

set(t,'FontSize',14);

set(t,'FontName','Times');

xt=xlabel('Radar Reflectivity in dBZ');

set(xt,'FontSize',12);

set(xt,'FontName','Times');

yt1=ylabel(refS(1),'Averaged Minimum RMSE in
mm/15min');

set(yt1,'FontSize',12);

set(yt1,'FontName','Times');

yt2=ylabel(refS(2),'Occupancy Ratio for Strong Reflectivity
Group');

set(yt2,'FontSize',12);

set(yt2,'FontName','Times');

lt=legend('Averaged minimum RMSE','Group
Ratio','Optimised Reflectivity');

set(lt,'FontSize',12);

set(lt,'FontName','Times');

subplot(1,2,2);

pOptS=plot(ErrorS(:,1),ErrorS(:,2),ErrorS(:,1),ErrorS(:,
3),'plot');

box on;

grid on;

set(pOptW,'LineWidth',3);

hold on;

plot(115:145,'black');

box on;

grid on;

subplot(1,2,2);

pOptS=plot(ErrorS(:,1),ErrorS(:,2),'red');

set(pOptS,'LineWidth',3);

box on;

grid on;

figure;

subplot(1,2,1);

plot(0:30,0:30,'black');

box on;

grid on;

xlim([0 30]);

ylim([0 30]);

hold on;

svmStruct = svmtrain(dataW,RcaliW(:,5),'ShowPlot',true);

hold off;

subplot(1,2,2);

plot(0:30,0:30,'black');

hold on;

T2winLow=scatter(RcaliW(G1,1),Rlow(:,2),30,'d','filled','blu
e');

hold on;

T2winHigh=scatter(Rhigh(:,1),Rhigh(:,2),30,'d','filled','red');

hold on;

f=fit(conSVMadjust(:,1),conSVMadjust(:,2),'poly2');

plot(f);

xlim([0 20]);

ylim([0 20]);

box on;

grid on;

hold off;

subplot(2,4,6);

plot(0:30,0:30,'black');

```

```

box on;

grid on;

xlim([0 30]);

ylim([0 30]);

hold on;

T2adjust=scatter(RcaliS(:,1),RcaliS(:,3),25,'d','filled','black');

f=fit(RcaliS(:,1),RcaliS(:,3),'poly2');

hold on;

plot(f);

hold off;

t=title('Adjusted Summer Tropical');

lt=legend('1:1 line','data','fitted curve');

set(t,'FontWeight','bold');

set(t,'FontSize',14);

set(t,'FontName','Times');

set(lt,'FontSize',12);

set(lt,'FontName','Times');

xt=xlabel('Gauge Rainfall in mm/15min');

set(xt,'FontSize',12);

set(xt,'FontName','Times');


%conSVMadjust=vertcat(G1,G2);

subplot(3,6,numFig);

box on;

grid on;

xlim([0 30]);

ylim([0 30]);

hold on;

plot(0:30,0:30,'black');

hold off;

svmStruct.SupportVectors

C=-svmStruct.Bias/w2;

w=-(w1/w2);

% or with line given as y = a*x + b

% using support vector machine....adjustment

for i=1:lenT2winter

    NWS=Rcali(i,2);

    HP=(0.1)*Rcali(i,1)+3.9;

    if NWS>=HP

        Rcali(i,7)=1;

    else

        Rcali(i,7)=-1;

    end

end

highREF=find(Rcali(:,7)==1);

lowREF=find(Rcali(:,7)==-1);

prevE=1000000000000000;

for i=0.1:0.1:20

    for j=10:10:1000

        %rainRate=(((10.^(Rcali(highREF,6)./10))/j).^(1/i))./4;

        %error=sum(abs(rainRate-Rcali(highREF,1)));

        rainRate=(((10.^(Rcali(lowREF,6)./10))/j).^(1/i))./4;

        error=sum(abs(rainRate-Rcali(lowREF,1)));

        if error<prevE

            a=j;

            b=i;

            XX=[a,b];

            disp(XX);

            prevE=error;

        end

    end

end

Rhigh=zeros(length(highREF),2);

Rlow=zeros(length(lowREF),2);

Rhigh(:,1)=Rcali(highREF,1);

Rhigh(:,2)=(((10.^(Rcali(highREF,6)./10))/50).^(1/1.1))./4;

Rlow(:,1)=Rcali(lowREF,1);

Rlow(:,2)=(((10.^(Rcali(lowREF,6)./10))/230).^(1/0.5))./4;

% concatenating two

conSVMadjust=vertcat(Rhigh,Rlow);

subplot(2,2,4);

plot(0:20,0:20,'black');

hold on;

T2winLow=scatter(Rlow(:,1),Rlow(:,2),30,'d','filled','blue');

hold on;

T2winHigh=scatter(Rhigh(:,1),Rhigh(:,2),30,'d','filled','red');

```

```

hold on;
f=fit(conSVMadjust(:,1),conSVMadjust(:,2),'poly2');
plot(f);
xlim([0 20]);
ylim([0 20]);
box on;
grid on;
hold off;
end
svmStruct = svmtrain(dataW,RcaliW(:,5),'ShowPlot',true);
hold on;
f=fit(dataW(:,1),dataW(:,2),'poly2');
plot(f);
xlim([0 20]);
ylim([0 20]);
box on;
grid on;
hold off;
t=title('SVM Classification by Reflectivity');
xt=xlabel('Gauge Rainfall in mm/15min');
yt=ylabel('Radar Rainfall in mm/15min');
lt=legend('weak reflectivity','strong reflectivity','Support Vector');
set(t,'FontWeight','bold');
set(t,'FontSize',14);
set(xt,'FontSize',12);
set(yt,'FontSize',12);
set(lt,'FontSize',12);
set(t,'FontName','Times');
set(xt,'FontName','Times');
set(yt,'FontName','Times');
set(lt,'FontName','Times');
% find a hyperplane line
% getting parameters for equation separation line from svmStruct
%  $w_1*x + w_2*y + bias = 0$ 
t=title('SVM Adjusted Winter Tropical');
xt=xlabel('Gauge Rainfall in mm/15min');
lt=legend('1:1 line','weak reflectivity','strong reflectivity','fitted curve');
set(t,'FontWeight','bold');
set(t,'FontSize',14);
set(xt,'FontSize',12);
set(lt,'FontSize',12);
set(t,'FontName','Times');
set(xt,'FontName','Times');
set(lt,'FontName','Times');
highREFs=find(RcaliS(:,6)>=32);
lowREFs=find(RcaliS(:,6)<32);
RcaliS(highREFs,5)=1;
RcaliS(lowREFs,5)=-1;
dataS=RcaliS(:,1:2);
% Create textbox
annotation('figure1','textbox',...
    [0.76087951807227 0.149655172413792
    0.120201563008811 0.0280172413793102],...
    'String',{'a=230, b=0.5'},...
    'FontWeight','bold',...
    'FontSize',12,...
    'FontName','Times New Roman',...
    'FitBoxToText','off',...
    'Color',[0 0 1]);
% open reflectivity
openConv=dlmread('Z:/DATA/KLVXoutput/ZRconvective.out',';');
%(RAIN)^(1.4)=(10^(REF/10))/300; % standard convective_CSO130 in mm/5min
lenScan=length(openConv);
format long;
% julian dates of selected gauge data
RjdG=zeros(lenData,2);
QuarterHr=1/24/4;
for i=1:lenData
    YYYY=openData(i,1);
    MO=openData(i,2);
    DD=openData(i,3);
    HH=openData(i,4);
    MM=openData(i,5);

```

```

JD=datetime(YYYY,MO,DD,HH,MM,0);
RjdG(i,1)=JD-QuarterHr;
RjdG(i,2)=JD;
disp(i);
end
% julian date of KLVX scan
RscanJD=zeros(lenScan,1);
for i=1:lenScan
    YYYY=openConv(i,1);
    MO=openConv(i,2);
    DD=openConv(i,3);
    HH=openConv(i,4);
    MM=openConv(i,5);
    SS=openConv(i,6);
    JD=datetime(YYYY,MO,DD,HH,MM,SS);
    RscanJD(i,1)=JD;
    disp(i);
end
% selection of weighted reflectivity
Rann=zeros(lenData,11);
Rann(:,1:6)=openData(:,1:6);
qtHrRain=openData(:,7);
Rann(:,7)=qtHrRain;

conv=find(openData(:,10)==1);
trop=find(openData(:,10)==2);
ecs=find(openData(:,10)==3);
strat=find(openData(:,10)==4);
Rann(conv,11)=1;
Rann(trop,11)=2;
Rann(ecs,11)=3;
Rann(strat,11)=4;
for i=1:lenData
    jdStart=RjdG(i,1);
    jdEnd=RjdG(i,2);
    TR=Rann(i,6);

    listScan=find(RscanJD(:,1)>jdStart &
RscanJD(:,1)<=jdEnd);

```

```

Rlist=zeros(length(listScan),5);
Rlist(:,1)=RscanJD(listScan,1); % julian data of scans
Rlist(:,2)=openConv(listScan,(6+TR)); % conv rain rate
scans
NAList=find(Rlist(:,2)==-999 | Rlist(:,2)>=500);
if length(NAList)>0
    Rlist(NAList,:)=[];
end
numScan=length(Rlist(:,1));
for j=1:numScan
    ref=10*log10(300*Rlist(j,2)^(1.4));
    Rlist(j,3)=ref; % reflectivity
    if numScan==1
        Rlist(j,4)=1; % weight
        prevInt=(Rlist(1,1)-jdStart)*24*60;
        postInt=(jdEnd-Rlist(1,1))*24*60;
        maxInt=max(prevInt, postInt); % max interval in the
15minute capsule
        Rlist(j,5)=maxInt;
    elseif numScan>1
        if j==1
            prevScan=(Rlist(j,1)-jdStart)*24*60;
            postScan=(Rlist((j+1),1)-(Rlist(j,1)))*24*60/2;
            scanRatio=(prevScan+postScan)/15;
            Rlist(j,4)=scanRatio;
            Rlist(j,5)=prevScan+postScan;
        elseif j==numScan
            prevScan=(Rlist(j,1)-Rlist((j-1),1))*24*60/2;
            postScan=(jdEnd-Rlist(j,1))*24*60;
            scanRatio=(prevScan+postScan)/15;
            Rlist(j,4)=scanRatio;
            Rlist(j,5)=prevScan+postScan;
        else
            prevScan=(Rlist(j,1)-Rlist((j-1),1))*24*60/2;
            postScan=(Rlist((j+1),1)-Rlist(j,1))*24*60/2;
            scanRatio=(prevScan+postScan)/15;
            Rlist(j,4)=scanRatio;
            Rlist(j,5)=prevScan+postScan;
        end
    end
end

```

```

end
end
if numScan==0;
    Rann(i,8:10)=-999;
else
    Rann(i,8)=numScan;
    Rann(i,9)=max(Rlist(:,5));
    ratioRef=sum(Rlist(:,3). *Rlist(:,4));
    Rann(i,10)=ratioRef;
    disp(i);
end
end
% check Rann
ndANN=find(Rann(:,10)==-999);
Rann(ndANN,:)=[];
figure;
%scatter(Rann(:,7),Rann(:,10),5,'filled','black');
subplot(1,2,1);
hist(Rann(:,10),12);
subplot(1,2,2);
hist(Rann(:,7),12);
dlmwrite('Z:/DATA/KLVXoutput/pair4ANN.out',Rann,'delimiter',' ','precision','%10.6f');
% bar plot of rainfall type
figure;
bar(Rvals4(:,1),'black');
xlim([0 1000]);
hold on;
bar(Rvals4(:,2),'red');
xlim([0 1000]);
hold on;
bar(Rvals4(:,3));
xlim([0 1000]);
hold on;
bar(Rvals4(:,4),'yellow');
xlim([0 1000]);
hold off;
t=title('Rainfall Type Variaton');

xt=xlabel('sequence');
yt=ylabel('rainfall accumulation in mm/15min');
lt=legend('convective','tropical','east cool stratiform','stratiform');
set(t,'FontWeight','bold');
set(t,'FontSize',14);
set(xt,'FontSize',12);
set(yt,'FontSize',12);
set(lt,'FontSize',12);
set(t,'FontName','Times');
set(xt,'FontName','Times');
set(yt,'FontName','Times');
set(lt,'FontName','Times');
Chapter5
a. Generate matched Hythograph and CSO130 overflow Hydrograph
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% Generate Event Hydrograph
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% input: R
% 1-5: time information - year,month,day,hour,minute
% 6:TR3, gauge Accumulation quarter hour in mm (15.3km away from study site)
% 7:TR5, gauge Accumulation quarter hour in mm (0.8km away from study site)
% 8:TR9, gauge Accumulation quarter hour in mm - No use
% 9:TR12, gauge Accumulation quarter hour in mm (4.9 away from study site)
% 10: total available scans of KLVX
% 11: max interval within 15min capsule in min
% 12: NWS type of rainfall - - 1(convective), 2(tropical), 3(ECS), 4(stratiform)
% 13: NWS KLVX radar accumulation mm/15min (quarter hour)
% 14: SVC type of rainfall - - 1(convective), 2(tropical), 3(ECS), 4(stratiform)
% 15: SVC KLVX radar accumulation mm/15min (quarter hour)
% 16: CSO overflow in MGD

```



```

rainBar=bar(Rcouple(:,1));
xlim([0 (lenCouple+1)]);
ylim([0 15]);
grid on;
box on;
set(rainBar,'FaceColor',[1 1
1],'EdgeColor','blue','barWidth',0.7);
hold on;
flowBar=bar(Rcouple(:,2),'red');
xlim([0 (lenCouple+1)]);
ylim([0 15]);
set(flowBar,'EdgeColor','red','barWidth',0.4);
hold off;
totalOverflow=openInfo(i,18);
titleName=strcat(num2str(i),',' , 'Overflow amount:
',num2str(totalOverflow),'mm');
t=title(titleName);
set(t,'FontWeight','bold');
set(t,'FontSize',12);
set(t,'FontName','Times');
if i==4
    yt=ylabel('rainfall/overflow amount in mm');
    set(yt,'FontSize',14);
    set(yt,'FontName','Times');
elseif i==8
    xt=xlabel('rainfall event timeflow in minute');
    set(xt,'FontSize',14);
    set(xt,'FontName','Times');
elseif i==9
    lt=legend('rainfall event','CSO130 overflow');
    set(lt,'FontSize',13);
    set(lt,'FontName','Times');
end
end
meanConn=mean(openInfo(:,21));

```

CURRICUM VITAE

Jin-Young Hyun

PhD

Civil and Environmental Engineering

Speed School of Engineering

University of Louisville, Louisville, KY 40208

Mobile: (319) 855-2270, Email: sojibi78@gmail.com; j0hyun01@cardmail.louisville.edu

QUALIFICATION

- Strong expertise in analysis of extreme hydrologic field such as flood frequency and magnitude evaluation and prediction in urban or sub-urban complex areas
- In depth knowledge of remote sensing – radar, satellite rainfall estimation and their validation with ground based rainfall monitoring and urban application for hydrologic modeling – hydro informatics.
- Strong expertise in storm water analysis/design for urban hydrologic structures
- In depth knowledge hydro-meteorological field and climate change
- In-depth knowledge of statistical methods and evaluation for data analysis (stochastic statistics, supervised or unsupervised clustering, artificial neural network, support vector machine)
- Hands on experience in implementing data mining models for big data analysis
- In-depth knowledge of advanced spatial and temporal statistics
- Analysis of sustainable building energy efficiency and green design
- Strong expertise in computational and statistical modeling and data visualization for in-depth system analysis

EDUCATION

University of Louisville- Louisville, KY

Ph.D. (finished all requirement to graduate; passed final oral presentation on September 29th)

August 2012- December 2016

- Civil Engineering with emphasis on Water Resource Engineering
 - Dissertation Title: Spatiotemporal rainfall structure and optimized local radar rainfall application to urban watershed, Louisville, Kentucky, 2010-2014
 - Advisor: Professor Mark N. French

University of Iowa- Iowa City, IA

M.S., Department of Civil Engineering

January 2010- May 2012 Water Resource Engineering (remote sensing and climatology) emphasis

- Advisor: Professor Witold F. Krajewski
- Major Contribution: Water resource data assimilation from rain gauge, weather radar and satellite

University of Kentucky- Lexington, KY

B.S., Department of Civil Engineering

August 2008- December 2010

- Water Resource Engineering emphasis

University of Louisville- Louisville, KY

B.S., Department of Civil Engineering

August 2006- August 2008

- General Civil and Environmental Engineering

PUBLICATIONS

Peer-Reviewed Published Journal Papers

- [1] Jin-young Hyun, Thomas D. Rockaway, and Mark N. French. (2016). "Ground-Level Rainfall Variation in Jefferson County, Kentucky." *Journal of Hydrologic Engineering – American Society of Civil Engineering (ASCE)* DOI: 10.1061/(ASCE)HE.1943-5584.0001438
- [2] Jonghoon Kim, Jin-young Hyun, Wai K. O. Chong and Samuel T. Ariaratnam. (2016). "Understanding the Effects of External Factors on Building Energy Efficiency Designs and Credits – Case Studies using Data mining and Real-time Data." *Journal of Engineering, Design and Technology*. DOI: 10.1108/JEDT-12-2015-0082

Peer-Reviewed Journal Papers In-Revision

- [1] Jin-young Hyun, Thomas D. Rockaway, and Mark N. French. "Rainfall event characterization with cluster and variogram analysis" *Journal of Hydrometeorology – American Meteorological Society (AMS)*, October 19th, 2016
- [2] Jin-young Hyun, Thomas D. Rockaway, and Mark N. French. "Urban sewershed overflow analysis using super-resolution weather radar rainfall" *Hydrology and Earth System Sciences (HESS)*, November 16 2016, DOI: 10.5194/hess-2016-362

Peer-Reviewed Journal Papers In-Preparation

- [1] Jin-young Hyun, Thomas D. Rockaway, and Mark N. French. "Multi-radar merging and optimization in micro/meso-scale urban area" as a Technical Paper, plan to submit to *Journal of Hydrology-Elsevier*
 - [2] Jin-young Hyun, Thomas D. Rockaway, and Mark N. French. "Rainfall event encapsulation by using DBSCAN" as a Technical Paper, plan to submit to *Journal of Hydroinformatics*
 - [3] Jin-young Hyun, Thomas D. Rockaway, and Mark N. French. "Downscaling of spatial analysis of rainfall event in urban area" as a Technical Paper, plan to submit *Journal of Atmospheric and Oceanic Technology, AMS*
-

- [4] Milad Ibrahimi, Jin-Young Hyun, Thomas D. Rockaway. “Quantitative approach on the salinity of the aquifer in Shiraz, Iran”

PEER-REVIEWER OF JOURNALS

-
- [1] *Journal of Hydrometeorology, AMS* (2016-present) requested by Dr. Steven A. Margulis (Editor)
- [2] *Journal of Atmospheric and Oceanic Technology, AMS* (2016-Present) requested by Dr. Willian Emery (Editor-in-Chief)
- [3] *Journal of Hydrologic Engineering - ASCE* (2016-Present) requested by editors who serve for Journal of Hydrologic Engineering – ASCE

RESEARCH EXPERIENCE

University of Louisville, Louisville, KY

(August 2013- Present)

Graduate Research Assistant

- Research in Progress
 - Incentives for green infrastructure on private property: lessons learned : writing proposal: in proposal status.
- Completed Research:
 - Green Infrastructure Validation using Super-Resolution Radar Rainfall for Combined Sewer Overflow (CSO130) area, Louisville Metropolitan (2013-current): Development for modeling of optimization of the merged radar and gauge data product using data mining method, design of database for super resolution radar rainfall, KLVX and MSD rain gauge networks, spatial analysis of rain gauge variation over Louisville Metro., Evaluation Green Infrastructure of storm water mitigation
 - Validation of the Satellite Rainfall Estimation in Kentucky, Tennessee, West Virginia area, 2003-2012: validation performed using radar rainfall estimation, pattern recognition over the area, develop supervised/unsupervised data analysis algorithms, developing ensemble algorithms for optimization of the satellite based products
 - Relationship between LEED building energy efficiency and atmospheric temperature (2014-2015): building energy data preprocessing –filtering noises, statistical analysis to evaluate LEED building scores

TEACHING EXPERIENCE

University of Louisville- Louisville, KY

Teaching Assistant

Mechanics I: Statics (CEE 205)

- Role: Teaching Assistant
 - Date Taught: Fall 2015, Spring 2016
-

- Instructor: Dr. Mark N. French
- Course Length: 3 hours per week, 16 weeks
- Class Size: 50 undergraduate
- Topic Covered: examine forces, equilibrium, frictions, centroids and moment of inertia .

Guest Speaker

Water Supply and Sewerage (CEE 470)

- Role: Guest Lecturer
- Date Taught: March 21st – Apr 1st , 2016
- Instructor: Dr. Nageshwar Bhaskar
- Course Length: 75 minute
- Class Size: 25 senior undergraduate students
- Topic Covered: Storm water management.

University of Iowa- Iowa City , IA

Teaching Assistant

Fluid Mechanics I: Hydraulics (ENGR 2510)

- Role: Instructor of Hydraulic Experimental Lab
- Date Taught: Fall 2010, Spring 2011 and Spring 2012
- Instructor: Dr. Jacob Odgaard
- Course Length: 150 minute
- Class Size: 8 lab groups (8-10 junior undergraduate students per group)
- Topic Covered: Falling velocity and viscosity, pipe flow (closed loop), air foil momentum loss/lifting force in air tunnel, particle image velocimetry (PIV).

WORK EXPERIENCE

University of Louisville, Louisville, KY

(August 2012-current)

Graduate Researcher

- Evaluation of Green Infrastructure which mitigates combined storm/sewer water overflow.
 - Assisting the City of Louisville meet their compliance reporting obligations for their federally mandated \$850 million stormwater overflow consent decree
 - Project Investigator: Dr. Thomas D. Rockaway, Director of Center for Infrastructure Research
 - Assimilation of rain gauge network operated by Metropolitan Sewer District (MSD), Louisville, and local weather radar, National Weather Service
- Satellite weather forecasting over State of Kentucky (multi-channel of geostationary and passive microwave satellites)

University of Iowa, Iowa City, IA

(January 2010- Spring 2012)

Graduate Researcher in Iowa Institute of Hydro-science Research

- Analyze, validate, and predicts rainfall estimation

- Coding intensive data retrieving, preprocessing of multiple rainfall data from rain gauge networks, weather radar and satellite

Researcher in Iowa Flood Center

- Data collection and spatial statistics of the spread of the hydro-stations
- Coding intensive data retrieving, preprocessing of multiple rainfall data from rain gauge networks, weather radar and satellite
- Restatement of FEMA flood manual for state of Iowa
- Index the hydro information form stations over Iowa

Chungigy-Corporation, Kyunggi-do, South Korea

(September 2000- March 2003)

- Project manager in advertising and human resource
- Manager of web based advertising and public relation with customers and government
- Survey strategist of web or letter based customer feedback
- Human resource

Korean Army Corps of Engineer- Kangwon-Do, South Korea

(July 1998- September 2000)

Assistant Field Engineer/Surveyor

- Involved in construction and surveying activities for various military facilities.
- Performed surveying and mapping include utility location using the total station, levels, and office equipment as well as mapping and plotting.
- Infantry and administrative clerk

TECHNICAL SKILLS

- MATLAB: Expert
- ArcGIS (spatial statistics): Expert
- SPSS: Expert
- R-Statistics: Expert
- SAS: Moderate
- MS-EXCEL POWERPIVOT and POWERVIEW: Expert
- C# using .NET framework: Moderate
- C, C++: Expert
- Transactional SQL: Moderate

LANGUAGE SKILLS

- English and Korean

PROFESSIONAL SOCIETY MEMBERSHIP

- Member of American Society of Civil Engineers, ASCE, USA (2016-Present)

- Member of Environmental and Water Resources Institute, EWRI, USA (2016-Present)

HONORS AND AWARDS

- University of Louisville – Doctoral Dissertation Completion Award at Graduate School, University of Louisville for Fall 2016
- University of Louisville - Spotlight Graduate Researcher of the month, October 2015
- University of Louisville - Speed School Fellowship for the Academic Year 2012-2013
- University of Louisville - Speed School Fellowship for the Academic Year 2013-2014
- University of Kentucky - Fall Honor of Deans List
- University of Louisville - Vice President of Korean Student Association